Online Learner Engagement: Opportunities and Challenges with Using Data Analytics

Robert Bodily
Charles R. Graham
Michael D. Bush

This article describes the crossroads between learning analytics and learner engagement. The authors do this by describing specific challenges of using analytics to support student engagement from three distinct perspectives: pedagogical considerations, technological issues, and interface design concerns. While engaging online learners presents a major challenge to the educational community, the affordances of online learning environments and new advances in data analytics present many opportunities that can lead to improvements in our ability to fully engage students. The authors' work specifically underscores the fact that no single tool or system will be able to address all of the challenges involved. As a result, tools and content need to be open, interoperable, and modular. The important field of learning analytics will develop much more quickly when tools and work are shared across organizations and projects.

Robert Bodily is a graduate researcher in the Instructional Psychology and Technology department at Brigham Young University. His dissertation research focuses on student-facing learning analytics reporting systems, which includes recommender systems, dashboards, and feedback systems. His research interests include learning analytics, educational data mining, learning analytics dashboards, and student engagement (e-mail: bodilyrobert@gmail.com). **Charles R. Graham** is a Professor at Brigham Young University who studies technology-mediated teaching and learning, with a focus on the design and evaluation of blended and online learning environments. He also researches the use of technology to enhance traditional teaching and learning (e-mail: charles.graham@byu.edu). **Michael D. Bush,** a Contributing Editor, recently retired from Brigham Young University as Associate Professor of French and Instructional Psychology and Technology. During his last 15 years there he directed a lab that developed architectures for delivering and studying online learning as well as produced video and award-winning online, multimedia applications. Support for eLearning standards for such as SCORM, the Experience API, and LTI has been central to his work (e-mail: michael_bush@byu.edu).

Introduction

Instructional methods informed by the developing field of data science offer great promise "for increasing the effectiveness of teaching, learning, and schooling" (Dede, 2016). In this digital age of big data, it is often assumed there is a wealth of educational data ready to be analyzed and used for data-driven decision making (Office of Educational Technology, 2013). However, from personal experience, we have found this is rarely the case. There are many challenges in: (1) selecting what data to capture, (2) collecting the data, (3) aggregating the data, (4) reporting the data to stakeholders, and (5) incorporating meaningful educational constructs into the data analysis. We provide an overview of how we have addressed these issues in a project that is investigating ways that big data and data analytics can be useful in researching online learner engagement and its relationship to learning outcomes. We first present a brief review of student engagement along with how it can be useful in our data analytics efforts.

It is not difficult to find research that references the benefits of student engagement. Engaged learning is associated with activity, energy, time on task, and many other positive characteristics. Researchers have found positive relationships between student engagement and outcomes such as academic achievement (Hughes, Luo, Kwok, & Loyd, 2008; Ladd, & Dinella, 2009), student learning (Carini, Kuh, & Klein, 2006), student satisfaction (Filak & Sheldon, 2008; Zimmerman & Kitsantas, 1997), persistence (Berger & Milem, 1999; Kuh *et al.*, 2008), and even student health and wellbeing (Van Ryzin, Gravely, & Roseth, 2009). Unfortunately, lack of engagement has been identified as a contributor to lower completion rates in online learning courses (Kizilcec, Piech, & Schneider, 2013) with some research showing even greater challenges with disengagement among at-risk and minority students (Rovai, 2003). This is particularly problematic as online learning continues to grow and to take on a more significant role in mainstream K–12 and higher education.

While engaging online learners presents a major challenge to the educational community, the affordances of online learning environments and new advances in data analytics also present many opportunities that can lead to improvements in our ability to fully engage students. This article will attempt to describe the crossroads between learning analytics and learner engagement. We will do this by describing specific challenges of using analytics to support student engagement from three distinct perspectives: pedagogical considerations, technological issues, and interface design concerns.

Technical Terms and Definitions

- **Experience API (xAPI, or formerly known as Tin Can API):** An e-learning data format specification, which allows learning applications to pass information back and forth.

- **Learning Tools Interoperability (LTI):** An e-learning specification allowing e-learning applications to integrate with learning management systems.

- **Learning Management System (LMS):** An e-learning platform that allows instructors to manage course materials, assess learners, and monitor student participation.

- **Learning Analytics:** The process of selecting, collecting, analyzing, and reporting data about learners and their interactions with online learning resources to improve teaching and learning.

- **Data Mining:** Using statistical techniques and machine learning to find hidden patterns in data.

Pedagogical Considerations

In the traditional face-to-face classroom, instructors are constantly monitoring the engagement levels of their students and adjusting their instruction accordingly. In this way, instruction is being adapted in real-time to meet the needs of students. The ability to monitor and adjust instruction based on student performance is significantly diminished once the student leaves the classroom. At best, traditional instructors know how students are engaging with content outside of the class through the delayed lens of course assessments or through interviews and surveys. However, these do not happen in real-time and remove students' time and focus away from the instruction.

As online and technology-mediated instruction become more prevalent, it is possible to capture data that can help educators understand the engagement level of students in the learning that takes place outside of the face-to-face classroom. If done effectively, this knowledge could allow instructors, parents, personalized learning systems, and even students themselves to make adjustments to instruction when it is apparent that students are not engaging fully.

However, before the potential benefits of monitoring learner engagement using analytics can be realized, there are several significant challenges that need to be understood and addressed.

1. **Defining Engagement:** We need a better understanding of what engagement is.
2. **Measuring Engagement:** We need to understand how to measure engagement.
3. **Acting on Engagement:** We need to understand how instructors should adjust online instruction based on engagement data.

Defining Engagement: What Is Student Engagement?

One might imagine that because the term student engagement is commonly used, there might also be a common definition or understanding of what is meant when we say *engagement*. It turns out that there is not agreement among educational researchers as to what engagement means (Christenson, Reschly, & Wylie, 2012; Henrie, Halverson, & Graham, 2016 in press; Reschly & Christenson, 2012; Sinatra, Heddy, & Lombardi, 2015). One prominent model of student engagement breaks down engagement into areas of *emotional* engagement, *behavioral* engagement, and *cognitive* engagement (Fredricks, Blumenfeld, & Paris, 2004). Other researchers include additional constructs like *agentive* engagement (Reeve & Tseng, 2011; Reeve, 2012). Upon closer inspection, many of these models have overlapping definitions for their constituent parts (Fredricks, Blumenfeld, Friedel, & Paris, 2005; Henrie *et al.*, 2016 in press). A common understanding of the definition of engagement is essential to addressing the next challenges of knowing how to best measure it and how to use the engagement data to guide interventions.

Measuring Engagement: How Do We Quantify and Report on It?

The level at which engagement is being investigated also has implications for how engagement is conceptualized, operationalized, and measured. Skinner and Pitzer (2012) outline four different levels at which engagement is defined: (1) institutional level, (2) school level, (3) classroom level, and (4) activity level. Engagement is measured differently at each level. A few examples that are used to look at student engagement at the institutional or program level include Purdue's Course Signals system (Arnold & Pistilli, 2012) and the Open Academic Analytics Initiative (Lauria, Moody, Jayaprakash, Jonnalagadda, & Baron, 2013). These systems typically provide data to administrators and are not helpful to instructors in making day-to-day decisions about their teaching activities. At the course and activity levels, engagement is often measured through the use of self-report surveys that are completed at the end of the course or activity. While this information is often useful, it also presents some significant limitations. First, as you move down to the activity level, it becomes less scalable, as it is not practical to give surveys after each activity.

Second, the data is being collected retrospectively, and therefore, the ability to make adjustments to instruction mid-implementation is not possible. Finally, self-report data has limitations because participants do not always remember their in-the-moment engagement accurately after the fact.

Analytics collected by the LMS or other instructional system has the potential to provide student engagement profiles in real-time, eliminating the need to interrupt instruction to collect survey data or wait until the end of instruction for reporting. However, more research is needed to develop a reliable model of using click data to measure student engagement (Henrie, Bodily, Manwaring, & Graham, 2015; Macfadyen & Dawson, 2012).

Additionally, because of the time-intensive nature of collecting engagement data by conventional means, engagement has often been measured at the course level rather than the activity level, limiting its usefulness for making activity-specific interventions that are based on the findings.

Acting on Engagement Data: What Do We Do with It?

There are various ways to report engagement data to stakeholders to try to motivate them to act: visualizations, recommendations, push notifications, or feedback. We will discuss data reporting in more detail in the Interface Design section below. In this section, we will address the challenge of helping stakeholders take action when presented with data.

Most analytics systems either visualize data or provide recommendations, not both. Visualizing data helps stakeholders easily see trends to understand *what has been happening*. Many LMS provide data reports for instructors, but these can be overwhelming because visualizations assume instructors know what to do because of the data represented in the visualization. Providing recommendations allows stakeholders to *know what to do*, but does not help them know *why* they should do it. Analytics systems that provide visualizations and recommendations can help stakeholders know what they should do and why they should do it. Examples of these systems include a visual recommender system that promotes self-reflection and sense-making about collaboration (Anaya, Luque, & Peinado, 2016), an infographic that both informs students of major course events and provides recommendations based on previous student behavior (Ott, Robins, Haden, & Shephard, 2015), and the Student Activity Meter that supports learner awareness and reflection (Govaerts, Verbert, Duval, & Pardo, 2012). Future analytics systems should be sure to incorporate both visuals and recommendations in their system to better motivate stakeholders to make data-driven decisions.

Technological Issues

Once a theoretically sound rationale has been made for the data to be collected and analyzed, attention must then turn to the technological issues that will enable the storage and retrieval of the data. Current thinking regarding "big data" and "learning analytics" is often based on the key assumption that the data is already being collected by some combination of learning management systems (LMS) and other information technology infrastructures of learning settings (Office of Educational Technology, 2013). Unfortunately, research efforts are revealing that this assumption often does not hold up (Santos, Verbert, Klerkx, Charleer, & Duval, 2015). Our efforts at Brigham Young University (BYU) with two learning management systems confirmed this to be the case. We worked first with Canvas from Instructure and later with BYU Learning Suite, the LMS created and used at our institution. Unfortunately, neither system could collect the data that was needed for real-time analysis and reporting.

We explored these systems in the context of a project to investigate how big data and data analytics can provide explanations for how online learner engagement relates to learning outcomes. The instructional approach that is being implemented in the project involves making videos available to learners via digital streaming and, then, having them respond to questions about the information contained in the videos. The first issue we encountered was that neither system could collect the sort of data needed for the analyses required by our research questions. Specifically, we needed detailed information on the viewing experience of each student, which proved to be impossible to collect with each of the two systems we explored. The second issue involved connecting learner scores from the assessments to the detailed viewing information. Each system could deliver the assessments, but neither could make the information available to whom and when it would be needed.

The problems relevant to data collection and the assessment process are subsumed in two concepts, with each having its own specific challenges. We first provide explanations of each before proposing technological solutions:

1. **Data Granularity:** Tracking click-level student interaction data within a traditional learning management system (LMS).
2. **Data Aggregation and Reporting:** Aggregating student engagement data to the appropriate level and having real-time access to it.

Data Granularity: The Problem with LMS and Proprietary System Analytics

The typical LMS does not collect data at the level of granularity that is sufficiently detailed for developing the sort of conclusions that a well-defined pedagogical theory base on learner engagement suggests is possible. In addition, rather than depending on the design of the learning materials of interest, the design of the typical LMS determines which data is to be tracked, how to store that data, and how to give users access to that data. Often, there are limits as to how often system users can access their data, which complicates real-time analysis and data reporting. The end result is that some portion of the software with which users interact must be either replaced or modified to ensure that the necessary data is available for analysis.

Data Aggregation and Reporting: Getting Data to the Right People at the Right Time and in a Usable Format

The second problem has to do with how the data is collected, but more specifically where it is stored. For example, once data is collected it is to be easily accessible, and available in real-time.

There are not many learning applications that have this functionality, so to solve this problem we: (1) modified an existing Web application *(ayamel.byu.edu)* to make it xAPI compliant, and (2) implemented an xAPI and analytics backend to an open source learning application *(openassessments.org)*. These solutions will be further explained in the next section.

Solutions to Technological Challenges

Fortunately, two organizations are pursuing solutions that provide the means to address each of these issues. The first is the Advanced Distributed Learning Initiative (ADL) of the Federal Government, which is working under the aegis of the Department of Defense. Its effort on the Experience API (xAPI), formerly the Tin Can API, has provided the foundation for solving the types of problems we are raising here. The second effort involves the development of the Caliper Learning Analytics Framework, led by the IMS Global Learning Consortium (IMS). IMS is comprised of several technology industry and publishing companies interested in solving the sort of problem that Caliper addresses.

In addition, another project underway at BYU has explored related areas with instructional modules created in Adobe Captivate. Building specifically on xAPI, the team used Learning Suite to launch the activities, and then relied on xAPI calls to send learner performance data back to the LMS. Unfortunately, the information that could be returned was constrained by the capabilities of Captivate and was limited to data such as the name of the page viewed, the time spent on a page, and the number of page views, etc.

To address the two technological challenges we have identified, members of our research team determined that it would be possible to use Ayamel, a streaming video system for education created at BYU, to collect the needed information. Because we had control of the delivery software, we were able to include the required functionality to collect the sort of data that would be needed. The remaining part of the solution involved finding the means to store and retrieve the learner performance data in a way that would make it available when needed and within the tools that would help reduce the data to some form of reporting infrastructure and usable display.

After investigating Caliper and xAPI, we determined that at that point in time (April, 2015) xAPI was better supported by existing software. As a result, we instructed our developers to insert xAPI calls into Ayamel to send learner performance data to the learning record store (LRS). In this case we used Learning Locker, an open-source LRS (See: *http://learninglocker.net/*) running on Amazon Web Services (AWS). Learning Locker is a database that will store all analytics data and make it available as needed, immediately or at a later time, for reporting to learners, teachers, and administrators. Please note that in order to have real-time access to this data, you may need to host your own instance of an LRS.

Once we had designed a system for delivering the instructional content, we needed something for administering assessments. To address this issue, we investigated various open-source assessment tools and quickly focused our attention on two tools that seemed to best fulfill our needs, Open Assessments (See: *https:// www.openassessments.com*) and Tao Testing (See: *http://www.taotesting.com*).

A key requirement for all of our efforts was to select tools that were conformant to existing standards. In the case of assessment, those standards were Learning Tools Interoperability (LTI) and Question and Test Interoperability Specification (QTI), both from IMS.

While both of the assessment tools we identified support LTI and QTI, we ultimately decided on Open Assessments. This assessment tool did not track the kind of data we were interested in capturing, so once again we turned to xAPI. We implemented a backend on the open-source quiz tool that would use xAPI calls to send click-level events to our LRS. We also implemented a few additional features in the quiz tool, such as hint and show answer buttons.

Our analytics system is configured as shown in ***Figure 1***. Using LTI, we can easily authenticate users with our learning applications.

We then use xAPI to send data in a common format to our instance of the Learning Locker LRS. This solution allows us to have real-time, click-level data from all learners in our analytics system. We then report this data, initially to students, and eventually to teachers and administrators, in our learning analytics dashboards.

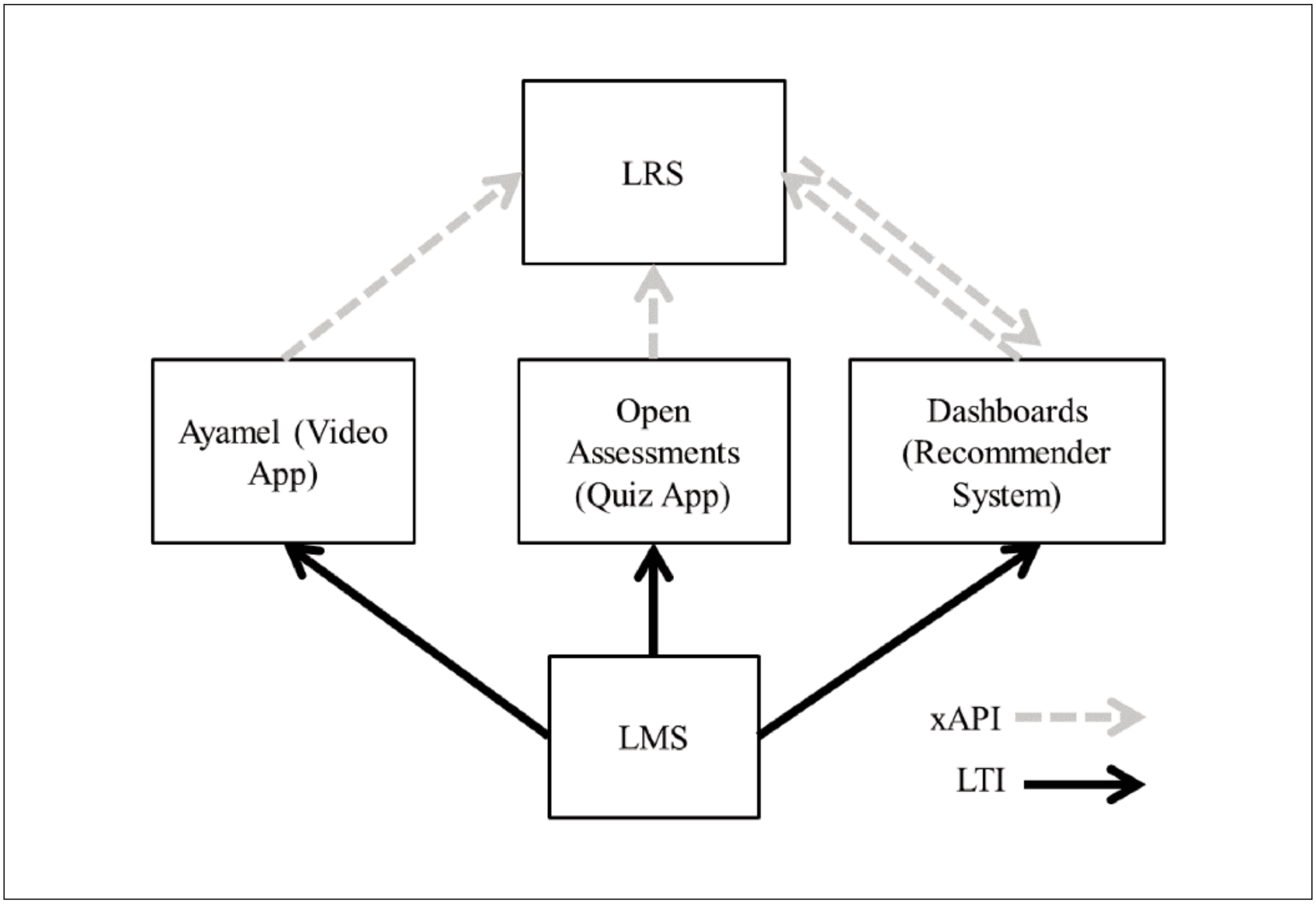

Figure 1. Learning analytics system using LTI and xAPI.

Interface Design Concerns

The reporting stage is one of the final and most important stages of the learning analytics process. No matter how well or how early in the course we can predict student engagement, if we cannot inform the appropriate stakeholders to intervene accordingly, then the rest of the system's functionality is for naught. There are a number of challenges associated with the reporting stage, namely:

1. **Goal Achievement:** How can student data help common education stakeholders achieve their goals of improving student engagement?
2. **Information Selection:** What data are useful and important to these stakeholders?
3. **Visual Representation:** In what way and how often should this data be reported to stakeholders?

These three challenges along with some potential solutions will be addressed in the following sections.

Goal Achievement: Helping Stakeholders with Engagement Data

Every stakeholder in education has unique challenges and goals, so each stakeholder should have their own personalized data reporting mechanism. ***Table 1*** shows common goals for each class of stakeholders in education.

While learning systems are often designed to support instructors and administrators in their respective goals, the data needed to inform the goals of designers, parents, and students are often not collected or reported. Learning analytics systems should focus on collecting the right data for each stakeholder in education in order to improve the entire educational system. For example, parents would like to know how their children are doing, and instructors and administrators need to know how well the system is working. With respect to designers, it is crucial to close the feedback loop in order to ensure that materials are improved over time.

Information Selection Identifying Useful Data

Because each stakeholder has different goals, different data will be necessary to help them achieve their goals. ***Table 2*** shows stakeholders along with potential data sources that might be useful to them.

Table 1. Common stakeholders and goals.

Stakeholders	Goals
Administrators	How can we achieve our goals and mission as a university?
Instructors	Who are the disengaged students? Which concepts are students struggling with?
Designers	Which parts of the course are less engaging? How can we improve them?
Parents	How can I help my child be a more engaged learner?
Students	Where are my knowledge gaps? How can I get an A in the course?

Table 2. Common stakeholders and data sources.

Stakeholders	Data Sources
Administrators	Finance, admissions, enrollment, student outcomes (graduation rate, retention rate, completion rate), student engagement, academic information, satisfaction, research, and external ratings (Terkla, Sharkness, Cohen, Roscoe, & Wiseman, 2012)
Instructors	Student-content interaction data, student assessment data, student outcomes data, and student engagement data (aggregated by student and concept)
Designers	Class averages for resource use, assessment data at the resource level, and student engagement data
Parents	Student assessment data, metrics indicating improvement over time, and student engagement data
Students	Student-content interaction data, student assessment data, student outcomes data, student engagement data, compare to class, and compare to A students

Referring to ***Table 2***, you can see why there are different names for analytics at different levels—they use completely different data sources! Academic analytics or supporting administrator goals with data is completely different from micro analytics, or supporting student goals with data.

Visual Representation: How and When to Report

The final consideration in data reporting is determining the best way to report information to assist stakeholders in achieving their goals. To best support stakeholders, it is important to consider: (1) the benefits of including both visualizations and recommendations in a dashboard, and (2) the differences in how often data should be updated based on which stakeholder is viewing the information.

Table 3. Stakeholders and the frequency at which they need data reports to be updated.

Stakeholders	Update Frequency
Administrators	Every semester
Instructors	Every day
Designers	Every semester
Parents	Every day
Students	Real-time

There are two general ways to report data: (1) in visualizations or dashboards, or (2) with recommendations or feedback. Dashboards visualize data for stakeholders so they can identify insights in a single glance. Recommendations and feedback provide stakeholders with specific suggestions on what they can do to achieve their goal based on what has happened in the past. In many cases, stakeholders view a dashboard, but do not know what to do as a result of the information they see. Likewise, many systems only provide recommendations of what stakeholders should do, but without understanding why, stakeholders may lack the motivation to follow the recommendation. Because of this, it is important to include both data visualizations and recommendations in reporting tools for stakeholders.

Another important consideration is how often the data should be updated for stakeholders. ***Table 3*** shows the frequency with which it makes sense for stakeholder data to be updated. Administrators and designers generally make decisions based on entire semester or multi-semester data sets, while instructors and parents might want daily updates to see how students are doing.

For more information on dashboard design principles in a business context, see Few (2006). For more information on learning analytics dashboards in an education context, see Verbert *et al*. (2014).

Solution to Interface Design Issues

Our solution to interface design issues is realized through the use of a student-facing learning analytics dashboard. This dashboard consists of a *content recommender dashboard* and a *skills recommender dashboard*.

Content recommender dashboard. The content recommender dashboard (see ***Figure 2***) visualizes assessment analytics data to easily show students what

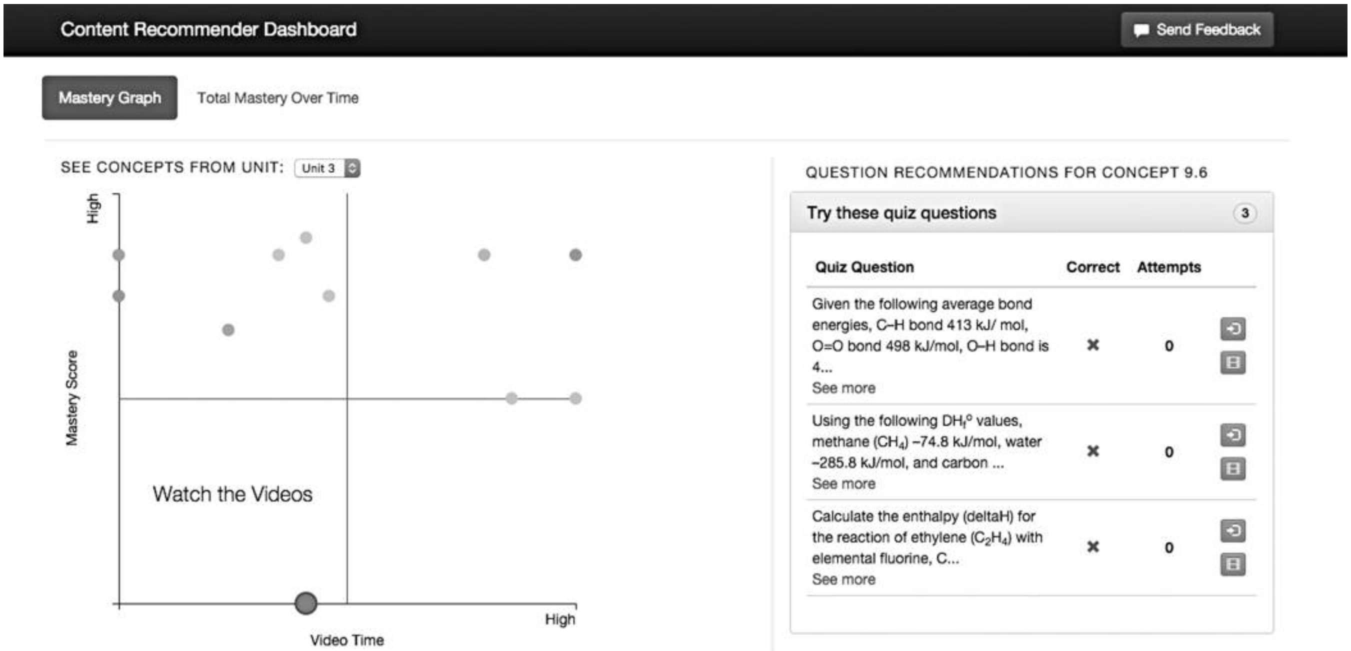

Figure 2. The Content Recommender Dashboard.

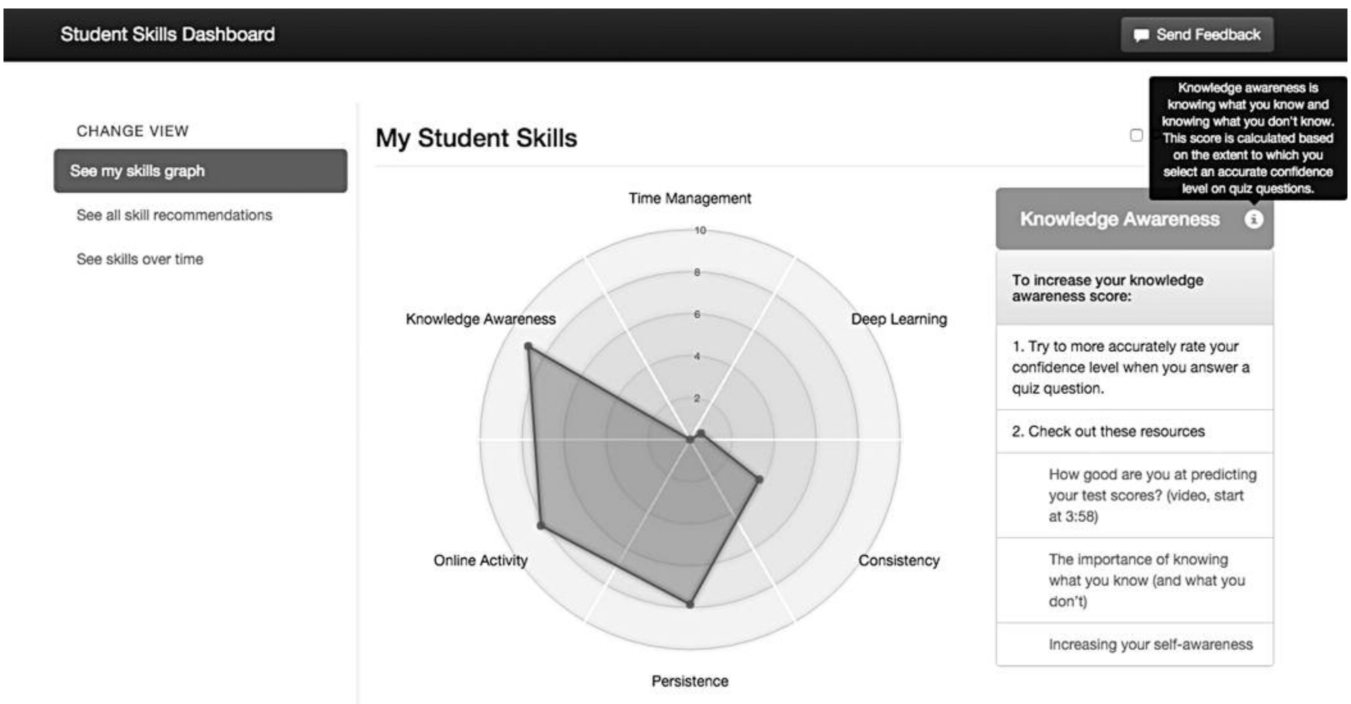

Figure 3. The Skills Recommender Dashboard.

knowledge gaps they have. When a student selects a concept, it provides clickable recommendations based on the student's activity and performance in the course. These recommendations include completing additional problems, watching related videos, contacting a friend or teaching assistant, and viewing additional information on that topic online.

Skills recommender dashboard. The metacognitive skills dashboard (see ***Figure 3***) uses click-level data to calculate metacognitive skill scores for each student.

These skills include consistency, persistence, gaming the system, procrastination, knowledge awareness, and online activity. Each score is calculated in real-time for each student. In addition, each score is graphed over time so a student can see how they are improving or getting worse throughout the course.

Conclusion

In this article, we have addressed a number of pedagogical, technological, and interface design challenges in implementing and using data analytics in a learner engagement context to improve online learning. We have also presented solutions to these problems based on our experiences in this area. Our work underscores the fact that no single tool or system will be able to address all of the challenges involved. As a result, tools and content need to be open, interoperable, and modular, a collection of features that has been described as "tool and content malleability" (Bush & Mott, 2009, p. 8). Indeed, we cannot overemphasize the importance of using open-source solutions (where possible) and sharing the developed solutions with others.

The work we have mentioned here supports the notion that the burgeoning and important field of learning analytics will develop much more quickly and effectively when tools and work are shared across many organizations and projects. As Dede (2016) points out, there is much to be gained by working together to support the infrastructure needed for curating and sharing data as well as developing analytic tools that can be shared as widely as possible among organizations. □

Reference

Anaya, A. R., Luque, M., & Peinado, M. (2016). A visual recommender tool in a collaborative learning experience. *Expert Systems with Applications, 45*, 248–259; *http://doi.org/10.1016/j.eswa.2015.01.071* .

Arnold, K. E., & Pistilli, M. D. (2012, April). Course signals at Purdue: Using learning analytics to increase student success. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 267–270). New York: ACM.

Berger, J. B., & Milem, J. F. (1999). The role of student involvement and perceptions of integration in a causal model of student persistence. *Research in Higher Education, 40,* 641–664; *doi:10.1023/A:1018708813711* .

Bush, M. D., & Mott, J. D. (2009). The transformation of learning with technology: Learner-centricity, content and tool malleability, and network effects. *Educational Technology, 49*(2), 3–20.

Carini, R. M., Kuh, G. D., & Klein, S. P. (2006). Student engagement and student learning: Testing the linkages. *Research in Higher Education, 47*(1), 1–32; *doi:10.1007/s11162-005-8150-9* .

Christenson, S. L., Reschly, A. L., & Wylie, C. (Eds.). (2012). *Handbook of research on student engagement*. New York: Springer Science & Business Media.

Dede, C. (2016). Next steps for "big data" in education: Utilizing data-intensive research. *Educational Technology, 56*(1), 37–42).

Few, S. (2006). *Information dashboard design* (pp. 120–206). Sebastopol, CA: O'Reilly.

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

Fredricks, J. A., Blumenfeld, P., Friedel, J., & Paris, A. (2005). School engagement. In *What do children need to flourish?* (pp. 305–321). New York: Springer US.

Filak, V. F., & Sheldon, K. M. (2008). Teacher support, student motivation, student need satisfaction, and college teacher course evaluations: Testing a sequential path model. *Educational Psychology, 28,* 711–724; *doi:10.1080/01443410802337794* .

Govaerts, S., Verbert, K., Duval, E., & Pardo, A. (2012). The student activity meter for awareness and self-reflection. In *CHI '12, Austin, Texas* (pp. 869–884); *http://doi.org/10.1145/2212776.2212860* .

Henrie, C. R., Bodily, R., Manwaring, K. C., & Graham, C. R. (2015). Exploring intensive longitudinal measures of student engagement in blended learning. *The International Review of Research in Open and Distributed Learning, 16*(3).

Henrie, C. R., Halverson, L. R., & Graham, C. R. (2016, in press). Measuring student engagement in technology-mediated learning: A review. *Computers & Education; doi:10.1016/j.compedu.2015.09.005* .

Hughes, J. N., Luo, W., Kwok, O.-M., & Loyd, L. K. (2008). Teacher-student support, effortful engagement, and achievement: A 3-year longitudinal study. *Journal of Educational Psychology, 100*(1), 1–14; *doi:10.1037/0022-0663.100.1.1* .

Kizilcec, R. F., Piech, C., & Schneider, E. (2013). Deconstructing disengagement: Analyzing learner subpopulations in Massive Open Online Courses. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 170–179). Leuven, Belgium: ACM.

Kuh, G. D., Cruce, T. M., Shoup, R., Kinzie, J., Gonyea, R. M., & Gonyea, M. (2008). Unmasking the effects of student engagement on first-year college grades and persistence. *The Journal of Higher Education, 79,* 540–563; *doi:10.1353/jhe.0.0019* .

Ladd, G. W., & Dinella, L. M. (2009). Continuity and change in early school engagement: Predictive of children's achievement trajectories from first to eighth grade? *Journal of Educational Psychology, 101*(1), 190–206; *doi:10.1037/a0013153* .

Lauria, E. J., Moody, E. W., Jayaprakash, S. M., Jonnalagadda, N., & Baron, J. D. (2013, April). Open academic analytics initiative: Initial research findings. In *Proceedings of the Third International Conference on Learning Analytics and*

Knowledge (pp. 150–154). New York: ACM.

Macfadyen, L. P., & Dawson, S. (2012). Numbers are not enough. Why e-learning analytics failed to inform an institutional strategic plan. *Educational Technology and Society, 15*(3), 149–163; *http://doi.org/10.1.1.441.9712* .

Office of Educational Technology. (2013). Expanding evidence approaches for learning in a digital world. *U.S. Department of Education; http://tech.ed.gov/wp-content/uploads/2014/11/Expanding-Evidence. pdf* .

Ott, C., Robins, A., Haden, P., & Shephard, K. (2015). Illustrating performance indicators and course characteristics to support students' self-regulated learning in CS1. *Computer Science Education, 25*(2), 174–198; *http://doi.org/10.1080/08993408.2015.1033129* .

Reeve, J. (2012). A self-determination theory perspective on student engagement. In *Handbook of research on student engagement* (pp. 149–172). New York: Springer US.

Reeve, J., & Tseng, C. M. (2011). Agency as a fourth aspect of students' engagement during learning activities. *Contemporary Educational Psychology, 36*(4), 257–267.

Reschly, A. L., & Christenson, S. L. (2012). Jingle, jangle, and conceptual haziness: Evolution and future directions of the engagement construct. In *Handbook of research on student engagement* (pp. 3–19). New York: Springer US.

Rovai, A. P. (2003). In search of higher persistence rates in distance education online programs. *The Internet and Higher Education, 6*(1), 1–16; *doi:10.1016/S1096-7516(02)00158-6* .

Santos, J. L., Verbert, K., Klerkx, J., Charleer, S., & Duval, E. (2015). Tracking data in open learning environments. *Journal of Universal Computer Science, 21*(7), 976–996.

Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. *Educational Psychologist, 50*(1), 1–13.

Skinner, E. A., & Pitzer, J. R. (2012). Developmental dynamics of student engagement, coping, and everyday resilience. In *Handbook of research on student engagement* (pp. 21–44). New York: Springer US.

Terkla, D. G., Sharkness, J., Cohen, M., Roscoe, H. S., & Wiseman, M. (2012, Winter). Institutional dashboards: Navigational tool for colleges and universities. *Professional File, Number 123,* Association for Institutional Research.

Van Ryzin, M. J., Gravely, A. A., & Roseth, C. J. (2009). Autonomy, belongingness, and engagement in school as contributors to adolescent psychological well-being. *Journal of Youth and Adolescence, 38*(1), 1–12; *doi:10.1007/s10964-007-9257-4* .

Verbert, K., Govaerts, S., Duval, E., Santos, J. L., Van Assche, F., Parra, G., & Klerkx, J. (2014). Learning dashboards: An overview and future research opportunities. *Personal and Ubiquitous Computing, 18*(6), 1499–1514; *http://doi.org/10.1007/s00779-013-0751-2* .

Zimmerman, B. J., & Kitsantas, A. (1997). Developmental phases in self-regulation: Shifting from process goals to outcome goals. *Journal of Educational Psychology, 89*(1), 29–36; *doi:10.1037//0022-0663.89.1.29* .

Contemporary Research Discourse and Issues on Big Data in Higher Education

Ben Daniel

The increasing availability of digital data in higher education provides an extraordinary resource for researchers to undertake educational research, targeted at understanding challenges facing the sector. Big data can stimulate new ways to transform processes relating to learning and teaching, and helps identify useful data, sources of evidence to support decision-making initiatives. However, in order to fully harness the potentials of big data, researchers must be able to make sense of this incredibly complex data, and pursue relevant questions, uncover patterns of interest, and identify and correct errors inherent in the processing and interpretation of big data. Furthermore, data governance, privacy, security, statistical algorithms, and analysis are processes that require contextualized human judgment. Researchers, therefore, must be able to tease out the specific significance of data and make valid and realistic interpretations of what the data offer. This article describes conceptual, technical, managerial, and educational opportunities and limitations associated with the use of big data in higher education. The article also identifies elements of a research agenda that could further the theoretical understanding of the relevance of big data within higher education discourse.

Background

The higher education sector is facing unprecedented pressure to adapt to an increasingly complex and dynamic

Ben Daniel is Educational Technology Coordinator and Senior Lecturer at the University of Otago, New Zealand. He teaches Research Methodology (Quantitative, Qualitative, and Mixed Methods), and Educational Technology in Higher Education. Currently, he is investigating the role of big data, learning analytics, and social media in enhancing learning, teaching, and research. He is also researching into the praxis and theories of teaching Research Methodology, and its development as a distinct discipline. Previously, he worked in the health care sector as a Research Methodologist and Innovation Analyst. He also worked in the software industry as a Knowledge Management Software Architect and Learning Specialist (e-mail: ben.daniel@otago.ac.nz).