For the Record: Learning Analytics and Student Engagement

James Moir

Abertay University

Abstract

Learning analytics and data mining are being used with students as an aid to engagement with their studies. Typically, these kind techniques rely upon mobile phone applications to register attendance in classes as well as tracking student electronic touchpoints within an institution. For example, the JISC-devised *Study Goal App* allows students to monitor their virtual learning activity, set study targets and track their engagement in relation to specific modules as well as norm-referenced indices of study patterns. While this is useful for students own private use, institutions are also using this data to monitor students with the aim of predicting those who may be at risk of failing or withdrawing from their studies. It is also the case that student engagement with different aspects a programme, such as individual units or modules, can also be analysed using this kind of data.

The benefits of using learning analytics is that they offer students a way of tracking their own study habits and setting targets and goals. This is very much based on the sort of self-tracking and big data usage that is commonly associated with fitness apps, devices and regimes. However, while the ethos of self-tracking may be appropriate for personal fitness regimes, it is another matter when applied to higher education where ethical, social and educational issues are thrown into relief. The paper discusses these issues and considers the meanings, assumptions, and values that are embedded in the use of data mining as an index of student engagement.

Introduction

Learning analytics and data mining are increasingly used in higher education with students to help them engage with their studies and monitor their own patterns of learning. Student interactions with their online learning activities, as well electronic attendance registrations in class, are captured and stored. These digital traces (i.e. log data) can then be 'mined' and analysed in order to identify patterns in students' learning behaviour. Thus, the study of learning analytics has been defined as the "measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs" (Siemens & Gašević, 2012).

Learning analytics draws upon a variety of research, method and techniques (e.g. learning sciences, data mining, information visualisation, and psychology) in order to reveal patterns in learning behaviour. Within this collection of approaches, data mining presents a different outlook for the management of learning and student engagement higher education. The ability to predict individual outcomes allows institutions to take advantage of the data that students generate and to use it to put in place interventions, or permit students themselves to use their own data to guide their learning. As Johnson (2014: 4) notes:

> "The aim of data mining is to identify relationships among variables that may not be immediately apparent using hypothesis-driven methods. Having identified those relationships it is possible to take action based on the fact that the relationships predict a given outcome."

Prediction followed by intervention are therefore the central features of this approach, built upon the identification of previously unseen data relationships. This predictive work can be used for a variety of purposes, for example: predicting student academic performance and retention, evaluating student learning within course management systems, monitoring student attendance; and evaluating different kinds of adaptive and personalised support.(see e.g., Ayesha et al. , 2010; Kumar and Chadha, 2011; Daniel, 2015; Gibson and Ifenthaler, 2017; Dawson et al. 2018).

While there is clearly much work going on in using student generated data, questions remain about the educational value of learning analytics and data mining (Gašević and Dawson, 2015). This paper considers a recent example of learning analytics and data mining as applied to student engagement through the Joint Information Systems Committee (JISC) *Study Goal App* (JISC, 2017). Four different questions are examined in relation to the use of this software: (1) What is engagement – activity versus social practice? (2) Why count log-ins as learning? (3) Why use a comparison model? (4) What about ethical concerns? I first begin by outlining the nature of the software and then proceed to take each question in turn.

The JISC Study Goal App

The JISC *Study Goal App* builds on some of the motivational aspects that are present in fitness trackers. In other words, it tracks learning activities, draws attention to patterns of behaviour that are not readily apparent, allows for comparisons with the activities and progress of other learners, enables the user to set study Goal and targets, prompts action through alerts in order to keep goals in mind, and provides a reward mechanism through congratulatory messages. It therefore seeks to combine data about what students are studying with data on the activities they are engaged in order to boost awareness of learning and motivation. The mobile app also has a feature that permits students to register their attendance for a given learning activity (e.g. a class at a set time) by logging in and entering a four-digit code generated through the system and enabled by the class tutor. Typically, this code is only available for the first ten minutes of the class and can be tied to a geolocation function, although students may opt out of this.

In effect the app records all electronic log-ins and time spent across activities such as time spent in class, or on a given part of the virtual learning environment VLE), as well as a student's own recorded activity (e.g. periods of time spent reading). In this way the system enables the recording of electronic 'touch points' as well as individualised recorded activity. Students can look at their own VLE activity for a specific module and then see a comparison with a module average or compare with their identified friends. They can also see in graphical form recorded weekly attendance over the past four weeks as well as setting targets for study activities (e.g. 'read for two hours per day', 'work on assignment 1 for two hours a day' etc.). Students can be also be awarded 'activity points' by their institution for their engagement in various ways (e.g. individual use of the VLE or attendance at lectures).

It is also the case that higher education institutions who are using the system also have access to the electronic touch point data and can use or mine it for a number of purposes. Institutions can access student data on attendance, module grades and use of their VLE but do not have any access to sensitive personal data such as disability, gender, or ethnicity. Staff can access overall module attendance statistics on a week-by-week basis in graphical form or can select individual students and examine their pattern of attendance for a given module.

In summary, the app operates on a number of levels. It has a comparative element for students (e.g. time spent on a learning activity with the module average). There is also a social aspect in being able to select friends with whom to compare with, and this can be seen as adding another level of engagement. It is gamified in terms of a rewards and points system. And finally, there is a privacy feature that enables students to disable location data.

Having given an outline of the app, I now turn to consider how it relates to the four questions set out earlier. In so doing my aim is evaluate its situated use in practice rather than how it is presented in, for example, the user's guide. My aim in doing this is to consider its operationalisation and what this means for both students and institutions.

What is engagement – activity versus social practice?

Student engagement is key to student achievement and retention (Krause & Coates, 2008) and is widely recognised as being crucial to success in higher education. However, the nature of what constitutes engagement and why students disengage with their studies poses complex questions with a range of theories being brought to bear upon the issue (see e.g., Trowler & Trowler, 2010; Zepke & Leach, 2010). Kahu (2013) identified three approaches to engagement: linkages between student behaviours and teaching practices (Astin, 1984), psychosocial process with behavioural, cognitive and affective dimensions (Fredricks et al., 2004), and the social context of engagement related to sociocultural factors (Mann, 2001). Other approaches emphasize the interplay between student, institutional, and sociocultural factors (Bryson, 2014; Kahu & Nelson 2018).

In encouraging student to consider their engagement through the *Study Goal App* in terms of log-ins and comparisons with others, there is clearly a focus on student behaviours. Indeed, it could be argued that this focus on statistical information betrays a crude scientism in treating engagement as an objective matter, writ large in the visualisations presented on screen. Focusing on engagement in functional terms therefore places activity first and foremost rather than helping students acquire a reflexive stance on their learning. Activity is emphasized as the expense of action, where engagement is considered in terms of time spent doing something rather as a social practice. This is important, for not only does the *Study Goal App* reflect what students do but also constructs engagement as time spent doing. Thus, students and teaching staff become concerned with engagement as a matter of counting time spent on activities. While there is little doubt that time spent attending classes or on the VLE is in one sense and index of engagement, there is a concern when this emphasised as being mostly what engagement is about.

As (Barnett, 2007: 38) reminds us, students experience higher education as ongoing 'transformation of being'. In this regard, engagement is meaningful and involves the active and ongoing navigation of a student's habitus and the culture (Thomas, 2002), and the academic practices of particular disciplines and institutions. Learning as a set of inter-related experiences is more than just time spent doing something. In this sense, engagement with learning offers the potential to challenge students' ways of being and thinking; it is a social practice that requires students to relate their learning experiences and identities through new knowledge and understanding in a transformational manner (Mezirow, 2018).

Why count log-ins as learning?

A point related that above is the use of log-ins to the VLE as a proxy for learning activity. This is typical of a learning analytics approach in considering the effects of operations performed by using proxy measures of learning derived from trace data. In the case of the *Study Goal App,* this is typically units of learning data as measured through interaction with the VLE. However, using log-ins and time spent on the VLE as in index of effective self-regulated learning is contentious. This is because it adopts a 'one-size-fits-all' strategy and does not take into account different uses that teaching staff use the VLE for, depending upon a specific module, discipline or students' preferences for how they learn. There is also a more subtle process going on here in terms of the assumption that VLEs are integral to students' learning.

Gašević et al. (2015: 67) note:

> "Learners are active agents in their learning process. This simple statement has many significant implications. Learner agency implies that even when learners receive the same instructional conditions, they may choose to adopt alternate study methods."

In other words, if learners adopt different methods of learning, for example in using the VLE to a greater or lesser extent, then this will show up in the log count as simply more or less learning activity. However, some learners may opt to read more books (e.g. where they have purchased copies of their own) and learn in this way rather than say using electronic resources via the VLE.

In a study with undergraduate students of education in a blended course, Lust et al. (2013) identified four types of students based on their use of learning tools. The groups were classified as: (i) non-users, or low level adoption of any tool in the learning management system (LMS) suggested (e.g., quizzes, web lectures, and discussion forums); (ii) intensive active users who used all tools in the LMS; (iii) selective users who used only used a selected number of tools offered to them; and (iv) intensive superficial users who used all the tools and spent more time than other groups on predominantly passive activities such as reading discussion posts rather than contributing to the forum. This study reminds us that time spent online and the usage of various tools can mean different things to different learners, and that activity logs need to be interpreted with some caution.

Why use a comparison model?

The *Study Goal App* is based on fitness tracking apps that allow users to upload and then compare their physical activity (e.g. running a set route or distance) with other users (e.g. through averages or percentiles). On the face of it, the mapping across of this model to study and learning activity, seems reasonable as learners can see how they fair in comparison with other students and, in particular if they are spending less time than the average. It might seem, therefore, that a little element of comparison and even competition might be conducive to good study habits.

However, there are several problems with this approach, not least the fact that it decontextualizes a student's learning activity by simply presenting comparisons with aggregated data. Students are individuals with different life circumstances. Some come to higher education straight from school, others through articulation routes through college and yet others as mature students after several years away from formal education. Some have family or other personal commitments and others do not. The point being made here is that a students' personal circumstances and the extent of their socialisation within the world of education can have a major impact on the amount of time they are able, or find necessary, to devote to study activity. Moreover, the comparison model used in the *Study Goal App* simply does not take into account socio-cultural factors related to a students' cultural and social capital (Bourdieu, 1979) in relation to their learning. Higher education preferentially favours certain forms of knowledge and understanding (Thomas, 2002) and for students those whose embodied practices are not equally valued, then the institutional habitus of the academy can lead to alienation and withdrawal (Mann, 2001).

A study conducted by Roberts et al. (2016) gives a student perspective on this issue. In their qualitative study of students' knowledge, attitudes, and concerns about big data and learning analytics their thematic analysis yielded a theme of "impeding independence", amongst others. Student expressed an appreciation for the additional support provided by learning analytics but valued being in charge of their own learning. There appeared to be concern that the use of learning analytics would have an adverse effect on the expectation to be self-reliant and that this could create an environment where students are no longer treated as independent adults. Furthermore, students commented that much of the learning analytics information, such as comparison to peers, was already available to them through other means, making learning analytics redundant.

What about ethical considerations?

The main ethical issues that relate to the use of learning analytics and data mining are privacy, consent, and how data is used, stored, and protected and acted upon (Alexander and Brown, 1998; Cumbley and Church, 2013; Rubel and Jones, 2016). In the *Study Goal App,* it is possible for an institution to draw down data on students who appear to be failing to engage, or in weak manner relative to other students, on a given programme of study or module. For example, the attendance recording feature can be used detect absences and, when combined with data on time spent on the VLE, can be used to spot students 'at risk' from disengaging with their studies. This in turn can be used or notify such 'disengaged' students that they may not be able to tackle a given assessment due to missing a particular class, or to invite them to a counselling session to discuss any issues that are impeding their engagement. This use of data poses important ethical questions, not least what students told about the utility of the app versus what teaching staff are told. If student are not explicitly told that *their* data can be mined and used by an institution to contact those 'at risk' from disengaging then this is simply ethically dubious.

With regard to this issue, Slade and Prinsloo (2015) hosted an online forum that elicited discussions. Most posts were concerned with the issue of transparency with students pointing out that their university could make more of an effort to inform them of what data is collected, for what purpose, how it is used and who would have access to it. Students expressed a desire to be informed on a regular basis about the use of their data and for this to be governed by a strong ethical case. They also viewed their data as both valuable and vulnerable with a need for institutional mechanisms of protection such as opt in/out options and informed choices. They also were concerned about the use of learning analytics alongside personal tutor support and that (mis)labeling and bias could impact negatively upon tutor-student relationships.

Although Slade and Prinsloo (2015) acknowledge that views expressed in the forum are not a representative sample, their research nonetheless throws up issues from a student perspective concerning the ethics of learning analytic and data mining. If higher education is to remain true to the rhetoric of the 'student-at-the-centre' then the use of *their* data must be based on express consent (Kruse and Pongsajapan, 2012; Slade and Prinsloo, 2013; Gaševic´ et al., 2015). It is all too easy for data collected for one purpose to be used by institutions for another. The European Union General Data Protection Regulation (GDPR) legislation (2018) should ensure that students are fully informed and required to give active consent to the use of their data for the purpose of learning analytic and data mining.

Conclusion

While higher education systems are now populated with vast amounts of data and the digital traces of learners, very little is related to the capture of conditions under which students learning takes place. It has been noted in this paper that social context, previous educational history, and students' sense of being independent learners all point to treating mined data with some degree of circumspection and caution. On the other hand, there may be some value in functional terms for student to use the Study Goal App as a means to organising their learning and study habits. This more modest usage does not require to be linked to data mining but instead can be used on a 'take-it-or-leave-it' basis by students as a potential aid to the *organisation* of their study.

The focus on event activities ignores an examination of the nature and meaning of learning for students. This is very much a problem inherent in the use of 'big data'; appearance is all built on the metrics of aggregation. The counting of engagement with certain types activities is not without its uses but these must be very carefully set within an understanding of the way they can construct (mis)understanding of what engagement constitutes. Perhaps it is best to conclude this paper with the words of the sociologist, William Bruce Cameron, who wrote in 1963 (p.13) that "not everything that can be counted counts, and not everything that counts can be counted". This observation still hold true today and perhaps more so when considering the ways in which learning analytic and data mining makes certain data on student learning appear objective and yet at the same time it can obscure the actual nature of learning that takes place in all its shades and hues.

References

Alexander, P., and Brown, S. (1998). "Attitudes toward information privacy: differences among and between faculty and students," in *AMCIS Proceedings*, 17 (Baltimore, MA). http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1444&context=amcis1998
(Accessed 18 May, 2018)

Astin, A. W. (1984). Student involvement: A developmental theory for higher education. *Journal of College Student Personnel*, 25(4), 297–308.

Ayesha, S., Mustafa, T., Sattar, A. R., & Khan, M. I. (2010). Data mining model for higher education system. *Europen Journal of Scientific Research*, *43*(1), 24-29.

Barnett, R. (2007). *A will to learn: Being a student in an age of uncertainty*. Maidenhead: Society for Research into Higher Education & Open University Press.

Bourdieu, P. (1997). The forms of capital. In A. H. Halsey, H. Laudner, P. Brown, & A. Stuart Wells (Eds.), *Education: Culture, economy, and society* (pp. 46–58). Oxford: Oxford University Press.

Bryson, C. (2014). Clarifying the concept of student engagement. In C. Bryson (Ed.), *Understanding and developing student engagement* (pp. 1–22). Abingdon: Routledge.

Cameron, W.B. (1963) *Informal sociology: A casual introduction to sociological thinking*. New York: Random House.

Cumbley, R., & Church, P. (2013). Is "Big Data" creepy? *Computer Law & Security Review*, 29 (5), 601–609.

Daniel, B. (2015). Big Data and analytics in higher education: Opportunities and challenges. *British journal of educational technology*, *46*(5), 904-920.

Dawson, S., Poquet, O., Colvin, C., Rogers, T., Pardo, A., & Gasevic, D. (2018). Rethinking learning analytics adoption through complexity leadership theory. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (pp. 236-244).

European Union (2018) General Data Protection Regulation (GDPR) https://www.eugdpr.org/ (Accessed 18 May, 2018)

Fredricks, J. A., Blumenfeld, P., & Paris, A. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74(1), 59–109.

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*, *59*(1), 64-71.

Gibson, D. C., & Ifenthaler, D. (2017). Preparing the next generation of education researchers for big data in higher education. In *Big data and learning analytics in higher education* (pp. 29-42). Springer, Cham.

JISC Study Goal App https://docs.analytics.alpha.jisc.ac.uk/docs/study-goal/Home (Accessed 18 May, 2018)

Johnson, J. A. (2014). The ethics of big data in higher education. *International Review of Information Ethics*, *21*(21), 3-10.

Kahu, E. R. (2013). Framing student engagement in higher education. *Studies in Higher Education*, 38(5), 758–773.

Kahu, E. R., & Nelson, K. (2018). Student engagement in the educational interface: understanding the mechanisms of student success. *Higher Education Research & Development*, *37*(1), 58-71.

Krause, K., & Coates, H. (2008). Students' engagement in first-year university. *Assessment & Evaluation in Higher Education*, 33(5), 493–505.

Kruse, A. N. N. A., & Pongsajapan, R. (2012). Student-centered learning analytics. *CNDLS Thought Papers*, 1-9. https://cndls.georgetown.edu/m/documents/thoughtpaper-krusepongsajapan.pdf (Accessed 18 May, 2018)

Kumar, V., & Chadha, A. (2011). An empirical study of the applications of data mining techniques in higher education. *International Journal of Advanced Computer Science and Applications*, *2*(3).

Lust, G., Elen, J., & Clarebout, G. (2013). Students' tool-use within a web enhanced course: Explanatory mechanisms of students' tool-use pattern. *Computers in Human Behavior*, 29(5), 2013–2021.

Mann, S. (2001). Alternative perspectives on the student experience: Alienation and engagement. *Studies in Higher Education*, 26(1), 7–19.

Mezirow, J. (2018). Transformative learning theory. In K. Illeris, (Ed.). *Contemporary theories of learning: learning theorists... in their own words*. London & New York: Routledge. (pp. 114 -128)

Roberts, L. D., Howell, J. A., Seaman, K., & Gibson, D. C. (2016). Student attitudes toward learning analytics in higher education:“ The fitbit version of the learning world”. *Frontiers in psychology*, *7*, 1959.

Rubel, A., & Jones, K. M. (2016). Student privacy in learning analytics: An information ethics perspective. *The Information Society*, *32*(2), 143-159.

Siemens, G., & Gašević, D. (2012). Special Issue on Learning and Knowledge Analytics. *Educational Technology & Society, 15*(3), 1–163.

Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist*, *57*(10), 1510-1529.

Slade, S., & Prinsloo, P. (2015). Student perspectives on the use of their data: between intrusion, surveillance and care. *European Journal of Open, Distance and E-learning*, *18*(1).

Thomas, L. (2002). Student retention in higher education: The role of institutional habitus. *Journal of Education Policy*, 17(4), 423–442.

Trowler, V., & Trowler, P. (2010). *Student engagement evidence summary*. York: The Higher Education Academy.

Zepke, N., & Leach, L. (2010). Improving student engagement in post-compulsory education: A synthesis of research literature. *A report prepared for Teaching and Learning Research Initiative, Wellington, New Zealand.* *http://tlri.org.nz/sites/default/files/projects/9261-Literature-review.pdf* (Accessed 18 May, 2018)