An overview of learning analytics

Journal Item

An Overview of Learning Analytics

Doug Clow

Institute of Educational Technology, The Open University, Walton Hall, Milton Keynes, MK7 6AA, UK

doug.clow@open.ac.uk

An overview of learning analytics

Learning analytics, the analysis and representation of data about learners in order to improve learning, is a new lens through which teachers can understand education. It is rooted in the dramatic increase in the quantity of data about learners, and linked to management approaches that focus on quantitative metrics, which are sometimes antithetical to an educational sense of teaching. However, learning analytics offers new routes for teachers to understand their students, and hence to make effective use of their limited resources. This paper explores these issues, and describes a series of examples of learning analytics to illustrate the potential. It argues that teachers can and should engage with learning analytics as a way of influencing the metrics agenda towards richer conceptions of learning, and to improve their teaching.

Keywords: learning analytics; analytics; metrics

Introduction

There is a tension between the framing of education as an economic activity and conceptions of education and learning that are concerned with the development of meaning and the transformation of understanding. These difficulties are far from purely theoretical concerns: they increasingly have very practical, concrete consequences for teachers and learners, notably around resource constraints, class sizes, and time pressures. Within this constrained environment, teachers are subject to accountability processes based on and enabled by the deployment of quantitative metrics of their practices.

Quantitative metrics are increasingly used not only because of theoretical framings that support them, but also because of a substantial and dramatic change in their practicability over the last ten or twenty years. This change is often referred to as Big Data: the quantity, range and scale of data that can be and is gathered has increased exponentially (or close to exponentially). Accompanying this explosion of data is a series of rapid advances in computational techniques for managing, processing and analysing these large volumes of data in ways that are actionable. These developments are transforming enquiry. The scale of data is greatest in science - for instance, the Large Hadron Collider at CERN produced 23 petabytes (23 million gigabytes) of information in 2011 (CERN, 2012). The effect is not restricted to science - for instance, the ability to manage and integrate textual and geographic data is changing scholarly practice in the classics (see e.g. Project HESTIA: the Herodotus Encoded Space-Text-Imaging Archive, http://www.open.ac.uk/Arts/hestia/index.html). New approaches become possible: for instance, rather than sampling, an entire population can be captured. The volume and scope of data can be so large that it is possible to start with a dataset and apply computational methods to produce results, and only subsequently to seek an interpretation or meaning.

Big data is by no means restricted to the academy. Technology companies such as Google and Facebook make managing staggeringly large datasets their core business, but even companies such as grocery retailers are increasingly deploying big data techniques to capture, understand, model, predict and influence consumer behaviour. The growing field of Business Intelligence is concerned with the management and processing of data to support corporate practice, including performance metrics.

The framing of education as an economic activity supports the view of educational institutions as businesses Business Intelligence is increasingly applied in higher education, in areas such as outreach and advertising, enrolment, management, and fund-raising, but also in more academic areas. 'Dashboards' showing performance metrics against targets are increasingly popular with senior managers, and political pressures such as the current focus on college completion in the US reinforce this direction.

These developments are not always welcomed by teachers. Two examples are illustrative. Texas A&M University introduced a system that calculated dollar amounts for each individual faculty member, ostensibly accounting for that person's net contribution to – or subtraction from – the university's financial position. This was not uniformly welcomed by all faculty, who argued that the figures were inaccurate and unfair (Simon and Banchero, 2010). In the United Kingdom, successive Research Assessment Exercises and the forthcoming Research Excellence Framework seek to calculate numeric values for research performance – again, to considerable controversy about validity and the effect on practice.

But what about the learners?

Learning analytics is the application of these Big Data techniques to improve learning. Learning analytics is currently a fixture in educational horizon-scanning reports (see e.g. Johnson et al 2011; Johnson, Adams and Cummins 2012; Sharples et al 2012) and in a raft of other publications aimed at practitioners and aspiring practitioners from organisations concerned with technology in education, such as Educause (http://www.educause.edu/library/analytics), JISC (http://jisc.cetis.ac.uk/topic/analytics) and SURF (http://www.surf.nl/en/themas/InnovationinEducation/learninganalytics/Pages/default.aspx). Vendors of learning technology are providing analytics packages: for instance, Blackboard, Desire2Learn, Instructure and Tribal have all released analytics tools, and there is also activity in the Moodle community. The high-profile providers of Massively Open Online Courses (MOOCs) - Coursera, Udacity and edX - are all using analytics tools to inform their practice.

There is also a growing research community around the topic. An annual conference, Learning Analytics and Knowledge, has been organised (Long et al 2011;

Buckingham Shum, Gasevic and Ferguson 2012), a special issue on the topic has been published recently (Siemens and Gasevic 2012), and an international research network set up: SoLAR, the Society for Learning Analytics Research (http://www.solaresearch.org/).

This increasing activity has a range of drivers and facilitators. Firstly, there is the pressure towards performance management, metrics and quantification. Secondly, there is an increasing volume of data available about learners and learning, particularly as more learning takes place online in Learning Management Systems or Virtual Learning Environments (LMS/VLEs). Every page visited, every interaction, every click can in theory be recorded and stored. Thirdly, statistical and computational tools to manage large datasets and to facilitate interpretation have become available as a result of Big Data activity.

The most commonly-cited definition of learning analytics emerged from an open online course on learning and knowledge analytics (LAK11, http://www.learninganalytics.net/?p=28) and was adopted by the associated First International Conference on Learning Analytics and Knowledge in 2011 (Long et al 2011):

> "the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs"

As with any field of activity, particularly new ones, drawing clear distinctions between related endeavours is problematic, contested and liable to change; however, broad outlines can be drawn. Two other emerging areas have significant overlap with learning analytics. The first is academic analytics, which is the use of business intelligence in education. This tends to focus more at the institutional and national level, rather than on individual students and courses (Long and Siemens, 2011). The second is educational data mining (EDM), which seeks to develop methods for analysing educational data, and tends to focus more on the technical challenges than on the pedagogical questions (Ferguson, 2012). Learning analytics is first and foremost concerned with learning.

A key concern in learning analytics is the need to use the insights gathered from the data to make interventions to improve learning, to generate 'actionable intelligence' (Campbell, DeBlois and Oblinger, 2007) which informs appropriate interventions. This is addressed in accounts of the learning analytics process address. Campbell and Oblinger (2007) set out five steps: Capture, Report, Predict, Act, Refine. Clow (2012) places this as the central idea in his Learning Analytics Cycle (figure 1). The cycle starts with learners, who generate data, which is processed in to metrics, which are used to inform interventions, which in turn affect learners. The learners may be students in a traditional higher education setting, or in less formal contexts. The data can include demographic information, online activity, assessment data, and final destination data. The metrics can be presented in a wide range of ways: from a simple indication of whether learning progress is being made, to a comparison of outcomes to some benchmark, or a visual picture of activity in an online forum. The interventions again range widely, from students taking action in the light of metrics showing their activity compared to that of their peers, to teachers contacting students identified as requiring some particular extra help.

[INSERT FIGURE 1 HERE]

Figure 1: The Learning Analytics Cycle, from Clow (2012).

Learning analytics is not so much a solid academic discipline with established methodological approaches as it is a 'jackdaw' field of enquiry, picking up 'shiny' techniques, tools and methodologies, including web analytics (the analysis of logs of activity on the web), social network analysis, predictive modelling, natural language processing, and more (examples and explanations of these are given below). This eclectic approach is both a strength and a weakness: it facilitates rapid development and the ability to build on established practice and findings, but it - to date - lacks a coherent, articulated epistemology of its own.

Having set out learning analytics and its context in broad terms, this paper presents a set of more concrete examples of learning analytics practice, to provide a more grounded view of the field. The examples are not intended to be exhaustive, but were selected to give an overview of the range of possibilities opened up by learning analytics. They are presented in a rough order of maturity and deployment, starting with approaches that are widely deployed and validated in use with real students, and ending with more speculative ideas under active development but not yet proven in practice.

Predictive modelling

The first example of learning analytics - in this paper, and indeed in the field - is predictive modelling. The basic concept of predictive modelling is fairly straightforward: a mathematical model is developed, which produces estimates of likely outcomes, which are then used to inform interventions designed to improve those outcomes.

Predictive modelling can be applied to education in a wide range of ways. The best-established application is estimating how likely it is that individual students will complete a course, and using those estimates to target support to students to improve the completion rate. Sophisticated mathematical techniques like factor analysis and logistic regression are applied to a large dataset containing information about previous students on the course. This information includes things that are known at the start of the course

- such as the students' previous educational experience and attainment, demographic information (such as age, gender, socio-economic status, etc), and things that become known during the course - data about their use of online course-related tools (how often they log in, how many postings that make) and formative and summative assessment data. The final key piece of information is whether the students went on to complete the course. A model is developed from this data, and then applied to the information available for current students, to give a quantified prediction of whether each student will complete the course. These predictions are typically displayed in some way to teachers, department heads, administrators and so on in a 'dashboard'.

In principle, predictive modelling is not profoundly different from a traditional teacher noticing which students are struggling in class and giving them extra help; predictive modelling could be seen as simply extending this ability to the online learning world. However, there are important practical differences. Firstly, the output of predictive modelling is a set of estimated probabilities, and it is widely established that many people struggle to correctly understand probabilities and to make consistent decisions based on probabilistic information. Secondly, the output is not (typically) restricted to a student's teacher: the information is readily made available to others beyond the immediate learning context. Thirdly, the output can be used directly to trigger actions and interventions without involving a teacher at all.

It is important to stress that the predictive power of these models is far from perfect. Not only do they produce probabilities, but those probabilities suffer from significant error: it is not possible to perfectly and accurately predict the chances of a student completing a course based on the data available. However, they are significantly more often right than wrong, and it is possible to use them to improve student completion.

Course Signals at Purdue

The Course Signals project at Purdue University (http://www.itap.purdue.edu/studio/signals/) is the most prominent and arguably the most successful application of predictive modelling to student completion in higher education.

The predictive model at the heart of Signals was first developed by Campbell (2007), and is based on four components: demographic characteristics, previous academic history, interaction with the LMS/VLE during the course, and performance on the course to date (Arnold 2010). The predictions from the model are translated in to a signal: green, denoting a high chance of success; yellow, denoting potential problems; or red, denoting a high chance of failure.

Teachers run the model and generate signals for the students on their course. The teacher can then choose what interventions to trigger: sending a personalised email or text, posting the signal on the LMS (where the student alone can see it), referral to support services, or arranging a face to face meeting.

The first pilot deployment of Signals was in 2007, and it is not applied to all courses at Purdue. Results so far are impressive (Arnold and Pistilli 2012). Overwhelmingly, students' signals tend to improve over a course, rather than worsen. This in-course improvement is reflected in improved grades: the increases vary between courses, but all see an improvement on previous semesters when Signals was not used, with an average of 10 percentage-point increase in grades A and B, and a 6 percentage-point decrease in grades D, F and withdrawals. There is increased overall retention: of the 2007 cohort, 69% of students with no exposure to Signals are retained, compared to 87% of students with exposure to at least one course using Signals. Qualitative feedback is very largely positive too, with students reporting that they perceive the emails as personal contact, and faculty reporting that the tool helps them provide help to students, and that Signals leads to students becoming more pro-active in seeking support. There is at least anecdotal evidence that students carry the support-seeking behaviours from one course to another, even where the subsequent course does not use Signals (John Campbell, personal communication, May 2012). Importantly, the Course Signals are not used in a decontextualised environment: the teacher is central to the process, and uses their judgement to direct students to appropriate existing resources within the university.

Other implementations

Predictive modelling has been used in many different universities (see e.g. Campbell, DeBlois and Oblinger 2007), often with powerful results. It is quite possible to transfer the overall approach between contexts, models themselves cannot be transferred, and significant work is needed to develop an implement a successful system: there will be variation not only in what data is available, but in its predictive power. As an example, a project at the UK Open University (RETAIN, http://retain.open.ac.uk/) found that the level of activity itself was not predictive of success or failure, but a fall-off in activity was a clear indicator of trouble (Wolff and Zdrahal 2012): students could be successful without being active online, but if a previously-active student stopped being so, they were unlikely to complete.

Far from all modelling efforts are written up and made available to the research community, particularly where the tools used are part of a proprietary system. One notable exception is Desire2Learn, which in addition to being a high-tier sponsor of the first two Learning Analytics and Knowledge conferences (LAK11 and LAK12) has also published details of the approach to predictive modelling it uses in its products (Essa and Ayad 2012).

Social network analysis

Social network analysis (SNA) is a set of methods for analysing the connections between people in a social context, using techniques from the computer science field of network analysis. Individual people (or, more technically, actors) in the social context are called nodes, and the connections between them are called ties or links. A map (a social network diagram, or sociogram) can be drawn by treating the nodes as points and the connections as lines between them as lines. So, for instance, in an online forum, the nodes might be the individual participants, and the ties might indicate replies by one participant to another's post. These diagrams can be interpreted simply by eye (for example, you can see whether a network has lots of links, or whether there are lots of nodes with few links). Alternatively, they can be interpreted with the aid of mathematical analysis of the network.

SNAPP

Social Networks Adapting Pedagogical Practice (SNAPP, http://www.snappvis.org/; Bakharia and Dawson 2011) is a social network analysis tool specifically developed for online learning contexts (Dawson 2010).

SNAPP allows teachers to track learner activity in the forums of a LMS/VLE over time, displaying a social network diagram with the individual learners indicated by a red circle, and the links between them as lines. SNAPP makes it easy for teachers to identify, for instance, learners who are entirely disconnected from the network (and hence are not fully participating), or learners who are central to the network (and hence are key enablers of the conversation). It also helps teachers to identify the pattern of interaction in the forum - whether, for example, it is largely teacher-centric, or is more diffuse with stronger peer interaction. Another use is to identify self-contained groups, or cliques, who interact with each other but not with those beyond the group. SNAPP is designed to be easy to use, but this does mean that SNAPP is not as flexible and powerful a tool for analysing social networks as the more general tools.

More advanced analysis

SNAPP lies at one end of a scale of complexity of social network analysis: it focuses on a single forum (at any one time), and the links between the nodes are simply whether a person has replied to another person's forum posting. It is possible to use SNA in more complex educational contexts. For instance, Suthers and Chu (2012) used SNA to explore the Tapped In community for educational professionals (http://tappedin.org). Their approach, inspired by Actor-Network Theory, was much more detailed and rich, based on an 'associogram', rather than a simple social network diagram: a complex multidirectional mapping of the participants, the artifacts they created (e.g. messages in chatrooms, postings in discussions, shared files), and the actions taken by the participants on those artifacts (e.g. writing/posting, and reading). Essentially, they were able to identify real communities purely from their online activity on the site, without directly using information about their affiliation, geographic location, and so on. This approach could be applied, for example, to identify communities among student populations, which could be used to better inform decisions about group work, placements, project or assignment topics, and so on.

Usage tracking

The data for learning analytics can come from student activity in the LMS/VLE, or in similar online community environments. It can also come from students' use of any application on a computer. Many tools exist to capture what a user does on a computer over time, and these can be used as a source of data about student activity when the learning task requires them to use something beyond the LMS/VLE.

For example, Santos et al (2012) developed a dashboard for students on a software development course at the Katholieke Universiteit Leuven. The students use a range of applications (word processor, programming environment, and web browser) to carry out a software development assignment, working in groups. Their activity was logged by time-tracking software (including RescueTime, http://rescuetime.com), and this data was presented in a dashboard. They could see, for instance, whether they were spending more or less time on email or writing code or looking things up than their peers, and how their web browsing compared. Feedback from students was on balance positive, but not very strongly so.

It is not yet clear how valuable this sort of approach can be for improved student learning. It also raises questions about what sorts of feedback and information are helpful, and also ethical concerns around privacy and monitoring, to which this paper will return in the Discussion.

Content analysis and semantic analysis

The examples discussed so far have concerned essentially quantitative data generated by students. However, advances in computation, in fields such as natural language processing and latent semantic analysis, make it possible to analyse qualitative, textual data - not just in terms of simple frequency counts (how many times particular words are used), but in richer, more meaningful ways.

For example, Lárusson and White (2012) have developed the Point of Originality tool, which enables teachers to track how students develop originality in their use of key concepts over the course of a series of writing assignments. The data in this context is the students' writing itself, analysed using a sophisticated database of English (WordNet, http://wordnet.princeton.edu/). The teacher types in the key words they want to explore, selects which student's work they want to examine, and the tool displays a series of coloured markers for each assignment, with bigger and 'hotter'-coloured markers indicating more original use of the key words. Clicking on a marker displays the writing sample in question. A trial on an introductory general course on computing showed a strong correlation between originality scores in the Point of Originality tool and the grades achieved for the final assessment, and also between the originality of their writing and the quantity of their contributions online.

A more speculative example is automated feedback to students about the nature of their online writing, with the aim of improving the quality of educational dialogue. Several frameworks for analysing and characterising the nature of educational dialogue have been developed, including the work of Neil Mercer and colleagues on exploratory talk in classrooms (see e.g. Mercer and Littleton 2007). This work has been applied to the analysis of online educational discussion (Ferguson and Buckingham Shum 2011) to identify places where exploratory talk took place, which could for learners visiting archived discussions to find the most useful material. These methods could be used to analyse students' contributions to an online forum, giving them feedback about the degree to which their online talk is exploratory (or matches other criteria for constructive educational dialogue), and offering suggestions for ways in which they might contribute more effectively.

The learning analytics community does not at present encompass the growing field of automated assessment, but there are many strong parallels, and one could argue that automated assessment is a particular form of learning analytics. In particular, any of the many tools under development to support marking of summative written assignments could be deployed in a formative way as cues for intervention.

Recommendation engines

Recommendation engines (or recommenders) are computational tools that provide suggestions to individuals for items they may be interested in, based on analysis of the behaviour of many users. The most famous example is Amazon's 'Customers Who Bought This Item Also Bought' feature; Amazon also uses a recommendation engine to suggest purchases based on a customer's purchase history, and on the ratings they have given to other products. The same techniques can be applied in an educational context. So, for example, a system could suggest learning resources to a student based on what resources they have previously used or found helpful, and on other students' behaviour and ratings.

However, it may be problematic to apply this approach in the context of a conventional university with a set curriculum: students typically are offered relatively little choice about the direction of their study, and so have less need for an automated system to suggest learning resources that might be helpful. It may have more application at higher levels of study, and perhaps the greatest potential benefit lies in more open-ended and less formal learning contexts.

Discussion

These examples show some of the potential benefits of learning analytics. They raise a series of implications for teachers in higher education.

The first and perhaps most obvious area is the ethics of personal data. Foucault (1991) uses Bentham's Panopticon as a symbol of how institutions and power structures enforce self-surveillance and control through the belief that scrutiny may occur at any time. The nightmare vision of Big Data for individuals is that the system does not rely on self-surveillance to enforce a disciplinary regime: all actions are captured, logged, and analysed, so transgressions will be noted regardless of whether the jailer happens to be looking in the right direction. A more positive vision of widespread disclosure is Brin's (1999) conception of a transparent society, where surveillance by those in power is held in check by openness and 'sousveillance' by those not in power, using increasingly widely-available tools for capturing and analysing data (e.g. political demonstrators streaming video live online from their phones). These radical visions of little or no privacy, and of highly-informed and capable sousveillance, are some way from the current situation.

In practice, surveillance is far from complete, and is circumscribed by a regime of policies on the ethical capture and use of personal information. In almost all developed world jurisdictions outside the US, there is comprehensive data protection legislation that requires that issues of informed consent, data accuracy, appropriateness of use, and access to information by the individual are addressed. Universities themselves typically have policies on data governance, and in a research context, any learning analytics activity will have to pass scrutiny by a body such as an Institutional Review Board or Ethics Committee. However, learning analytics is often applied outside an explicit research context; practitioners then have the responsibility to ensure that their practice meets those ethical standards.

Being open about learning analytics with students can improve their perceptions of the activity (as with Signals), but openness need not and arguably should not be complete in learning contexts. The opportunity to learn by making mistakes in a safe context can be a powerful learning experience, and far from all learners are happy to have their mistakes kept on record for all time.

Students typically know and care more about their own learning situation than even the most dedicated teacher. In numerate disciplines many students are quite capable of making intelligent use of data about their learning. Using learning analytics, they can be encouraged to take personal responsibility for their own situation - making use of the feedback available about what they're doing, and making appropriate decisions about support.

Teachers too have responsibilities. Educators have a professional responsibility to use tools and methods that can improve student learning, and learning analytics offers potentially powerful ways of doing this. A learning analytics system can reveal information about students, which leads to new ethical challenges. If you know before they start that a potential student is extremely unlikely to complete, should you admit them? Or will that simply reinforce existing power structures that put them in that position? Feeding back a negative view of a student's prognosis needs to be handled sensitively and appropriately. Learning analytics offers the possibility of focusing resources on where they are most needed. However, if resources are directed entirely towards students who are in danger of failure, there is a risk of short-changing the experience of stronger students. The experience of Signals at Purdue suggests that this need not be the case - as described above, there was a greater improvement in high grades than there was a reduction in fail grades.

As a field, learning analytics is data-driven, and is often atheoretical, or more precisely, is not explicit about its theoretical basis. Several authors have sought to ground learning analytics in theory (e.g. Clow 2012; Suthers et al 2008; Dawson 2008; Atkisson and Wiley 2011), but this is not universal, running the risk of treating the data that has been gathered as the data that matters. The choice of what is measured - in learning analytics terms, the selection of metrics - is critical. If an educational system is designed to optimise metrics that do not encompass learning, it is likely that learning will be optimised away. For those who care about learning, the choice is to attempt total resistance to the regime of metrics, or to take a more pragmatic course and insist on the inclusion of appropriate metrics that do reflect learning.

This raises the crucial question of assessment. If assessment does not reflect and reward those aspects of learning that are valued, a learning analytics system that improves assessment scores will not improve learning. Concerns about the appropriateness and reliability of assessment practices are far from new (e.g. Rowntree 1987), but analytics places a new weight and scrutiny on assessment.

Conclusion

The promise of learning analytics is the empowerment of teachers and students to understand the wealth of data that relates to their learning. Engaging in this process is a way of taking control of the agenda, so that the economic framing can be at least supplemented with a concern for learning. It is not a simple or straightforward process, and a focus on the data alone is not sufficient: to achieve institutional change, learning analytics data need to be presented and contextualised in ways that can drive organisational development (Macfadyen and Dawson 2012).

Learning analytics is a new technology, which affords new social actions. The question of the nature of technology and its relationship to existing power relationships and structures is well beyond the scope of this article, but it seems clear that educational data and can and will be used in attempts to reinforce the status quo. Ewing (2011), in a comprehensive demolition of the use of value-added modelling for evaluating teacher and school performance, argues that mathematics is often misused:

> "as a rhetorical weapon - an intellectual credential to convince the public than an idea or process is "objective" and hence better than other competing ideas or processes. This is mathematical intimidation. It is especially persuasive because so

> many people are awed by mathematics and yet do not understand it - a dangerous combination."

This neatly captures the main risks of the use of analytics in difficult times. The process of data gathering and interpretation is proceeding apace in higher education, often driven by the demands and worldview of managers and the economic framing of education. There is value - and not just in the economic sense - for teachers in more information about their students. The opportunity afforded by learning analytics is for educators to refuse to be overawed by the process, to understand the tools and techniques, their strengths and limitations, and to use that understanding to improve teaching and learning.

References

Arnold, Kimberly. 2010. "Signals: Applying Academic Analytics." *Educause Review*, 33 (1).

Arnold, Kimberly and Matthew Pistilli. 2012. "Course Signals at Purdue: Using Learning Analytics to Increase Student Success." In Buckingham Shum, Gasevic and Ferguson (2012), 267-270.

Atkisson, Michael and David Wiley, 2011. "Learning analytics as interpretive practice: applying Westerman to educational intervention." In Long et al (2011), 117-121.

Bakharia, Aneesha and Shane Dawson. 2011. "SNAPP: a bird's-eye view of temporal participant interaction." In Long et al (2011), 168-173.

Brin, David. 1999. *The Transparent Society: Will Technology Force Us to Choose Between Privacy and Freedom?* New York: Basic Books.

Buckingham Shum, Simon, Dragan Gasevic and Rebecca Ferguson. 2012. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK12)*, Vancouver, BC, Canada — April 29 - May 02, 2012. New York: ACM.

Campbell, John P. 2007. *Utilizing student data within the course management system to determine undergraduate student academic success: An exploratory study*. PhD diss., Purdue University.

Campbell, John P., and Diana Oblinger. 2007. "Academic Analytics". *Educause Quarterly*, October 2007. http://net.educause.edu/ir/library/pdf/pub6101.pdf

Campbell, John P., Peter B. De Blois, and Diana G. Oblinger. 2007. "Academic Analytics: A New Tool for a New Era". Educause Review, 42 (4): 40-57.

CERN (European Organization for Nuclear Research), 2012. *CERN Annual Report 2011: Fifty-seventh Annual Report of the European Organization for Nuclear Research*. Geneva: CERN. http://library.web.cern.ch/library/content/ar/yellowrep/varia/annual_reports/2011/AnnualReport2011-en.html

Clow, Doug. 2012. "The Learning Analytics Cycle: Closing the Loop Effectively". In Buckingham Shum, Gasevic and Ferguson (2012), 134-138.

Coffield, Frank, David Moseley, Elaine Hall, and Kathryn Ecclestone. 2004. *Learning Styles and Pedagogy in Post-16 Learning: A Systematic and Critical Review*. London: Learning and Skills Research Centre.

Dawson, Shane. 2008. "A study of the relationship between student social networks and sense of community." *Educational Technology and Society*, 11 (3): 224–238.

Dawson, Shane. 2010. "'Seeing' the learning community: An exploration of the development of a resource for monitoring online student networking." *British Journal of Educational Technology*, 41 (5): 736–752.

Deakin Crick, Ruth and Guoxing Yu. 2008. "Assessing learning dispositions: is the Effective lifelong learning inventory valid and reliable as a measurement tool?" *Educational Research*, 50 (4): 387–402.

Essa, Alfred and Hanan Ayad. 2012. "Student Success System: Risk Analytics and Data Visualization Using Ensembles of Predictive Models". In Buckingham Shum, Gasevic and Ferguson (2012), 158-161.

Ewing, John. 2011. "Mathematical Intimidation: Driven by the Data". *Notices of the AMS*, 58 (5): 667-673.

Ferguson, Rebecca and Simon Buckingham Shum. 2011. "Learning analytics to identify exploratory dialogue within synchronous text chat." In Long et al (2011), 99-103.

Ferguson, Rebecca. 2012. *The State of Learning Analytics in 2012: A Review and Future Challenges.* Technical Report KMI-12-01. Milton Keynes: Knowledge Media Institute, The Open University. http://kmi.open.ac.uk/publications/techreport/kmi-12-01

Foucault, Michel. 1991. *Discipline and Punish: The birth of the prison*. Translated by Alan Sheridan. London: Penguin.

Johnson, L., S. Adams, and M. Cummins. 2012. *The NMC Horizon Report: 2012 Higher Education Edition*. Austin, TX: The New Media Consortium. http://www.nmc.org/pdf/2012-horizon-report-HE.pdf

Johnson, L., R. Smith, H. Willis, A. Levine, and K. Haywood. 2011. *The 2011 Horizon Report.* Austin, TX: The New Media Consortium. http://net.educause.edu/ir/library/pdf/hr2011.pdf

Lárusson, Jóhann Ari and Brandon White. 2012. "Monitoring student progress through their written "point of originality"". In Buckingham Shum, Gasevic and Ferguson (2012), 212-221.

Long, Phil and George Siemens. 2011. "Penetrating the Fog: Analytics in Learning and Education". *Educause Review,* 46 (5): 31-40.

Long, Philip, George Siemens, Gráinne Conole and Dragan Gašević (2011). *Proceedings of the 1st International Conference on Learning Analytics and Knowledge (LAK11)*, Banff, AB, Canada, Feb 27 - Mar 01, 2011. New York: ACM.

Macfadyen, Leah and Shane Dawson. 2012. "Numbers Are Not Enough. Why e-Learning Analytics Failed to Inform an Institutional Strategic Plan." *Educational Technology and Society*, 15 (3): 149–163.

Mercer, Neil and Karen Littleton. (2007). *Dialogue and the development of children's thinking: a sociocultural approach*. London: Routledge.

Rowntree, Derek. 1987. *Assessing Students: How Shall We Know Them?* London: Routledge.

Santos, Jose Luis, Sten Govaerts, Katrien Verbert, and Erik Duval. 2012. "Goal-oriented visualizations of activity tracking: a case study with engineering students". In Buckingham Shum, Gasevic and Ferguson (2012), 143-152.

Sharples, Mike, Patrick McAndrew, Martin Weller, Rebecca Ferguson, Elizabeth FitzGerald, Tony Hirst, Yishay Mor, Mark Gaved, and Denise Whitelock. 2012. *Innovating Pedagogy 2012: Open University Innovation Report 1.* Milton Keynes: The Open University. http://www.open.ac.uk/personalpages/mike.sharples/Reports/Innovating_Pedagogy_report_July_2012.pdf

Siemens, George and Dragan Gasevic, eds. (2012). "Learning and Knowledge Analytics." Special issue, *Journal of Educational Technology and Society*, 15 (3).

Simon, Stephanie and Stephanie Banchero. 2010. "Putting a Price on Professors". *Wall Street Journal*, October 22, 2010.

Suthers, Dan and Kar-Hai Chu. 2012. "Multi-mediated community structure in a socio-technical network". In Buckingham Shum, Gasevic and Ferguson (2012), 43-53.

Suthers, Daniel, Ravi Vatrapu, Richard Medina, Samuel Joseph, and Nathan Dwyer. 2008. "Beyond Threaded Discussion: Representational Guidance in Asynchronous Collaborative Learning Environments." *Computers and Education*, 50, 4, 1103-1127.

Wolff, Annika and Zdenek Zdrahal (2012). "Improving Retention by Identifying and Supporting "At-Risk" Students." *Educause Review Online*, http://www.educause.edu/ero/article/improving-retention-identifying-and-supporting-risk-students