Research/Scholarship/Creative Activity—Candidate's Statement

I study the use of data in education, especially in STEM education contexts. I work in the context of two disciplines, *educational data science*, a research methods-focused discipline using newer data sources and methods to study teaching, learning, and educational systems, and *data science education*, a discipline focused on supporting educators and learners to use digital data sources and technologies. Within these disciplines, my research has focused on 1) documenting the positive and negative aspects of digital technology-based networks and communities, 2) using fine-grained (experience sampling method) methods and multi-level models for understanding learners' motivation and engagement in formal and informal STEM learning environments, and 3) advancing new ways for science learners to analyze scientific data.

My work has resulted in 32 articles since beginning my appointment in 2018, many in journals that have among the highest citation impact in educational research, including *Computers & Education,* with an impact factor (IF) of 11.1, the *British Journal of Educational Technology* (IF: 5.3), and the *Journal of Research in Science Teaching* (IF: 3.9). I have also published a book, *Data Science in Education Using R* (Routledge), with an open-access version that has been accessed more than 100,000 times. Further, I have published four chapters in edited volumes, and two conference proceedings papers in proceedings considered to be as competitive as leading journals, the *International Computing Education Research* conference. Lastly, I was invited to speak at and write one of four commissioned papers for the first event on data science education coordinated by the National Academies of Sciences, Engineering, and Medicine. This work has been highly-cited in my disciplines: according to Google Scholar, the *H*-index associated with my publications is 19 (with my work being cited over 2,500 times in total). Regarding other components of impact, my research has been funded by federal and other agencies (totaling more than $5.4*M* as PI or Co-PI) and recognized with awards, including the Early Career Award from *AERA's* SIG *Technology as an Agent of Change in Teaching and Learning*. Below, I detail the contributions I have made in my three areas of focus.

1. Studying the positive and negative aspects of digital technology-based networks and communities. My interest in this area began with studies of state-based Twitter communities (Rosenberg et al., 2016)—thriving among educators but under-acknowledged by researchers. Building on this initial experience studying an informal community using a large, digital data source (tweets), I set out to understand the nature of the conversations taking place online about a large, nationwide educational improvement effort, the *Next Generation Science Standards*. Having participated in the #NGSSchat community on Twitter as a science teacher, I used social network analysis techniques to document how this social media-based community connected science education stakeholders from diverse roles (i.e., teachers, researchers, administrators, and organizations) in substantive conversations about the reform (Rosenberg et al., 2020b). In other words, this work demonstrated that there were meaningful (for the reform) conversations about science teaching and learning taking place in an unexpected place, a social media-based community. This work on educators' uses of social media also informed a publication in the first special issue of an educational research journal (*AERA Open*) on educational data science. In this publication, my colleagues and I used a data set consisting of more than ten years of posts about the Next Generation Science Standards and multi-level modeling methods to document how public sentiment toward the NGSS has been different—is far more positive—than that toward other standards-based reform initiatives (i.e., the Common Core State Standards; Rosenberg et al., 2021). This study provided one of the first looks into how this nationwide, growing reform may be finding success in terms of how both the public and science teachers perceive its implementation.

In addition to these and other studies that have documented the positive sides of (or patterns of use in) social media and educational technology use (Greenhalgh et al., 2020, 2021; Rutherford et al., 2022), I have also researched how educational institutions use social media (Michela et al., 2021; Kimmons et al., 2020, 2021). In this research, I shifted my emphasis to documenting not only the positive aspects but also identifying the negative features associated with the information shared in educational technology and social media spaces—especially privacy risks (Rosenberg & Staudt Willet, 2021). I have focused on threats to students' privacy posed by the posting of student photos and names by their schools in publicly accessible ways (Burchfield et al., 2021), finding that millions of photos of students have been inadvertently released into the public domain through schools' and school districts' posts on their public Facebook pages. This work was recently covered in an article in Wired on balancing the beneficial aspects of social media against potential privacy risks to students. An extension of this work that documented how many photos of students accompanied by personally identifiable information have been shared will be published in *Educational Researcher* (IF: 6.3; Rosenberg et al., 2021, in press). My future research in this area will continue to document the two sides of technology-based communities, using a critical lens to understand the trade-offs involved in the use of educational technologies.

2. Developing more contextualized ways to study learners' motivation, engagement, and learning. Motivation, engagement, and learning are often conceptualized as dynamic and context-dependent but are often studied in ways that treat them as more or less static (such as end-of-course surveys). In my research, I have used and advanced new ways of collecting data through the Experience Sampling Method (ESM)—a survey technique that involves asking learners about their experiences using brief, frequent questions—and analyzing these data in ways that reflect the contextual variation in students' experiences. Specifically, I used Bayesian methods to carry out many of these analyses. I initially focused on students' engagement in K-12 science classes, building on prior work by using advanced statistical methods to demonstrate that students demonstrate distinct profiles of engaging in high school science classes that vary depending on the nature of the classroom activity (i.e., laboratory-based vs. lecture; Schmidt et al., 2018). This work involved developing open analytic tools, including the *tidyLPA* R package (Rosenberg et al., 2019), for other researchers to conduct similar analyses; this R package has been used in more than 250 publications. I have also explored student interest and engagement in informal STEM learning environments (Beymer et al., 2020; Schmidt et al., 2020). Supported by the National Science Foundation (NSF), I am presently investigating undergraduates' experiences in computer science (Lishinski & Rosenberg, 2021; Lishinski et al., 2022) and biology classes using a text message-based approach to carry out experience sampling method studies.

Another line of research in this area involves combining existing methods or advancing ways of using new methods. For instance, I used a *computational grounded theory* approach to analyze students' written responses (Rosenberg & Krist, 2021). I expanded on this approach to analyze audio and video from mathematics classrooms through a project supported by more than $1M in funding from the NSF. I have also used this work as a basis to constructively criticize the predominant way machine learning has been used in science education, arguing that machine learning can be used in concert with traditional methods (Kubsch et al., 2022). In the future, I plan to use the experience sampling method in concert with the digital data sources from within educational technologies to develop further contextualized measures of students' experiences in STEM learning environments. I have recently submitted a grant to the NSF for 2.5M with collaborators at the University of North Carolina, Chapel Hill, and the University of Delaware to validate the interpretation of the digital data recorded through course learning management systems using the text message-based approach to carry out experience sampling method studies I developed.

3. Supporting teachers and students in using data learning. Finally, I turn the research methods-focused experience I have toward a pedagogical aim, supporting teachers and students in their use of data. One way I have done this is by collaborating with practicing science teachers to design (and iteratively revise) new approaches for students to use data to make sense of scientific phenomena, such as by using newly-developed statistical software tools for learners, especially the Concord Consortium's Common Online Data Analysis Platform (Rosenberg & Lawson, 2019; Rosenberg et al., 2020a). Another way I have worked to support teachers to use data more authentically is by asking science teachers about what they need to engage their students in the kinds of data analyses called for in the Next Generation Science Standards. Through a survey of 300 science teachers interested in using technological tools to support their students' work with data, I found that teachers are eager to use advanced statistical software tools, but that the cost and the time it would take to learn a new tool were key barriers (Rosenberg et al., 2022b). This research highlights the importance of knowing about what teachers already do and value before advancing new approaches to students' analysis of scientific data.

A second way I have supported the use of data for learning is to identify and design data analysis approaches that do not necessarily require sophisticated and challenging-to-learn digital tools, but that still enable sophisticated forms of reasoning about data. My colleagues and I integrated research in developmental psychology, science education, and statistics to argue for the utility of Bayesian data analytic methods, especially in light of the challenges that professionals and learners both face when using conventional (i.e., frequentist) statistical methods. I identified straightforward principles (e.g., update what we know in light of new evidence) and have developed an easy-to-use web-based application to make Bayesian data analysis and interpretation accessible for grades K-12 science learners (Rosenberg et al., 2022a). This work was recently covered in an EdWeek article on the roles of uncertainty in science. Premised on this work and my earlier and ongoing work with teachers to support their students in analyzing data in meaningful ways, as noted earlier, I was invited to write a commissioned paper on what is known about data science learning for the first National Academies of Sciences, Engineering, and Medicine workshop on data science education in grades K-12 (Rosenberg & Jones, 2022). My future research in this area will explore how science teachers support their students in using data in more meaningful and creative ways. For instance, a proposed project for the NSF that is currently under review involves forming a partnership with a science education organization—the Great Smoky Mountains Institute at Tremont—and a company that has created a digital tool for science learners to analyze data—DataClassroom—to explore how a place-based approach can fruitfully be combined with the Bayesian approach of updating what one knows. The place-based element highlights the accessibility and importance of learning opportunities that are tied to where students are, while the Bayesian component provides a framework for learners to express what they know and find within data.

In sum, my work examines data in education from a methodological and pedagogical perspective, drawing insights about important questions from my experiences working with teachers and highlighting opportunities for new data sources and methods from my use of research methods. Looking ahead, my plan is to continue to document the positive and negative elements of digital technology use, often using the educational data science research methods I have used and developed to do so. For instance, with a student mentee, I am using social media data to understand district messaging about race and racism in the wake of the murder of George Floyd (Thomas et al., 2022). I will also continue to make statistical and technical contributions to the study of learners' motivation, engagement, and learning by combining an experience sampling method approach with the analysis of the complex kinds of digital data with which I have developed expertise. Last, I will continue to work with educators and students, especially around science, the subject I taught as a high school teacher before entering educational research. To do this, I will work to bridge the goals of teachers and learners with newer, data science methods that may help them to overcome the limitations of existing approaches. Beyond these specific plans for future research, an over-arching but tacit aim of my research is to recognize and build on the strengths of the individuals—students, teachers, educational leaders, and researchers—to use data in ways that achieve their (varied) goals. My work at the intersection of using and supporting others to use new data sources and methods provides me with a strong foundation to work toward this goal in the years to come.

References

Anderson, D. J., Rowley, B., Stegenga, S., Irvin, P. S., & Rosenberg, J. M. (2020). Evaluating content-related validity evidence using a text-based, machine learning procedure. *Educational Measurement: Issues and Practice, 39*(4), 53-64. https://onlinelibrary.wiley.com/doi/abs/10.1111/emip.12314

Beymer, P. N., Rosenberg, J. M., & Schmidt, J. A. (2020). Does choice matter or is it all about interest? An investigation using an experience sampling approach in high school science classrooms. *Learning and Individual Differences, 78*(2), 1-15. https://doi.org/10.1016/j.lindif.2019.101812

Estrellado, R. A., Freer, E. A., Mostipak, J., Rosenberg, J. M., & Velásquez, I. C. (2020). *Data science in education using R*. London, England: Routledge. Nb. All authors contributed equally. http://www.datascienceineducation.com/

Greenhalgh, S. P., Rosenberg, J. M., Koehler, M. J., Akcaoglu, M., & Staudt Willet, K. B. (2020). Identifying multiple learning spaces within a single teacher-focused Twitter hashtag. *Computers & Education, 148*(4). https://doi.org/10.1016/j.compedu.2020.103809

Greenhalgh, S. P., Rosenberg, J. M., & Russell, A. (2021). The influence of policy and context on teachers' social media use. *British Journal of EducationalTechnology, 52*(5), 2020-2037. https://bera-journals.onlinelibrary.wiley.com/doi/full/10.1111/bjet.13096

Kessler, A. M., Likely, R., & Rosenberg, J. (2022). Open for whom? The need to define open science for science education. *Journal of Research in ScienceTeaching, 58*(10), 1590-1595. https://onlinelibrary.wiley. com/doi/epdf/10.1002/tea.21730

Kubsch, M., Krist, C., & Rosenberg, J. M. (2022). Distributing Epistemic Functions and Tasks - A framework for augmenting human analytic power with machine learning in science education research. *Journal of Research in Science Teaching* https://onlinelibrary.wiley.com/doi/full/10.1002/tea.21803

Lishinski, A., & Rosenberg, J. (2021). All the pieces matter: The relationship of momentary self-efficacy and affective experiences with CS1 achievement and interest in computing. *Proceedings of the 17th ACM Conference on International Computing Education Research*, 252–265. https://doi.org/10.1145/3446871.3469740. *Nb.* In computing education, proceedings papers for this conference are considered to be on par with articles in refereed journals.

Lishinski, A., Narvaiz, S., & Rosenberg, J. M. (2022, August). Self-efficacy, Interest, and Belongingness–URM Students' Momentary Experiences in CS1. In *Proceedings of the 2022 ACM Conference on International Computing Education Research-Volume 1* (pp. 44-60).

Rosenberg, J. M., & Jones, R. S. (2022). *K-12 data science learning.* Paper commissioned by the National Academies of Sciences, Engineering, and Medicine for Foundations of Data Science for Students in Grades K-12: A Workshop. https://www.nationalacademies.org/event/09-13-2022/docs/DD667E469D0EC5DD91A7D85BC839A9852491A3CF9F15

Rosenberg, J. M., & Krist, C. (2021). Combining machine learning and qualitative methods to elaborate students' ideas about the generality of their model-based explanations. *Journal of*

Science Education and Technology, 30(2), 255-267. https://link.springer.com/article/10.1007%2Fs10956-020-09862-4.

Rosenberg, J. M., & Lawson, M. J. (2019). An investigation of students' use of a computational science simulation in an online high school physics class. *Education Sciences, 9*(49), 1-19. https://www.mdpi.com/2227-7102/9/1/49

Rosenberg, J. M., & Staudt Willet, K. B. (2021). Balancing participants' privacy and open science in the context of COVID-19: A response to Ifenthaler & Schumacher (2016). *Educational Technology Research & Development, 69*(1), 347-351

Rosenberg, J. M., Beymer, P. N., Anderson, D. J., Van Lissa, C. J., & Schmidt, J. A. (2019). tidyLPA: An R package to easily carry out Latent Profile Analysis (LPA) using open-source or commercial software. *Journal of Open Source Software,3*(30), 978. https://joss.theoj.org/papers/10.21105/joss.00978.pdf

Rosenberg, J. M., ^Borchers, C., Dyer, E., Anderson, D. J., & Fischer, C. (2021). Advancing new methods for understanding public sentiment about educational reforms: The case of Twitter and the Next Generation Science Standards. *AERA Open, 7,* 1-17. https://journals.sagepub.com/doi/10.1177/23328584211024261.

Rosenberg, J. M., Edwards, A., & Chen, B. (2020a). Getting messy with data: Tools and strategies to help students analyze and interpret complex data sources. *The Science Teacher, 87*(5). https://learningcenter.nsta.org/resource/?id=10.2505/4/tst20_087_05_30

Rosenberg, J. M., Kubsch, M., Wagenmakers, E.-J., & Dogucu, M. (2022a). Making sense of uncertainty in the science classroom: A Bayesian approach. *Science & Education,* 1-24. https://link.springer.com/article/10.1007/s11191-022-00341-3

Rosenberg, J. M., Reid, J., Dyer, E., Koehler, M. J., Fischer, C., & McKenna, T. J. (2020b). Idle chatter or compelling conversation? The potential of the social media-based #NGSSchat network as a support for science education reform efforts. *Journal of Research in Science Teaching, 57*(9), 1322-1355. https://onlinelibrary.wiley.com/doi/10.1002/tea.21660

Rosenberg, J. M., Schultheis, E., Kjelvik, M., Reedy, A., & +Sultana, O. (2022b). Big data, big changes? A survey of K-12 science teachers in the United States on which data sources and tools they use in the classroom. *British Journal of Educational Technology, 53*(5), 1179-1201. https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13245

Rosenberg, J., Borchers, C., Burchfield, M. A., Anderson, D., Stegenga, S. M., & Fischer, C. (in press). Schools share students' personally identifiable information on Facebook. *Educational Researcher.* https://doi.org/10.35542/osf.io/j89pf

Schmidt, J. A., Rosenberg, J. M., & Beymer, P. (2018). A person-in-context approach to student engagement in science: Examining learning activities and choice. *Journal of Research in Science Teaching, 55*(1), 19-43. https://dx.doi.org/10.1002/tea.21409

Schmidt, J. A., Beymer, P. N., Rosenberg, J. M., Naftzger, N. J., & Shumow, L. (2020). Experiences, activities, and personal characteristics as predictors of engagement in STEM-focused summer programs. *Journal of Research in Science Teaching,57*(8), 1281-1309. https://onlinelibrary.wiley.com/doi/full/10.1002/tea.21630