Expanding the Frame: Designing a Learning Analytics System Using a Theory of Learning

Philip J. Piety, Ph.D. University of Maryland College Park, ppiety@umd.edu

Abstract: This essay presents the case of designing a learning analytics system using a theory of learning. Learning analytics systems are often institutional artifacts using data collected from and to support educational practice and practitioners including learning, teachers, and administrators. There is a substantial and growing body of work under the learning analytics banner. Much of it framed technically around data harvested from digital tools and presentation mechanisms called dashboards. Using a specific case involving a collaborative game-based education research project, this paper provides a broad, sociotechnical design perspective through three frame expanding sections: the educational game's learning theory-driven approach, the information architecture of the learning analytics system, and the activity system that the information from learning analytics information are used within. This paper illustrates a portion of the conceptual landscape that can guide the design, development, and research for these data systems that are potentially consequential for students and educators.

Learning analytics: A significant movement to use data to support learning

Learning analytics is one of the most popular terms for activities that collect data out of educational contexts to then be used to inform and guide the learning processes from which the data came. This cycle of information collection and use has the potential to be consequential for students and educators as these human actors have their work of learning and portions of their lives represented through these increasingly popular analytic engines.

The consequential potential of learning analytics comes from not only the authority given to quantification and the labels of analytics or data science that carry cultural currency. The power of these tools is also in their nature. Many analytic efforts slope towards reductionism. Classification brings with it convenience in analysis. While an area might be complex and interdependent as we know learning is, from an analytic perspective the translation into data may distort the complex and interdependent nature of organized learning work creating challenges for the analytic endeavor's value proposition. In that translation could also be lost the knowledge gained from theories and research into how learning works in social settings. This paper's exploration of the landscape of learning analytics within an educational program designed using situated and sociocultural models of learning will show the depth that learning analytics design may involve and ways to understand the complexity.

The field of learning analytics is not clearly defined and discrete. Indeed, there are related movements known as *educational data mining* (Romero, & Ventura, 2007; Baker & Yacef, 2009) and *data driven decision-making* or *data use* (Ikemoto & Marsh, 2007; Marsh, 2012) that share much in common with learning analytics. There are additional names for these fields that emphasize various aspects such as *education data science* (Agasisti, & Bowers, 2017; Buckingham Shum, et al., 2013; Piety, Behrens, and Pea 2013; Piety, Hickey, and Bishop, 2014) and *education leadership data analytics* (Bowers, et al., 2019). Although these other names may suggest different aspects of the analytic program, they share the idea of using data from practice to inform practice. While different in their formulations, these areas have a number of characteristics in common, including little attention to theories of learning. Although a few learning analytics and education data mining programs have made contributions to research on learning as well as on research on data tools and practices (Piety, 2019), learning analytics studies tend to focus on technologies first.

The space between these fields and learning theory has been noted by various scholars in recent years (Papamitsiou & Economides, 2014; Piety, 2019). Gašević, Dawson, & Siemens, (2015) in a broad review of the field found both an emergent research base and opportunities for impact observing,

> The limited empirical research to date has revealed some significant issues that the field needs to consider and address in the future. The most significant is that learning analytics tools are generally *not* (emphasis added) developed from theoretically established instructional strategies (P. 65).

In the area of data driven decision-making, Penuel & Shepard (2017) identified a challenge may be how the field began using corporate rather than educational or learning theories noting that DDDM "does not have a theory of learning but rather derives from theories of organizational change …." (P. 795). Piety (2013) sketched out these origins in a sociotechnical review of the educational data movement highlighted expectations that education might be similar to the way other fields have been transformed by collecting and using data. Understanding these expectations that core organizational practices such as a learning environment rather than external outcome measures can be seen through data. This desire to use data to see into the "black box" of education have driven various accountability schemes and policies aimed at establishing the use of data as an organizational practice in education. While felt most strongly in the K-12 education space in the United States as a result of many years of federal policy, the interest in and press for data use exist in higher education and internationally as well.

It may be worth noting that learning analytics has large technological roots. Much of the early work in this field was in response to the voluminous data coming from massive online open courses (MOOCs) and similar distance education technology. Education data mining, similarly grew out of work with cognitive tutors and other learning technologies with a more individualized focus. The production of data from these learning tools fueled the initial work this area that has grown substantially to include a journal and conferences as has learning analytics.

There are some examples of recent work that pushes the boundaries of learning analytics. Blikstein & Worsley (2016) develop the concept of *multimodal learning analytics* that connect a field largely focused on data harvested from technical environments and testing towards the nature of learning as it occurs with students in rich semiotic spaces. Marzouk, et al. (2016) provide an example of how this can be done drawing off the literature of self-regulated learning, while also noting this kind of connection between learning theory and data research is rare. The project this paper is written about can be seen as somewhere between the broad view of Blikstein & Worsley (2016) and the focus of Marzouk, et al. (2016). It looks specifically at the relationship between learning theory and analytics involving an education technology. It draws off of a design study from a research grant into the use of an alternate reality game (ARG) to teach students about team-based, role-specific collaboration as part of a science, technology, engineering, and math (STEM) program.

The project this paper is based on calls for a multi-leveled and differentiated approach to learning analytics. It uses a sociocultural theory of learning known as *expansive framing* (Engle, et al., 2004). Expansive framing posits that the nature of the tasks students are given will improve their ability to transfer what they learn in classrooms to other settings. It is an approach strongly rooted in a situated and sociocultural tradition of learning and aligned with *the productive disciplinary engagement* framework developed by Engle and Conont (2002). While this project uses these specific models in its learning design, that can be seen as an example of a broader movement of designing educational approaches based on how students learn as articulated in the National Academies *How People Learn* (Bransford, Brown, and Cocking, 2000) and *How People Learn II* (National Academies of Sciences, Engineering, and Medicine. 2018).

In contributing to the important yet sparse literature on the relationships between learning theory and learning analytics, this paper aims for a comprehensive account of learning analytics. A recent review of the literature in learning analytics and related education data science activities found that systemic accounts were rare. Typically, learning analytics studies were of either technical components, relationships in the data, or visualization tools (1) (Piety, 2019). In addition, very few learning analytics studies used design-based research (DBR) that is often appropriate to understanding the emergent qualities of learning technologies (Barab & Squire, 2004; Brown, 1992; Collins and Bielaczyc, 2004). The project that forms the case study for this exploration is a DBR project that supports understanding an evolving learning analytics program.

Learning analytics as an embedded design science

Learning analytics can be seen as an embedded *design science* (Piety and Pea, 2018). Design sciences also known as *artificial sciences* such as architecture, medicine, and engineering solve human goals within constrains. This contrasts with *natural scientists* who endeavor to understand the fundamental and universal character of the physical or biological world. Both the designer of a bridge and a doctor are design scientists trying to accomplish human goals with a range of resources—including natural sciences—at their service. While this connection to design sciences is not common in the learning analytics literature, it accomplishes some important goals. It provides a language for the kinds of choices that learning analytics designers have to make in terms of what data to collect and how reliable that data must (or can) be, what to store, and then what to distribute and share.

Understanding the constraints that a design science operates under is fundamental to understanding its character. The design of bridges for example must work within the constraints of physical laws including gravity and the characteristics of various materials and construction approaches as well as project constraints of time and cost. Learning analytics is similar in that it must be done within some constraints characteristic of information systems that involve the use of various technical tools and the process of conversion (reduction) of real-world events into digital data as well as constraints specific to the educational process that shape learning analytics. Learning analytics constraints are both technical and social. Collecting data related to learning, transforming it to be useful in analysis are some of the technical constraints. Social constraints include the ways in which the data can be used; the time allowed for use and conditions of culture and praxis that may influence how information is used.

The constraints common with information systems, include classification (Bowker and Star, 2000) is information is placed into categories, often predetermined; and resemiotization (Idema, 2003) as one kind of information artifact (ex: the product of some student task) is inscribed in a digital code and then perhaps later rendered into a dashboard for use by student, teacher, or both. This process involves coding and classifying that invariably filter the information that transits the digital infrastructures (Rhibes and Finholt, 2009). If the product of student tasks are not encoded, then they are invisible in subsequent digital analysis. If the way they are encoded is coarse and reductionist then the nuance and detail of those events is filtered out of the system. The designers of information systems such as those that support learning analytics must make choices about what to encode and how to encode it. All of this entails implicit learning theories as illustrated in Figured 1a.

a. Learning analytics viewed structurally **b. Learning analytics as a programmatic process**

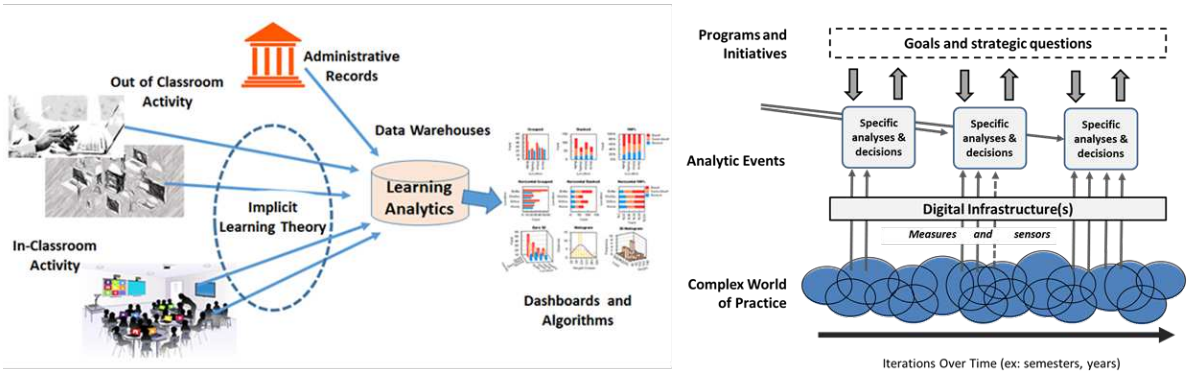


Figure 1. Learning analytics as an embedded design science.

Among educationally specific constraints are temporal, precision, and cultural factors. The temporal constraints are significant in that any program of learning analytics can only work within, the times(s) that the education is offered. The opportunities to collect information are typically available once a semester or year depending on the length of the course. As Figure 1b (adapted from Piety and Pea, 2018) illustrates, educational practice is an ongoing complex world where various measures and sensors with varying degrees of precision can collect data. That data that might include student assessment responses, in-class tasks, digital records of work in technical environments—*telemetry*—that are commonly recorded in logfiles.

Precision is a major issue for any educational measurement. While healthcare has many biomedical indicators such as heart rate, blood pressure, and body chemistry where high precision is achieved through *physical* sensors. Education mostly uses cognitive sensors—assessments and other tasks that can be distorted because of language, culture, circumstances. Educational assessments of even the highest quality have an imprecision that is fundamental. Defining what learning is and then developing a meaningful task that can approximate it is a great challenge across the curricular spectrum and more challenging in areas that entail student creativity and requiring interpretation. Additionally, this information once coded and presented will need to be interpreted by people whose cultural perspective and perceptions will influence the meaning they are given and the kinds of actions that may be taken as a result. Across the range of activities in the learning analytics process as illustrated in Figure 1a is an implicit learning theory. If learning is conceived of as direct and mechanic then the kinds of data collected will be narrow (ex: multiple choice test items) and the uses of that data similarly framed around sorting students and identifying knowledge gaps for remediation. A broader learning theory that focuses on development of practice competencies within a social system that sees learning as a multileveled and mediated activity will need to collect different data and use a more developed sociocultural approach to its evaluation and interpretation.

The learning theory we use in this project influences many of its design decisions. It is different from a simplified model of learning that considers knowledge as decomposable into small skills that can be measured using specific tasks such as a test item with a limited number of choices. That kind of learning model would feature scores from test items that could be subjected to an array of statistical analyses. However, the learning theory this project uses is highly situative and sociocultural. It prioritizes connections between students and between what happens in the classroom and in the real world. It puts students in situations where the kinds of tasks they encounter are broad from the simple test and survey items to richer team-based collaborative moves.

Expanding the frame: Three perspectives on a learning analytics project

To help explore the systemic nature of learning analytics, this paper presents three expansive frames. The first frame focuses on the learning environment that uses a combination of technologically mediated and facilitated classroom activities to improve students' understanding of science, technology, engineering, and math (STEM) careers. This learning innovation uses a situated model known as expansive framing (Engle and Conont, 2002). The second frame focuses on the specific information architecture of this project that takes data from a range of sources and uses it in an analytics process. The third frame looks at the activity systems (Engeström, 1999) of the use of that data that include individual students, groups of students, and instructors and teaching assistants.

Frame 1: Transmedia mediated learning environment

This learning analytics system is based on is a National Science Foundation funded project to use technology to support the development of collaboration skills and awareness of STEM careers. STEM can mean either any of these fields individually or a combination of those skills in service of a common goal. This combinatorial approach to STEM is reflective of many careers where professionals with an expertise work with other professionals with related expertise to solve real-world problems. It is built around this concept of a problem solving STEM. This project works with undergraduates in two different institutions: one a public university outside of an east coast urban center and the other a private university in a western state. Positing that students from more and less resources homes may have different experiences regarding STEM careers: those well-resources homes are more likely to have exposure STEM professions that those from less resourced homes, this project using a technology-assisted simulation designed to show students what working in a STEM job is like and then assess the extent to which the simulation supports student STEM career awareness and intent.

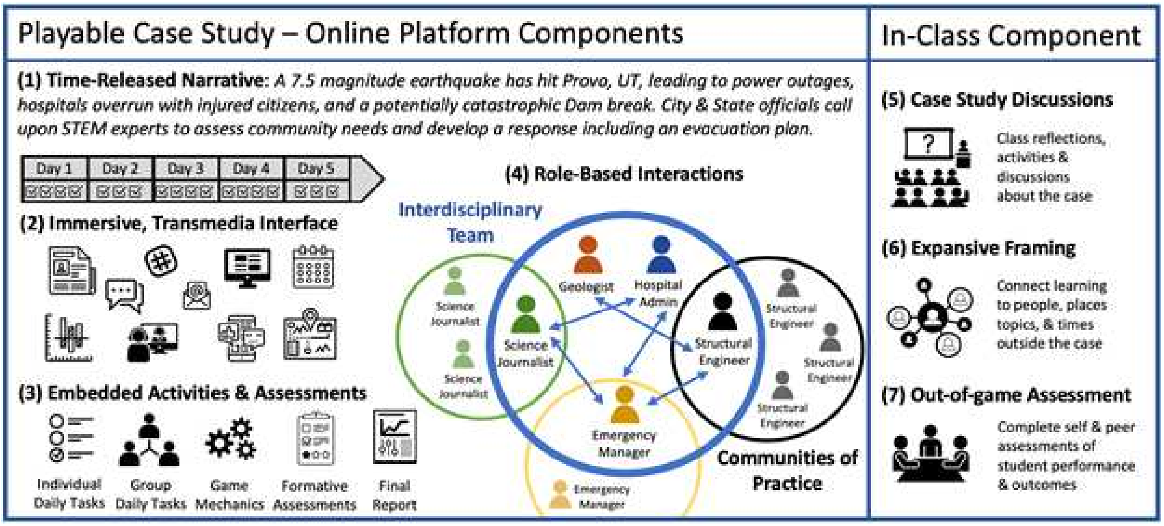


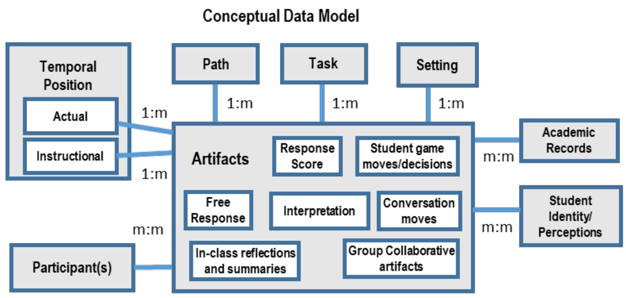
Figure 2. Playable Case Studies schematic.

The simulation used to teach students about STEM careers is a kind of epistemic game called a Playable Case Study (PCS). A PCS is a form of alternate reality game that provides students with a coherent and realistic narrative supported by transmedia experiences including video, chat, and fictional entities that the players interact with (Balzotti, et al., 2017; Bonsignore, et l., 2016). The scenario or narrative unfolds over a series of fictional days as illustrated in Figure 2. The students play unique team roles with other students & fictional characters in an immersive, transmedia narrative that unfolds over virtual days. The students take on fictional STEM roles and then work in teams. Each team will then play through fictional days in the technology environment and then have discussions in class that are facilitated to use expansive framing. The in-class components and indeed the conception of the activities leverages concepts of expansive framing (Engle and Conont, 2002; Engle et al., 2012) that push learners to frame their engagement by referencing people, places, topics, and times that are personally relevant and come from beyond the boundaries of a specific learning simulation to increase transfer. The program operationalized collaborative STEM work on the scale of communities with use two curricular versions: cybersecurity and disaster response (ex: epidemics, earthquakes, and floods).

As the students engage with the narrative, working as a team to solve the challenges presented to them, they are provided with role-specific information. In the cybersecurity version, the security analyst role will have access to certain kinds of information and the public safety role will have access to other information. Similarly, in the disaster response version the roles are different including emergency manager and hospital specialist. In both cases, the learning activities are designed around information access and translation. The PCS technology has the ability to support text and emails between players and this information along with other game telemetry is directly recorded in the learning analytics system. Information from both the game play and also from their classroom settings, including surveys and quizzes, is sent to the learning analytics platform for subsequent use by teachers and students.

Frame 2: Information architecture as theory

Learning analytics systems create an informational lens onto practice. The view the learning analytics information system provides is not the only one available to teachers and students. It is one that can privilege and prioritize those elements that have been encoded into their designs. As with other infrastructures, they enable and constrain actions through what they make convenient as they become invisible (Rhibes & Finholt, 2009). Their information architecture reflects a theoretical perspective privileging certain representations of knowledge by collecting and retaining that information over others not represented in the architecture (Piety & Palincsar, 2006).



Artifact Category	Examples	Source	Inscription Method
Student choices	Responses on multiple choice items	Tests through game or LMS	Downloaded from LMS Transmitted from PCS
	In-game moves	Game tasks	Transmitted from PCS
Student free responses	Responses to open-ended items	Tests through game or LMS	Downloaded from LMS Transmitted from PCS
Conversational moves between students/agents	Email threads Student to chat-bot conversations	Game platform	Transmitted from PCS
Student decisions reflected in game moves	Students allocate resources to respond to game conditions	Game platform	Transmitted from PCS
Student in-class reflections and summaries	Student paper	Graded papers entered in LMS	Downloaded from LMS
Collaborative artifacts	Team-developed summary paper	Graded papers entered in LMS	Downloaded from LMS
Student academic records	Student major and GPA	Administrative systems	Loaded from administrative system
Student personal history	Ethnicity, SES, Family academic history	Optional survey	From LMS

Figure 3. Conceptual information architecture of the learning analytics engine.

The design and development of this learning analytics system followed a common path for information systems of data analysis leading to design of some underlying structures. Unlike other information domains, this process began with artifacts generated in learning contexts and from other sources that describe learners and learning-relevant activity. These artifacts are then translated into abstractions in the form of database tables, columns (variables), and relationships. These abstractions can be used for different kinds of analyses, both quantitative and qualitative, as inferences are made about what these data show about what occurred in the learning process. The resulting design used generalized/reusable data structures as Figure 3 illustrates (2).

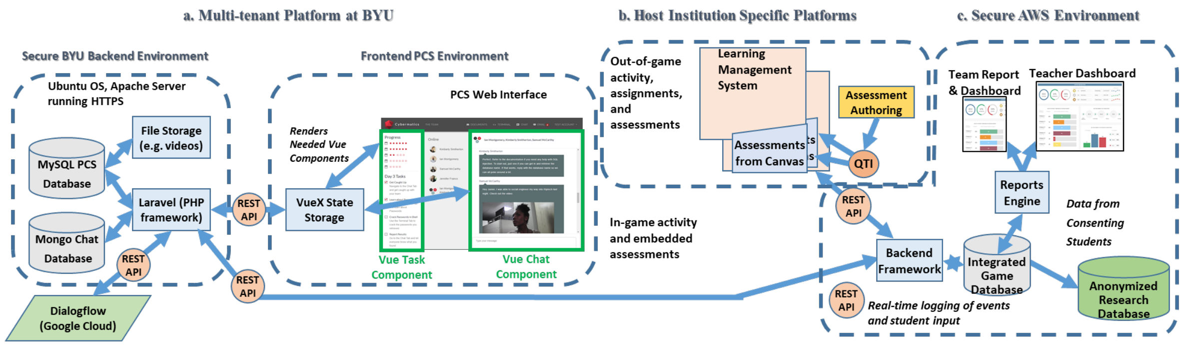


Figure 4. High-level conceptual schematic of the technical architecture.

Additional practical and technical concerns motivated the design. Since the PCS technology was to be separate from the secure analytics environment two cloud environments were developed: a semi-proprietary for the PCS support that used a model that allowed multiple kinds of concurrent uses (multi-tenant) and an open source code running in a secure cloud environment running on Amazon Web Services (AWS). Additionally, the structure allows for different host institutions that might extend beyond the initial grant collaborators to participate using their own learning management systems. The design needed to account for both secure management of student data that could be collected from multiple locations/institutions as well as only allowing research to be conducted on data collected from students who had given permission.

Frame 3: Activity systems for interpretation and action

One of the three research questions for this project involves understanding the role and impact of a learning analytics system embedded in instructional practice. Rather than an external evaluation engine, the information collected from the learning environment will be presented to the teachers and students in some summarized way that could be called dashboards, although the actual form is being designed presently. The design process being undertaken is to construct the information presentation in a way to reflect the nature of the learning process and what the narrative is to elicit from students a responsive to their learning processes. This is an area that is currently undergoing design and will be the topic of a separate conference paper addressing the visual affordances (Gibson, 1977; Greeno, 1994) of various displays such as quantitative dashboards for influencing desired learning outcomes.

Independent of the form of these data displays—a matter that the project plans to experiment with suspecting that traditional quantitative composite displays known as dashboards may not be optimal—the presentation of information to students and teachers will be done as part of activity structures where the information presented is seen as a mediating tool towards a particular outcome (Engeström, 1999) that for students will be integrated in their alternate reality. During the PCS enactment and subsequent to it, teams and individuals will be presented with information about their collaboration. These opportunities—instructional affordances—may occur both within the context of the PCS virtual day system as well as within the instructional calendar whereby at specific points in the PCS narrative and afterwards when the teams are being debriefed. Following the portion of the course the PCS is used in, typically four class sessions, students are provided with individual and group assignments to reinforce the goals of the program to develop an understanding of the collaborative nature of STEM work. These assignments will include an after-action report that allows students to gain credit for their learning even if they have not "won" the game. These activities after the PCS and some that will occur before it in the class are part of a design where the PCS is not seen as the totality of the research, but rather one element of an instructional ecosystem that is being studied. Similarly, the learning analytics tools are also seen as not standing on their own, but rather as situated within this curricular and research context.

Conclusions and implications

Much of learning analytics research has focused on the technical domain of collecting data and producing dashboards (Piety, 2019). Little attention has been paid to the designs of these systems and how those designs may inscribe values and encode implicit learning theories. This paper shows how a situated and sociocultural approach applies not only to the site of instruction, but the entire evidentiary and iterative process. This paper contributes to an important and underdeveloped portion of the knowledge base involving the design of learning analytics systems. Designs of information systems involve making choices that are inscribed into technology (Friedman, Kahn, & Borning, 2008; March & Smith, 1995). These choices are reflected in the form of instruments and data capture technology, in the underlying data storage (ex: databases), and in the presentation systems (reports and dashboards) that use that information. These design choices can be aligned in that they span the system, although it is possible to collect information that is then filtered out by the design of the storage approach or is collected and stored, but not shown. Advances in technology make this easier, but this kind of disjointed design is typically not employed and designers of all information systems, including learning analytics systems are generally guided by common viewpoints and assumptions. The viewpoints and assumptions can be encoded in technological models (ex: data models) that become infrastructural. The decisions about what to abstract and how to do this are not automatic or value neutral. Rather they involve choices about what to value. In learning analytics the models that undergird these institutional artifacts are theories of learning, albeit often implicit.

The way that learning is conceived of can influence how the learning analytics system can be designed. Where learning is viewed as a process of small bits of decomposed knowledge being transmitted from teacher to learner then the learning analytics process is akin to an inventory system tracking what learning is being registered in what learning categories. This view aligns with simple visualizations such as dashboards composed of colored bars that can readily show data in discrete categories. Where learning is seen as more complex, interconnected, and consisting of actions and artifacts difficult to reduce then the learning analytics design will need to evolve to reflect this contextual complexity.

It is perhaps worth noting that during the early stages of the various data movements that little impact on actual practice was found. There may be many reasons for this lack of impact demonstrated through research, but one possibility might be that the designs of these systems are oriented around what is technically straightforward rather than what makes sense for learning. Assessments of student learning have long been dominated by what Behrens and Dicerbo (2014) call a *correctness paradigm* that requires student data to be normatively framed. Teachers and other educational stakeholders seeking to use these systems will almost certainly seek information that relates to the learning that is specific to that setting rather than some technical abstraction that does not relate to the kind of learning that is occurring.

This project provides a richer view. This paper has developed a model for looking at how learning analytics systems can be informed by a sociocultural theory of learning. In expanding the frame of learning analytics from the technical matter of data and dashboards that dominate the current literature to a broader sociotechnical view, this paper has focused on three areas. These three areas of the learning environment, the information architecture, and the activity systems for use cover important parts of the life-cycle of educational programs. These three areas are not the totality of educational programs and significant processes that might relate to learning analytics and other forms of data use to guide educational work, but three of the most important features of this complex landscape.

Endnotes

(1) See Piety (2019) for a taxonomy and more details, but the largest category of pieces in the learning analytics literature is conceptual models with empirical studies having a smaller share.
(2) The figure is simplified for an educational audience. The actual implementation utilizes a de-normalized structure with several cloud-based data stores.

References

Agasisti, T., & Bowers, A. J. (2017). 9. Data analytics and decision making in education: Towards the Educational Data Scientist as a key actor in schools and higher education institutions. In Handbook of contemporary education economics (p. 184). Edward Elgar Publishing.

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. The Journal of the Learning Sciences, 13(1), 1-14.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. The journal of the learning sciences, 2(2), 141-178.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. The Journal of the learning sciences, 13(1), 15-42.

Balzotti, J., Hansen, D., Ebeling, D., & Fine, L. (2017, January). Microcore: A playable case study for improving adolescents' argumentative writing in a workplace context. In Proceedings of the 50th Hawaii International Conference on System Sciences.

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

Behrens, J. T., & DiCerbo, K. E. (2014). Harnessing the currents of the digital ocean. In Learning analytics (pp. 39-60). Springer, New York, NY.

Bonsignore, E., Hansen, D., Pellicone, A., Ahn, J., Kraus, K., Shumway, S., … Koepfler, J. (2016). Traversing Transmedia Together: Co-designing an Educational Alternate Reality Game For Teens, With Teens. In Proceedings of the 15th International Conference on Interaction Design and Children IDC 2016 (pp. 11–24). NY, NY: ACM.

Bowers, A. J., Bang, A. H., Pan, Y., & Graves, K. E. (2019). Education Leadership Data Analytics (ELDA): A White Paper Report on the 2018 ELDA Summit.

Bowker, G. C., & Star, S. L. (2000). Sorting things out: Classification and its consequences. MIT press.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn*. Washington, DC: National academy press.

Buckingham Shum, S., Hawksey, M., Baker, R. S., Jeffery, N., Behrens, J. T., & Pea, R. (2013, April). *Educational data scientists: a scarce breed*. In Proceedings of the third international conference on learning analytics and knowledge (pp. 278-281). ACM.

Engeström, Y. (1999). *Activity theory and individual and social transformation.* Perspectives on activity theory, 19(38).

Engle, R. A., & Conant, F. R. (2002). *Guiding principles for fostering productive disciplinary engagement: Explaining an emergent argument in a community of learners classroom.* Cognition and instruction, 20(4), 399-483.

Engle, R. A., Lam, D. P., Meyer, X. S., & Nix, S. E. (2012). *How does expansive framing promote transfer? Several proposed explanations and a research agenda for investigating them.* Educational Psychologist, 47(3), 215-231.

Friedman, B., Kahn, P. H., & Borning, A. (2008). *Value sensitive design and information systems*. The handbook of information and computer ethics, 69-101.

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. TechTrends, 59(1), 64-71.

Gibson, J. J. (1977). *The theory of affordances.* Hilldale, USA, 1(2).

Greeno, J. G. (1994). *Gibson's affordances.* Psychological Review, 101(2), 336–342.
Greller, W., & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. Education Technology & Society, 15, 3, 42–57.
Iedema, R. (2003). Multimodality, resemiotization: Extending the analysis of discourse as multi-semiotic practice. *Visual communication*, 2(1), 29-57.
March, S. T., & Smith, G. F. (1995). *Design and natural science research on information technology.* Decision support systems, 15(4), 251-266.
Marzouk, Z., Rakovic, M., Liaqat, A., Vytasek, J., Samadi, D., Stewart-Alonso, J., ... & Nesbit, J. C. (2016). What if learning analytics were based on learning science? Australasian Journal of Ed. Technology, 32(6).
National Academies of Sciences, Engineering, and Medicine. (2018). *How people learn II: Learners, contexts, and cultures*. National Academies Press.
Papamitsiou, Z., & Economides, A. A. (2014). Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Journal of Educational Technology & Society*, 17(4), 49-64.
Penuel, W. R., & Shepard, L. A. (2017). Social models of learning and assessment. The Wiley handbook of cognition and assessment: Frameworks, methodologies, and applications, 146-173.
Piety, P. J., Behrens, J., & Pea, R. (2013). *Educational data sciences and the need for interpretive skills*. American Educational Research Association Annual Meeting, 27.
Piety, P. J., Hickey, D. T., & Bishop, M. J. (2014, March). *Educational data sciences: Framing emergent practices for analytics of learning, organizations, and systems*. In Proceedings of the fourth international conference on learning analytics and knowledge (pp. 193-202). ACM.
Piety, P. J., & Palincsar, A. S. (2006, June). *How do we see?: information architecture as theory*. In Proceedings of the 7th international conference on Learning sciences (pp. 536-542). International Society of the Learning Sciences.
Piety, P. J., & Pea, R. D. (2018). Understanding learning analytics across practices. *Learning analytics in education*, Information Age 215-232.
Piety, P. J. (2019). Components, Infrastructures, and Capacity: The Quest for the Impact of Actionable Data Use on P–20 Educator Practice. *Review of Research in Education,* 43(1), 394-421.
Ribes, D., & Finholt, T. A. (2009). The long now of infrastructure: Articulating tensions in development.
Romero, C., & Ventura, S. (2007). *Educational data mining: A survey from 1995 to 2005. Expert systems with applications*, 33(1), 135-146.
Siemens, G., & Baker, R. S. (2012, April). *Learning analytics and educational data mining: towards communication and collaboration*. In Proceedings of the 2nd international conference on learning analytics and knowledge (pp. 252-254). ACM.
Simon, H. A. (1969). *The sciences of the artificial.* MIT press.

Acknowledgments
This work is supported by the NSF Division of Undergraduate Education award 1915563.

