The ASU Learning at Scale (ASU L@S) Digital Learning Network Platform

Danielle S. McNamara, Tracy Arner
Department of Psychology
Arizona State University
Tempe, AZ USA
{dsmcnama, tarner}
@asu.com

Elizabeth Reilley, Paul Alvarado
University Technology Office
Arizona State University
Tempe, AZ USA
{ereille1, palavara2}
@asu.com

Chani Clark, Thomas Fikes, Annie Hale, Betheny Weigele
EdPlus
Arizona State University
Tempe, AZ USA
{cclark8, tgfikes, aehale1, bweigele}
@asu.com

ABSTRACT

Accounting for complex interactions between contextual variables and learners' individual differences in aptitudes and background requires building the means to connect and access learner data at large scales, across time, and in multiple contexts. This paper describes the ASU Learning@Scale (L@S) project to develop a digital learning network platform with the capacity to connect, access, and examine undergraduate student data and courses. Arizona State University (ASU) collectively serves over 100,000 undergraduate students and 40,000 K-12 students within traditional in-person courses, online and blended courses, Earned Admission, and ASU Prep Online (K-12). Combined, these programs at ASU provide access to the large-scale data needed to improve the generalizability of learning sciences research. Foremost, we are challenged by this wealth of data at ASU currently being housed in multiple locations, with varying restrictions, and means of access. The ASU L@S team is developing the data and policy infrastructure necessary to enable data access while supporting multiple types of research (e.g., historical data analyses, rapid A/B testing, efficacy studies, design studies). Our objective is to render learner data and educational contexts available to both internal and external researchers, and to facilitate researchers' ability to conduct research more efficiently on ASU student learning while maintaining the safety of all stakeholders' personally identifiable information. Here we describe the three primary datasets currently being compiled within the ASU L@S data warehouse (i.e., Student Profile, Student Trajectory, Course Profile) and two datasets leveraging natural language produced by learners in various contexts: discussion board posts and written assignments. Combining and integrating these datasets within a single data warehouse sets the stage to enable impactful embedded research at ASU that enhances student outcomes and in turn contributes to theories of how people learn.

Author Keywords

A/B testing; Digital Learning Network; big data; learning analytics

Current Landscape of Learning Sciences Research

Education research has widely adopted the Randomized Controlled Trial (RCT) as the 'Gold Standard' to evaluate interventions grounded in the learning sciences. Learning sciences research conducted using RCTs has yielded powerful insights regarding the impact of various factors related to learning context and individual differences. While the RCT is considered the optimal method for evaluating the impact of learning interventions, there are inherent limitations such as constrained sample sizes and less diverse samples. When combined with the natural complexity associated with learning and the individual differences of learners, reproducing the results of individual RCTs remains challenging [1]. Further, translating the insights garnered from individual RCTs to improved outcomes in real-world contexts, particularly across diverse learner populations has been limited. Consequently, modifications in pedagogy, instructional design, and student support or interaction fail to demonstrate effectiveness beyond their initial learner populations, contexts, and scales [2-8]. Despite the failure to replicate, many instructional interventions or methods are adopted with the hope that addressing dosage, fidelity, or practice will lead to the improved outcomes demonstrated in original research [6]. This failure to replicate is among the most significant and complex challenges to the scientific study of learning and its practical applications. It also poses a hindrance to embracing diversity and inclusion, which further complexifies efforts to attain equitable outcomes. Such complexities, in combination with the multivariate and multidimensional nature of learning, create a combinatorial explosion that traditional experimental approaches cannot accommodate [9].

Overcoming this challenge requires a different approach that includes fine-grained, representative data about learners and the contexts in which learning happens. Further, this work cannot be done in isolation, necessitating a partnership between learning and data scientists to maximize the collection of meaningful data and their effective use. One answer to this call emerged from the Institute for Education Sciences creation of the Digital Learning Network to support digital learning platforms in conducting rigorous education research and support researchers as they study new ideas and seek to replicate studies more efficiently across multiple sites and a wide range of student populations. ASU Learning@Scale (L@S) is one of the platforms funded within this IES initiative (https://seernet.org). The ASU L@S platform is one of several digital learning platforms conceived to address many of the issues encountered in current methods of conducting educational research. The platform, and the context in which it is being developed, provide a broader array of data from a diverse population on a much larger scale than typically available to researchers. Ultimately, ASU L@S is geared toward enabling agile studies of education and learning, including not only large-scale data analyses, but also experimental studies situated within ASU classrooms as well as the larger ecosystem. Achieving this objective necessitates first devising the means for researchers to access student and course data within a single platform. This paper describes the ASU L@S data warehouse and the research opportunities that it will afford.

ASU L@S Project

The purpose of the ASU L@S project is to develop a digital learning network platform with the capacity to connect, access, and examine undergraduate student data and courses within the scope of ASU online and digital classrooms (ASU Online). The ASU L@S platform focuses on how students and courses are represented within the ASU ecosystem and how those data and research opportunities can be surfaced to the widest range of researchers. Our long-term vision is to provide the infrastructure to promote innovation in research, emerging technologies, and community outreach to enhance universal, lifelong learning to students around the world. The ASU L@S project will develop foundational infrastructure and protocols to connect a wide range of data available in the ASU data ecosystem on student achievement, learning, and persistence, make those data accessible to researchers across and beyond ASU in ways that honor institutional and individual privacy, so that it can examined through exploratory and experimental methods. The platform will leverage diverse types of data from large learner populations and facilitate experimental studies that examine the impact of a wide range of learning tools and approaches to enhance learning. Furthermore, inclusion of demographic data from a highly diverse population of students will empower researchers to advance inclusion and equity-oriented research to benefit all students.

A primary aim of this project is to develop the data and policy infrastructure necessary to enhance features, dashboards, datasets, data-sharing tools, and other system components needed for researchers without access to identified data to conduct collaborative research that focuses on enhancing academic success and achievement for ASU students and courses. The platform is ultimately envisioned to support multiple types of research including analyses of existing data, efficacy studies, replication studies, rapid A/B testing, and design studies with the overarching objective of contributing to theories of how people learn, both holistically and at scale. ASU L@S will iteratively develop the capacity to collect multiple types of user activity and interaction data (e.g., interaction log data, response accuracy, homework completion, LMS/learning tool activity time, natural language input), rich demographic and user data (e.g., gender, ethnicity, age, socio-economic status proxies, part-time/full-time status, transfer or other admit type), as well as multiple types of educational outcomes (e.g., learning, achievement, persistence, progress in postsecondary education, literacy, retention, performance).

Our Context: ASU Online and Beyond

The current focus of the L@S project is ASU Online with the long-term goal of expanding to leverage and serve all ASU courses and students, including traditional in-person courses, online and blended courses, Earned Admission, and ASU Prep Online (K-12), which collectively serve over 100,000 undergraduate students and 40,000 K-12 students. Combined, these programs at ASU provide access to the large-scale data needed to improve the generalizability of learning sciences research. In addition to the benefit of scale, the diversity of the ASU student body has steadily increased over the last 20 years to represent a broad range of backgrounds and cultures. For example, over 50% of first-year students registering in 2021 identified as students of color and 33% of undergraduate students were eligible for a Pell Grant. Indeed, since 2002, enrollments by first-year students with household incomes of less than $60,000 per year have quadrupled. Similarly, enrollments by first-generation students have nearly quadrupled in the last 18 years. Students who possess these characteristics are often overlooked in traditional methods of data collection such that results obtained in traditional lab based or RCT research designs are not generalizable to these populations. In addition to a large, diverse student body, ASU currently offers over 300 online degrees and over 1,800 online undergraduate courses alone. In light of the increased importance of virtual education worldwide, these numbers continue to grow.

Our Challenges: Connecting People and Data at ASU

ASU has an unparalleled ecosystem of online student (e.g., ASU Online, Open Scale, ASU Prep K-12) and learning data (e.g., learning management system, student information system, Integrated Learning Tools) combined with a uniquely diverse venue to conduct a plethora of experimental studies, including rapid A/B studies. However, we face multiple challenges. A significant challenge in addressing the lack of large-scale, diverse data is locating and connecting the sources. Foremost, student and course data are currently housed in multiple locations, under the auspices of various, sometimes hard-to-find data sentinels, with ambiguous, unstandardized means of accessing data or efficient means to conduct experimental studies. Prior to the inception of the ASU L@S project, conducting a study on student learning at ASU (and most institutions) might take multiple years to obtain the permissions, largely depending on 'who you know'. Most often, student data are only available to internal administrative and research staff for isolated purposes such as institutional improvement or specialized longitudinal analysis (e.g., enrollment tracking). Further, gathering these data requires implementing numerous data calls, connecting each data type, and thoroughly deidentifying data to remove all personally identifiable information (PII).

Protecting student data is paramount. Thus, the critical need to protect student and institutional privacy generally limits the ability for educational institutions to make data accessible to researchers outside the offices of administrative staff. While these administrators, dedicated to enhancing educational experiences and improving learning outcomes, gain valuable insights, they are often restricted by privacy guidelines, time, or priorities in their ability to publish their empirical findings. As such, like many digital learning platforms in private industry, many potential insights into learning at academic institutions remain behind closed doors.

The ASU L@S platform will serve as a mediator between education stakeholders (e.g., administrators, instructors, faculty, students), researchers, and ASU research staff to facilitate and enhance research capacity. It is envisioned to render these diverse, yet restricted, data and educational venues available to both internal and external researchers, and to facilitate researchers' ability to more efficiently conduct research on ASU student learning while maintaining the safety of all stakeholders' PII. To that end, the ASU L@S team is developing the capacity to connect and access multiple types of user data (e.g., interaction log data, response accuracy, homework completion, LMS/learning tool activity time, natural language input), from multiple data sources (e.g., LMS, SIS, LTI integrated tools), as well as multiple types of educational outcomes (e.g., learning, achievement, persistence, progress in postsecondary education, literacy, retention, performance). For example, the ASU L@S platform team is currently working toward integration of multiple data systems, including Peoplesoft student information system data which stores students' academic history and demographics, Salesforce client relationship data for interactions between students and support staff, and Canvas learning management system data on learner actions/interactions. Integration of these data sources, contained in one data warehouse, affords internal and external researchers the ability to investigate how numerous independent data inputs may interact and affect student outcomes. The L@S infrastructure will contribute to work geared toward enhancing students' learning experiences (e.g., dashboards) as well as system components needed to conduct collaborative research (e.g., datasets, data-sharing tools). In turn, the platform will support multiple types of research including rapid-cycle A/B experiments, efficacy studies, replication studies, and design studies with the overarching objective of advancing theories of how people learn.

ASU L@S Data Implementations

Building out our capacity to (more facilely) conduct experimental studies, particularly rapid cycle A/B studies necessitates first devising the means for researchers to access student and course data within a single platform. During our first year of funding within the IES Digital Learning Network, the ASU L@S platform team has identified three datasets, each with different structural qualities, that collectively provide the means to construct a more complete picture of students' learning experiences at ASU. The following describes the three primary datasets currently being compiled within the ASU L@S data warehouse.

Student Profile. Basic information about students' academic history such as students' GPA or Advanced Placement (AP) courses are important considerations in predicting academic success [10, 11]. It is also well known that gender, race, and socioeconomic background play major roles in success [12]. Demographic information about the students available in the Student Information System (SIS) includes variables such as: age, gender, ethnicity, first generation status, Pell eligibility, admit type (e.g., transfer, first-time/first-year), and academic load. Prior performance measures include high school GPA, high school rank, ACT/SAT scores, and Advanced Placement (AP) courses. Current performance measures include overall GPA, number of courses enrolled per semester, progress, and so on. The ASU L@S Platform will be developed such that data provided to researchers are deidentified, safe, secure, and do not violate FERPA regulations. Notably, one important consideration, and a challenge, will be to limit information released about a student if the combination of certain variables allows identification of particular students. Alternatively, we will limit direct access to the dataset, such that analyses are mediated within data enclaves while still allowing for intersectionality of data.

Student Trajectory. A student's academic trajectory (i.e., course history and outcomes) includes each course in which the student enrolls as well as their performance and participation within those courses. Course completion and grade-based metrics are standard measures of student performance as they demonstrate students' ability to progress toward degree completion. More nuanced measures of achievement may also be extracted from the LTI tools used in the LMS. In-course persistence, activity, participation, and performance data are collected via the LMS and stored as raw and processed activity log files and gradebook tables (e.g., weekly summaries of assignments submitted and points earned). These measures are valuable as mediators between course-level input variables and success outcomes such as course grade, retention, and degree completion, and are valuable proxies of learning [13-16]. Multiple metrics for evaluating longer-term persistence and retention (term-to-term, year-to-year) are computed, stored, and tracked to understand how students move from term-to-term, which is especially important for part-time and non-traditional students who do not plan to attend each term (i.e., planned stopouts), but for whom the probability of dropout (non-retention) increases with the duration of stopout [17, 18].

Course Profile. Each student is embedded within their own trajectory as well as the history and characteristics of each course. Courses differ on various dimensions such as number of students enrolled, passing rates, failure rates, number of assignments, average grades on assignments, use of discussion board, and the extent to which learning tools are incorporated within the course (e.g., InScribe, Gradarius, iClicker). Information about each course can be combined with student characteristics and a student's course trajectory in order to address research questions related to attainment gaps and statistical relationships between course characteristics and student outcomes.

Natural Language Processing (NLP)

One unique contribution of the ASU L@S project is a focus on the integration of NLP tools within the ASU L@S platform. The use of NLP is envisioned in two use cases currently under development: discussion board posts and written assignments. Discussion board posts can provide insight into students' learning processes, behaviors, comprehension and individual differences via analysis of linguistic and semantic indices extracted from the text and forum discourse (i.e., original posts and replies). For example, NLP analyses of discussion board posts [13, 14, 19, 20] have been used as predictors of course completion in Massive Open Online Courses (MOOCS) [14, 21-24] and online courses [25]. Beyond the ability to extrapolate students' course success, NLP analysis of discussion board text can shed light on students' individual differences [26] and reading skill [27-30].

The second NLP use case in the ASU L@S project is the extraction of text from written assignments uploaded in Canvas LMS. Prior NLP analyses of written assignments (i.e., essays) have been used to predict students' reading skill as measured by the Gates-MacGinitie Reading Test [31] and evaluate students' comprehension of complex science texts [26, 32, 33]. Likewise, cognitive measures such as performance on working memory tasks can also be predicted using linguistic features of essays [33]. The application of NLP in these use cases provides stealth assessment of natural language that can predict a wide range of individual differences and outcome measures.

The long-term vision of the integration of NLP within the ASU L@S Platform is to provide a model for other platforms and universities to connect their systems to NLP. Discussion Board and Written Assignment data serve as proofs of concept for NLP analyses, setting the stage for future applications that might be performed in the future. To that end, a team will analyze these data to evaluate the degree to which the discourse from these sources can be released to researchers for analysis while also maintaining the safety of PII. In anticipation of situations wherein deidentification (e.g., deletion of names) is not deemed sufficient in ensuring PII, we are also exploring an alternative approach of providing researchers solely with the linguistic and semantic *features* of language extracted from assignments or discussion boards. For example, a researcher might be provided indices on the familiarity, diversity, or breadth of words used within essays written by STEM majors or indices specifying the amount of overlap between source documents and research reports. There is a wealth of information that can be gleaned from linguistic and semantic features of students' discourse, without necessarily revealing all of the actual words used by students [34].

Conclusion

The ASU L@S project is developing foundational infrastructure and protocols to *connect* a wide range of data available in the ASU ecosystem on student achievement, learning, and persistence. Further, the project is developing the necessary mechanisms to make those data *accessible* to researchers across and beyond ASU in ways that honor institutional and individual privacy, so that resulting datasets can be *examined* through exploratory and experimental methods. The overarching objective of the ASU L@S platform is to promote and support research and technological innovations by faculty, research scientists, and industrial partners to address questions regarding the impact of various factors on learning.

The potential studies range from analyses that leverage existing data based on student and course information within the SIS and LMS to experimental studies that incorporate surveys and measures or varying contextual factors [9]. While LMS data include information such as activity (clicks, minutes, logins) and performance (e.g., assignments, exams), richer measures derived from interactivity tools will potentially include social network linkages, natural language, and behavior patterns in problem-solving activities. Recent increases in enrollment in online degree programs, as well as increases in hybrid learning environments render these analyses particularly useful in supporting students' persistence and improving learning outcomes through better understanding of students' behaviors but also as a mechanism to provide just-in-time support to struggling learners.

There are innumerable questions within the domain of education that remain to be answered and replicated at scale. ASU L@S will provide a venue and source of data for thousands of students. The value of this research will follow from the use of data-intensive science to consider the multidimensional impact of various factors across multiple contexts and individuals - affording inferences leading to more targeted interventions for course instructors, administrators/support staff, and individuals nested within classrooms. Achieving this goal will require not only access to data, but also the means to coordinate and collaborate with course instructors, staff, and university administrators.

Such studies can empower scientists to consult with instructors and university administrators to identify the interventions and learning tools that offer the strongest combinations of probable impact, feasibility, and advancement of learning science. We envision a multitude of studies and experiments that might be conducted within the ASU L@S Platform, in collaboration with ASU researchers, Digital Learning Network participants, and other IES funded researchers. The current project lays the groundwork necessary to realize our overarching objective to better understand the impact of individual and contextual factors on educational outcomes, and in turn, improve outcomes for all students.

ACKNOWLEDGMENTS

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305N210041 to Arizona State University. The opinions expressed are those of the authors and do not represent views of the institute or the U.S. Department of Education.

REFERENCES

[1] Hugues Lortie-Forgues, Hugues, and Matthew Inglis. 2019. Rigorous large-scale educational RCTs are often uninformative: Should we be concerned?. Educational Researcher 48, 3 (March, 2019), 158-166. https://doi.org/10.3102/0013189X19832850

[2] Kelli A. Bird, Benjamin L. Castleman, Jeffrey T. Denning, Joshua Goodman, Cait Lamberton, and Kelly Ochs Rosinger. 2019. Nudging at Scale:

Experimental Evidence from FAFSA Completion Campaigns. NBER Working Paper No. 26158. National Bureau of Economic Research.

[3] Christopher R. Dobronyi, Philip Oreopoulos, and Uros Petronijevic. 2019. Goal setting, academic reminders, and college success: A large-scale field experiment. Journal of Research on Educational Effectiveness, 12, 1 (April, 2019), 38-66. https://doi.org/10.1080/19345747.2018.1517849

[4] Paul Hanselman, Christopher S. Rozek, Jeffrey Grigg, and Geoffrey D. Borman. 2017. New evidence on self-affirmation effects and theorized sources of heterogeneity from large-scale replications. Journal of Educational Psychology. 109, 3 (April, 2017), 405-424. https://doi.org/10.1037/edu0000141

[5] René F. Kizilcec, Justin Reich, Michael Yeomans, Christoph Dann, Emma Brunskill, Glenn Lopez, Selen Turkay, Joseph Jay Williams, and Dustin Tingley. 2020. Scaling up behavioral science interventions in online education. Proceedings of the National Academy of Sciences, vol. 117. 14900-14905. (June, 2020). https://doi.org/10.1073/pnas.1921417117

[6] Matthew C. Makel, and Jonathan A. Plucker, J. A. 2014. Facts are more important than novelty: Replication in the education sciences. Educational Researcher. 43, 6 (August, 2014), 304-316. https://doi.org/10.3102/0013189X14545513

[7] Philip Oreopoulos and Uros Petronijevic. 2019. The remarkable unresponsiveness of college students to nudging and what we can learn from it (No. w26059). (July, 2019) National Bureau of Economic Research. https://doi.org/10.3386/w26059

[8] Yeager, David S., Paul Hanselman, Gregory M. Walton, Jared S. Murray, Robert Crosnoe, Chandra Muller, Elizabeth Tipton et al. 2019. A national experiment reveals where a growth mindset improves achievement. Nature, 573, 7774 (August, 2019), 364-369. https://doi.org/10.1038/s41586-019-1466-y

[9] Danielle S. McNamara, Tracy Arner, Reese Butterfuss, Debshila Basu Mallick, Andrew S. Lan, Rod D. Roscoe, Henry L. Roedinger III, and Richard G. Baraniuk. 2022. Situating AI (and big data) in the learning sciences: Moving toward large-scale learning sciences. In A. Alavi, & B. McLaren (Eds.), Artificial intelligence in STEM education: The paradigmatic shifts in research, education, and technology. CRC Press.

[10] Xin Ma and Willis Johnson. 2008. Mathematics as the critical filter: Curricular effects on gendered career choices. In H. M. G. Watt & J. S. Eccles (Eds.), Gender and occupational outcomes: Longitudinal assessments of individual, social, and cultural influences. American Psychological Association. 55–83. https://doi.org/10.1037/11706-002

[11] National Academies of Sciences, Engineering, and Medicine. 2017. The economic and fiscal consequences of immigration. National Academies Press.

[12] Xueli Wang. 2013. Modeling entrance into STEM fields of study among students beginning at community colleges and four-year institutions. Research in Higher Education, 54, 6 (February, 2013), 664-692. https://doi.org/10.1007/s11162-013-9291-x

[13] Scott Crossley, Tiffany Barnes, Collin Lynch, and Danielle S. McNamara. 2017. Linking language to math success in a blended course. In X. Hu, T. Barnes, A. Hershkovitz, & L. Paquette (Eds.), Proceedings of the 10th International Conference on Educational Data Mining. International Educational Data Mining Society. (June, 2017) 180–185.

[14] Scott Crossley, Mihai Dascalu, Danielle S. McNamara, Ryan Baker, and Stefan Trausan-Matu. 2017. Predicting success in massive open online courses (MOOCs) using cohesion network analysis. In Proceedings of the 12th International Conference on Computer-Supported Collaborative Learning (July, 2017). International Society of the Learning Sciences. 103-110. https:dx.doi.org/10.22318/cscl2017.17

[15] Josh Gardner and Christopher Brooks. 2018. Student success prediction in MOOCs. User Modeling and User-Adapted Interaction, 28, 2 (May, 2018). 127-203. https://doi.org/10.1007/s11257-018-9203-z

[16] Wanli Xing and Dongping Du. 2019. Dropout prediction in MOOCs: Using deep learning for personalized intervention. Journal of Educational Computing Research, 57, 3 (March, 2019). 547-570. https://doi.org/10.1177/0735633118757015

[17] Stephen R. Porter. 2003. Understanding retention outcomes: Using multiple data sources to distinguish between dropouts, stopouts, and transfer-outs. Journal of College Student Retention: Research, Theory & Practice, 5, 1 (May, 2003). 53-70.
https://doi.org/10.2190/NV6H-55NG-8EYW-EKGP

[18] Sherry Woosley, Katie Slabaugh, Aimee E. Sadler, and Gary W. Mason. 2005. The mystery of stop-outs: Do commitment and intentions predict reenrollment? NASPA Journal, 42, 2 (March, 2005). 188-201. https://doi.org/10.2202/1949-6605.1472

[19] Scott A. Crossley, Kristopher Kyle, Jodi Davenport, and Danielle S. McNamara, D. S. 2016. Automatic assessment of constructed response data in a chemistry tutor. Proceedings of the 9th International Educational Data Mining Society Conference (EDM 2016). International Educational Data Mining Society. 336-340.

[20] Maria-Dorinela Sirbu, Mihai Dascalu, Scott Crossley, Danielle McNamara, and Stefan Trausan-Matu. 2019. Longitudinal analysis of participation in online courses powered by cohesion network analysis. In K. Lund, G. Niccolai, E. Lavoué, C. Hmelo-Silver, G. Gweon, & M. Baker (Eds.), A Wide Lens: Combining Embodied, Enactive, Extended, and Embedded Learning in Collaborative Settings. Proceedings of the 13th International Conference on Computer Supported Collaborative Learning (June, 2019). International Society of the Learning Sciences. 640–643. https://doi.dx.org/10.22318/cscl2019.640

[21] Juan Miguel L. Andres, Ryan S. Baker, Dragan Gašević, George Siemens, Scott A. Crossley, and Srećko Joksimović. 2018. Studying MOOC completion at scale using the MOOC replication framework. Proceedings of the 8th International Conference on Learning Analytics and Knowledge (March, 2018). 71-78. https://doi.org/10.1145/3170358.3170369

[22] Ryan S. Baker, Yuan Wang, Luc Paquette, Vincent Aleven, Octav Popescu, Jonathan Sewall, Carolyn Rosé et al. 2016. Educational data mining: a MOOC experience. Data Mining and Learning Analytics: Applications in Educational Research. (September, 2016). 55-66. https://doi.org/10.1002/9781118998205.ch4

[23] Scott Crossley, Danielle S. McNamara, Ryan Baker, Yuan Wang, Luc Paquette, Tiffany Barnes, and Yoav Bergner. 2015. Language to completion: Success in an educational data mining massive open online class. Proceedings of the 8th International Conference on Educational Data Mining. (June, 2015). 388–391.

[24] Crossley, Scott, Luc Paquette, Mihai Dascalu, Danielle S. McNamara, and Ryan S. Baker. 2016. Combining click-stream data with NLP tools to better understand MOOC completion. In D. Gašević & G. Lynch (Eds.), Proceedings of the 6th International Conference on Learning Analytics and Knowledge (April, 2016). ACM. 6–14. https://doi.org/10.1145/2883851.2883931

[25] Dascalu, Maria-Dorinela, Stefan Ruseti, Mihai Dascalu, Danielle S. McNamara, Mihai Carabas, Traian Rebedea, and Stefan Trausan-Matu. 2021. Before and during COVID-19: A cohesion network analysis of students' online participation in moodle courses. Computers in Human Behavior. (March, 2021). 121, 106780. https://doi.org/10.1016/j.chb.2021.106780

[26] Laura K. Allen and Danielle S. McNamara. 2015. You Are Your Words: Modeling Students' Vocabulary Knowledge with Natural Language Processing Tools. (June, 2015). International Educational Data Mining Society.

[27] Laura K. Allen, Erica L. Snow, and Danielle S. McNamara. 2015. Are you reading my mind? Modeling students' reading comprehension skills with Natural Language Processing techniques. In Proceedings of the fifth international conference on learning analytics and knowledge. (March, 2015). 246-254. https://doi.org/10.1145/2723576.2723617

[28] Laura K. Allen, Matthew E. Jacovina, and Danielle S. McNamara. 2016. Cohesive Features of Deep Text Comprehension Processes. In Proceedings of the 38th Annual Meeting of the Cognitive Science Society. (CSS 2016). Cognitive Science Society. 2681-86.

[29] Kathryn S. McCarthy, Laura K. Allen, and Scott R. Hinze. 2020. Predicting Reading Comprehension from Constructed Responses: Explanatory Retrievals as Stealth Assessment. In International Conference on Artificial Intelligence in Education. Springer, Cham. (June, 2020). 197-202. https://doi.org/10.1007/978-3-030-52240-7_36

[30] Laura K. Varner(Allen), G. Tanner Jackson, Erica L. Snow, and Danielle S. McNamara. 2013. Does size matter? Investigating user input at a larger bandwidth. In The Twenty-Sixth International FLAIRS Conference. (May, 2013). 546-549.

[31] Walter H. MacGinitie and Ruth K. MacGinitie. 2006. Gates-MacGinitie reading tests (4th ed.). Iowa City: Houghton Mifflin.

[32] Laura K. Varner (Allen), Erica L. Snow, and Danielle S. McNamara. 2014. The long and winding road: Investigating the differential writing patterns of high and low skilled writers. In J. Stamper, Z. Pardos, M. Mavrikis, & B. M. McLaren (Eds.), In Proceedings of the 7th International Conference on Educational Data Mining. (January, 2014) International Educational Data Mining Society. 304-307.

[33] Laura K. Allen, Cecile Perret, and Danielle S. McNamara. 2016. Linguistic Signatures of Cognitive Processes during Writing. In Proceedings of the 38th Annual Meeting of the Cognitive Science Society in Philadelphia, PA. Cognitive Science Society. (2016). 2483-2488.

[34] Danielle S. McNamara, Laura K. Allen, Scott A. Crossley, Mihai Dascalu and Cecile Perret. 2017. Natural language processing and learning analytics. In G. Siemens & C. Lang (Eds.), Handbook of Learning Analytics and Educational Data Mining. Society for Learning Analytics Research. (2017). 93-104.

Optional ERIC Coversheet — Only for Use with U.S. Department of Education Grantee Submissions

This coversheet should be completed by grantees and added to the PDF of your submission if the information required in this form **is not included on the PDF to be submitted**.

INSTRUCTIONS

- Before beginning submission process, download this PDF coversheet if you will need to provide information not on the PDF.
- Fill in all fields—information in this form **must match** the information on the submitted PDF and add missing information.
- Attach completed coversheet to the PDF you will upload to ERIC [use Adobe Acrobat or other program to combine PDF files]—do not upload the coversheet as a separate document.
- Begin completing submission form at https://eric.ed.gov/submit/ and upload the full-text PDF with attached coversheet when indicated. Your full-text PDF will display in ERIC after the 12-month embargo period.

GRANTEE SUBMISSION REQUIRED FIELDS

Title of article, paper, or other content

The ASU Learning at Scale (ASU L@S) Digital Learning Network Platform

All author name(s) and affiliations on PDF. If more than 6 names, ERIC will complete the list from the submitted PDF.

Last Name, First Name	Academic/Organizational Affiliation	ORCID ID
McNamara, Danielle	Arizona State University	
Arner, Tracy	Arizona State University	
Reilley, Elizabeth	Arizona State University	
Alvarado, Paul	Arizona State University	
Clark, Chani	Arizona State University	
Fikes, Thomas	Arizona State University	

Publication/Completion Date—(if *In Press,* enter year accepted or completed) 2022

Check type of content being submitted and complete one of the following in the box below:

- ☐ If article: Name of journal, volume, and issue number if available
- ☑ If paper: Name of conference, date of conference, and place of conference
- ☐ If book chapter: Title of book, page range, publisher name and location
- ☐ If book: Publisher name and location
- ☐ If dissertation: Name of institution, type of degree, and department granting degree

ACM Conference on Learning at Scale, 2022, Roosevelt Island, New York City

DOI or URL to published work (if available) doi.org/10.1201/9781003181187-23

Acknowledgement of Funding— Grantees should check with their grant officer for the preferred wording to acknowledge funding. If the grant officer does not have a preference, grantees can use this suggested wording (adjust wording if multiple grants are to be acknowledged). Fill in Department of Education funding office, grant number, and name of grant recipient institution or organization.

"This work was supported by U.S. Department of Education **[Office name]** Institute of Education Scier through **[Grant number]** R305N21004 to **Institution]** Arizona State U. The opinions expressed are those of the authors and do not represent views of the **[Office name]** Institute of Education Scie or the U.S. Department of Education.