Data Literacy for Social Justice

Camillia Matuk (co-organizer), Anna Amato, Kayla DesPortes, Marian Tes, and Veena Vasudevan
cmatuk@nyu.edu, ada437@nyu.edu, kayla.desportes@nyu.edu, marian.tes@nyu.edu, vv2052@nyu.edu
New York University

Susan Yoon (co-organizer), Jooeun Shim, Amanda Cottone, Kate Miller, Blanca Himes, Ryan Urbanowicz, and Michael Gonzalez
yoonsa@upenn.edu, jshim@gse.upenn.edu, amandaco@upenn.edu, kmmiller@upenn.edu
bhimes@pennmedicine.upenn.edu, ryanurb@pennmedicine.upenn.edu, gonzalezm2@email,chop.edu
University of Pennsylvania

Joseph Polman (discussant) and Stephen Sommer
joseph.polman@colorado.edu, stephen.sommer@colorado.edu
University of Colorado Boulder

Jacob Barton, Nicole Bulalacao, Krista Cortes, Kris Gutiérrez, Kathryn Lanouette, M. Lisette Lopez, Thomas Philip, Edward Rivero, Collette Roberto, and Michelle Hoda Wilkerson
jcbarton@berkeley.edu, nbulalacao@berkeley.edu, kcortes@berkeley.edu, gutierrkd@berkeley.edu, kathryn.lanouette@berkeley.edu, mlisettelopez@gmail.com, tmp@berkeley.edu, edward_rivero@berkeley.edu, roberto_c@berkeley.edu, mwilkers@berkeley.edu
University of California, Berkeley

Francesco Cafaro and Milka Trajkova
fcafaro@iu.edu, milktraj@iupui.edu
Indiana University-Purdue University Indianapolis

Tim Erickson, eepsmedia@gmail.com, eeps media
William Finzer, wfinzer@concord.org, Concord Consortium
Cynthia Graville, cynthia.graville@slu.edu, Saint Louis University
Lina Haldar, lchaldarconsulting@gmail.com,LCHaldar Consulting, LLC
Katie Headrick Taylor, kht126@uw.edu, University of Washington
Beth Herbel-Eisenmann, bhe@msu.edu, Michigan State University

Hollylynne Lee and David Stokes
hstohl@ncsu.edu, djstokes@ncsu.edu
North Carolina State University

Vivian Lim, vivian.liu@guttman.cuny.edu, City University of New York (Guttman College)
Leilah Lyons, llyons@nysci.org, New York Hall of Science
Dan Milz, dmilz@hawaii.edu, University of Hawaii at Manoa
Maria C. Olivares, mariaco@bu.edu, Boston University
Elizabeth Osche, eosche@endicott.edu, PERG
Tapan S. Parikh, tsp53@cornell.edu, Cornell University
Tony Petrosino, anthony.petrosino@gmail.com, University of Texas Austin
Jessica Roberts, jessica.roberts@cc.gatech.edu, Georgia Institute of Technology
Laurie H. Rubel, LRubel@edu.haifa.ac.il, University of Haifa
Andee Rubin, Traci Higgins, andee_rubin@terc.edu, traci_higgins@terc.edu, TERC
Joey Shelley, joeytrshelley@gmail.com, Common Sensing
Megan Silander, MSilander@edc.org, Education Development Center
Ralph Vacca, rvacca2@fordham.edu, Fordham University
Sarah Van Wart, sarah.wart@northwestern.edu, Northwestern University
Peter J. Woods, pwoods2@wisc.edu, University of Wisconsin-Madison
Moira L. Zellner, mzellner@uic.edu, University of Illinois at Chicago

Abstract: The projects in this interactive poster symposium explore ways of engaging learners with social justice issues through the design and study of data literacy interventions. These interventions span classroom to museum contexts, and environmental to social sciences domains. Together, they illustrate research and practice approaches for engaging learners with data to promote emancipatory activity.

Keywords: community-based design, data literacy, K-12 classrooms, museums, social justice, teacher professional development

Overview and significance

The Big Data revolution has brought numerous opportunities to better understand our world and ourselves. At the same time, the biases inherent in predictive algorithms have the potential to leave the most vulnerable populations more vulnerable (O'Neill, 2016). Just as when Postman (1992) cautioned against surrendering human values to the primacy of new technology, this session calls for education to prepare data literate citizens who can use, and not be used, by Big Data. It responds to recent calls to harness the data revolution (NSF, no date), and to provide much needed curriculum, professional development, and research on developing learners' data literacy (Lee & Wilkerson, 2018).

Data literacy involves an understanding of how data are produced, an ability to formulate relevant questions that can be answered with data, to select and use appropriate analyses, and to make data-based inferences and arguments with attention to the role of context in which data were produced (Common Core State Standards Initiative, 2010; Franklin et al., 2007; NGSS Lead States, 2013). These abilities can empower learners to engage with issues of social, environmental, and economic justice, such as by promoting awareness of, and prompting activism to close gaps in resource access. However, promoting data literacy poses a number of challenges, including creating personal relevance, developing cross-disciplinary knowledge and representational competency, and building educator capacity.

This structured poster session features twelve diverse initiatives to improve learners' data literacy through the design of social justice-centered learning environments. The projects broadly examine how engagement with data can develop learners' awareness of, and agency in the face of inequities in their local communities. Contexts range from museums to classrooms to communities; focal domains span finance, art, journalism, and environmental sciences; and data visualizations include analog maps, visual art, and large-screen interactive displays. Across the projects, social justice is addressed either through the focal datasets themselves, or by targeting learners who are typically underrepresented in the data science professions. Together, the projects address the nature of data literacy; how it should be supported and assessed; and what it affords in terms of inquiry into, and action on social justice issues.

Attendees will have opportunities to discuss these projects with presenters in a two-round interactive poster format (six posters per round). Following synthesizing remarks by our discussant, session organizers will lead an audience discussion of opportunities, challenges, and open questions for promoting data literacy for social justice. Attendees will find this session complementary to a related panel discussion that examines critical intersections between data literacy and social justice in the learning sciences (Arastoopour Irgens, Knight & Wise, submitted to ICLS 2020).

Embedding environmental justice data literacy curricula in the science classroom: The need for social, cultural, material, and symbolic resource activation

Susan A. Yoon, Jooeun Shim, Amanda Cottone, Blanca Himes, Ryan Urbanowicz, and Michael Gonzalez

We take an environmental justice perspective (e.g., Morales-Doyle (2017) to develop data literacy knowledge and skills for critical decision making and action in students' local communities. We documented the affordances and constraints of implementing PD and curriculum in high school science classrooms in light of urban public school mandates and teacher understanding. The topic of asthma reduction and disease control (with high rates in urban situated African American populations) was examined through a confluence of air quality data collection with mobile devices, examination of the field of bioinformatics, and problem-based learning. Through PD with teachers and locally adapted curricular experiences, we supported students to collect CO and particulate matter (PM) data through mobile sensors, to apply central tendency analyses, and to make inferences about the quality of their air compared to air quality collected by the EPA in the other parts of the country. Students present their findings through evidence displayed in graphical representations, and make decisions about actions they can take to

improve the air quality in their communities. We discuss affordances in terms of student interest, empowerment, and ownership as well as consequential constraints in terms of the need to activate social, cultural, material, and symbolic resources (Maulucci, 2010) for achieving environmental justice goals within school science curricula.

How positioning youth as data journalists affords and constrains a justice orientation

Joseph L. Polman, Cynthia Graville, and Stephen Sommer

This poster examines the work of high school aged youth positioned as data journalists, with opportunity to author science news infographics for an online newsmagazine. A considerable number of educators who adapted this STEM literacy program, and youth who participated, pursued justice agendas. Using a lens of phronesis, or practical wisdom/wise action (Flyvbjerg, 2001), we examine the interplay of social and scientific values at play in cases where youth pursued a quantitative analysis of social justice issues, either related to public health, social inequity, or environmentalism/stewardship, while researching and authoring news infographics for public audiences. From a naive standpoint, one might expect there to be a tension between "scientific rigor" and crafting "activist messages." But analysis shows that processes which look less rigorous (like looking for data and striking out) can turn out to be highly rigorous (larger understanding of systemic social justice issues, like a dearth of certain mental health or gun ownership data). In addition, motivation to expose injustice (for example in health issues with a particular fast food chain's offerings) often leads to movement from specific issues to broader trends, and to development of greater understanding of scientific and social complexity on the authors' parts.

Challenges and opportunities in designing to support students' art-based social justice inquiry

Anna Amato, Peter J. Woods, Veena Vasudevan, Ralph Vacca, Marian Tes, Megan Silander, Camillia Matuk, and Kayla DesPortes

Visual art offers unique, but underexplored opportunities to engage learners with social issues (e.g., Bhargava et al., 2016). Moreover, integrating art-based perspectives into data science reflects its multidisciplinary nature, and can allow students to approach data in ways that promote their identification *with* data, and *as* data scientists. Yet, teachers need support to create effective disciplinary-integrated curriculum. Building on research on co-design (DiSalvo et al., 2017) and interdisciplinary teacher teams (Vangrieken, et al., 2015), we describe our partnerships with 2 pairs of art and math teachers from different diverse urban middle schools, to engage their students with social justice issues (e.g., mental health stigmas in social media; representation in online music streaming) using art as a means for expression and social change. Qualitative analyses of interviews, observations, and artifacts from our co-creation and implementation highlight opportunities and challenges for professional learning of our cross-domain approach. For example, co-design partners must negotiate logistical hurdles (e.g., making art and math activities that are consistent both internally and across one another given that students may not be enrolled in each teacher's course). They must also negotiate, teach, and learn from one another's disciplinary values in, and perspectives on data sensemaking. These values and perspectives are furthermore reflected in curriculum design (e.g., whether art or math is emphasized, and how one domain serves the other). Findings contribute strategies for co-designing with teachers, and an understanding of the learning affordances of arts-integrated, data science inquiry into social justice issues.

Data clubs for middle-school youth: Data as mirror and window

Andee Rubin, Traci Higgins, and Elizabeth Osche

We introduce middle-school youth to the ubiquity and power of data through experiences analyzing both large, publicly available data sets and related data they collect themselves. Our modules comprise ten hours of experiences, delivered in afterschool clubs or summer camps; we have worked with both urban (of diverse racial backgrounds, with an emphasis on girls) and rural youth (who are generally underserved because of geography). Our chosen data topics (tick-borne disease and teens' use of technology and time) are both mirrors and windows for students: they can see their own experience in the data and are simultaneously introduced to a broader perspective. Having the skills to use data to make arguments about personally relevant issues is a social justice issue. We have developed pedagogical goals for learners, which have guided the development of materials and assessments. In particular, we find it important to pay close attention to measurement issues, e.g.: how were attributes defined?, when, why, how and from whom were data collected? in looking at already-collected data.

We have developed an interview that investigates, among other things, whether participants know which questions they could answer with a dataset and which they could not.

"I see where we at": Sense-making through finding self

Laurie H. Rubel, Vivian Lim, and Beth Herbel-Eisenmann

With *Cash City,* high-schoolers investigated the spatial distribution of their city's financial system, comprised of banks and alternative financial institutions (AFIs). Digital and analog maps displayed locations of institutions; shaded distributions of various demographic variables; and allowed exploration of various data normalizations, some standard (e.g., banks per square mile) and others nonstandard (e.g., banks per AFI). These maps communicate an issue of (in)justice in the spatial patterns of bank locations relative to more expensive AFIs. We recorded on-screen activity of six learners and of several whole-class sessions. Previously, we showed how this engagement—through a set of practices—furthered learners' political formation (Rubel et al., 2017). We termed a baseline practice, "finding self," to mean the spatial positioning of oneself in relation to the data and the given spatial model. Here we ask: What strategies do people use to find themselves on maps? How are various forms of data supportive or inhibitive? Conceptualizing finding-self as a type of perspective-taking (Roberts & Lyons, 2019), we found that students use various expected and unexpected strategies. We saw students recruit data with their senses of place to interpret and at times, challenge, these shared models of space. Ways of "finding-self" attest that ubiquitous traditional birds'-eye views of maps are too limited. We need new forms of data visualizations that better correspond with peoples' experiences in navigating place.

Writing data stories: Reauthoring scientific data through syncretic computational investigations in middle school science

Kathryn Lanouette, Edward Rivero, Jacob Barton, Nicole Bulalacao, M. Lisette Lopez, Krista Cortes, Collette Roberto, Kris Gutiérrez, Michelle Hoda Wilkerson, Hollylynne Lee, David Stokes, William Finzer, Tim Erickson, Tony Petrosino, and Lina Haldar

Writing Data Stories (WDS) addresses a problematic narrative that "objective data" and "subjective experience" are fundamentally distinct. Students author multimodal texts that put their out-of-school knowledge in direct conversation with scientific data. Our goal is for students to use connections and disconnections between everyday practices and patterns found in data about socioscientific issues to imagine more socially just futures. We draw from Gutiérrez and Jurow's (2016) notion of syncretism to adopt a syncretic approach to data literacy, whereby students move across everyday and scientific genres to historicize texts within individual, local, cultural, and social contexts. Students then use data moves (Erickson, Wilkerson, Finzer, & Reichsman, 2019) to re-author scientific datasets to reflect their experiences and allow new paths of investigation. In year one of this three-year project, we are working with two 6th grade science teachers and eight classrooms with a high proportion of students designated as English Learners. This year, we are completing four multi-week WDS units focused on food practices and nutrition in homes, advertisements, and public datasets. We are collecting classroom and focal group video and teacher, student, and curricular artifacts. Ongoing analysis focuses on teacher learning about syncretic approaches to data in science; and student learning through re-authoring data.

Racial data literacies: The need for specificity in the emergent intersections of justice and data

Thomas M. Philip and Maria C. Olivares

For data literacy to address issues of social justice, we emphasize the need to attend to intersections of new forms of data literacy and specific processes of variegated forms of oppression. We adopt a racial data literacy lens to make a case for supporting students in becoming racially literate about data and data literate about race (Philip, Olivares-Pasillas, & Rocha, 2017). We employ interaction analysis to highlight and analyze micro-level contestations of racial ideologies in classroom discussions (Philip et al., 2018; Philip et al., 2017). These methods allow us to attend to how "heterogeneous meaning-making practices" in racial literacy and data literacy "come into contact—explicitly and implicitly, intentionally and emergently—to generate new understandings, extend navigational possibilities, and adapt meaning-making practices to new forms and functions" (Rosebery et al., 2010, p. 324). The introductory high school data science classes we studied over 6 years had stated commitments to social justice, but did not have collective processes to explicitly attend to histories, systems, structures, and interactional processes of race and racism in the curriculum, professional development, and implementation.

Within this context, our analysis demonstrates that the goal of cultivating data literacy for social justice inadvertently operated to reinscribe deficit notions of students, their families, and their communities.

Supporting participatory planning with personalized data visualizations

Leilah Lyons, Moira Zellner, and Dan Milz

Our multidisciplinary team of urban planners, learning scientists, and HCI researchers addressed the following research challenge: how can community stakeholders meaningfully engage in simulation-based participatory planning processes, such that they can (1) propose plan details, (2) make sense of the simulation's projected outcomes for a given plan, (3) revise plans in response to that feedback, and (4) ensure their values are captured by plans. Our sociotechnical planning system supports social justice by giving non-expert community members, who are often left out of the critical "devil-is-in-the-details" phase of plan design, an opportunity to develop plans that will directly impact their lives. We performed controlled lab studies with high school and college students, and in situ with local stakeholders in planning meetings. We collected log data on what simulation data participants consulted and how plans were altered in response; and transcribed audio to capture how participants' understanding of (and interest in) outcome variables shifted via discussion. Here, data literacy involves the situated use of visual analytics, and the ability to recognize when additional information is needed. With this work we developed the "Zone of Proximal Interest" theory, which suggests a mechanism for helping novices to engage with rich, multivariate problem spaces.

Scripts and counterscripts in community-based participatory mapping

Sarah Van Wart, Kathryn Lanouette, and Tapan S. Parikh

We examine the idea of scripts in data science and specifically how they shape data literacy efforts. Focusing on two educational initiatives, an air quality study and a civic planning project, we examine how scripts around data were invoked and navigated as high school students engaged in various data practices supporting local advocacy. We show how three "social valences of data" (Fiore-Neff & Gartland, 2013)—discovery, actionability, and truthiness—constituted the normative data science script that data science can help produce and discover knowledge that can in turn be used to inform principled decision-making and action. However, we show that this script did not necessarily reflect students' experiences and lived realities. Student participants, all of whom belonged to non-dominant social categories, sometimes challenged overly optimistic ideas about the value of data-centric ways of knowing, and the authority that data would presumably give to students' research and ideas. Given these tensions, we argue that when there are mismatches between students' realities and how reality is described to work in data science, making the time and space to examine these contradictions can lead to robust engagement with data science and its applications. Findings have design implications for tools and learning activities that connect data science with the broader contexts of youth's lives.

Off the map: Data for community transformation

Katie Headrick Taylor

I report on a community-based design project (e.g., Bang, Faber, Gurneau, Marin, & Soto, 2016) called *Off the Map* (OTM). I joined a group of public high school students, teachers, librarians, and city officials to build an educational mobile app about the community's forgotten histories. We wanted to support students to create data-driven, place-based stories that emphasized the relationship between schools in the area and larger social issues (war, religion, white supremacy). OTM participants thus collected data from public library archives, oral history interviews, and guided community walks to author historical, digital walking tours of stories not otherwise memorialized in public space. Histories of Black leaders and artists, of civil rights struggles, of Native American communities, and of women activists are but a few that have disappeared from the community-wide narrative over time. My analysis focuses on different uses of data toward community change: *personalizing, historicizing, and enacting*. Data here are not intended to shed light on what is wrong with an individual person in order to fix him (like a low standardized test score), but on collectively imagining and working toward an outcome that potentially benefits a larger whole, perhaps still years away (e.g., Jurow & Shea, 2015).

Something in the air: Contextualizing air quality data for public engagement

Jessica Roberts

Air pollution such as fine particulate matter ($PM_{2.5}$) is tied directly to health and economic impacts. $PM_{2.5}$ is associated with premature mortality (Laden et al., 2006), even at concentrations below current EPA standards (Di et al. 2017). Di et al. find that the health effects of exposure to $PM_{2.5}$ are particularly pronounced among racial minorities and at low incomes. The study of the causes and impacts of air pollution is a heavily data-driven area of scientific research, and increasingly the data from sensor networks is being shared online through open data portals. However, these data portals are typically inaccessible to audiences who lack fluency in decoding complex visualizations and the significant scientific background required to interpret them. While community members may be able to correctly identify overall trends in data, the lack of context provided by these data portals limits the ability to understand what those trends mean. Here I will describe early efforts to design Contextualized Air Quality Timelines (ContAQT), a data interaction platform linking abiotic sensor data such as $PM_{2.5}$ with other more relatable forms of data such as soot buildup on bird feathers, human heart attack rates, and GDP. These alternative datasets demonstrate what sensor readouts mean for lived experiences of plants, animals, and humans in an environment with the goal of empowering citizens to use available air quality data to ask and answer questions, make decisions, and craft and support narratives about air pollution in their local area.

Designing interaction to aid reasoning about correlation and causation

Francesco Cafaro and Milka Trajkova

Our work explores how the design of the interaction with a data visualization impacts the way in which museum visitors discuss causation and correlation. In our experimental setup, people control a data visualization using three interaction modalities: with gesture and body movements, using tangible objects, or with a traditional gamepad. We analyzed visitors' conversations around a 75" screen that displays large sets of data on two interactive globes. The datasets were selected to be provocative and to spark conversations around social issues, including, for example, gun violence, climate change, and unemployment. In this context, the ability to describe causation and correlation is an essential part of data literacy: it relates to people's competence to identify and interpret patterns in data and, in turn, to their attitude towards the social issues that those data exemplify. Causation and correlation can, however, be difficult concepts. Using the lenses of embodied interaction (Dourish, 2004) and embodied schemata (Lakoff & Johnson, 1999), we hypothesized that the use of gestures and body movements will help people who tend to conflate correlation with causation, while tangible interaction may be beneficial for visitors who tend to disregard correlation without seeing it as a possible hint to causation. In this symposium, we will discuss design recommendations that we distilled from a series of user studies at Discovery Place, a science museum in Charlotte, NC.

References

Anderson, J. Q., & Rainie, H. (2014). *Digital life in 2025*. Pew Research Center [Internet & American Life Project].

Bang, M., Faber, L., Gurneau, J., Marin, A., & Soto, C. (2016). Community-based design research: Learning across generations and strategic transformations of institutional relations toward axiological innovations. *Mind, Culture, and Activity*, *23*(1), 28-41.

Bhargava, R., Kadouaki, R., Bhargava, E., Castro, G., & D'Ignazio, C. (2016). Data murals: Using the arts to build data literacy. *The Journal of Community Informatics*, *12*(3).

Common Core State Standards Initiative. (2010). Common Core State Standards for Mathematics. Washington, DC.: National Governors Association Center for Best Practices & Council of Chief State School Officers.

Di, Q., Wang, Y., Zanobetti, A., Wang, Y., Koutrakis, P., Choirat, C., Dominici, F., … Schwartz, J. D. (2017). Air Pollution and Mortality in the Medicare Population. *The New England Journal of Medicine, 376*(26), 2513-2522

Dourish, P. (2004). Where the action is: the foundations of embodied interaction. MIT press.

Erickson, T., Wilkerson, M., Finzer, W., & Reichsman, F. (2019). Data moves. *Technology Innovations in Statistics Education*, *12*(1). https://escholarship.org/uc/item/0mg8m7g6

Flyvbjerg, B. (2001). Making social science matter: Why social inquiry fails and how it can succeed again. New York: Cambridge University Press.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report*. Alexandria: American Statistical Association.

Gutiérrez, K. & Jurow, S. (2016). Social design experiments: Toward equity by design. *Journal of the Learning Sciences, 25*(4), 565-598. https://doi.org/10.1080/10508406.2016.1204548

Laden, F., Schwartz, J., Speizer, F. E., & Dockery, D. W. (2006). Reduction in fine particulate air pollution and mortality: Extended follow-up of the Harvard Six Cities study. *American Journal of Respiratory and Critical Care Medicine, 173*(6), 667–672. https://doi.org/10.1164/rccm.200503-443OC

Lakoff, G., & Johnson, M. (1999). Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought. New York, NY, USA: Basic Books.

Lee, V. R., & Wilkerson, M. (2018). Data use by middle and secondary students in the digital age: A status report and future prospects. Commissioned Paper for the National Academies of Sciences, Engineering, and Medicine, Board on Science Education, Committee on Science Investigations and Engineering Design for Grades 6-12. Washington, D.C.

Morales-Doyle, D. (2017). Justice-centered science pedagogy: A catalyst for academic achievement and social transformation. *Science Education*, *101*, 1034–1060.

Maulucci, M.S. (2010). Resisting the marginalization of science in an urban school: Coactivating social, cultural, material, and strategic resources. *Journal of Research in Science Teaching, 47*,7, 840–860.

NGSS Lead States. (2013). *Next generation science standards: For states, by states*. Washington, DC: The National Academies Press.

NSF - National Science Foundation. (no date). NSF's 10 Big Ideas - Special Report. Retrieved November 14, 2019 from nsf.gov/news/special_reports/big_ideas/harnessing.jsp

O'Neill, C. (2016). Weapons of math destruction: How big data increases inequality and threatens democracy. New York, NY: Crown Publishing Group.

Postman, N. (1992). Technopoly: The surrender of culture to technology. New York, NY: Vintage Books.

Roberts, J. & Lyons, L. (2019). Examining spontaneous perspective taking and fluid self-to-data relationships in informal open-ended data exploration. *Journal of the Learning Sciences*. Advance online publication.

Rubel, L., Hall-Wieckert, M., & Lim, V. (2017). Making space for place: Mapping tools and practices to teach for spatial justice. *Journal of the Learning Sciences, 26*(4), 643-687.

Vangrieken, K., Dochy, F., Raes, E., & Kyndt, E. (2015). Teacher collaboration: A systematic review. *Educational Research Review, 15*, 17–40.