13

Learning Analytics and Educational Data Mining

Ryan S. Baker and George Siemens

Introduction

In recent years, the use of *analytics* and *data mining* – methodologies that extract useful and actionable information from large datasets – has become commonplace in science (i.e. Jing et al., 2018) and commerce (Erevelles et al., 2016; Wang et al., 2016). When applied to education, these approaches are referred to as *learning analytics* (LA) and *educational data mining* (EDM).

This work now appears in specialized journals and conferences specifically dedicated to work in this area – the *International Conference on Educational Data Mining,* the *International Learning Analytics and Knowledge Conference,* the *Journal of Educational Data Mining,* and the *Journal of Learning Analytics*, as well as more general educational research journals. The impact of learning analytics has grown – adaptive learning platforms have grown considerably in their user bases (Ubell, 2019), dropout prediction has emerged in K-12 education at even greater scale than its earlier emergence in higher education (Coleman et al., 2019), and the widespread adoption of analytics and its use to enhance learning systems has rapidly scaled in countries where learning analytics was rare even when the previous volume of this Handbook was published (i.e. Cui et al., 2018; Alkhalisi, 2019).

In this chapter we'll define LA and EDM and give many examples of how they have been used. For example, one can analyze the impact of the design of learning environment features on learner behavior (cf. Cheng et al., 2017; Harpstead et al., 2019); one can conduct fine-grained, even second-by-second, analysis of phenomena that occur over long periods of time (Yeung & Yeung, 2018; Hutt et al., 2019a); one can study how differences between groups influence the impact of behaviors on outcomes (cf. Karumbaiah et al., 2019); one can assess the impact that adaptive feedback has on subsequent learner behavior (Pardo et a., 2018); and one can study how individual behaviors impact the dynamics of groups (cf. Martinez-Maldonado et al., 2017; Järvelä et al., 2019). The data used in these analyses has broadened beyond interactive learning environment or learning management system data (i.e. Koedinger et al., 2010) to include multimodal sensor data (Blikstein & Worsley, 2016; Schneider & Radu, this volume) and a wide range of other information collected by schools and educational agencies (Bowers et al., 2010, 2019; Agasisti et al., 2019), from grade data, to disciplinary and attendance data, to principal and teacher employment data. In the sections that follow, we discuss how learning analytics can be used to increase the sophistication of models of learning, contributing both to theory and practice.

Developing Research Communities

The two research communities we review in this chapter, educational data mining and learning analytics, have adopted complementary perspectives on the analysis of educational data. The two communities have considerable overlap (in terms of both research topics and researchers), and the two communities believe in conducting research that has applications that benefit learners as well as informing and enhancing the learning sciences (Siemens & Baker, 2012). But although there is substantial overlap, there are also some relevant differences (Siemens & Baker, 2012):

1) Researchers in EDM are more interested in automated methods for discovery within educational data; researchers in LA are more interested in supporting people in better exploring educational data. This difference approximately tracks the relationship between data mining and exploratory data analysis, in the wider scientific literature.

2) Researchers in EDM emphasize modeling specific constructs (such as creating a model that can infer when a student is bored) and the relationships between them; researchers in LA emphasize a more holistic, systems understanding of constructs. This difference parallels long-standing differences in approaches among learning sciences researchers. EDM research in this fashion more closely ties to theoretical approaches such as the Pittsburgh Science of Learning Center Theoretical Framework (Koedinger, Corbett, & Perfetti, 2012); LA research more closely aligns with theory that attempts to understand systems as wholes or understand students as part of communities (Scardamalia & Bereiter, this volume).

3) Researchers in EDM tend to focus on applications of their work to automated adaptation, where educational software identifies a need and automatically changes to personalize learners' experience (cf. Ritter et al., 2016; DeFalco et al., 2018). Researchers in LA tend to focus on informing and empowering instructors and learners, such as informing instructors about ways that specific students are struggling so that the instructor can contact the learner (cf. Wise & Jung, 2019).EDM and LA methods are each suited to both types of use; the differences in focus are primarily due to the applications that were historically of interest to researchers in each community.

EDM and LA are fundamentally based on computational data analysis, but researchers in these two areas have emphasized theorical implications and analysis, drawing from and contributing to both learning sciences and educational theory more broadly. Most researchers that publish at the EDM and LA conferences use theoretical frameworks to guide their choice of analyses and aim to contribute back to theory with the results of their analyses. The theory-oriented perspective marks a departure of EDM and LA from technical approaches that use data as their sole guiding point (see, for example, Anderson's argument in 2008 that big data will render the scientific method obsolete: "But faced with massive data, this approach to science – hypothesize, model, test – is becoming obsolete."). In this regard, LA and EDM are part of broader movement that emphasizes computational methods in addressing questions in social processes (Lazer et al, 2019; also see Jacobson & Wilensky, this volume).

Key Methods and Tools

The methodologies used in EDM and LA have come from a number of sources, but the largest two sources of inspiration for the area have been methods from data mining and analytics in general, and from psychometrics and educational measurement. In many cases, the specific characteristics of educational data have resulted in different methods playing a more prominent role in EDM/LA than in data mining in general, or have resulted in adaptations to existing psychometric methods. In this section, we survey some of the key methodologies in the field and discuss a few examples of how these methodologies have been applied. This review draws on early reviews of these fields (cf. Baker & Yacef, 2009; Ferguson, 2012; Romero & Ventura, 2010; Siemens & Baker, 2012), but extends them to incorporate more recent developments.

Prediction Methods

One of the most prominent categories of EDM methodogies at the time of the review by Baker and Yacef (2009), and continuing to this day, is *prediction*. In prediction, the goal is to develop a model that can infer a single aspect of the data (the *predicted variable*, similar to dependent variables in traditional statistical analysis) from some combination of other aspects of the data (*predictor variables*, similar to independent variables in traditional statistical analysis). Developing a prediction model depends on knowing what the predicted variable is for a small set of data; a model is then created for this small set of data and statistically validated so that it can be applied at greater scale. For instance, one may collect data on whether 1,000 students dropped out of college, develop a prediction model to predict whether a specific student will drop out of college, validate it on subsets of the 1,000 students that were not included when creating the prediction model, and then use the model to make predictions about new students. As such, prediction models are commonly used either to predict future events (cf. Feng & Roschelle, 2016; Fancsali et al., 2018; Coleman et al., 2019) or to predict variables that are not feasible to directly collect in real time – for example, collecting data on affect or engagement (Jarvela & Renninger, this volume) in real time often requires expensive observations or disruptive self-report measures, whereas a prediction model based on logs (records) of student interaction with learning software can be largely nonintrusive (cf. Hutt et al., 2019b; Paquette & Baker, 2019).

These methods have been successfully used within interventions to improve student outcomes. For example, Civitas uses prediction models to identify students who are at risk for dropout in courses and university programs and uses this information in partnership with universities to increase student persistence and graduation rates (Milliron et al., 2017). If we can identify students at risk earlier, then we have more opportunities to develop interventions that may reduce student dropout. Additionally, platforms such as Civitas use these predictions to provide students with feedback directly, which has the potential to help students learn reason to about their own learning processes (see Winne & Azevedo, this volume).

Three types of prediction models are common in EDM/LA: *classifiers*, *regressors*, and *latent knowledge estimation*. In classifiers, the predicted variable can be either a binary (e.g., 0 or 1) or a categorical variable. Popular classification methods in educational domains include decision trees, random forest, logistic regression, support vector machines, and increasingly, neural network variants such as recurrent neural networks, long short-term memory networks, and convolutional neural networks. In regressors, the predicted variable is a continuous variable, for example a number.

In latent knowledge estimation (which is actually just a special type of classifier), a student's knowledge of specific skills and concepts is assessed by their patterns of correctness on those skills (and occasionally other information as well). The models used in online learning typically differ from the psychometric models used in paper tests or in computer-adaptive testing (see Pellegrino, this volume), because with an interactive learning application, the continual change in student's knowledge is visible to the system. A wide range of algorithms exist for latent knowledge estimation. There is currently an ongoing debate between whether it is better to use more classic, interpretable algorithms such Bayesian Knowledge Tracing (BKT – Corbett & Anderson, 1995), Performance Factors Analysis (PFA – Pavlik, Cen, & Koedinger, 2009), and ELO (Pelanek, 2016), or less interpretable algorithms that are better at predicting immediate student performance within a learning system, such as Deep Knowledge Tracing (DKT -- Piech et al., 2015) and Dynamic Key-Value Memory Networks (DKVMN – Zhang et al., 2017). Knowledge estimation algorithms increasingly underpin intelligent tutoring systems and adaptive learning systems, such Cognitive Tutors (cf. Koedinger & Corbett, 2006) and ALEKS, and form the foundation of what is called *adaptive learning* – systems where learners are provided with learning activities based their performance, rather than requiring all students to follow the same curriculum at the same pace.

Structure Discovery

Structure discovery algorithms attempt to find structure in the data without focusing in advance on a specific variable. This is a very different goal than in prediction. In prediction, the EDM/LA researcher attempts to create a model to predict a specific variable. By contrast, structure discovery has no specific variable of interest. Instead, the researcher attempts to determine what structure emerges naturally from the data; different approaches to structure discovery find different types of structure. Common approaches to structure discovery in EDM/LA include *clustering/latent class analysis*, *factor analysis*, *domain structure discovery*, and *network analysis*.

In clustering, the goal is to find data points that naturally group together, splitting the full dataset into a set of clusters. Clustering is particularly useful in cases where the most common categories within the dataset are not known in advance. If a set of clusters is well selected, each data point in a cluster will generally be more similar to the other data points in that cluster than to data points in other clusters. Clusters are most frequently used to group students in terms of their behaviors (i.e. Jovanovic et al., 2017; Khalil & Ebner, 2017), but are used for other purposes in learning analytics as well. A related approach, latent class analysis*,* is used to model groups in the data more statistically and can be used to track shifts between groups over time. It has been used to study school principals' leadership strategies (Agasisti et al., 2019), to study the difference between schools' academic offerings (Vaval et al., 2019), and to study how students approach gameplay differently within educational games (Slater et al., 2017).

In factor analysis, a closely related method, the goal is to find variables that naturally group together, splitting the set of variables (as opposed to the data points) into a set of latent (not directly observable) factors. Factor analysis is frequently used in psychometrics for validating or determining scales. In EDM/LA, factor analysis is used for dimensionality reduction (e.g., reducing the number of variables) for a wide variety of applications. For instance, Fincham and colleagues (2019) used factor analysis to understand how different student behaviors in learning management systems interrelate. Factor analysis is often also used for domain structure discovery, discussed below.

Domain structure discovery consists of finding the structure of knowledge in an educational domain (e.g., how specific content maps to specific knowledge components or skills across students). This could consist of mapping problems in educational software to specific knowledge components required in a subject area in order to group the problems effectively for latent knowledge estimation and problem selection (cf.

Cen, Koedinger, & Junker, 2006) or of mapping test items to skills (cf. Tatsuoka, 1995). There has been considerable work on this topic, both using approaches that take an existing model and attempt to improve it (Liu & Koedinger, 2017), and in a fully-automated bottom-up approaches that start from scratch (Desmarais, 2011; Matsuda et al., 2015; Vie & Kashima, 2019).

In network analysis, models are developed of the relationships and interactions between individual actors or elements, as well as the patterns that emerge from those relationships and interactions. The most common use of network analysis is within social network analysis (SNA), where the social relationships between people are studied. Used in learning settings long before learning analytics or educational data mining emerged as research areas (i.e. Haythornthwaite, 2001), social network analysis has been used by researchers to study the relationship between student patterns of usage of discussion forums and their academic outcomes (Joksimovic et al., 2016) and to identify sub-communities of learners (Jan & Vlachopolous, 2019). One criticism of SNA has been that it shows how people connect to each other, but not what the content of those connections are, or how they connections impact them. This has led to social network analysis being paired with other methodologies that more closely focus on meaning, social relationships, and linguistic interaction patterns; for example, SNA might be coupled with discourse analysis (see Enyedy & Stevens, Chapter 10, this volume; Buckingham, Shum, & Ferguson, 2012).

In recent years, the use of network analysis in education research has broadened far beyond social networks, with a larger community of researchers emerging who study a range of phenomena through *epistemic network analysis* (ENA --Shaffer, 2017), a method for representing the relationships in hand-coded or other categorical data. ENA has been used to study the arguments made by pre-service teachers (Bauer et al., 2019), to study group contributions in makerspaces (Espino et al., 2019), and to study identity exploration in video games (Barany & Foster, 2019).

Relationship Mining

In relationship mining, the goal is to discover relationships between variables in a dataset with a large number of variables. Relationship mining has historically been the most common category of EDM research (Baker & Yacef, 2009), and remains prominent to this day. It may take the form of attempting to find out which variables are most strongly associated with a single variable of interest, or may take the form of attempting to discover which relationships between any two variables are strongest. Broadly, there are four types of relationship mining: *association rule mining*, *sequential pattern mining*, *correlation mining*, and *causal data mining*.

In association rule mining, the goal is to find if-then rules of the form that "*if* some set of variable values is found, *then* another variable will generally have a specific value." For instance, Grawemeyer and colleagues (2017) used association rule mining to study how different behaviors and contexts within exploratory learning environments influenced changes in students' affect. Though this specific analysis involves behavior over time, association rule mining does not necessarily need to involve changes over time. A different method -- sequential pattern mining – captures complex patterns over time. In one study, Venant and colleagues (2017) used this method to study how students' behaviors changed over time when working in an inquiry learning environment.

In correlation mining, the goal is to find positive or negative linear correlations between variables (using post hoc corrections or dimensionality reduction methods when appropriate to avoid finding spurious relationships). An example can be found in Owen (2014), which looked at how different patterns of gameplay correlated to learning gains.

In causal data mining, the goal is to find whether one event (or observed construct) was the cause of another event (or observed construct), for example to predict which factors will lead a student to do poorly in a class (Fancsali, 2012; Holstein et al., 2017) or to understand the impacts of interventions within a learning context (Andor et al., 2018). All of these methodologies share the potential to find unexpected but meaningful relationships between variables; as such, they can be used for a wide range of applications, generating new hypotheses for further investigation or identifying contexts for potential intervention by automated systems.

Distillation of Data for Human Judgment (aka Visualization)

For data to be useful to educators, it has to be presented in a manner that readily allows the viewer to understand what they are seeing. When educators have immediate access to visualizations of learner interactions or misconceptions that are reflected in students' writing and interaction, they can incorporate those data quickly into pedagogical activity or plan a needed intervention. For this reason, one methodology that is common in LA is the *distillation of data for human judgment* (often referred to, more simply, as visualization). There has been a rich history of data visualization methods, which can be leveraged to support both basic research and practitioners (teachers, school leaders, and others) in their decision making. For example, visualizations of student trajectories through the school years can be used to identify common patterns among successful and unsuccessful students or to infer which students are at risk sufficiently early to drive intervention (Bowers, 2010). Some visualization methods specific to education include *learning curves* (which show performance over time – i.e. Koedinger et al., 2010; Peddycord-Liu et al., 2018) and visualizations of classroom layout (Holstein et al., 2017). There has also been increasing recent interest in the use of augmented reality (Schneider & Radu, Chapter 23, this volume; also see Holstein et al., 2018) and ambient displays (An et al., 2019) for visualizing information for teachers.

Discovery with Models

In *discovery with models* (Baker & Yacef, 2009; Hershkovitz, Baker, Gobert, Wixon, & Sao Pedro, 2013), the results of one data mining analysis are utilized within another data mining analysis. Most commonly, a model of some construct (such as boredom, or asking for help, or self-explaining) is obtained, generally through prediction methods. This model is then applied to data in order to identify where the construct appears within the data. The predictions of the model are then used as input to another data mining method. For example, the output of one prediction model may be used within another prediction model or in a relationship mining analysis. In such a situation, the initial model's predictions (which represent predicted variables in the original model) become predictor variables in the new prediction model. In this way, models can be composed of other models or based on other models, sometimes at multiple levels. For instance, prediction models of student robust learning (cf. Baker, Gowda, & Corbett, 2011) have been developed on top of models of student metacognitive behaviors (cf. Aleven, McLaren, Roll, & Koedinger, 2006), which have in turn depended on assessments of latent student knowledge (cf. Corbett & Anderson, 1995), which have in turn depended on models of domain structure (cf. Koedinger, McLaughlin, & Stamper, 2012).

Tools for Conducting EDM/LA Methods

In recent years, dozens of tools (software applications or packages within programming languages) have emerged for data mining and analytics, from both the commercial and academic sectors. We review some of the key trends here – a fuller treatment of the tools available for EDM and LAK can be found in (Slater et al., 2017).

Many EDM and LAK researchers have moved from using software applications that offer libraries of data mining algorithms to using packages implemented in Python and R. These packages support general-purpose machine learning algorithms, and also algorithms specific to learning data. For example, there are several packages that implement different latent knowledge estimation algorithms, several packages for epistemic network analysis and social network analysis, and several packages and tools for domain structure discovery. Tools are also found in repositories such as the Pittsburgh Science of Learning Center DataShop (Stamper et al., 2019). While the growing availability of tools, within an integrated "platform" such as R or Python, removes a technical layer for end users, it raises the importance of researchers having clear conceptual understanding of the problems they are trying to address.

Commercial software applications drive the administrative use of analytics in many schools and districts. Enterprise (sold to companies rather than individuals) software such as IBM Cognos Analytics, SAS Enterprise Miner, analytics offerings by learning management system providers such as Blackboard and Canvas, and analytics offered by student information system providers such as Ellucian and Infinite Campus enable an integrated research/application approach. The questions administrators and educators ask, however, can sometimes differ in focus from those EDM researchers ask. For example, a researcher may look for patterns in data and test algorithms or develop analytics models to understand what contributed to learning success. In contrast, when analytics is used by schools or colleges, there is likely to be a direct focus on improving learner success, on generating efficiencies in institutional systems such as course registration, and providing support services and programs. Increasingly, these goals are also supported by special-purpose student success analytics platforms such as BrightBytes Clarity/Early Warning System, Civitas, and Clever.

Impacts on Learning Sciences

Educational data mining and learning analytics have had several impacts on the learning sciences. In this section, we give two brief examples.

Engagement and disengagement

One area where these methods have been particularly useful is in research on engagement and disengagement within educational software (see Jarvela & Renninger, this volume). Prior to the development of analytics, disengagement was difficult to measure (Corno & Mandinach, 1983), but EDM and LA methods have produced models that can infer multiple forms of disengagement in a fine-grained (even second-by-second) fashion, including detectors of disengaged behaviors such as gaming the system

(Baker, Corbett, & Koedinger, 2004; Paquette & Baker, 2019), and detectors of negative emotions (Hutt et al., 2019). Surprisingly, work within the educational data mining community has demonstrated that some engaged behaviors are associated with worse outcomes. For example, – in "wheel-spinning" behaviors, a student works hard and is very engaged with the software, but nonetheless does not make progress on mastering the skill they are working on (Beck & Gong, 2013). Correspondingly, some seemingly disengaged behaviors may in fact reflect engagement, such as clicking through hints, obtaining the answer, but then self-explaining that answer (Shih, Koedinger, & Scheines, 2008). Recent work has demonstrated that detectors of engagement are predictive of learner outcomes over a decade later (Almeda & Baker, in press). Detectors have also been embedded into intelligent tutors that adapt based on student disengagement, such as DeFalco's (2018) work, which automatically detected and responded to frustration, leading to better learning outcomes.

Collaboration and learning

EDM and LA methods have been useful in understanding student learning in various collaborative and group-based settings. Collaborative learning behaviors have been analyzed in multiple contexts to determine which behaviors are characteristic of more successful groups and more successful learners, both in fully computer-mediated learning activities (McLaren, Scheuer, & Mikšátko, 2010; Zhang et al., 2019) and in collaboration activities taking place in face to face learning, such as collaboration around physical displays (Martinez et al., 2016) and nursing education involving simulated patients (Martinez et al., 2017). For instance, Jarvela and colleagues (2019) studied how self-regulated learning manifested in group learning activities, using multi-model learning data to fill in the gaps in a theoretical model of self-regulated learning. In another example, Dyke and colleagues (2012) found that off-topic discussions during collaborative learning are more harmful to learning during some parts of the learning process than during other parts – specifically, off-topic discussion is more harmful when learning basic facts than during discussion of problem-solving alternatives. This work has been embedded into learning systems that use animated characters or chatbots to scaffold more effective collaboration (McLaren et al., 2010; Dyke et al., 2012) and into tools to support instructors in scaffolding their students' collaboration (Martinez, Yacef, Kay, & Schwendimann, 2012; van Leeuwen & Rummel, 2017), providing an illustration of how fine-grained insights by LA researchers impact subsequent tools and technology development.

Impacts on Practice

There has often been a mutually reinforcing cycle between research and practice in EDM and LA – with research discoveries leading to changes in practice, which in turn lead to the possibility of studying new research questions. One example of this can be seen in research over the years on student knowledge in cognitive tutors (cf. Koedinger & Corbett, 2006). In the mid-1990s, mastery learning (where a student keeps receiving problems of a certain type until he or she successfully demonstrates mastery) was introduced, based on assessments of student mastery from the EDM Bayesian Knowledge Tracing algorithm (Corbett & Anderson, 1995). This technology led to better outcomes for students and has now scaled up to a wide variety of learning systems.

This technology also provided new opportunities to research further questions and impact other aspects of the learning experience. A prerequisite structure –an ordering of which content must be learned prior to other content because it is needed to understand the later content – was developed for the content in cognitive tutors and applied in the design of tutor lessons. However, instructors were free to deviate from the planned prerequisite structure in line with their pedagogical goals – in other words, if an instructor thought that students did not need to study a prerequisite topic before advancing to a later topic, they could skip the prerequisite topic. This enabled later researchers to study whether using the prerequisite structure benefitted students (Vuong, Nixon, & Towle, 2011). These results found that students learn less when instructors ignore prerequisite structure. This result has been replicated in more recent research on learning systems that recommend what students should work on based on prerequisite structure (Zou et al., 2019).

Another key area of application of learning analytics and educational data mining has been the reporting of important information back to instructors and teachers in a way that non-experts can understand. Dashboards— displays that provide key information to end users, typically using tables and simples graphs--have played a key role in communicating student risk assessments (Milliron et al., 2017) and info on what students know assessments of student knowledge (Xhakaj, Aleven, & McLaren, 2017; Ahn et al., 2018; Bywater et al., 2019). There has also been increasing focus on communicating learning analytics findings not just to instructors, but to students as well, to help student reflect on their learning and progress (see Winne & Azevedo, this volume; Teasley, 2017).

Ongoing Considerations

As the use of learning analytics increases, both in research and practice, new considerations move to the forefront. Perhaps the largest question that has emerged over the last few years is what uses of data, and what uses of analytics, are ethical. Data availability has decreased in Europe over the last few years, due to European privacy regulations, although the implementation and interpretation of these regulations has varied considerably from country to country. Learning providers in the United States and other countries who develop learning platforms used worldwide have had to adapt their practices in response to the EU's "right to be forgotten". There is an ongoing debate (see discussions in Lynch, 2017; Kitto & Knight, 2019) about what data is ethical to use, what data is ethical to share, and how researchers and practitioners can protect student privacy while being able to benefit learners with analytics models. On the other hand, some researchers have argued that there is an ethical "obligation to act", if learning analytics can improve student outcomes (Prinsloo & Slade, 2017). Students also vary considerably in how they want their data to be used, even for their own direct benefit (Arnold & Sclater, 2017). Beyond student privacy, there has also been increasing concern about algorithmic bias in education (Holstein & Doroudi, 2019), as in many other uses of artificial intelligence. For instance, research has raised concerns that some learning analytics models may not work equally well for historically underrepresented learners (Kai et al., 2017; Paquette et al., 2020).

Questions on what data sources are appropriate to use becomes increasingly relevant as learning analytics methods are applied to an ever-widening range of data sources. Much of the early work in EDM was conducted within intelligent tutoring systems (as described in Koedinger & Corbett, 2006) and much of the work in LA began in Web-based learning and social learning environments. In recent years, this has extended to a wider variety of educational situations, including data from a broader range of online learning environments (Slater et al., 2017; Gobert et al., 2018; Barany & Foster, 2019; Steinkuehler & Squire, Chapter 17, this volume) and collaborative learning environments (Dyke et al., 2012; Martinez et al., 2016; Zhang et al., 2019; Stahl, Koschmann, & Suthers, Chapter 21, this volume), more contextual data from school districts and universities (Bowers, 2010; Milliron et al., 2017; Coleman et al., 2019), as well as teacher learning (Fishman, Davis, & Chan, Chapter 32, this volume), and increasingly involves multimodal data sources (Martinez et al., 2016, 2017; An et al 2019; Järvelä et al., 2019). Emerging interest in psychophysiological data (Spann et al., 2017) and wearable technologies provide additional data, but also raise new ethical considerations.

As EDM and LA continue to be used in more and more educational research and practice, it takes its place as a key method for the 21st-century learning sciences. Finer-grained and more-detailed understanding of learners and learning, and more precisely targeted learning opportunities, are following accordingly.

Acknowledgments

We would like to thank Lisa Rossi and Keith Sawyer for helpful comments and suggestions.

References

Agasisti, T., Bowers, A. J., & Soncin, M. (2019). School principals' leadership types and student achievement in the Italian context: Empirical results from a three-step latent class analysis. *Educational Management Administration & Leadership*, *47*(6), 860-886.

Ahn, J., Campos, F., Hays, M., & DiGiacomo, D. (2019). Designing in Context: Reaching beyond Usability in Learning Analytics Dashboard Design. *Journal of Learning Analytics*, *6*(2), 70-85.

Aleven, V., McLaren, B., Roll, I., & Koedinger, K. (2006). Toward meta-cognitive tutoring: A model of help seeking with a cognitive tutor. *International Journal of Artificial Intelligence and Education*, **16**, 101–128.

Alkahlisi, Z. (2019) Abu Dhabi startup is using AI to transform how kids learn. Article published on *CNN Business,* March 5, 2019. Available from https://www.cnn.com/2019/03/05/tech/alef-education-ai-uae/index.html

Almeda, M.V.Q., Baker, R.S. (in press) Predicting Student Participation in STEM Careers: The Role of Affect and Engagement during Middle School. To appear in *Journal of Educational Data Mining.*

An, P., Bakker, S., Ordanovski, S., Taconis, R., Paffen, C. L., & Eggen, B. (2019). Unobtrusively enhancing reflection-in-action of teachers through spatially distributed ambient information. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1-14).

Anderson, C. (2008). The end of theory: The data deluge makes the scientific method obsolete. *Wired*.

Andor, M. A., Fels, K. M., Renz, J., & Rzepka, S. (2018). *Do planning prompts increase educational success? Evidence from randomized controlled trials in MOOCs* (No. 790). Ruhr Economic Papers.

Arnold, K.E., Sclater, N. (2017). Student Perceptions of their Privacy in Learning Analytics Applications. Proceedings of the Seventh International Learning Analytics & Knowledge Conference.

Baker, R. S., Corbett, A. T., & Koedinger, K. R. (2004). Detecting student misuse of intelligent tutoring systems. *Proceedings of the 7th International Conference on Intelligent Tutoring Systems*, 531–540.

Baker, R. S. J. d., Gowda, S. M., & Corbett, A. T. (2011). Automatically detecting a student's preparation for future learning: Help use is key. *Proceedings of the 4th International Conference on Educational Data Mining*, 179–188.

Baker, R. S. J. d., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining*, **1**(1), 3–17.

Barany, A., & Foster, A. (2019). Examining Identity Exploration in a Video Game Participatory Culture. In *International Conference on Quantitative Ethnography* (pp. 3-13). Springer.

Bauer, E., Sailer, M., Kiesewetter, J., Schulz, C., Pfeiffer, J., Gurevych, I., ... & Fischer, F. (2019). Using ENA to Analyze Pre-service Teachers' Diagnostic Argumentations: A Conceptual Framework and Initial Applications. In *International Conference on Quantitative Ethnography* (pp. 14-25). Springer.

Beck, J. E., & Gong, Y. (2013). Wheel-spinning: Students who fail to master a skill. In *International conference on artificial intelligence in education* (pp. 431-440). Springer, Berlin, Heidelberg.

Blikstein, P., & Worsley, M. (2016). Multimodal learning analytics and education data mining: Using computational technologies to measure complex learning tasks. *Journal of Learning Analytics*, *3*(2), 220-238.

Bowers, A. J. (2010). Analyzing the longitudinal K-12 grading histories of entire cohorts of students: Grades, data driven decision making, dropping out and hierarchical cluster analysis. *Practical Assessment, Research & Evaluation (PARE)*, **15**(7), 1–18.

Buckingham Shum, S., & Ferguson, R. (2012). Social learning analytics. *Educational Technology and Society*, **15**(3), 3–26.

Bywater, J. P., Chiu, J. L., Hong, J., & Sankaranarayanan, V. (2019). The Teacher Responding Tool: Scaffolding the teacher practice of responding to student ideas in mathematics classrooms. *Computers & Education*, *139*, 16-30.

Cheng, M. T., Rosenheck, L., Lin, C. Y., & Klopfer, E. (2017). Analyzing gameplay data to inform feedback loops in The Radix Endeavor. *Computers & Education*, *111*, 60-73.

Coleman, C., Baker, R., Stephenson, S. (2019) A Better Cold-Start for Early Prediction of Student At-Risk Status in New School Districts. *Proceedings of the 12th International Conference on Educational Data Mining*, 732-737.

Corbett, A. T., & Anderson, J. R. (1995). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction*, **4**, 253–278.

Corno, L., & Mandinach, E. B. (1983). The role of cognitive engagement in classroom learning and motivation. *Educational Psychologist*, **18**(2), 88–108.

Cui, W., Xue, Z., & Thai, K. P. (2018). Performance comparison of an AI-based Adaptive Learning System in China. In *2018 Chinese Automation Congress (CAC)* (pp. 3170-3175). IEEE.

DeFalco, J.A., Rowe, J.P., Paquette, L., Georgoulas-Sherry, V., Brawner, K., Mott, B.W., Baker, R.S., Lester, J.C. (2018) Detecting and Addressing Frustration in a Serious Game for Military Training. *International Journal of Artificial Intelligence and Education, 28* (2), 152-193.

Desmarais, M. C. (2011). Conditions for effectively deriving a q-matrix from data with non-negative matrix factorization. In C. Conati, S. Ventura, T. Calders, & M. Pechenizkiy (Eds.), *4th International Conference on Educational Data Mining, EDM 2011* (pp. 41–50). Eindhoven, Netherlands.

Dyke, G., Adamson, D., Howley, I., & Rosé, C. P. (2012). Towards academically productive talk supported by conversational agents. *Intelligent Tutoring Systems*, 531–540.

Erevelles, S., Fukawa, N., & Swayne, L. (2016). Big Data consumer analytics and the transformation of marketing. *Journal of Business Research*, *69*(2), 897-904.

Espino, D. P., Lee, S. B., Van Tress, L., & Hamilton, E. R. (2019). Examining the Dynamic of Participation Level on Group Contribution in a Global, STEM-Focused Digital Makerspace Community. In *International Conference on Quantitative Ethnography* (pp. 55-65). Springer.

Fancsali, S. (2012). Variable construction and causal discovery for cognitive tutor log data: Initial results. *Proceedings of the 5th International Conference on Educational Data Mining*, 238–239.

Fancsali, S. E., Zheng, G., Tan, Y., Ritter, S., Berman, S. R., & Galyardt, A. (2018). Using embedded formative assessment to predict state summative test scores. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (pp. 161-170).

Feng, M., & Heffernan, N. T. (2006). Informing teachers live about student learning: Reporting in the assistment system. *Technology Instruction Cognition and Learning*, *3*(1/2), 63.

Feng, M., & Roschelle, J. (2016) Predicting Students' Standardized Test Scores Using Online Homework. In *Proceedings of the Third (2016) ACM Conference on Learning @ Scale,* ACM Press, New York, NY, 213-216.

Ferguson, R. (2012). The state of learning analytics in 2012: A review and future challenges. *Technical Report KMI-12–01*, Knowledge Media Institute, The Open University, UK. http://kmi.open.ac.uk/publications/techreport/kmi-12-01.

Fincham, E., Whitelock-Wainwright, A., Kovanović, V., Joksimović, S., van Staalduinen, J. P., & Gašević, D. (2019). Counting clicks is not enough: Validating a theorized model of engagement in learning analytics. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge* (pp. 501-510).

Gobert, J. D., Moussavi, R., Li, H., Sao Pedro, M., & Dickler, R. (2018). Real-time scaffolding of students' online data interpretation during inquiry with Inq-ITS using educational data mining. In *Cyber-physical laboratories in engineering and science education* (pp. 191-217). Springer.

Halevy, A. Y., Norvig, P., & Pereira, F. (2009). The unreasonable effectiveness of data. *IEEE Intelligent Systems*, **24**(2), 8–12.

Harpstead, E., Richey, J. E., Nguyen, H., & McLaren, B. M. (2019). Exploring the subtleties of agency and indirect control in digital learning games. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge* (pp. 121-129).

Haythornthwaite, C. (2001). Exploring multiplexity: Social network structures in a computer-supported distance learning class. *The Information Society: An International Journal*, **17**(3), 211–226.

Hershkovitz, A., Baker, R. S. J. d., Gobert, J., Wixon, M., & Sao Pedro, M. (2013). Discovery with models: A case study on carelessness in computer-based science inquiry. *American Behavioral Scientist*, *57*(10), 1479-1498.

Holstein, K., & Doroudi, S. (2019). Fairness and equity in learning analytics systems (FairLAK). In . In *Companion Proceedings of the Ninth International Learning Analytics & Knowledge Conference (LAK 2019)*.

Holstein, K., Hong, G., Tegene, M., McLaren, B. M., & Aleven, V. (2018, March). The classroom as a dashboard: co-designing wearable cognitive augmentation for K-12 teachers. In *Proceedings of the 8th international conference on learning Analytics and knowledge* (pp. 79-88).

Holstein, K., McLaren, B. M., & Aleven, V. (2017). SPACLE: investigating learning across virtual and physical spaces using spatial replays. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 358-367).

Hutt, S., Gardner, M., Duckworth, A. L., & D'Mello, S. K. (2019a). Evaluating Fairness and Generalizability in Models Predicting On-Time Graduation from College Applications. *Proceedings of the International Conference on Educational Data Mining*.

Hutt, S., Grafsgaard, J. F., & D'Mello, S. K. (2019b). Time to scale: generalizable affect detection for tens of thousands of students across an entire school year. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1-14).

Jan, S. K., & Vlachopoulos, P. (2019). Social network analysis: A framework for identifying communities in higher education online learning. *Technology, Knowledge and Learning*, *24*(4), 621-639.

Järvelä, S., Järvenoja, H., & Malmberg, J. (2019). Capturing the dynamic and cyclical nature of regulation: Methodological Progress in understanding socially shared regulation in learning. *International Journal of Computer-Supported Collaborative Learning*, *14*(4), 425-441.

Jing, Y., Bian, Y., Hu, Z., Wang, L., & Xie, X. Q. S. (2018). Deep learning for drug design: an artificial intelligence paradigm for drug discovery in the big data era. *The AAPS journal*, *20*(3), 58.

Joksimović, S., Manataki, A., Gašević, D., Dawson, S., Kovanović, V., & De Kereki, I. F. (2016). Translating network position into performance: importance of centrality in different network configurations. In *Proceedings of the sixth international conference on learning analytics & knowledge* (pp. 314-323).

Jovanović, J., Gašević, D., Dawson, S., Pardo, A., & Mirriahi, N. (2017). Learning analytics to unveil learning strategies in a flipped classroom. *The Internet and Higher Education*, *33*(4), 74-85.

Kai, S., Andres, J.M.L., Paquette, L., Baker, R.S., Molnar, K., Watkins, H., Moore, M. (2017) Predicting Student Retention from Behavior in an Online Orientation Course. *Proceedings of the 10th International Conference on Educational Data Mining*, 250-255.

Karumbaiah, S., Ocumpaugh, J., Baker, R.S. (2019) The Influence of School Demographics on the Relationship Between Students' Help-Seeking Behavior and Performance and Motivational Measures. *Proceedings of the 12th International Conference on Educational Data Mining*, 99-108.

Khalil, M., & Ebner, M. (2017). Clustering patterns of engagement in Massive Open Online Courses (MOOCs): the use of learning analytics to reveal student categories. *Journal of Computing in Higher Education*, *29*(1), 114-132.

Kitto, K., & Knight, S. (2019). Practical ethics for building learning analytics. British Journal of Educational Technology.

Koedinger, K. R., Baker, R. S. J. d., Cunningham, K., Skogsholm, A., Leber, B., & Stamper, J. (2010). A data repository for the EDM community: The PSLC DataShop. In C. Romero, S. Ventura, M. Pechenizkiy, & R. S. J. d. Baker (Eds.), *Handbook of educational data mining* (pp. 43–56). Boca Raton, FL: CRC Press.

Koedinger, K. R., & Corbett, A. T. (2006). Cognitive tutors: Technology bringing learning science to the classroom. In K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 61–78). New York: Cambridge University Press.

Koedinger, K. R., Corbett, A. T., & Perfetti, C. (2012) The Knowledge-Learning-Instruction (KLI) framework: Bridging the science-practice chasm to enhance robust student learning. *Cognitive Science*, **36**(5), 757–798.

Koedinger, K. R., McLaughlin, E. A., & Stamper, J. C. (2012). Automated student model improvement. *Proceedings of the 5th International Conference on Educational Data Mining*, 17–24.

Lazer, D., Pentland, A. S., Adamic, L., Aral, S., Barabasi, A. L., Brewer, D., ... & Jebara, T. (2009). Life in the network: the coming age of computational social science. *Science (New York, NY)*, *323*(5915), 721.

Liu, R., & Koedinger, K. R. (2017). Closing the Loop: Automated Data-Driven Cognitive Model Discoveries Lead to Improved Instruction and Learning Gains. *Journal of Educational Data Mining*, *9*(1), 25-41.

Lynch, C. F. (2017). Who prophets from big data in education? New insights and new challenges. *Theory and Research in Education*, *15*(3), 249-271.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute.

Martinez, R., Yacef, K., Kay, J., & Schwendimann, B. (2012). An interactive teacher's dashboard for monitoring multiple groups in a multi-tabletop learning environment. *Proceedings of Intelligent Tutoring Systems*, 482–492. Springer.

Martinez-Maldonado, R., Goodyear, P., Kay, J., Thompson, K., & Carvalho, L. (2016, May). An actionable approach to understand group experience in complex, multi-surface spaces.

In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 2062-2074).

Martinez-Maldonado, R., Pechenizkiy, M., Buckingham Shum, S., Power, T., Hayes, C., & Axisa, C. (2017). Modelling embodied mobility teamwork strategies in a simulation-based healthcare classroom. In *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization* (pp. 308-312).

McLaren, B. M., Scheuer, O., & Mikšátko, J. (2010). Supporting collaborative learning and e-discussions using artificial intelligence techniques. *International Journal of Artificial Intelligence in Education (IJAIED)*, **20**(1), 1–46.

Milliron, M., Kil, D., Malcolm, L., & Gee, G. (2017). From innovation to impact: How higher education can evaluate innovation's impact and more precisely scale student support. *Planning for Higher Education*, *45*(4), 125-136.

Owen, V. E. (2014). *Capturing In-Game Learner Trajectories with ADAGE (Assessment Data Aggregator for Game Environments): A Cross-Method Analysis* (Doctoral dissertation). University of Wisconsin-Madison, Madison, WI.

Paquette, L., & Baker, R. S. (2019). Comparing machine learning to knowledge engineering for student behavior modeling: a case study in gaming the system. *Interactive Learning Environments*, *27*(5-6), 585-597.

Paquette, L., Ocumpaugh, J., Li, Z., Andres, J.M.A.L., Baker, R.S. (2020) Who's Learning? Using Demographics in EDM Research. *Journal of Educational Data Mining, 12* (3), 1-30.

Pardo, A., Bartimote, K., Shum, S. B., Dawson, S., Gao, J., Gašević, D., ... & Moskal, A. C. M. (2018). OnTask: Delivering data-informed, personalized learning support actions. *Journal of Learning Analytics*, *5*(3), 235-

Pavlik, P. I., Cen, H., & Koedinger, K. R. (2009). Performance factors analysis – A new alternative to knowledge tracing. *Proceedings of AIED2009*.

Peddycord-Liu, Z., Harred, R., Karamarkovich, S., Barnes, T., Lynch, C., & Rutherford, T. (2018). Learning Curve Analysis in a Large-Scale, Drill-and-Practice Serious Math Game: Where Is Learning Support Needed?. In *International Conference on Artificial Intelligence in Education* (pp. 436-449). Springer, Cham.

Pelánek, R. (2016). Applications of the Elo rating system in adaptive educational systems. *Computers & Education*, *98*, 169-179.

Piech, C., Bassen, J., Huang, J., Ganguli, S., Sahami, M., Guibas, L. J., & Sohl-Dickstein, J. (2015). Deep knowledge tracing. In *Advances in neural information processing systems* (pp. 505-513).

Prinsloo, P., & Slade, S. (2017). An elephant in the learning analytics room: the obligation to act. Proceedings of the Seventh International Learning Analytics & Knowledge Conference

Ritter, S., Yudelson, M., Fancsali, S. E., & Berman, S. R. (2016). How mastery learning works at scale. In *Proceedings of the Third (2016) ACM Conference on Learning@ Scale* (pp. 71-79).

Romero, C., & Ventura, S. (2010). Educational data mining: A review of the state-of-the-art. *IEEE Transaction on Systems, Man and Cybernetics, part C: Applications and Reviews*, **40**(6), 610–618.

Scardamalia, M., & Bereiter, C. (2014). Knowledge building and knowledge creation: Theory, pedagogy, and technology. *Cambridge handbook of the learning sciences*, *2*.

Shaffer, D. W. (2017). *Quantitative Ethnography.* Madison, WI: Cathcart Press.

Shih, B., Koedinger, K., & Scheines, R. (2008). A response time model for bottom-out hints as worked examples. *Proceedings of the 1st International Conference on Educational Data Mining*, 117–126.

Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, *57*(10), 1380-1400.

Siemens, G., & Baker, R. S. J. d. (2012). Learning analytics and educational data mining: Towards communication and collaboration. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*.

Slater, S., Bowers, A., Kai, S., & Shute, V. J. (2017). A typology of players in the game physics playground. In *Proceedings of the Digital Games Research Association DiGRA Conference*.

Slater, S., Joksimovic, S., Kovanovic, V., Baker, R.S., Gasevic, D. (2017) Tools for Educational Data Mining: A Review. *Journal of Educational and Behavioral Statistics*, 42 (1), 85-106.

Spann, C. A., Schaeffer, J., & Siemens, G. (2017). Expanding the scope of learning analytics data: Preliminary findings on attention and self-regulation using wearable technology. In *Proceedings of the seventh international learning analytics & knowledge conference* (pp. 203-207).

Stamper, J., Carvalho, P., Moore, S., & Koedinger, K. (2019) Tigris: An Online Workflow Tool for Sharing Educational Data and Analytic Methods. Gsqtermsr$Tvsgiihmrkw$=xl$Mrxivrexmsrep$ GsrjivirgisrPievrmrk$Erep}xmgw$*$Ors{pihki

Tatsuoka, K. K. (1995). Architecture of knowledge structures and cognitive diagnosis: A statistical pattern recognition and classification approach. In P. D. Nichols, S. F. Chipman, & R. L. Brennan (Eds.), *Cognitively diagnostic assessment* (pp. 327–359). Hillsdale, NJ: Lawrence Erlbaum Associates.

Teasley, S. D. (2017). Student facing dashboards: One size fits all?. *Technology, Knowledge and Learning*, *22*(3), 377-384.

Ubell, R. (2019) The Adaptive Learning Market Shakes Out. Article published in *Inside Higher Ed,* June 12, 2019. https://www.insidehighered.com/digital-learning/views/2019/06/12/explaining-shakeout-adaptive-learning-market-opinion

van Leeuwen, A., & Rummel, N. (2017). Teacher regulation of collaborative learning: research directions for learning analytics dashboards. *Making a Difference: Prioritizing Equity and Access in CSCL*, *2*(805-806), 1939-1382.

Vaval, L., Bowers, A. J., & Snodgrass Rangel, V. (2019). Identifying a typology of high schools based on their orientation toward STEM: A latent class analysis of HSLS: 09. *Science Education*, *103*(5), 1151-1175.

Venant, R., Sharma, K., Vidal, P., Dillenbourg, P., & Broisin, J. (2017). Using sequential pattern mining to explore learners' behaviors and evaluate their correlation with performance in inquiry-based learning. In *European Conference on Technology Enhanced Learning* (pp. 286-299). Springer, Cham.

Vie, J. J., & Kashima, H. (2019). Knowledge tracing machines: Factorization machines for knowledge tracing. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 33, pp. 750-757).

Vuong, A., Nixon, T., & Towle, B. (2011). A method for finding prerequisites within a curriculum. *Proceedings of the 4th International Conference on Educational Data Mining*, 211–216.

Wang, G., Gunasekaran, A., Ngai, E. W., & Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications. *International Journal of Production Economics*, *176*, 98-110.

Wise, A. F., & Jung, Y. (2019). Teaching with analytics: Towards a situated model of instructional decision-making. *Journal of Learning Analytics*, *6*(2), 53-69.

Xhakaj, F., Aleven, V., & McLaren, B. M. (2017). Effects of a teacher dashboard for an intelligent tutoring system on teacher knowledge, lesson planning, lessons and student learning. In *European conference on technology enhanced learning* (pp. 315-329). Springer, Cham.

Yeung, C. K., & Yeung, D. Y. (2018). Incorporating features learned by an enhanced deep knowledge tracing model for stem/non-stem job prediction. *International Journal of Artificial Intelligence in Education*, 1-25.

Zhang, X., Meng, Y., de Pablos, P. O., & Sun, Y. (2019). Learning analytics in collaborative learning supported by Slack: From the perspective of engagement. *Computers in Human Behavior*, *92*, 625-633.

Zhang, J., Shi, X., King, I., & Yeung, D. Y. (2017). Dynamic key-value memory networks for knowledge tracing. In *Proceedings of the 26th international conference on World Wide Web* (pp. 765-774).

Zou, X., Ma, W., Ma, Z., Baker, R. (2019) Towards Helping Teachers Select Optimal Content for Students. *Proceedings of the 20th International Conference on Artificial Intelligence in Education*, 413-417.