Enhancement Themes

Evidence for Enhancement:
Improving the Student Experience

Learning Analytics: Working with and understanding data

Heather Gibson and Alan Clarke
The Open University

About this paper

This paper has been written for institutional managers and academics who are using, or wish to use, learning analytics to support the enhancement of the student experience. The aim of the paper is to help inform conversations with learning analytics experts in their institutions about some of the issues and challenges that are emerging from the learning analytics research field that may impact on institutional activities.

An overarching trend is the need to increase capacity for institutional staff and students to engage with ethics, design, understanding and using learning analytics. Where this has previously been the concern of a relatively small number of experts, it is becoming increasingly important that a broader community is equipped to participate in the conversation.

This is one part of a longer discussion paper[1] based on an adaptation of Clow's 2012 cycle of learning analytics. The main paper includes four key sections:

- data creation and collection
- **working with and understanding data (this paper)**
- using data to enhance the student experience
- implementing learning analytics in institutions.

The main paper also includes a series of 'hot topics', which we have made available as separate factsheets.

---

[1] www.enhancementthemes.ac.uk/docs/ethemes/evidence-for-enhancement/learning-analytics-discussion-paper.pdf.

Working with and understanding data

Vast quantities of information are produced by students engaging with online systems. Summarising and presenting that information in easy-to-understand and compelling formats can help users and decision-makers to interact with the data. Making the data more accessible can motivate users (including students) to design and implement interventions because it is easier to see where these may be most effective (Sclater, 2017, p 99).

The most commonly used tools by higher education institutions to visualise data are dashboards, which sometimes use predictive models. As most institutional policy managers and academics will be familiar with dashboards and predictive models, these are the tools discussed in this paper.

Dashboards

Data is extracted from the various data systems in the institution, analysed, processed (perhaps through a predictive model) and the results summarised as tables, graphs and other data representation methods. What kind of data is extracted, how it is presented and whether it is used should depend on the identified needs of the user. Dashboards have been designed to summarise institution-wide data and key performance indicators, but they have also been designed to summarise data at module or individual level. Sclater (2017) discusses a tentative taxonomy for dashboards:

- module (performance indicators such as retention rates, demographic monitoring)
- pastoral (tutors looking at data for individual student performance and support)
- central support (institutional level, KPI).

It might also be apposite to add:

- student/learner (for students to monitor own performance).

Figure 1 illustrates an example of a dashboard summarising data for an individual student.

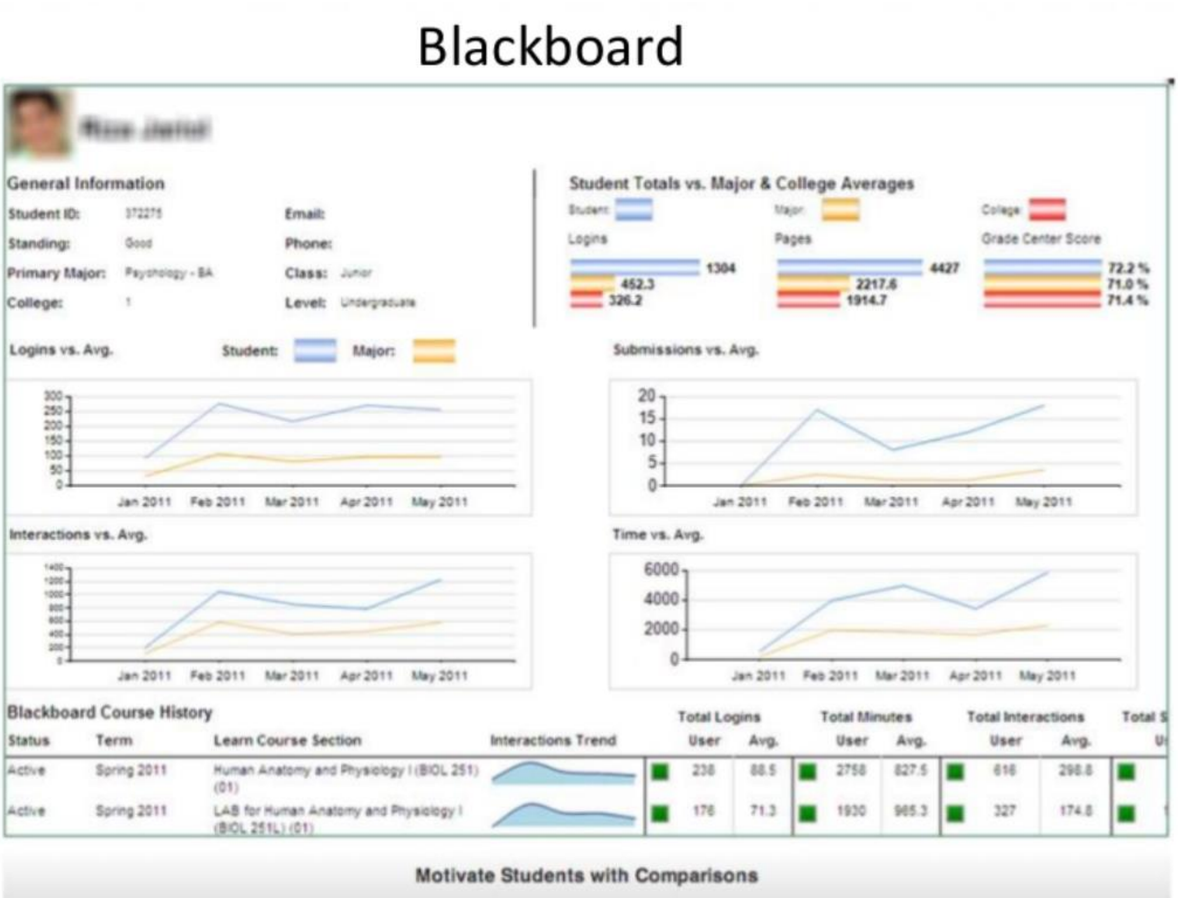

Figure 1: Learning Analytics Dashboard (Vozniuk, Govaerts, & Gillet, 2013)

Good review

Schwendimann et al (2017) carried out a review of research on learning dashboards that investigated the contexts dashboards were being used in, how well they were being evaluated, types of dashboards being used, and identifying any future directions research into learning dashboards might take. It is a very useful summary of work being carried out into learning dashboards.

Look out for: Ioana Jivet (Open Universiteit).

Predictive models

A very common use of data is to develop models that aim to predict student performance. Using predictive modelling can help institutions work with their data, identify those students particularly at risk, and target interventions that might help those students.

Predictive models are often complex, and are developed by using data collected about students' behaviours and performance. This data can be collected via VLEs, student record systems (including assessment), and interactions with other institutional systems such as library usage. Data may be static (demographic data, previous educational qualifications) or dynamic (engagement with the VLE and achievement (assessment, quiz scores)).[2] Different

[2] library.educause.edu/resources/2015/10/the-predictive-learning-analytics-revolution-leveraging-learning-data-

types of data have different strengths: because static data is relatively stable, it is used in many predictive models. Achievement data is used to measure student performance and dynamic data can yield valuable information about how students engage with course content and each other.

Figure 2 below details the data that was collected to form the probability model for the Open University's Early Alert Indicators Project (Gilmour, Boroowa, & Herodotou, 2018).

Student Probabilities Model	
Student Factors	Demographic factors including Disability, Occupation, Index of Multiple Deprivation.
Module study behaviour	Total numbers of credits being studied, late registration, proportion of assignments submitted to date.
Student's previous study	Credit transfer, previous educational qualification, sponsored or not
Student's previous progress at the OU	OU credit already achieved, number of previous OU passes, withdrawals or fails
Module within a qualification	Current module, study intensity of module, number of assignments due by milestone.

Figure 2: probability model developed by the Open University

Raw data like this is then used to produce metrics, quantitative measures that act as proxies for more complex behaviours. Metrics can be relatively simple, or more complex, but what is clear is that the more knowledge that informs a metric the more accurate proxy it will be. Therefore, metrics development may include qualitative information, such as pedagogical knowledge about curriculum requirements: a particular learning activity might support the acquisition of particular skills, or attendance at the library might be necessary to complete a task.

Metrics are then often subjected to statistical methods as part of creating the predictive model. Sclater (2017, p 88) notes three common methods: linear regression (relationship between two or more variables); logistic regression (relationship between two or more variables with the aim of calculating a probability of a student being at risk); and naïve Bayes (a type of probability analysis that assumes there is no relationship between variables, but that these variables 'contribute independently' to a probability that a student would be at risk[3]).

The outputs of predictive models are very often presented in dashboards, such as the traffic light system used at Purdue.

for-student-success
[3] blog.aylien.com/naive-bayes-for-dummies-a-simple-explanation.

Figure 3: course signals system used at Purdue (Educause)

Sclater and Mullan (2017) identified evidence that suggests that predictive models do work to improve student outcomes and their report details some examples of this. The report noted that VLE engagement appears to predict student success much more effectively than student demographic factors.

Good review

Sclater (2017) provides an excellent and accessible summary of describing metrics and predictive modelling, including examples and explaining how these three common statistical methods work. Alhadad et al (2015), for Educause, have also produced a good summary.[4]

[4] https://library.educause.edu/~/media/files/library/2015/10/ewg1510-pdf (1.03MB).

References

Alhadad, S., Arnold, K., Baron, J., Bayer, I., Brooks, C., Little, R., . . . and Whitmer, J. (2015). *The Predictive Learning Analytics Revolution: Leveraging Learning Data for Student Succes.* EDUCAUSE. Retrieved from https://library.educause.edu/resources/2015/10/the-predictive-learning-analytics-revolution-leveraging-learning-data-for-student-success.

Clow, D. (2012). The learning analytics cycle: closing the loop effectively. *Second International Conference on Learning Analytics and Knowledge*, (p. 134).

Gilmour, A., Boroowa, A., & Herodotou, C. (2018). *Using predictive analytics to support students.* Quality Assurance Agency for Higher Education in Scotland.

Schwendimann, B. A., Rodriguez-Triana, M. J., Vozniuk, A. P., Boroujeni, M. S., Holzer, A., Gillet, D., & Dillenbourg, P. (2017). Perceiving Learning at a Glance: A Systematic Literature Review of Learning Dashboard Research. *IEEE Transactions on Learning Technologies, 10*(1), 30-41.

Sclater, N. (2017). *Learning Analytics Explained.* New York: Routledge.

Sclater, N., & & Mullan, J. (2017). *Learning analytics and student success – assessing the evidence.* Jisc.

Vozniuk, A., Govaerts, S., & and Gillet, D. (2013). Towards portable learning analytics dashboards. *EPFL Switzerland.* Beijing: ICALT.