

Exploring Learning Analytics as Indicators of Study Behaviour

Rob Phillips

Educational Development Unit, Murdoch University, Western Australia, Australia
r.phillips@murdoch.edu.au

Dorit Maor

School of Education, Murdoch University, Western Australia, Australia

Greg Preston

School of Education, The University of Newcastle, New South Wales, Australia

Wendy Cumming-Potvin

School of Education, Murdoch University, Western Australia, Australia

Abstract

In this paper we describe the use of learning analytics to provide indicators of students' behaviour in technology-enhanced learning environments. We first provide an overview of the emerging field of learning analytics, and then describe a learning-analytic tool which reports student use of the Lectoria lecture capture system across the weeks of a semester, both collectively and individually. Observation of the data provided by the system enabled us to develop a set of proto-theoretical categories of behaviour, with associated algorithms to numerically identify these categories. Finally, we describe the use of the learning-analytic tool in a mixed methods approach to investigate in depth how students, with different characteristics, engage with and learn from technology-enhanced learning environments.

Introduction

The university teaching landscape has changed over the past decade due to online educational technologies. In these blended-learning environments, the 'classroom' is now a mix of live and/or online lectures, interactions and activities. Research is lacking on understanding how students engage with these environments and how well they learn using these new technologies. Students today are leaving a data trail while they are interacting with others, with information, and with institutions through different technologies (Siemens, Gasevic, Haythornthwaite et al., 2011), and Learning Analytics is an emerging approach for interrogating that data (Johnson, Smith, Willis et al., 2011). Learning analytics includes a variety of data-gathering tools and analytical techniques that can be used to examine students' engagement, performance, and progress on tasks. These data-mining tools enable researchers to identify and analyse the complexity and diversity of learning behaviours that can take place in technology-enhanced learning environments (Calvo, Howard, & Markauskaite, 2010; Campbell & Oblinger, 2007; Chang, 2007; Choquet, Luengo, & Yacef, 2009).

Furthermore, learning analytics can connect the power of data mining to an understanding of teaching and learning to improve curriculum, learning and assessment. A main aim of this technique is to improve learning outcomes and reduce attrition, especially for at-risk students. Through its visual display of data, learning analytics enables educators to 'zoom in' on particular students who need more specific support (Siemens, et al., 2011). Learning analytics can provide data about the impact of various interactions to provide educators with the necessary information to assist learners. In addition, learners can take greater responsibility for their learning by becoming aware of their behaviour. Thus, the use of learning analytics should improve completion rates and increase learning outcomes.

The terminology around 'analytics' in education is evolving. Despite calls for its use in the early 1990s (Kozma, 1994; Salomon, 1991), usage logs from e-learning applications have been under-utilised in e-learning research. The use of this automatically-captured data, which records who accessed what and when to study student behaviour, is now termed *learning analytics* (Goldstein & Katz, 2005; Oblinger & Campbell, 2007). This can be contrasted with *academic analytics*, which considers similar data at an institutional level (van Barneveld, Arnold, & Campbell, 2012).

The use of learning analytics to diagnose and improve student learning has begun to emerge in recent years (Dawson, Macfadyen, & Lockyer, 2009; Dawson, McWilliam, & Tan, 2008). Early examination of usage logs began with custom-built interactive multimedia learning systems with built-in usage tracking (Judd & Kennedy, 2001; Kennedy & Judd, 2004; Thornton & Phillips, 1997). As

web-based learning management systems evolved, researchers began to collect usage log data to understand how students engaged with e-learning environments, for example in Biology (Phillips, Baudains, & van Keulen, 2002). Further work involved analysing institutional usage patterns of the myriad tools available through learning management systems (Phillips, 2006). Subsequent work has focused on tools to monitor and evaluate the formation and ongoing development of student social networks by extracting data from online discussion forums and visually displaying the resultant social networks (Dawson, 2006a, 2006b; Dawson, Bakharia, & Heathcote, 2010). Learning analytics provides direct evidence of student learning behaviour, in contrast with other approaches (e.g., surveys and interviews) which provide indirect evidence, filtered by the perceptions of the student (Salomon, 1991).

With the increase in technological capabilities and our ability to develop learning-analytic systems, we can focus on the learner and gather data from learning management and student information systems to enhance student success. In addition, the system can detect signs of a learner's reduced interaction or participation in the learning environment. This enables university educators not only to monitor students' behavior but also to act upon it and intervene to improve individual students' learning. However, the continued advancement and acceptance of learning analytics in academia is still a work in progress to which this paper seeks to add.

Learning analytics for lecture capture

Lecture capture is a generic term for the process of automated recording of audio (and video) from live lectures. The Echo360 Lectoria application (<http://echo360.com/>) is one such product, which is in wide use in Australia, where it is being superseded by the EchoSystem product. The Lectoria system automatically records all access to the system in various database tables. Relevant information includes:

- The student user name, passed through an authentication system
- The date and time of access to a unique lecture recording (a *hit*)
- Details specific to a particular course, or unit of study (unit code, unit name, lecturer, etc.)
- Information about the format of the recording (streamed or downloaded; audio or video; bit rate; etc.)

Our particular interest was week-by-week patterns of behaviour across a semester. The access dates and times were converted into the week of the teaching period in which they occurred. The number of hits per week was graphed against the week of semester for the entire class and for individual students.

The process was automated through the development of a server-side PHP script with a simple web-based interface (Phillips, et al., 2010). This script works on a single unit of study (course), and displays all relevant data in tabular (and downloadable) form, together with a summary graph. The web interface continues with a listing of all students in the unit, ordered according to a heuristic which displays the heaviest users first. The reporting tool also provides alternative ways of drilling down into the data, so that the educator can see at a glance the nature of each 'hit' (or access to a Lectoria event):

- the format of each hit (download/stream/MP3/MP4, etc.)
- the timing of each hit (hits in the first day/within 7 days/after seven days)
- the type of each hit (whether it is an initial hit or a repeat hit on a recording)
- the number of the lecture that is accessed in which week

Overall patterns of behaviour

The learning-analytic tool reveals different patterns of behaviour in different units. Figure 1 shows one representation of the overall data for all students who accessed Lectoria in one unit. This shows the total number of hits on the Lectoria system (y-axis) against the week of the semester (x-axis). In this case, the semester is structured as 10 weeks of formal teaching (Weeks 1-3, 6-9 & 12-14), with four non-teaching weeks (shown on Fig. 1). The study and exam period extends from weeks 15-17. Assessment dates are also overlaid on Figure 1. In addition, the time between the lecture recording and the time recordings were accessed are shown as an extra dimension in Figure 1, with three intervals: on the day of the lecture, during the first week, and after the first week.

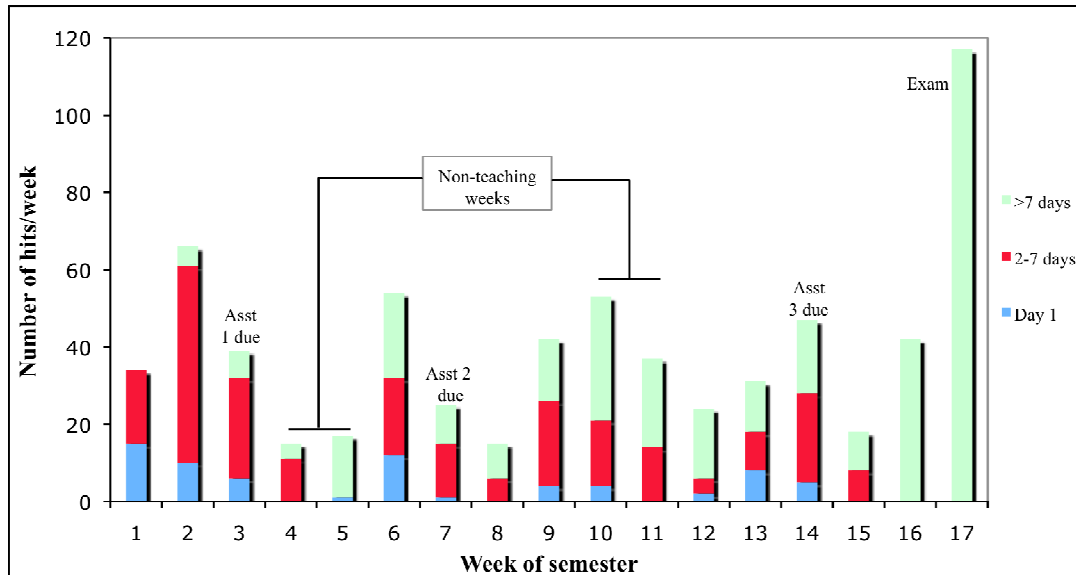


Figure 1. Representation of the delay in listening to recordings against the week of the semester for all students in a unit.

Individual patterns of behaviour

The Lectopia tool also enables us to drill down into the behaviour of individual students within a unit. Observation of the data made it clear that there were very different patterns of use across students, as illustrated in Fig. 2. Some students used Lectopia regularly, some used it rarely, and some used it in bursts. We set out to categorise this behaviour and developed ten different usage categories which are summarised in Table 1. Conscientious and high-achieving students access Lectopia regularly. Good-intentioned and repentant students have some weeks of regular use, at either the beginning or end of the teaching period. Other students access recordings in blocks – they are binging users. A sub-category of bingers are the free-timers, who access recordings during non-teaching weeks. Crammers leave their engagement with recordings until just before the examination period. Other

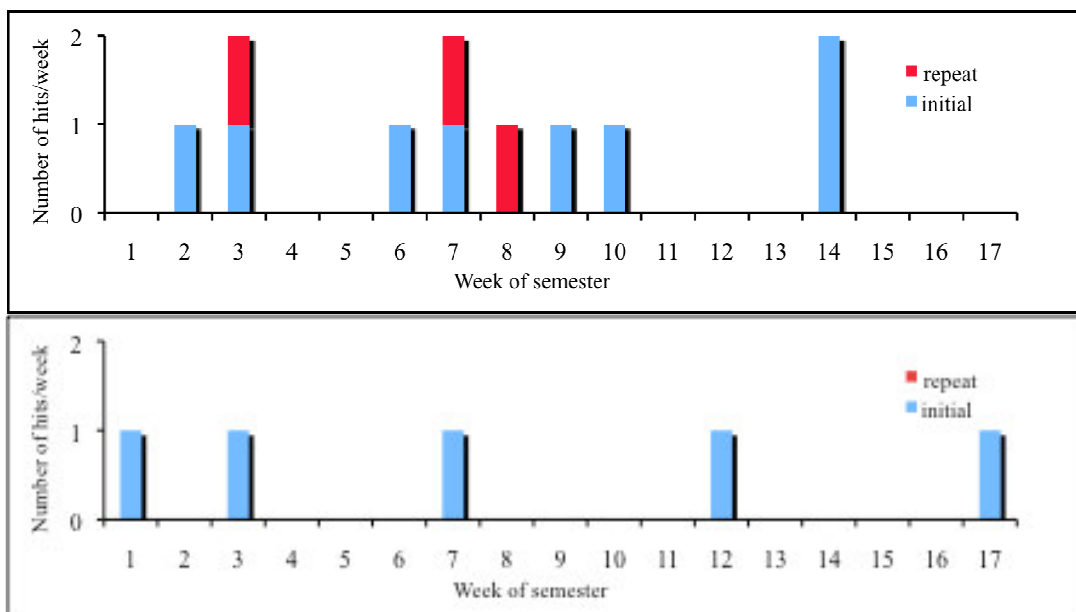


Figure 2. Lecture recording access patterns for two students in the same unit against the week of the semester.

students may access recordings once or not at all, or their pattern of use may not fit any of the other categories.

The third column of Table 1 defines algorithms which we propose to use to numerically categorise behaviour patterns. A drawback of our current approach is that each student graph needs to be observed individually to determine the behaviour patterns. The algorithms proposed here, once implemented, will enable us to see at a glance broad trends in student behaviour, which we can delve into as required. However, they are yet to be tested, and the definitions may be revised as they are trialled on live data.

Table 1. Potential categories of study behaviour for individual students in Lectopia-supported learning environments, and their algorithmic representations.

Category	Typical Profile	Algorithm definition
Conscientious	Students access the current lecture in the majority of weeks where a lecture is posted.	75% of lectures are accessed within 7 days of the lecture being posted.
High-achieving	Students access the current lecture in the majority of weeks where there is a lecture posted and 'revisit' most of the lectures.	As for Conscientious, but >50% of lectures are accessed more than once.
Good intentioned	Students start with a regular, weekly access pattern for the first part of the unit - which reduces during the semester.	The mean week of usage of lecture recordings is in the first half of the unit.
Repentant users	A systematic profile, or extended activity is recorded sometime after week 5 of the semester, with little or no activity before this.	The mean week of usage of lecture recordings is in the second half of the unit.
Bingers	Students access multiple lecture recordings in a single week followed by weeks with no access.	Students will have at least 3 weeks where they access multiple weeks of lectures, with at least 1 week of no access between these occurrences.
Free-timers	The majority of the hits fall during weeks where there are no new lectures posted, e.g. mid-semester breaks. (sub-category of Binger).	>75% of lectures will be accessed during non-teaching weeks.
Cramming	Students have the majority of their usage in the two weeks immediately prior to final exam/assessment tasks.	The mean week of usage of lecture recordings is in the last 2 weeks of the unit.
One-hit wonders	Students have only a single successful access of a single lecture.	There is only one unique hit.
Random	No typical profile.	All others who have more than one hit.
Non-user	No Lectopia activity - Student Number present on enrolment list with no hits on the Lectopia system	Enrolment =Yes & Hits = 0

Challenges and limitations of behaviour categorisation

Learning-analytic tools provide objective data about what a student clicked on at a particular point in time (Salomon, 1991). "At their most basic level, audit trails [learning analytics] measure the behavioural responses and activities of users" (Kennedy & Judd, 2004, p. 19). However, care should be taken in analysing and interpreting this data. The categories shown in Table 1 provide indicators of behaviour, but they do not explain that behaviour. For example, bingeing students could be very effective in balancing their study, work and family commitments, doing concentrated study when they have the opportunity. On the other hand, a bingeing student could be falling behind in their work because of poor time-management and prioritisation skills, and their efforts could be ineffective. In other words, Learning-analytic tools of the type reported here are insufficient to explain student behaviour on their own.

Developing a learning-analytic methodology to investigate study behaviour

Given the preceding caveat, we have developed a mixed-methods approach to combine learning-analytic tools with more traditional e-learning research methods to create rich descriptions of student behaviour in particular units of study. An illustration of our methodology is shown in Fig. 3, and an initial trial is reported in Phillips, et al. (2011). In a nutshell, survey and learning-analytic data are used to identify students with different characteristics who engage in different ways with technology-enhanced learning environments. A sample of 8-10 students is then selected for each case, and interviewed to discuss their recorded usage behaviour, as well as their specific studying behaviours.

We will then combine grades, interview transcripts and usage behaviours to describe in depth how each student is engaging with the unit and the level to which they are successful, in order to develop design principles about student behaviour. Continued refinement and validation of the methodology is currently underway, with projects commencing to look at both student behaviour and the role of unit/course design on that behaviour. We also intend to develop 'plug-ins' to our reporting tool to access data from other e-learning systems, such as Echosystem, Moodle and Blackboard.

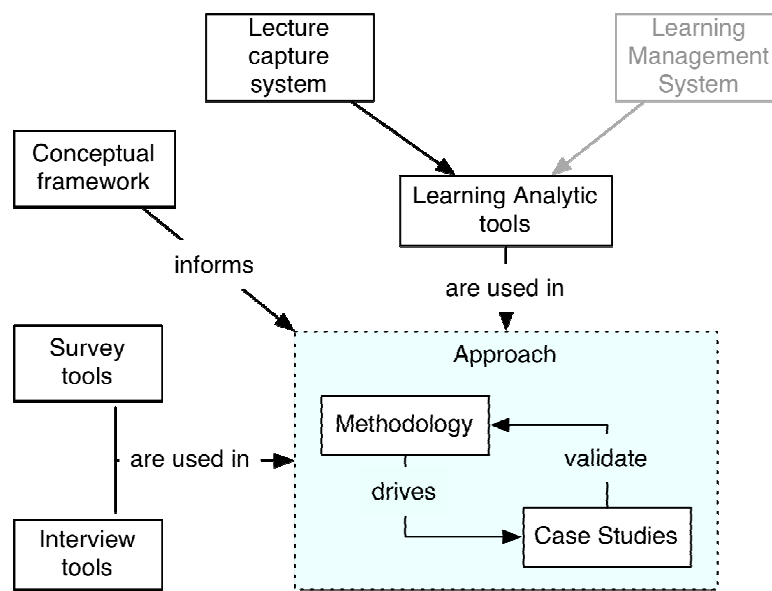


Figure 3. Mixed-methods approach which utilizes learning analytic data.

Conclusion

As part of a trend in increasing the use of learning analytics for educational purposes in higher education, we have developed a learning-analytic tool to observe students' behaviour in a lecture-capture system through the data trail they left. The tool can be used to observe different patterns of use across units of study and also identify individual students' use within units.

This paper describes the conceptual basis of the next iteration of the learning-analytic tool, by proposing a set of algorithms to numerically categorise students' usage patterns. We also describe a mixed methods approach to investigate students' study behaviour, which includes learning-analytic tools.

The learning-analytic tool provides an objective measure of students' behaviour rather than relying on the students' or educators' perceptions that could have been gathered through surveys or interviews.

In this relatively new trend of using learning analytics in educational research to analyse students' learning behaviour, we added a significant contribution by designing an empirical study which tracked students' patterns of behaviour over a semester. This type of study could be applicable to a variety of learning technologies and could help in assessing the impact of different types of technologies or learning environments. While conceptual frameworks are still being developed (van Barneveld, et al., 2012), a closer look at the possible usage of the system as demonstrated in our study potentially creates a better understanding of how well students interact in their learning environments while using specific new technologies.

References

- Calvo, R. A., Howard, & Markauskaite, L. (2010). Factors affecting students's experiences and satisfaction about teaching quality in engineering. *Australasian Journal of Engineering Education*, 16(2), 139-148.
- Campbell, J. P., & Oblinger, D. G. (2007). Academic Analytics. Retrieved 30 June, 2010, from <http://net.educause.edu/ir/library/pdf/PUB6101.pdf>
- Chang, S. (2007). *Academic perceptions of the use of Lectoria: A University of Melbourne example*, Singapore.
- Choquet, C., Luengo, V., & Yacef, K. (2009). *Usage Analysis in Learning Systems*: AACE.
- Dawson, S. (2006a). Online forum discussion interactions as an indicator of student community. *Australasian Journal of Educational Technology*, 22(4), 495-510.
- Dawson, S. (2006b). Relationship between student communication interaction and sense of community in higher education. *Internet and Higher Education*, 9(3), 153-162.
- Dawson, S., Bakharia, A., & Heathcote, E. (2010). SNAPP: Realising the affordances of real-time SNA within networked learning environments. In H. V. Dirckinck-Holmfeld L, Jones C, de Laat M, McConnell D & Ryberg T (Ed.), *Proceedings of the 7th International Conference on Networked Learning*.
- Dawson, S., Macfadyen, L., & Lockyer, L. (2009). Learning or performance: Predicting drivers of student motivation *Same places, different spaces*. *Proceedings ascilite*. Auckland: ascilite.
- Dawson, S., McWilliam, E., & Tan, J. P. L. (2008). Teaching smarter: How mining ICT data can inform and improve learning and teaching practice *Hello! Where are you in the landscape of educational technology?* *Proceedings ascilite*. Melbourne.
- Goldstein, P. J., & Katz, R. N. (2005). Academic Analytics: The Uses of Management Information and Technology in Higher Education, ECAR Research Study 8 Available from <http://net.educause.edu/ir/library/pdf/ers0508/rs/ers0508w.pdf>
- Johnson, L., Smith, R., Willis, H., Levine, A., Haywood, A., & Haywood, K. (2011). *The 2011 Horizon Report*. Austin: Texas: The New Media Consortium.
- Judd, T., & Kennedy, G. (2001). Extending the Role of Audit Trails: A Modular Approach. *Journal of Educational Multimedia and Hypermedia*, 10(4), 377-396.
- Kennedy, G. E., & Judd, T. S. (2004). Making sense of audit trail data. *Australasian Journal of Educational Technology*, 20(1), 18-32.
- Kozma, R. B. (1994). Will media influence learning? Reframing the debate. *Educational Technology Research and Development*, 42(2), 7-19.
- Oblinger, D. G., & Campbell, J. P. (2007). Academic Analytics. Retrieved 30 June, 2010, from <http://connect.educause.edu/library/abstract/AcademicAnalytics/45275>
- Phillips, R. A. (2006). Tools used in Learning Management Systems: analysis of WebCT usage logs. In L. Markauskaite, P. Goodyear & P. Reimann (Eds.), *The 23rd Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education: Who's Learning? Whose Technology?* (pp. 663-673). Sydney: Sydney University Press.
- Phillips, R. A., Baudains, C., & van Keulen, M. (2002). An Evaluation of Student Learning in a Web-Supported Unit on Plant Diversity. In A. Williamson, C. Gunn, A. Young & T. Clear (Eds.), *19th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education* (pp. 525-534). Auckland, New Zealand.
- Phillips, R. A., Maor, D., Cumming-Potvin, W., Roberts, P., Herrington, J., Preston, G., et al. (2011). Learning analytics and study behaviour: A pilot study. In G. Williams, P. Statham, N. Brown & B. Cleland (Eds.), *Changing domains, changing directions. Proceedings ascilite 2011* (pp. 997-1007). Hobart: ascilite.
- Phillips, R. A., Preston, G., Roberts, P., Cumming-Potvin, W., Herrington, J., Maor, D., et al. (2010). Using academic analytic tools to investigate studying behaviours in technology-supported learning environments. In C. Steel, M. Keppell, P. Gerbic & S. Housego (Eds.), *Curriculum, technology & transformation for an unknown future. Proceedings ascilite Sydney 2010* (pp. 761-771). Sydney: ascilite.
- Salomon, G. (1991). Transcending the qualitative-quantitative debate: The analytic and systemic approaches to educational research. *Educational Researcher*, 20(6), 10-18.
- Siemens, G., Gasevic, D., Haythornthwaite, C., Dawson, S., Shum, S. B., Ferguson, R., et al. (2011). *Open Learning Analytics: an integrated & modularized platform. Proposal to design, implement and evaluate an open platform to integrate heterogeneous learning analytics techniques*. .

- Thornton, D., & Phillips, R. A. (1997). Chapter 6, Evaluation *The Developer's Handbook to Interactive Multimedia - A Practical Guide for Educational Applications* (pp. 127-146). London: Kogan Page.
- van Barneveld, A., Arnold, K. E., & Campbell, J. P. (2012). Analytics in Higher Education: Establishing a Common Language. *ELI White Papers*, (1/2012). Retrieved from <http://www.educause.edu/Resources/AnalyticsinHigherEducationEsta/245405>

Acknowledgements

This work was funded by the Echo360 International Research Grants Program, and by a grant from the Murdoch University Strategic Research Fund. The programming was done by Reef Turner, with technical support from the Perth office of Echo360. We also acknowledge the technical assistance of the then Lectopia administrator at Murdoch, Zig Tan. We also acknowledge the contribution of Professor Jan Herrington to this work.