

Can learning analytics provide useful insights? An exploration on course level

Eva Heinrich

Massey University
New Zealand

This concise paper reports on an analysis of access logs of a first year university course that was delivered in a blended format. This analysis is an initial step in a wider project aimed at investigating if learning analytics can provide useful insights on course level, targeting both student learning and the needs of teachers. Preliminary findings show potential in noting when students need targeted help, a lack of correlation between access logs and grades, and insights into the degree by which course completion rates are affected by the lack of student engagement.

Keywords: Learning analytics, first year courses, blended learning.

Introduction

In 2011 Long and Siemens wrote about the potential of learning analytics to “improve the quality and value of the learning experience” (p.40). Since then learning analytics research has been published in a special edition of the *Journal of Educational Technology & Society* and in a dedicated journal published by the Society for Learning Analytics Research. Valuable contributions have been made, yet, it seems that learning analytics is not ready to make widespread contributions in the sense of providing teachers with solutions, conceptually and technically, that are easy to apply, in terms of knowledge and time requirements, and provide meaningful insights. The 2014 review of learning analytics research by Papamitsiou and Economides confirms the potential but also alerts to a considerable number of gaps. While research articles highlight strength in approaches, it is fairly easy to spot the missing links. As there is no room for a comprehensive review in this concise paper only a few examples are given. Sophisticated statistical methods for preparing and analyzing log data are used by Andergassen et al. (2014), yet the results are not correlated with the wider learning context and no insights are provided on how to use the results to assist student learning. Promising results are reported for efficiently identifying students with poor evaluation skills (Gunnarsson et al., 2014), but a specialized blogging tool is required. The eLAT (exploratory Learning Analytics Toolkit) described by Dyckhoff et al. (2012) looks encouraging in allowing teachers to analyze teaching data presented in graphical form. Important considerations, such as not requiring specialist data mining knowledge, have been made, yet it remains to be seen if the relative simplicity of the indicators presented will be powerful enough.

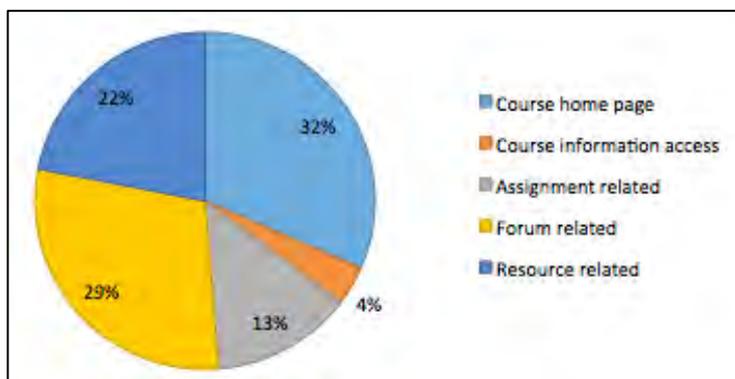
The research reported in this concise paper represents a first exploratory step in a wider research effort examining the potential of learning analytics in the context of first year university courses delivered in a blended format. The specific challenges of teaching first year courses are well documented (see, for example, Zepke, 2013). First year students are diverse in terms of readiness for study and subject knowledge. Engagement and retention are common problems. In this context the research presented here targets conventional applications of learning analytics around assisting teachers in helping students to learn better and in detecting problems early. Yet, the researchers are also interested in aspects less commonly emphasized. The current climate of higher education favors research outputs and makes even dedicated teachers question the wisdom of investing large amounts of time into preparing sophisticated teaching material (Ginns et al., 2010). Can learning analytics confirm the value of such material for student learning and provide re-assurance to teachers that their efforts are worthwhile? Governments (most certainly in New Zealand) set increasingly high completion rate targets. Institutions pass on the pressure to achieve these completion rates to teachers. These teachers work hard to provide learning opportunities to their students, but ultimately have no control over uptake. Learning analytics can help teachers document the level of engagement (or the lack of thereof) by students. In the longer run such data might be required to demonstrate the problem of widening the participation in higher education without higher resourcing but increasing completion rate expectations while supposedly not dropping standards.

This exploratory research step focused on a first year information technology course taught in a blended mode supported by the Learning Management System (LMS) Moodle and the multimedia-streaming platform Mediasite. The teacher who developed and taught the course is experienced in first year teaching, blended approaches and the use of technology for teaching. The course was delivered to one cohort of 59 students in the first semester of 2015 at a New Zealand university. The data analyzed were LMS records (access logs, marks for internal assessments), streaming records, exam marks and final grades. Data on student demographics were not accessed.

Areas investigated

Level of course participation by students

In the blended course design access to the LMS was vital for full participation in the course. The LMS provided access to organizational information, topic introductions, lecture slides, links to screencasts and assessment specifications. Discussion forum contributions were required and two assignments had to be submitted via the LMS. Access records stored by the LMS can therefore be regarded as indicators for the degree of course participation. In total 31,226 access logs were recorded by the LMS (up to the hour of the final exam for the course). At the end of the course, the official classroll stood at 59 students. Of those, five students withdrew during the semester and six did not complete the assessment requirements (they did not sit the final exam), leaving 48 students who received a grade for the course. Records of the non-graded students, as well as records of additional five students who started the course but withdrew very early in the semester formed part of the LMS access logs. Looking at graded students only, the average number of access logs per student was 520, the minimum number 147 and the maximum number 1,321. Figure 1 shows the distribution of log entries across course areas: Access to course homepage; access to information about the course (e.g., teaching team, contact times, assessment details); access to assignment specifications, submission and feedback (as relevant as part of the internal assessment occurred face-to-face); access to forum discussions (reading could also occur via email digests); access to resources (text entered directly into the LMS, links to PDF lecture slides; links to streamed video recordings). The differences in how students interacted with the course site can be exemplified by looking at the proportion of access to the course home page to total access. Across all students this proportion was 32%, the minimum 11% and the maximum 48%. A lower proportion indicates that a student has done more work on the course site per session than a student with a higher proportion (as an entry for the



home page would be triggered in most cases when starting a new session). Without additional data and much deeper investigation it is not possible to say how the different ways of interacting might be related to learning.

Figure 1: Distribution of access logs

The five students who started the course but withdrew within the first two weeks did so without financial or academic penalties and are of little concern to the teacher. Of the eleven students who remained on the classroll but did not complete all course requirements, three do not seem to have made any effort to study the course. These students had zero, eight and eleven log entries over a very short timespan. It is unlikely that the course design was a major factor in those students' behaviour. The remaining eight students had on average 147 log entries each over up to the first eight weeks of the semester. Only one of those students attempted one of the internal assessment components. Monitoring of the access logs during the semester might have been able to pick up that these students were in danger of dropping out.

The course had six internal assessment components. Most of those were practical tasks to be shown to the teaching staff during lab times. Based on staff feedback students could rework their solutions and present those again for marking. This approach led to many constructive conversations between staff and students and to an opportunity to gain full marks by taking on the feedback and persisting with the work. In the end-of-course survey conducted by the university, which had an above average return rate of close to 50%, the questions on appropriateness of workload, assessment turn-around times and value of feedback given received the highest ratings of all questions (5.0, 5.2, and 5.1 out of 6), indicating that the assessment design was suitable. From a teaching perspective it was therefore disappointing to see that twelve students severely affected their final grades (and missed out on learning opportunities) by not attempting several assessment components. On average these students lost 18 marks out of 100 (assuming they would have achieved average marks for the work). Eight of those students failed the course, four passed with a 'C' grade. By doing all the course work to an average standard only two of those students would have failed and two would have lifted their grades to 'A'. These numbers show the difficulties teachers face in achieving the completion rates expected. The twelve students who did not attempt substantial parts of the internal assessment amount to 20% of the classroll. The eleven students who according to LMS logs and assessment records did no to very little work for the course make up another 19%.

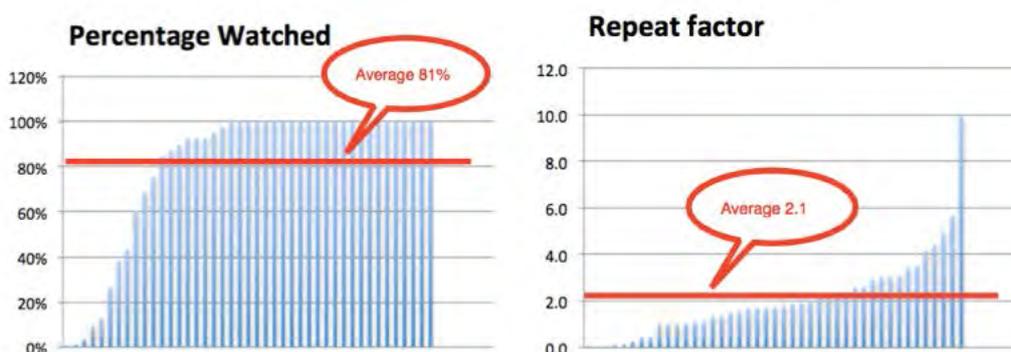
Access to static material – lecture slides and readings

Lecture slides and readings were made available via the LMS in PDF format. On average the graded students accessed 12.4 of the 17 sets of lecture slides available and 2.4 of the three readings. 48% of graded students accessed all 17 sets of lectures slides and 58% accessed all three readings. Access varied across the five topics of the course. All graded students accessed all lecture slides and the reading for Topic 3, a topic not directly linked to a practical course component. Some differences in total marks gained can be observed. The students who accessed all lecture slides had a mark average of 68, which was ten marks higher than the average across students who accessed less than half of the lecture slides. Students who accessed all three readings also had an average of 68, compared to 53 gained by students who accessed only one reading. Despite these differences, no statistically significant correlations could be observed.

While the LMS access logs for lecture slides and readings might be indicators for engagement with the study material, they seem to carry limited meaning. Access to the material does not show the level to which students might have engaged. While no access to the material should be a strong indicator for a lack of engagement (assuming students did not get copies from friend without going through the LMS), this does not seem to carry any predictive value in terms of grades. In this course, like in others observed before, there were several students who received high grades, despite not accessing the material. This might be a sign of the diversity of students taking first year courses, where some students enter with a high level of subject knowledge and independent study skills.

Access to dynamic material – screencasts

Links to 15 screencasts were available for streaming from the university's media server. The teacher used these screencast to demonstrate and explain how to use software to solve the practical course tasks. The content of the screencasts was directly relevant for the assessed work. The screencasts had an average length of 11 minutes 21 seconds and were accessed on average by 36 students. The students who accessed the screencasts watched 89% of the full lengths and repeated 1.7 times. Figure 2 shows the data on percentage watched and repeat factor for one of the screencasts used in



Week 3 of the course.

Figure 2: Access to one of the screencasts used in Week 3

There were no significant correlations between the streaming of the screencasts and the marks for the directly related assessment tasks. A number of students did not access the recordings yet still completed the tasks to a high level. An explanation might again be related to student diversity. Log entries that show that some students have only played a few seconds of each recording. It would be interesting to investigate if this was intended behaviour or if there are technical reasons for those log entries. Monitoring of the access logs during the semester could provide opportunities to identify students who might need help. If a student watches one screencast many times more than other students, this might be an indication that the student struggles with the material. If a student has watched all screencasts for a topic but has not submitted the matching assessment work, the student might require extra assistance. For the teacher of the course the access statistics on the screencasts provided re-assurance that their effort in creating these recordings was worthwhile. In the past students had asked for more screencasts. The access logs confirmed that the students made use of the material.

Knowledge, tool and time requirements

For this research the analysis was carried out with a spreadsheet program using descriptive statistics. The work was time consuming and had to be approached with care. The data came from multiple sources (downloads from the LMS and from the media streaming system, spreadsheets with assessment details). Access logs for the streaming data had to be downloaded separately per recording. The LMS records contained data for anyone with access to the course at some stage and had to be separated into the various student cohorts (graded and others) and teachers and administrators. The LMS and media streaming logs contained student identifiers in different formats, which had to be converted. While none of these steps was difficult, the work was time consuming and error prone. Few teachers would be able to invest the time to analyse the data for a course after its completion. Even fewer teachers will be able to do such analysis on an on-going basis while teaching a course.

Conclusions and recommendations

Based on the analysis of this single course preliminary conclusions can be drawn. Looking at log data might provide two ways of assisting students. First, access logs indicate when student engagement with a course drops off. Noting this in a timely fashion might provide opportunities to engage with students directly, learn of issues and help. Second, over-engagement with material might indicate that a student is trying hard but is struggling with a topic. Again it might be possible to offer targeted help. Both of these scenarios will require some flexibility in course structures, for example, allowing a student to catch up with assessment components. Based on the data analyzed for this course it seems unlikely that a below average access to study material, such as lecture slides, readings, or recordings, is a reliable indicator for a lower study performance. Using analytics to encourage student access to material needs to be treated with care, as nothing will be gained by students just clicking on resource links to increase their rankings in access logs.

In terms of assisting teachers the analytics gathered for this course showed promise. The degree to which students watched the screencast recordings was encouraging and confirmed to the teacher that the effort invested was worthwhile. While the data showed no statistically significant correlations between access logs and marks, they still allow a teacher to highlight when students fail to engage. As part of a more holistic assessment of a course, considering course design, student workloads and course evaluations, learning analytics can be a valuable part of a teacher's argument should their completion rates be criticized.

The next research steps will be to analyze several other first year blended courses, taking individual course characteristics into consideration. This will show if the preliminary findings can be confirmed. Should this investigation show how learning analytics can provide substantial benefits to students and teachers, further research steps will be to investigate an efficient toolset that makes it feasible to apply learning analytics. In this context bigger issues will have to be examined. Papamitsiou and Economides (2014) question if learning analytics should be the domain of individual teachers or of

specialists who have the required technical, statistical and pedagogical knowledge. Are there parallels to the areas of learning design and technology-enhanced learning, where research focused academics by large need specialist support?

References

- Andergassen, M., Mödritscher, F., & Neumann, G. (2014). Practice and repetition during exam preparation in blended learning courses: correlations with learning results. *Journal of Learning Analytics*, 1 (1), 48-74.
- Dyckhoff, A.L., Zielke, D., Bültmann, M., Chatti, M.A., & Schroeder, U. (2012). Design and implementation of a learning analytics toolkit for teachers. *Educational Technology and Society*, 15 (3), 58-76.
- Ginns, P., Kitay, J., & Prosser, M. (2010). Transfer of academic staff learning in a research-intensive university. *Teaching in Higher Education*, 15 (3), 235-246. <http://dx.doi.org/10.1080/13562511003740783>
- Gunnarsson, B.L. & Alterman, R. (2014). Peer promotions as a method to promote content. *Journal of Learning Analytics*, 1 (2), 126-150.
- Long, P., & Siemens, G. (2011). Penetrating the fog: analytics in learning and education. *EDUCAUSE Review*, September/October 2011.
- Papamitsiou, Z. & Economides, A.E. (2014). Learning analytics and educational data mining in practice: a systematic literature review of empirical evidence. *Educational Technology and Society*, 17 (4), 49-64.
- Zepke, N. (2013). Student engagement: A complex business supporting the first year experience in tertiary education. *The International Journal of the First Year in Higher Education*, 4(2), 1-14. <http://dx.doi.org/10.5204/intjfyhe.v4i2.183>

Heinrich, E. (2015). Can learning analytics provide useful insights? An exploration on course level. In T. Reiners, B.R. von Kinsky, D. Gibson, V. Chang, L. Irving, & K. Clarke (Eds.), *Globally connected, digitally enabled*. Proceedings ascilite 2015 in Perth (pp. CP:108-CP:112).

Note: All published papers are refereed, having undergone a double-blind peer-review process.



The author(s) assign a Creative Commons by attribution licence enabling others to distribute, remix, tweak, and build upon their work, even commercially, as long as credit is given to the author(s) for the original creation.