

Chapter 5

Analytics Through an Institutional Lens: Definition, Theory, Design, and Impact

Matthew D. Pistilli, James E. Willis III, and John P. Campbell

Only about a decade after initial projects emerged, learning analytics have already had a profound effect on how institutions view the combination and use of multiple data sets and their subsequent analysis on learners and instructors. The success or failure of these institutional projects lay in the confluence of the data that is available, the extent to which analysis is conducted on that data, and the ensuing action taken on the results. Some higher education institutions have begun using the power of analytics to affect positive outcomes in critical areas like learning, pedagogy, student retention, and institutional decision-making (Long and Siemens 2011). But what does analytics mean to an institution and how might an institution implement some form of analytics? Further, what can institutions expect from the successful implementation of analytics?

This chapter begins with a broad description of analytics from an institutional perspective, including the foundational theory from which institutions can build. From there, what happens to an institution's learning environment when a successful implementation occurs, as well as how various learning communities are affected, will be examined. Finally, a discussion directly of how institutions may be changed as a result of data-driven models to enhance levels of success will be presented.

5.1 Defining Analytics from an Institutional Perspective

Bichsel (2012) defines analytics as “the use of data, statistical analysis, and explanatory and predictive models to gain insights and act on complex issues” (p. 6). Campbell et al. (2007) liken this joining of institutional data sets to a marriage—one that allows for the institution to discern patterns of student behaviors, traits, or outcomes.

M.D. Pistilli (✉) • J.E. Willis III
Purdue University, West Lafayette, Indiana, USA
e-mail: mdpistilli@gmail.com

J.P. Campbell
West Virginia University, Morgantown, West Virginia, USA

The identification of patterns provides an institution with the opportunity to provide targeted actions (interventions) to improve student learning. As these scholars and others have noted, a necessary component in learning analytics is taking action; institutions cannot simply collect and report on data. Institutional investments must go beyond performing data analysis and actually take specific actions to enhance student success and institutional achievement. These actions can include emailing or sending letters to students with specific steps that can be taken to enhance course performance (Pistilli and Arnold 2010; Arnold and Pistilli 2012), providing tips to students as to what it took other students with similar academic backgrounds to be successful in a course (Bramucci and Gaston 2012; McKay et al. 2012), or developing programs designed to mitigate various behaviors (Frankfort et al. 2012; Taylor and McAleese 2012).

The concept of using data to take action, to address something, or to facilitate decision-making processes is not new. Corporations have long used data on consumers and their habits to determine marketing strategies, directions for product development, and predicting sales based on current buying habits. They call this practice business analytics, which is defined as “the practice of iterative, methodological exploration of an organization’s data with emphasis on statistical analysis. [It] is used by companies committed to data-driven decision making” (Rouse 2010).

Using data to drive decision-making processes is not new to the higher education sphere, either. Colleges and universities have begun using data to better understand and begin to address student success, retention and graduation rates, course offerings, financial decisions, hiring and staffing needs, or admissions models of admits, yield, and matriculation. Furthermore, the use of analytics allows institutions “to test ... assumptions [regarding theoretical, practice-based and/or evidence-based examples of sound educational design] with actual student interaction data in lieu of self-report measures,” making for much more compelling arguments for continuing a process once one has been shown to be effective (Lockyer and Dawson 2011, p. 155). Analytics should take these tests and the broad use of data one step further. Analytics moves an institution, and the realm of education, from simply understanding various data points and their intersections, to using them to create actionable intelligence—and then taking action on that intelligence as a means of positively affecting one or more behaviors or outcomes.

Early systems like Purdue University’s *Signals* demonstrate how an institution can use analytics to create actionable intelligence that give students tools for real-time change (Pistilli and Arnold 2010). The key term here is “actionable” because theoretical knowledge of analytics, while perhaps valuable administratively, does little to influence student behaviors; rather, specific direction to help struggling students is what really matters (Cooper 2012). Long and Siemens (2011) identify key ideas driving the “value” of learning analytics, noting that their “role in guiding reform activities in higher education, and ... how they can assist educators in improving teaching and learning” (p. 38).

Although learning analytics is fairly new, higher education institutions do not have to explore the use of analytics without a guide. Educational research has provided decades of studies waiting to be transformed into practice. For example, Google Scholar provides over 274,000 articles on “prompt feedback” and analytics may

become the basis by which institutions can bring the feedback literature to scale. Whether looking at foundational research from Astin and Tinto on student success or focusing on Chickering and Gamson's seven principles, analytics provides the opportunity for scaling decades of educational research into daily practice.

5.2 Theoretical Basis for Institutions Implementing Analytics

Institutions should not arbitrarily approach analytics to change student behaviors, activities, or outcomes, as these kinds of efforts are already overwhelming to some institutions based on organizational and technical challenges alone. Given this, institutions should consult existing education theory and research to determine where best to apply resources. For many institutions, use of analytics to improve retention remains the key focus. While retention of students is a necessary thing to examine, it is an outcome measure. Effort should be put into improving student success, however success is defined by an institution, in such a way that a more successful student is more likely to be retained (Tinto 2012). To guide institutional efforts towards student success and, ultimately, retention, we turn to three guiding theories: Tinto's (1975, 1993) theory of student departure, Astin's (1984, 1993, 1996) theory of student involvement, and Chickering and Gamson's (1987) principles for good practice in undergraduate education.

Tinto's theory of student departure is the most widely cited theory in retention circles, and according to Braxton (1999), has reached "near paradigmatic stature" for those in the field of higher education (p. 93). In *Leaving College*, Tinto (1993) proposed that institutions need to meet three main conditions in order to achieve student persistence. First, students need to have access to retention programs that put their welfare above the institution's goals. While many institutions provide retention programs, analytics may provide a basis for better understanding who is attending such retention programs and how an institution may encourage all students to take advantage of the existing programming.

Second, retention programs should not just focus on a particular population (e.g., minority students, low-income students, athletes), but, instead, need to be available to all students from all walks of life. An institution's use of learning analytics may provide an opportunity to reexamine the use of student success programs. Rather than focus on particular categories of students, analytics may allow an institution to identify particular behaviors that change over time—providing a new dynamic learning environment that progressively evolves rather than remaining fixed on a particular group.

Third, retention programming must work to provide a degree of student integration within an institution to be successful. In his theory of student departure, Tinto (1993) notes that it is critical that students become formally and informally integrated to both their academic and social lives while in college. Formal academic integration occurs when a student interacts with a professor in class, visits a professor

to discuss class concepts during that professor's office hours, or attends a tutoring session or resource room to gain a better understanding of topics covered in class. Informal academic integration can be seen through students' interactions about course content with one another outside of class, or their attendance at a voluntary field trip that further explores the topics discussed during normal course time. Formal social integration is a function of students pledging fraternities or sororities, joining clubs, or participating in student government. Informal social integration revolves around students simply interacting with one another, such as playing video games together in a residence hall or playing a pick-up game of basketball. Tinto (1993) argued that the likelihood of persistence for students is increased if they have positive experiences with all four of these types of integrations.

For an institution, the formal and informal social integration becomes the most promising and challenging aspect for analytics. The potential data sources from these activities could provide a new insight to students integration into campus—however the collection and analysis of such data is probably beyond the capability of most institutions. Metrics that would indicate the achievement of a level of integration would need to be built into the interventions employed by an institution so that appropriate data could be collected to show the efficacy of the effort. It is possible that proxies may need to be developed by an institution if the exact data that would indicate an outcome is difficult, if not impossible, to obtain.

Tinto's theory provides a means of understanding the principles behind a student deciding to remain enrolled at an institution or to seek to be enrolled elsewhere, if at all. It should be noted that the decision to stay or go is first predicated on the extent to which students have been successful in their coursework. An unsuccessful student is likely to leave an institution, possibly of their own volition or more likely as a result of being academically dismissed. Analytics, then, is a process that can influence students' behaviors to help them be more successful, thus leading to their retention on campus. To this end, then, Tinto's model can be used by institutions as a roadmap for analytics.

The results of an institutional project based on Tinto's work could come in the form of encouraging a student struggling in one or more courses to actually visit the professor or available help rooms. Students who haven't joined a student organization might be encouraged to do so. Professional advisors and student resident assistants might be employed to outreach to specific students. In short, by looking at the integration scenarios and decision-making points, collecting data, and executing analytics, an institution can directly apply the "actionable intelligence" from the analytics to ensure that students are receiving the feedback needed, are connected to the institution, and are on track to graduate from that university in a reasonable amount of time.

Astin's theory of involvement (1984, 1993, 1996) noted that the more involved students are with certain aspects of their collegiate lives, the more likely they are to succeed. This notion was incorporated in the Input-Environment-Outcome (IEO)

college impact model (Astin 1993). Greatly simplified, outcomes, or characteristics of students once they have experienced collegiate life, are thought to be based on the initial characteristics students bring with them to college (input; e.g., gender, ethnicity, socioeconomic status) and are impacted by the collegiate experience as a whole (environment; e.g., going to classes, seeking academic help, participating in clubs/organizations). With regard to inputs, Astin (1993) identified 146 characteristics in several different groupings, including demographic, past academic achievement, previous experiences, and self-perception. Institutions may look at these characteristics as potential data elements for their analytics efforts.

Outcomes range from very concrete, easy to measure things such as level of academic achievement, retention from one year of study to the next, and persistence to graduation. More abstractly, there are skills, pieces of knowledge, and specific behaviors that are also developed. These tend to be the things that most institutions want to develop in students in some form. In the end, however, an institution has no impact on the inputs students bring with them to college, and only have the ability to potentially influence the achievement or development of various outcomes.

However, an institution can have a direct effect on student outcomes—and that is through the manipulation and alteration of the environment provided to the students. The environment consists of factors that are directly related to students' experiences while in college, many of which are influenced, if not directly offered, by an institution's administration. Astin (1993) identified 192 variables across eight classifications of characteristics that are associated with environment: institutional (e.g., Carnegie classification, size); peer group (e.g., socioeconomic status, values instilled, attitudes portrayed); faculty (e.g., teaching methods, interaction opportunities); curriculum (e.g., existence of a core set of courses, course requirements, delivery method of courses); financial aid (e.g., types of aid provided, amount awarded); major field choice; place of residence (e.g., on/off campus, fraternity/sorority housing); and student involvement in his/her education (e.g., hours spent studying, number of courses taken in a specific field).

Astin (1996) also concluded that there were three degrees of involvement for students at a given college or university. The first, involvement with academics, involves looking at the amount and quality of time spent on homework, working on projects, or engaging in other course-oriented activities. These activities are often the focus of many current institutional analytics projects. The second surrounds students' involvement with faculty within their courses and outside of the classroom. From an analytics perspective, these activities are often difficult to track as they usually require faculty compliance in noting interactions or attendance or some other form of recording of the students' involvement. The third degree, involvement, is the extent to which students interact with peer groups during college. These are fostered out of the collegiate environment, but are influenced both by the inputs students bring with them as well as their individual desired outcomes. These factors are important to keep in mind, because their interplay with the environment will have distinct effects on the extent to which the application of analytics is successful.

Where Tinto provides some insight into some of the behaviors that can be influenced as well as key decision-making points for a student, Astin is more concerned with altering the environment in which students exist. As a result, the input-environment-output model integrates nicely with the application of analytics in that it provides the ability to achieve a desirable outcome (success in college, increased learning, progression towards and awarding of a degree, etc.) by taking any number of inputs from a multitude of experiences and actions, and providing suggestions for behaviors and environmental interactions that would allow students to interact with one another, with faculty, and with the university in a positive and supportive manner. Astin’s research becomes a foundation from which to begin an analytics project and allows for the analysis of many different pieces of data in coordination with each other to find meaningful relationships and areas to address. Using Astin’s model, a framework has been created that provides variables based on, but different from, Astin’s work to examine and behaviors to influence in an effort to achieve a desired outcome.

5.3 Institutional Analytics by Design

Tinto and Astin each provide a theoretical basis for student success and retention, but where does an institution turn to begin designing an analytics project?

The goal for an institution is to design a learning environment that leverages the institution’s data, the seamless integration of pedagogy, and the actions required to have an impact on student success and retention. The ideal interaction for these three aspects has each informing the other to create the optimal mix for a given institution. For example, a given pedagogy will direct the kind of action(s) to be taken, as well the analysis of certain data points. Both specific actions and pedagogies will drive the kinds of data to be collected. The extent to which data *can* be gathered will certainly influence actions taken, and also may influence the extent to which a given pedagogy is used by an instructor. In the end, however, individual institutions must determine what kinds of interactions are the most meaningful given their unique needs and priorities (Fig. 5.1).

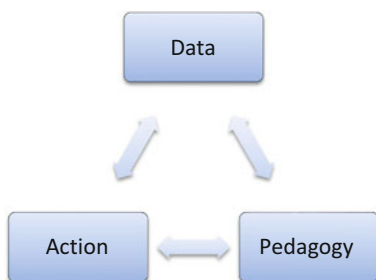


Fig. 5.1 The interactions that are needed in a learning environment to positively affect student success and retention. Each aspect informs the other to provide an institution with insight into what data to collect, analyze, and disseminate; what actions to influence or perform; and, which pedagogical practices to employ or examine

5.3.1 *Designing for Data*

For analytics to be successful, an institution must place an initial emphasis on collecting, organizing, and analyzing data that is meaningful, useful, and obtainable. As [Campbell and Oblinger \(2007\)](#) note, “data is the foundation of all analytics efforts” (p. 3); absent a strong foundation of good data, any analytics effort will likely fail. Therefore, an institution should place a considerable effort on determining what data is available—or could be obtained—that can provide meaningful insights into achieving the intended goal of the effort. Once institutions have established a practice of using analytics, attention may be focused on collecting new data identified by Tinto, Astin, or other researchers. Chickering and Gamson (1987) provide a foundation for what data might be meaningful through their analysis of 50 years of education research. The authors proposed seven principles of good practice for undergraduate education. These principles include developing educational environments that:

1. Encourage contact between students and faculty
2. Develop reciprocity and cooperation among students
3. Encourage active learning
4. Give prompt feedback
5. Emphasize time on task
6. Communicate high expectations
7. Respect diverse talents and ways of learning ([1987](#), p. 3)

For example, if an institution may wish to focus on the contact between students and faculty, the data might include the number of exchanges in emails, discussion postings, or participation in office hours.

While one could imagine a large number of potential data sets, it is important to balance the practicality of collecting said data with the number of students who will actually have their data recorded. To that end, institutions should not go out and create new data for the purposes of implementing analytics. Rather, institutions should strive to utilize ambient data—data that already exists or is collected as a matter of course—that may be used as a proxy to measure or represent student success. Institutions can identify potential data sets by utilizing the work of Tinto, Astin, Chickering, and Gamson—then test those data elements through statistical analysis. For example, while a computer or algorithm can easily determine if an email exchange was made between a faculty member and a student, the system is likely not able to determine the content of the exchange—thus not able to determine if the message was to provide prompt feedback, communicate expectations, or clarify the purpose of the assignment—or something different altogether.

Taking this approach of using ambient data is reinforced by Macfadyen and Dawson, who write that learning analytics should be derived from student data points that are “readily accessible, scalable, and non-intrusive, and provide sound lead indicators of eventual student achievement or failure” (2010, p. 598). Data sets may include interaction data derived from learning management system (LMS)

logs, demographic information obtained through the admissions process, or past academic performance garnered as a result of students simply being enrolled at an institution, to name a few examples.

While self-report data can be highly valuable, researchers are dependent on high degrees of compliance with regard to completion of surveys or reporting of behaviors in order to make conclusions that can lead to action being taken. This is not to say that self-report data should be ignored altogether; these data are extraordinarily valuable in their own right. However, the collection and maintenance of self-report data is usually not sustainable. Self-report data are useful for confirming output from an algorithm using existing data to ensure that the targeted students are being identified correctly. The use of self-report data can also be influenced by the student's recognition of how the data may be used.

The paradox contained in the wealth of information contained in self-report data and the challenges associated with collecting and analyzing that data does not go unnoticed. On one hand, the collection of this kind of data comes down to the feasibility and sustainability associated with its gathering, analysis, and subsequent use. On the other hand, by collecting self-report data, an institution is making students aware that their data is being collected. This process, then, could lead to students automatically changing their behaviors—but could also result in students ceasing to provide information at all. The process of collecting any data, be it ambient or otherwise, should be focused on being both scalable beyond a pilot collection at one or two events *and* sustainable so that the data collection effort isn't abandoned because of efforts or challenges associated with it.

If institutions choose to utilize self-report data, it is important for them to consider whether the return on the investment associated with these actions is worth the information that it may potentially yield. If not, institutions may wish to look for other means of determining student behavior or action, or develop variables that approximate the desired outcome. Ultimately, an institution's decision to collect non-ambient data should be done so in the ways they feel are best for their individual purposes.

Campbell and Oblinger (2007) suggest that the manner in which data is stored and the length of time it is retained should be examined when determining what data will be used and the extent to which historical data will be analyzed to make real-time predictions. In addition, they also indicate that the granularity of data be considered as well. Defined as “a balance between what the system provides, what questions the institution is attempting to answer, and the storage requirements of the data,” granularity refers to just how finite a set of data needs to be created (Campbell and Oblinger 2007, p. 5).

Differentiation needs to be made for current analytics projects and potential analytics projects. A handful of data points may accurately predict a given phenomenon equally as well as dozens of data points for the same event. To this end, then, an institution may strive for parsimony in its data sources for two reasons. First, institutions should create as sustainable and easy a process as possible for obtaining and analyzing data, and the fewer the number of points that need to be collected, the more likely it is that this end will be achieved. Second, in striving for something that

can be easily interpreted by faculty and students, the smaller the number of pieces of information utilized, the easier the institutions will be able to pinpoint exact areas that can be altered or addressed by either party. Finally, storing data for potential future use should be a separate function from current analytics projects.

5.3.2 *Designing for Pedagogy*

With the nature of today's electronic tools, designing for pedagogy is an interaction between two components—the process and the environment. From a process perspective, the success of an institution's analytics depends on the quality and quantity of the data that is utilized.

Institutions can design for analytics projects by integrating Chickering and Gamson's (1987) model as a basis for pedagogical design—ensuring that course design and analytics work synergistically. For example, the more that an instructor meaningfully includes tools that encourage active learning, provides prompt feedback, and increases communication between students that not only improves learning, but also provides a meaningful data set for analytics. Using a tool that allows faculty to email students regarding their performance and activities in which they can engage to enhance or maintain their current grades meets the first and fourth principles directly, and potentially the fifth, sixth, and seventh, depending on the content of the message (see Sect. 5.3.1 in this chapter for a numbered list of practices). By using analytics to determine teams or groups in classes, the course design fulfills the second principle. The third principle, use of active learning techniques, might have an analytic solution that provides students with tailored practice exams, readings, or exercises that meet their specific needs or deficiencies. The broad point here is that the application of analytics can provide an institution with a means of not only achieving certain metrics or learning gains, but also creating an environment that is best suited and supportive of students and faculty alike.

When institutions use analytics as a process, they provide a means for evaluation and enhancement of course design through the alteration and enhancement of pedagogical practice. Analytics processes that are implemented well should identify courses where students continually struggle, and, further, specific instructors who may need to alter the manner in which they deliver material. In turn, pedagogy can be altered and enriched; by offering feedback to students where none was offered in the past, the instructor can shift the focus of instruction to students' needs in an effort to help them improve. Over time, as students receive more and better feedback, and as faculty recognize areas for improvement in their own modes of instruction, the quality of these courses should improve. Instructors will have a better understanding of how material is being received, can address specific topics, and potentially move more quickly through the course. Students, then, are learning more and are better prepared for subsequent courses. (For insight into how the application of analytics can improve student success and other outcomes, see Arnold 2010; Arnold and Pistilli 2012; Baepler and Murdoch 2010; Essa and Ayad 2012; Ferguson

2012; Oblinger 2012; Salas and Alexander 2008; Taylor and McAleese 2012; or Tinto 2012.) The secondary courses are then able to cover more material, since students will not need to be remediated on core concepts. In all, the institutional application of analytics can result in a major shift for colleges and universities with regard to the culture fostered around undergraduate learning.

As mentioned above, the use of analytics can alter the structure of the classroom and, indeed, the institution itself. However, as noted by Bonfiglio et al. (2006) the transition to a student-focused environment from one that was more suited for the instructor takes a great deal of institutional analysis, with particular emphasis placed on the extent to which learning is taking place. This can be done in many ways, but is increasingly seen within the realms of computer-supported collaborative learning, social learning, and distance learning.

From an environmental perspective, the institutional culture will have a significant impact on how an individual faculty member designs courses with analytics in mind. The availability of data, willingness to allow faculty to experiment, and trust between data “holders” and data “users” will all have an effect on the extent to which faculty can utilize data to change their pedagogies—and the extent to which an institution is willing to exert effort to collect additional data to inform pedagogy. The interaction between those that collect, hold, and utilize the data will set the tone as to how analytics will be fostered—or abandoned—at an institution. An institutional culture that controls data tightly or seeks to use data as an evaluation hammer will limit the use of analytics to improve pedagogy. Conversely, an institutional culture that allows more liberal use of data may foster new pedagogical models based on analytics.

5.3.3 *Designing for Feedback*

As Lockyer and Dawson (2012) note, though, analytics and resulting actions are usually done “retrospectively—and often on an *ad hoc* basis” (p. 14). While Lockyer and Dawson’s reflective view is both effective and necessary, Chickering and Gamson (1987) suggest providing feedback at a point in a term where students have the opportunity to alter their academic behaviors, thus giving students the biggest chance of success. Lockyer and Dawson (2012) also posit that the application of learning analytics needs to be in synch with what instructors are planning to do inside of their courses. Data needs to be appropriately chosen for analysis, and provided in such a manner that it is easily interpreted by and broadly useful to the instructor – and that is ultimately meaningful to the student. Utilizing Chickering and Gamson’s framework, designing courses with feedback as an integrated piece would look for solutions that not only provide for prompt feedback, but also seek ways to encourage communication between the faculty and student, continue to set high expectations, and direct students towards more active learning approaches.

Schunk (1985) notes that students gain self-efficacy and knowledge from having the opportunity to work at a task and receive both positive and negative feedback.

Furthermore, he indicates that giving students training in various strategies that could be used to address a problem, and then allowing them to determine which one is most effective, provides opportunities for students to better understand the required task. Finally, feedback is necessary for students to fully understand their performance, their shortcomings, and their successes. All of these activities, particularly if they're done in an online environment, create large quantities of useful data that can be extracted and mined for patterns of performance, areas of weakness or strength, or prediction of future grades or outcomes. Another avenue for creating these data points would be through practice tests and quizzes offered through various technologies where data can be readily harvested. The point is not to create busy work, but, rather, to create data that can be analyzed and acted upon in a timely manner that will allow students to focus their studying efforts on areas of deficiency. When coupled with precise feedback and specific things to address, students will have the best opportunity to improve their performance.

5.4 Impact of Analytics

There are distinct differences between the types of communities that can be formed via brick and mortar classrooms or virtual learning environments, and the application of analytics can be useful in either one. In the end, however, the focus should be less on the mode of learning and content delivery, and more on the learning itself. This becomes readily apparent as the realized and potential effects that learning analytics has on learners, instructors, and institutions are examined.

5.4.1 *The Impact of an Institution's Use of Analytics on Learners*

If the driving force behind an institution's decision to delve into learning analytics is increased learning, then the learners must be the direct beneficiaries of such efforts. Assessing *how* learning is affected is best done thematically; though this list is not exhaustive, the current trends of learning analytics typically fall within examining predictive power, providing change opportunities with near live-time feedback, pulling together big data in ways that were impossible previously, and identifying knowledge gaps in learners.

One of the most promising aspects of learning analytics is the power of prediction based on historical and current data points (Elias 2011). Student predictions are not static because if the model is robust enough, constantly changing data will sharpen the prediction "and then feed those results back in order to improve the predictions over time ... as it relates to teaching and learning practices" (Elias 2011, p. 5). Though it may not be prudent to describe the process as somewhat parasitic, the metaphor works: if modeled correctly, data points help build and then clarify the

algorithm to statistically predict a learning outcome. Put into real-world terms, this means that seemingly disparate data on a given student may be brought together meaningfully to provide actionable information for the student.

Macfadyen and Dawson (2010) demonstrated this point by “providing data from an international research project investigating which student online activities accurately predict academic achievement” (p. 588). They found that providing instructors with a “dashboard-like interface that incorporates predictive models and network visualization tools” was able to provide statistically significant predictive student outcomes, such as increased engagement with peers and course material and “higher overall final grades” (p. 597). There is a sharp distinction between prediction and actuality; as the power of predictive learning analytics increases over time, all parties involved in the learning process must heed the differentiation between what a statistical model can predict as an outcome and how students’ *choice of action* once presented with the data affects the actual outcome.

Learners are ultimately responsible for how they take analytical predictions and create an opportunity for change. What drives this change mechanism is data mining, which, in this context, “discover[s] potential student groups with similar characteristics and reactions to a particular strategy” (Romero et al. 2008, p. 1). Amalgamating and assessing large data sets of student demography, history, performance, and numerous other data points have the potential to provide near real-time feedback. At the granular level, data mining can provide expeditious information that, among other things, helps “to identify learners with low motivation and find remedial actions to lower drop-out rates” (Romero et al. 2008, p. 1). Fostering the shift from using large data sets to real, quantifiable information that can nudge a student to make a real change are the algorithms that drive institutions’ LMSs; the information that institutions use for their formulae is already at their fingertips. The task in learning analytics is to take information from the LMS which functions in near real-time and convert it into a compelling interface to encourage student change; the process of converting LMS data into a student nudge has the potential “to offer students a truly flexible and rich learning experience” (Phillips et al. 2011, p. 1005).

The use of algorithms generated from large data sets for the purposes of learning analytics presents something of a contemporary phenomenon: the convergence of big data for actionable intelligence. The consequences of using this data, which can persuade students to make meaningful change in their courses of study, have directly measurable outcomes, including “retention, graduation in a timely manner, and preparation for the workforce and citizenship” (Chacon et al. 2012, p. 7). Making meaning out of large data sets, especially as they apply to student learning, is no easy task; this requires being able to think outside of LMS-generated data (Pardo and Kloos 2011). As Pardo and Kloos (2011) highlight, the “interaction” of data is what really matters because it is where the “learning experience” is measured (p. 163).

Learning analytics, especially as it is used as a measure of the “learning experience,” helps identify knowledge gaps in students. A student’s performance on an

assessment may be indicative of learning gaps of specific information; when coupled with performance on multiple assessments, patterns may emerge that show that student's learning deficiencies in an entire area. The difficulties of measuring these patterns, though, are akin to measuring students' learning with perceptions. As Phillips et al. (2011) point out, learning measured with perceptions "rarely indicate the causality of effects" (p. 998). Measuring patterns in student performance may not necessarily highlight the causality of the effects, therefore indicating actionable change may be a product of correlation, not causation.

The impact of learning analytics on students may also have some poignant challenges including not knowing what students will do with learning analytic data, suffering from self-fulfilling prophecies, and the effects of relinquishing data. There is an element of the unknown insofar as analytics are not necessarily the actual outcome of a given set of variables. Perhaps the more potent idea is that student-centered data, when constructed in a rigorous and ethically-sound model, can empower students to make responsible changes to enact more favorable outcomes. Limitations like the innate intrusiveness of gathering and using data present unique challenges; the priorities of helping individual students must be weighed against the potential problems of using personally-identifiable variables.

Predictive analytics works in a way that lives up to its name: using yesterday's and today's data to predict tomorrow's outcome. With learning analytics, institutions cannot know what students will do with information regarding performance, suggestions for improvement, or predictions of outcomes. Mattingly et al. (2012) draw this out explicitly in terms of "distance education" when they note "the lack of knowledge about the ways that students interact with learning materials" (p. 238).

Highlighting the dearth of information available on how students interact with learning materials, the parallel implication can be examined in terms of predicting how a student will use learning analytics. For example, Pardo and Kloos (2011) suggest today's students "rely less on the functionality offered by the LMS and use more applications that are freely available on the net" (p. 163). There is an element of Occam's razor here: students tend to use whatever resource is easiest to obtain and most expeditious in providing a result. The problem with learning analytics in this environment, then, is measuring how deviating from a set of metrics (often configured in an LMS) changes the model or the predictive power within a set of probabilities.

With any analytical process that relies on predictive power, there is a risk of impacting students for the worse via a potential self-fulfilling prophecy (Merton 1948). In other words, if students are given information that their prior performance on assessments indicates a real jeopardy of failing a course, they may simply give up and accept failure as a predetermined outcome (McKown et al. 2010). Even worse, if students interpret failure in one course as failure in general, they may drop out of school altogether. The challenge then lies in providing meaningful and actionable feedback to students in ways that will help them. Providing constructive information to the student helps overcome the communication barriers known to exist when encouraging a student to act on analytical knowledge (Tanes et al. 2011).

One of the consequences of using big data and learning analytics is the fact that the data being used has to be acquired. The interplay between large sets of group student activity (both academic and non-academic) coupled with granular student-level data is at once powerful, compelling, and worrisome. In this sense, “big data” can be that which is culled from numerous students or multiple data points from one student. What makes it “big” is not just the number of computable variables, but also the power of what that data might suggest through the power of statistical regression. This data can be obtained either through individuals giving it up (voluntarily or otherwise) or through the mining of existing data sets. (For a broader discussion of big data within the context of education and analytics, see also: [Campbell et al. 2007](#); [Ferguson 2012](#); [Picciano 2012](#); or, [Siemens and Long 2011](#).) Chacon et al. (2012) encourage “a clear policy framework in applying learning analytic tools and developing intervention strategies” because “institutional real-time monitoring of student progress ... might be perceived as intrusive” (p. 7). This is especially true as social media posts, behavioral interactions, and other lifestyle data become more widely available and are shown to have efficacy with regard to predicting student success.

Like measuring learning, there is no definitive way to measure the effect of analytics on students. Learning analytics has provided institutions the ability to take existing data, compute predictions of students’ performance, and intervene in near real-time to compel them to change their academic and/or social behaviors. Learning analytics seizes the power of data mining and applies data sets to quickly and efficiently identify how to help students improve their learning and performance on assessments. With this power of analysis, though, comes a bit of warning as to how such information may lead to self-fulfilling prophecy of failure as well as the relinquishment of what many may consider private data. One should consider the complexity of human learning; while data may provide some insights to student learning, most data today provides only a hint of current student progress. Presenting data as the one, authoritative source of student progress will cause unintended consequences – in the extreme, this may present itself as students or faculty giving up in the process. It is important to draw the potential connection between data sets that apply to cohorts of students as well as individual data: there is influence in metadata and individual data insofar as it affects the present and future of student success.

5.4.2 Impact on Instructors

Like learners, those who teach have gained immensely from learning analytics. In prior educational models, instructor feedback may have been measured in days and weeks; learning analytics has the ability to help instructors provide much quicker feedback. Additionally, learning analytics helps instructors reshape pedagogical practices because they have access to data-rich information on what works and what does not for a given class.

Perhaps one of the most difficult aspects of teaching is being able to assess if students are learning and, if so, to what extent. Greller and Drachsler (2012) provide a succinct summary of what learning analytics offers to instructors:

Teachers can be provided with course monitoring systems that inform them about knowledge gaps of particular pupils and thus enable them to focus their attention on those pupils. They can also harvest emergent group models that can lead to shared understanding of domain topics or processes for better curriculum design and on-the-fly adaptations (p. 47).

The information flow from the analysis of data takes the same form as for students: information alone does not change anything, but can equip the instructor to make time-sensitive, meaningful changes that affect student outcomes. This requires what Elias (2011) calls “faculty buy-in,” because it requires use of newer technologies, willingness to change long-held practices, and constructive criticism of methodologies of pedagogy. These elements are “paramount to the institution’s ability to build and sustain a culture of evidence-based action” (p. 16). The power to shape student outcomes, however, is realized in how it affects individual students. Where many universities have lecture halls filled with several hundred students at a time, having the granular data to affect individual student outcomes is a direct benefit to an instructor’s overall effectiveness.

When instructors use learning analytics to approach students with constructive criticism, effective communication is often what helps students succeed. If used properly, analytics enables instructors to “empower ... students to monitor their coursework and take greater responsibility for their learning” (EDUCAUSE 2010, p. 2). This responsibility is important. The onus of action is still on the student, but the instructor is better prepared to guide the student with feedback in an amount of time necessary for changes in behavior. Feedback generated via learning analytics goes further than in previous educational models because it has the capacity to “go beyond reinforcement and [provide] an elaborate picture of where a learner stands in reference to others, certain criteria or their previous performance” (Tanes et al. 2011, p. 2415). Feedback in a timely manner provides a boon for instructors to affect student outcomes in real-time.

Having a proverbial yardstick by which to measure student learning is a key benefit of analytics, if for no other reason than it allows instructors the flexibility to alter pedagogical practices in a timeframe that positively affects student outcomes. If an indicator of pedagogical practice was once measured by end-of-course surveys (where changing teaching practices would only benefit future students), learning analytics provides feedback for instructors so that they can make quick alterations in teaching using real-time data. Using feedback as an aid in the teaching process does not necessarily yield a perfect or ideal pedagogy (van Harmelen and Workman 2012). This is due to a “multiplicity of uncontrolled variables,” particularly because learning analytics is applicable to widely varying subject matters, teaching styles, and learning outcomes (van Harmelen and Workman 2012, p. 18). Learning analytics is not an elixir for ineffective teaching, nor does it reveal an ideal pedagogy; instead, it provides data-driven tools or suggestions to help instructors make changes that can be measured in terms of student outcomes.

Putting learning analytics into service for better teaching means thinking about the “learning environment” in terms of “flexible modalities for study” (Dawson et al. 2009, p. 185). Today, learning environments may be online, in traditional lecture or seminar with LMS integration, in massive open online courses (MOOCs), or in some hybrid combination of these modalities. What learning analytics does is help transform the various challenges of teaching in different modalities into actionable data that are primed for appropriate pedagogical changes. Dawson et al. (2009) discuss these recent transformations at length:

Regardless of the overall didactic story telling [sic] and engagement prowess of individual presenters a shift to online necessitates a re-configuration in learning design and a conceptual shift in pedagogical practice. While this transition has been at ease for some educators—others have found the change rife with new complexities surrounding technology usability and integration in a context where communication cues and notions of student engagement are largely invisible (p. 190).

To develop the tools necessary to make pedagogical changes, current research indicates that models ought to be developed that consider “informative feedback [because it] is more effective in teaching desirable outcomes, and is perceived as more valuable by learners” (Tanes et al. 2011, p. 2415). As an example of an informative feedback system currently utilized, Purdue University’s *Signals* operates in conjunction with the LMS to “provide both performance and outcome oriented feedback to students” (Tanes et al. 2011, p. 2415). This means that while instructors specify the parameters of performance for student feedback, that same data is used to assess whether the pedagogy being employed is effective for large groups of students. If an instructor has many “red lights” in the signal system (indicating students are in need of immediate performance alteration), then that instructor should realize that a pedagogical shift may be necessary in order for students to realize greater success. Pedagogical shifts that are made with “informed change” help instructors “provide evidence on which to form understanding and make informed (rather than instinctive) decisions” (van Harmelen and Workman 2012, p. 17). The benefit of learning analytics for instructors is the production and promulgation of hard data that allows for alterations in teaching method to be employed relatively quickly.

As a formalized system of research, learning analytics is relatively new. Dawson et al. (2009) acknowledge the “scarcity of resources available that can readily assist teachers in rapidly evaluating learning progress and behavior in order to better design learning activities to provide a more personalized and relevant learning environment” (p. 191). As more historical student performance data becomes available to researchers, better algorithms likely will be developed. Recent work in causal models have “identified links between certain measurable data attributes describing past student behavior and the performance of a student” but this, too, “is dependent on a body of historical data” (van Harmelen and Workman 2012, p. 17).

Preliminary results of measuring how learning analytics provide actionable data to instructors indicate that “student success was associated with instructional rather than motivational feedback, and type of rather than frequency of summative and formative feedback” (Tanes et al. 2011, p. 2420). The challenge of analytical data for these purposes is the sheer amount of “comprehensive” data needed to make the

case (Ali et al. 2012, p. 470). Multiple data points help bring out statistically significant patterns to refine algorithms relevant for feedback tools that impact pedagogy; the problem in the intermediary time, though, is the amount of data needed to compute such multivariate algorithms and a consensus on which data points are most useful.

As the ongoing work of learning analytics is used to help improve pedagogical practices, one of the important caveats to the research is ensuring that the data employed by instructors does not discourage students. Greller and Drachsler (2012) are quite emphatic on this point because they see that statistical modeling may box in “individual teachers or learners against a statistical norm” with the possible result of “strongly stifl[ing] innovation, individuality, creativity, and experimentation that are so important in driving learning and teaching developments...” (p. 47). Discouragement in students is an important measure when examining how algorithms compute pedagogical conclusions. Research in this area indicates that “positive feedback ... generally emphasized performance, while negative feedback emphasized outcome” (Tanes et al. 2011, p. 2420). This research helps alter pedagogy because it can assist faculty on how to form appropriate messages to students based on specific parameters of success or failure and help institutions develop instruction. Appropriate wording in and construction of messages to students help mitigate the problems of student discouragement.

Learning analytics has impacted instructors in terms of forming pedagogical practices that are current with the types of learning environments seen in education today. Analytics provides instructors with tools to provide quick feedback as well as make rapid changes in pedagogical practice to affect student outcomes positively. As more data becomes available for study, more precise algorithms will be shaped to help instructors avoid discouraging students by providing accurate and actionable feedback.

5.4.3 Impact on Institutions

Where the effect of learning analytics for students and instructors is best seen as a microcosm wherein the outcome is measured with individuals and groups, the impact on institutions is a look at the macrocosm. Variables such as learning environments, student retention and graduation rates, and pedagogical effectiveness as measured by the achievement of positive student outcomes are all considered when measuring the impact of learning analytics on institutions. This effect is best examined by how learning analytics aids student retention and how institutions are able to refocus resources once specific areas are identified. These measures are examined through the prism of the potential difficulties of learning the “right” way to think about and use analyzed data, as well as the legal and ethical issues of using data for analytics (Johnson 2013; Willis et al. 2013).

Learning analytics has a direct, quantifiable effect on institutions that can be seen in how they affect student success and retention. EDUCAUSE (2010) identifies the

major goals of improving “student achievement, retention, and graduation rates and to demonstrate institutional accountability” through the major initiatives of “harness[ing] the power of analytics to develop student recruitment policies, adjust[ing] course catalog offerings, determin[ing] hiring needs, or mak[ing] financial decisions” (p. 1). Retention often becomes an institutional focus for using analytics because it can be a sound measure of how effective curriculum changes, effective recruitment, and institutional accountability are institution-wide. The difficulty with an institutional view of retention is that the data is reaped after a student has dropped out, revealing “gaping holes of delayed action and opportunities for intervention” (Long and Siemens 2011, p. 32).

As with retention, many current institutional learning analytics projects are frequently focused on binary predictions or results—whether the student will be retained, if the student is at risk, whether the student understands a particular learning outcome, and other similar propositions. This binary focus of learning analytics is more a result of the current sophistication of the data, models, and interventions than the long-term potential of learning analytics.

The task for assessing retention data quickly is best accomplished by using the vast resources of data that are already present institutionally; these data points can be utilized with analytics to “serve as a foundation for systemic change” (Long and Siemens 2011, p. 32). A definitive institutional strength to using learning analytics is the methodological process of performing “hypothesis-driven [analysis], using a particular dataset to solve a practical academic problem, such as increasing student retention levels” (Baepler and Murdoch 2010, p. 2). While retention is certainly an important measure of institutional effectiveness, it is also illustrative of the interconnectedness of other measures of change; retention may be the computed number, the quantitative measure, but it is the byproduct of how well an institution focuses and redirects its energies to ensure students are successful.

Identifying areas of focus for institutions often demands the use of “granular level” analysis that is possible through statistical modeling, prediction, and analytics (Greller and Drachler 2012, p. 47). In an age of increasing accountability and tightening budgets, learning analytics provides data that “can support optimal use of both economic and pedagogical resources while offering a structure for improved educational outcomes” (EDUCAUSE 2010, p. 2). However, the art of designing an analytic system to account for system-wide variables in pointing out areas of concern means combining “principles of different computing areas (data and text mining, visual analytics and data visualization) with those of social sciences, pedagogy, and psychology” (Ali et al. 2012, p. 470). This is extremely difficult; it requires drawing “value from data in order to guide planning, interventions, and decision-making [as] an important and fundamental shift in how education systems function” (Siemens and Baker 2010, p. 253).

Granular data usage requires critical evaluation of those variables that are the most important for the model, the same data points that will lead to increased knowledge of where students are failing to gain “self-directedness, critical reflection, analytic skills, and evaluation skills” (Drachler and Greller 2012, p. 129). The influx of variables require diligent analysis because one of the most acute problems

in learning analytics is not only how to interpret the data, but to understand why “even . . . the best evaluative algorithms can result in misclassifications and misleading patterns” (EDUCAUSE 2010, p. 2). Once the variables that most closely align with the desired model are determined, institutions may take actionable change to impact students and instructors.

The actionable changes that institutions may take are often dependent upon knowing the “right” way to think about and use the analyzed data. van Barneveld et al. (2012) echo the warning that “analytics is not a one-size-fits-all endeavor and that one has to consider that analytics is a goal-directed practice” (p. 2). Rushing to make decisions can have unintended effects on institutions; rather, as Kay and van Harmelen (2012) argue, “the analytics ‘silver bullet’ lies in the potential to derive and to act upon pre-emptive indicators, on ‘actionable insights,’ stepping beyond the long haul reactive measures” (p. 5). Careful consideration of the various factors that lead to interpreted data is vital to institutions because it is integrated with learning, teaching, and administrative variables. Consideration of these variables also must include legal and ethical impacts on institutions.

The legal impacts on institutions must be considered carefully because all data sets have a certain amount of liability attached to them. The primary beneficiaries of learning analytics are students and instructors, but their data may be used to directly affect institutions for the better (Drachsler and Greller 2012, p. 123). Comparable institutions may find it beneficial to share data for mutual gain, but anonymization of data is important to prevent litigation and ethical breaches of conduct (Drachsler and Greller 2012, p. 127). Ensuring anonymity within data is important to data handling and transfer. To this end, Powell and MacNeill (2012) specifically describe the current need for more data handlers and data scientists—individuals who are able to work “across teaching and administrative domains, to ensure that relevant actionable insights from data can be identified and acted upon in meaningful, measurable ways” (p. 3). Proper training in the handling, use, and anonymization of data is important to safeguard institutions from litigation.

Beyond legal aspects of learning analytics, ethical uses of data are also of important concern for institutions. Even though students are increasingly “born digital” and thus have “new expectations” of the ways their data will be used, institutions have an ethical obligation to protect the data and work within accepted research methodologies (Kay and van Harmelen 2012, p. 5). The widespread use of big data begs questions of “data ownership and openness, ethical use and dangers of abuse, and the demand for new key competences to interpret and act on learning analytics results” (Drachsler and Greller 2012, p. 120). The problem with ethical analysis is keeping pace with the speed of technological development; though there are ethical models that are directly applicable to data use, the constantly changing environment demands an openness of ethical questioning. Though there are no definitive answers to ethical questions being asked in learning analytics, perhaps what is most important is to remember that analytics “is much more about a personal and organizational perspective on using data for decision-making and action-planning and less about how it is processed in a computer; evaluating, planning and doing are human activities” (Cooper 2012, p. 7).

5.5 An Example of Application

This chapter ends with a brief overview of how one institution chose to contextualize data and act on it accordingly. As mentioned above, Purdue University developed a system called *Signals* in 2007, which was a direct offshoot of Campbell's (2007) dissertation investigating the extent to which data derived from the LMS could be used to predict student performance. The challenge was to identify the student at risk of doing poorly in a class using only data that was readily available; to wit, current course grades (e.g., test/assignment grades), past academic history information (e.g., standardized test scores and high school or current cumulative GPA), demographic descriptors, and data indicating the extent to which a student was interacting with the LMS.

As these various data points were examined, it became clear to the researchers that the integration of these data could be done so in a way that would provide an outlet for meaningful feedback to be provided to students by their instructors. The goal became to assist students and “help them understand both their current grades in their classes and what they can do to earn a higher grade while there is still time to act” (Pistilli and Arnold 2010, p. 23). Throughout the development, care was taken to focus on the behaviors that could be addressed. Specifically, the intention was to get students information about the specific actions they could take to positively affect their standing in a course. This information was to be written by the instructor and delivered via brief email messages and postings within the LMS site for the course (Arnold 2010). Pistilli and Arnold (2010) note that the posting in the LMS is accompanied by a color—green, yellow, or red—which serves as a primary indication of how a student is doing in a course. Clicking on the light revealed a message containing substantive suggestions as to what a student could do to increase performance in the class.

The novelty of *Signals*, as compared to other early warning systems, is that it took students' effort into account by measuring the extent to which a student was interacting with the LMS *and* comparing that interaction to the interaction levels of other students in the same class (Arnold 2010; Pistilli and Arnold 2010). What this resulted in was a means for the instructor to directly tell students exerting less effort that they were, in fact, not expending as much energy online as the rest of their peers, and by taking more initiative they might be able to improve their performance in the class. In addition, the system provides a dashboard to both instructors and academic advisors, allowing both parties to see a student's progression of signals over the course of a semester. This allows for either person to directly intervene with students when they see a disturbing or downward trend in student performance in an effort to help them be more successful (Pistilli et al. 2012). Further, intervention can be made in an early and timely manner so that students, as discussed earlier in this chapter, have the opportunity to change how they interact with the course and the institution.

5.6 Conclusion

This chapter has provided insight into what analytics can mean to an institution, how one might go about implementing analytics, and some of the expected outcomes of the application of analytics. Ultimately, however, analytics is more than just a tool. It is a framework for a process that can drive other institutional activities. The systematic collection and analysis of data that drives predictions for student success that can be acted upon and have the process refined over time is how an ideal implementation should be envisioned.

The manner in which this can effect and alter an institution is undeniable. A well-coordinated analytics implementation allows institutions to use existing data to identify and interpret trends that result in increased student success, retention, and graduation. It facilitates the refocusing of efforts and resources to identify, remediate, and enhance programs and services offered to students. It has forced institutions to carefully and deliberately consider how they contextualize data and, subsequently, act on it both legally and ethically. Furthermore, decisions have had to be made surrounding the types of data to use, how to collect it, and the extent to which an institution involves students in both the determination of use and collection of the data itself.

References

- Ali, L., Hatala, M., Gašević, D., & Jovanović, J. (2012). A qualitative evaluation of evolution of a learning analytics tool. *Computers & Education*, *58*, 470–489.
- Arnold, K. E. (2010). Signals: Applying academic analytics. *EDUCAUSE Review Online*. Retrieved from <http://www.educause.edu/ero/article/signals-applying-academic-analytics>
- Arnold, K. E., & Pistilli, M. D. (2012). Course signals at Purdue: Using learning analytics to increase student success. In S. B. Shum, D. Gašević, & R. Ferguson (Eds.), *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 267–270). New York: ACM. Retrieved from http://dl.acm.org/ft_gateway.cfm?id=2330666. doi: [10.1145/2330601.2330666](https://doi.org/10.1145/2330601.2330666).
- Astin, A. W. (1984). Student involvement: A developmental theory for higher education. *Journal of College Student Personnel*, *25*, 297–308.
- Astin, A. W. (1993). *What matters in college: Four critical years revisited*. San Francisco: Jossey-Bass.
- Astin, A. W. (1996). Involvement in learning revisited: Lessons we have learned. *Journal of College Student Development*, *37*, 123–134.
- Baepler, P., & Murdoch, C. J. (2010). Academic analytics and data mining in higher education. *International Journal for the Scholarship of Teaching and Learning*, *1*(2), 1–9.
- Bichsel, J. (2012, August). *Analytics in higher education: Benefits, barriers, progress and recommendations* (Research Report). Louisville, CO: EDUCAUSE. Retrieved from <http://net.educause.edu/ir/library/pdf/ERS1207/ers1207.pdf>
- Bonfiglio, R., Hanson, G. S., Fried, J., Roberts, G., & Skinner, J. (2006). Assessing internal environments. In R. P. Keeling (Ed.), *Learning reconsidered 2: Implementing a campus-wide focus on the student experience* (pp. 43–49). Champaign, IL: Human Kinetics. Retrieved from <http://www.myacpa.org/pub/documents/LearningReconsidered2.pdf>

- Bramucci, R., & Gaston, J. (2012). Sherpa: Increasing student success with a recommendation engine. In S. B. Shum, D. Gašević, & R. Ferguson (Eds.), *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 82–83). Retrieved from <http://delivery.acm.org/10.1145/2340000/2330625/p82-bramucci.pdf>. doi: 10.1145/2330601.2330625.
- Braxton, J. M. (1999). Theory elaboration and research and development: Toward a fuller understanding of college student retention. *Journal of College Student Retention*, 1(2), 93–97.
- Campbell, J. P. (2007). *Utilizing student data within the course management system to determine undergraduate student academic success: An exploratory study*. Unpublished doctoral dissertation. West Lafayette, IN: Purdue University.
- Campbell, J. P., DeBlois, P. B., & Oblinger, D. G. (2007). Academic analytics: A new tool for a new era. *EDUCAUSE Review*, 42(4), 40–52. Retrieved from <http://www.educause.edu/EDUCAUSE+Review/EDUCAUSEReviewMagazineVolume42/AcademicAnalyticsANewToolforaN/161749>
- Campbell, J. P., & Oblinger, D. G. (2007, October). Academic analytics. Washington, DC: EDUCAUSE. Retrieved from <http://www.educause.edu/ir/library/pdf/PUB6101.pdf>
- Chacon, F., Spicer, D., & Valbuena, A. (2012, April 10). Analytics in support of student retention and success. *Research Bulletin 3, EDUCAUSE Center for Applied Research* (pp. 1–9). Retrieved from <http://www.educause.edu/library/resources/analytics-support-student-retention-and-success>
- Chickering, A. W., & Gamson, Z. F. (1987). Seven principles for good practice in undergraduate education. *AAHE Bulletin*, 39(7), 3–7. Retrieved from <http://www.eric.ed.gov/PDFS/ED282491.pdf>
- Cooper, A. (2012, November). What is analytics? Definition and essential characteristics. *JISC CETIS*, 1(5), 1–10. Retrieved from <http://publications.cetis.ac.uk/2012/521>
- Dawson, S. P., Macfadyen, L., & Lockyer, L. (2009). Learning or performance: Predicting drivers of student motivation. In R. Atkinson & C. McBeath (Eds.), *Ascilite 2009: Same places, different spaces* (pp. 184–193). Auckland, NZ: Ascilite.
- Drachsler, H. & Greller, W. (2012). The pulse of learning analytics: Understandings and expectations from the stakeholders. In S. B. Shum, D. Gašević, & R. Ferguson (Eds.), *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 120–129). New York: ACM. Retrieved from <http://dl.acm.org/citation.cfm?id=2330634> doi: 10.1145/2330601.2330634.
- EDUCAUSE. (2010, April). 7 things you should know about... analytics. *Creative Commons Attribution*. Retrieved from <http://net.educause.edu/ir/library/pdf/ELI7059.pdf>
- Elias, T. (2011). Learning analytics: Definitions, processes and potential. *Creative Commons*. Retrieved from: <http://learninganalytics.net/LearningAnalyticsDefinitionsProcessesPotential.pdf>
- Essa, A. & Ayad, H. (2012). Student success systems: Risk analytics and data visualization using ensembles of predictive models. In S. B. Shum, D. Gašević, & R. Ferguson (Eds.), *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 158–161). New York: ACM. Retrieved from <http://dl.acm.org/citation.cfm?id=2330641>. doi: 10.1145/2330601.2330641.
- Ferguson, R. (2012). Learning analytics: Drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5), 304–317. doi:10.1504/IJTEL.2012.051816.
- Frankfort, J., Salim, K., Carmean, C., & Haynie, T. (2012, July). Analytics, nudges and learner persistence. *EDUCAUSE Review Online*. Retrieved from: <http://www.educause.edu/ero/article/analytics-nudges-and-learner-persistence>
- Greller, W., & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Journal of Educational Technology and Society*, 15(3), 42–57.
- Johnson, J. A. (2013, May 8). Ethics of data mining and predictive analytics in higher education. *Association for Institutional Research Annual Forum, Long Beach, California, May 19-22, 2013*. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2156058
- Kay, D., & van Harmelen, M. (2012). Analytics for the whole institution: Balancing strategy and tactics. *JISC CETIS*, 1(2), 1–31. Retrieved from: <http://publications.cetis.ac.uk/wpcontent/uploads/2012/11/Analytics-for-the-Whole-Institutionvol1-No2.pdf>

- Lockyer, L. & Dawson, S. (2011). Learning designs and learning analytics. *LAK '11: Proceedings from the 1st International Conference on Learning Analytics and Knowledge* (pp. 153–156). New York: ACM. doi: [10.1145/2090116.2090140](https://doi.org/10.1145/2090116.2090140).
- Lockyer, L. & Dawson, S. (2012). Where learning analytics meets learning design (workshop summary). In S. B. Shum, D. Gašević, & R. Ferguson (Eds.), *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 14–15). New York: ACM. Retrieved from <http://delivery.acm.org/10.1145/2340000/2330609/p14-lockyer.pdf>. doi: [10.1145/2330601.2330609](https://doi.org/10.1145/2330601.2330609).
- Long, P., & Siemens, G. (2011, September/October). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, *46*(5), 31–40.
- Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an “early warning system” for educators: A proof of concept. *Computers & Education*, *54*, 588–599. doi:[10.1016/j.compedu.2009.09.008](https://doi.org/10.1016/j.compedu.2009.09.008).
- Mattingly, K. D., Rice, M. C., & Berge, Z. L. (2012). Learning analytics as a tool for closing the assessment loop in higher education. *Knowledge Management & E-Learning: An International Journal*, *4*(3), 236–247.
- McKay, T., Miller, K., & Tritz, J. (2012). What to do with actionable intelligence: E²Coach as an intervention engine. In S. B. Shum, D. Gašević, & R. Ferguson (Eds.), *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 88–91). New York: ACM. Retrieved from <http://delivery.acm.org/10.1145/2340000/2330627/p88-mckay.pdf>. doi: [10.1145/2330601.2330627](https://doi.org/10.1145/2330601.2330627).
- McKown, C., Gregory, A., & Weinstein, R. S. (2010). Expectations, stereotypes and self-fulfilling prophecies in classroom and school life. In J. L. Meece & J. S. Eccles (Eds.), *Handbook of research on schools, schooling and human development* (pp. 256–274). New York: Taylor & Rutledge.
- Merton, R. K. (1948). The self-fulfilling prophecy. *The Antioch Review*, *8*(2), 193–210. doi:[10.2307/4609267](https://doi.org/10.2307/4609267).
- Oblinger, D. (Ed.). (2012). *Game changers: Education and information technologies*. EDUCAUSE. Retrieved from <http://net.educause.edu/ir/library/pdf/pub7203i.pdf>
- Pardo, A. & Kloos, C. D. (2011). Stepping out of the box. Towards analytics outside the learning management system. *LAK '11: Proceedings from the 1st International Conference on Learning Analytics and Knowledge* (pp. 163–167). New York: ACM. doi: [10.1145/2090116.2090142](https://doi.org/10.1145/2090116.2090142).
- Phillips, R., Maor, D., Cumming-Potvin, W., Roberts, P., Herrington, J., Preston, G., et al. (2011). Learning analytics and study behaviour: A pilot study. *Proceedings Ascilite Hobart 2011: Full Paper*, 997-1007. Retrieved from: <http://www.ascilite.org.au/conferences/hobart11/downloads/papers/Phillips-full.pdf>
- Picciano, A. G. (2012). The evolution of big data and learning analytics in American higher education. *Journal of Asynchronous Learning Networks*, *16*(3), 9–20. Retrieved from <http://www.eric.ed.gov/PDFS/EJ982669.pdf>
- Pistilli, M. D. & Arnold, K. E. (2010, July–August). Purdue signals: Mining real-time academic data to enhance student success. *About Campus*, *15*(3), 22–24. doi: [10.1002/abc.20025](https://doi.org/10.1002/abc.20025).
- Pistilli, M. D., Arnold, K. E., & Bethune, M. (2012, July). Signals: Using academic analytics to promote student success. *EDUCAUSE Review Online*. Retrieved from <http://www.educause.edu/ero/article/signals-using-academic-analytics-promote-student-success>
- Powell, S. & MacNeill, S. (2012). Institutional readiness for analytics. *JISC CETIS 1*(8), 1–11. Retrieved from <http://publications.cetis.ac.uk/2012/527>
- Romero, C., Ventura, S., Espejo, P. G., & Hervás, C. (2008). Data mining algorithms to classify students. *Educational Data Mining 2008: The 1st International Conference on Educational Data Mining, Montréal, Québec, Canada, June 20-21, 2008 Proceedings*, 8-17. Retrieved from <http://www.educationaldatamining.org/EDM2008/uploads/proc/full%20proceedings.pdf#page=8>
- Rouse, M. (2010). Business analytics definition. Retrieved from <http://searchbusinessanalytics.techtarget.com/definition/business-analytics-BA>
- Salas, G., & Alexander, J. S. (2008). Technology for institutional enrollment, communication, and student success. In R. Junco & D. M. Timm (Eds.), *Using emerging technologies to enhance*

- student engagement, New Directions for Student Services (124)* (pp. 103–116). San Francisco: Jossey-Bass. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1002/ss.298/pdf>. doi: 10.1002/ss.298.
- Schunk, D. H. (1985). Self-efficacy and classroom learning. *Psychology in the Schools*, 22(2), 208–223. doi:10.1002/1520-6807(198504)22:2<208::AID-PITS2310220215>3.0.CO;2-7.
- Siemens, G., & Baker, R. S. J. D. (2010). Learning analytics and educational data mining: Towards communication and collaboration. In S. B. Shum, D. Gašević, & R. Ferguson (Eds.), *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 252–254). New York: ACM. Retrieved from <http://delivery.acm.org/10.1145/2340000/2330661/p252-siemens.pdf>. doi: 10.1145/2330601.2330661.
- Siemens, G. & Long, P. (2011, September/October). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, 46(5), 30–32.
- Tanes, Z., Arnold, K. E., King, A. S., & Remnet, M. A. (2011). Using *Signals* for appropriate feedback: Perceptions and practice. *Computers & Education*, 57(4), 2414–2422. doi:10.1016/j.compedu.2011.05.016.
- Taylor, L., & McAleese, V. (2012, July). Beyond retention: Using targeted analytics to improve student success. *EDUCAUSE Review Online*. Retrieved from <http://www.educause.edu/ero/article/beyond-retention-using-targeted-analytics-improve-student-success>
- Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, 45, 89–125.
- Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition* (2nd ed.). Chicago: University of Chicago Press.
- Tinto, V. (2012). Moving from theory to action: A model of institutional action for student success. In A. Seidman (Ed.), *College student retention: Formula for student success* (2nd ed., pp. 251–266). Lanham, MD: Rowman & Littlefield.
- van Barneveld, A., Arnold, K. E., & Campbell, J. P. (2012). Analytics in higher education: Establishing a common language. *EDUCAUSE Learning Initiative*, 1, 1–11. Retrieved from <http://net.educause.edu/ir/library/pdf/ELI3026.pdf>
- van Harmelen, M., & Workman, D. (2012, November). Analytics for learning and teaching. *JISC CETIS*, 1(3), 1–40. Retrieved from <http://publications.cetis.ac.uk/2012/516>
- Willis, J. E., III, Campbell, J. P., & Pistilli, M. D. (2013, May 6). Ethics, big data, and analytics: A model for application. *EDUCAUSE Review Online*. Retrieved from <http://www.educause.edu/ero/article/ethics-big-data-and-analytics-model-application>