

This paper was commissioned for the *Workshop on Foundations of Data Science for Students in Grades K-12*, whose work was supported by the Valhalla Foundation. Opinions and statements included in the paper are solely those of the individual authors, and are not necessarily adopted, endorsed, or verified as accurate by the *Workshop Planning Committee on Foundations of Data Science for Students in Grades K-12*, the Board on Science Education, the Board on Mathematical Sciences and Analytics, the Computer Science and Telecommunications Board, or the National Academy of Sciences, Engineering, and Medicine.

Critical data literacy: Creating a more just world with data

A paper prepared for the National Academy of Sciences' workshop on Foundations of Data Science for Students in Grades K–12, September 13–14, 2022

Josephine Louie
Education Development Center

We live in a society that has engineered the capacities to collect ever increasing amounts of data about our world and ourselves. These data open possibilities to develop greater understandings of every aspect of our existence, including how our universe formed, the ways in which pandemics may spread, and where we may find a job and an affordable place to live. Like any human artifact, however, data can be used for good or ill. The data that others collect about us, or that we intentionally or unintentionally provide to others, have tremendous power to shape our personal behaviors and the contours of society. Within this context have come urgent calls for all people to develop *critical data literacy*—to understand what we and others can do with data, what data can do to us, and what kind of world we can create with data. This paper examines existing literature to address the following questions: What is critical data literacy, and what social futures could such literacy help forge? What are examples of efforts to promote and examine critical data literacy among young people in formal and informal K–12 education settings? What have been outcomes and challenges of these efforts? To provide fodder for further discussion, the paper then ends with brief reflections and ideas for future directions.

What is critical data literacy, and what social futures could such literacy help forge?

Advances in computing, information, and communications technologies have thrown us into a world of Big Data, where vast amounts of data are now harnessed to inform business and governmental decision-making, and where large, open datasets have become more widely available for public use (e.g., Engel, 2016; Kitchen, 2014; Ridgway, 2015). Although the data deluge is relatively new, discussions about data literacy and the skills that people need to make sense of data have been around for a long time (e.g., Donoho, 2017; Rubin, 2020). Twenty years ago, Gal (2002) proposed that all adults in modern society should have a basic level of statistical literacy, which he defined as the abilities to “interpret and critically evaluate” information and arguments drawn from data, and to “discuss or communicate” reactions and opinions about this information, including “concerns regarding the acceptability of given conclusions” (pp. 2-3). Such abilities would allow people to make sense of data on important social and economic topics that affect their everyday lives, and support people in making informed decisions when participating in public life.

But as the prevalence of Big Data and publicly available open data have increased, people have endeavored to refine understandings of statistical literacy and to expand what it means to be literate with data. Building on a framework articulated by Wild and Pfannkuch (1999), Bargagliotti et al. (2020) emphasize that all students should develop the abilities not only to consume statistics and data with scrutiny, but also to think and work with data following a disciplined process of data inquiry. This process involves four cyclical steps: 1) formulating questions that can be answered with data, 2) collecting or assembling data to address one’s questions, 3) using statistical and other tools to analyze the data, and 4) interpreting results to address the original questions (Bargagliotti et al., 2020). These scholars suggest that

data literacy requires core levels of fluency with each of these steps. Such literacy also involves constant questioning throughout the data inquiry process. People must interrogate who collected the data, how data were collected, what types of data were gathered, how data were measured, and the tools and procedures used to analyze the data to assess the conclusions that one can draw from the data.

The growing complexity of today's data and the computational effort that is required to make sense of it have added to what scholars argue should be a part of data literacy. Along with Bargagliotti et al., Engel (2017) and Ridgway (2015) have suggested that data literacy now requires an awareness of the multivariable nature of large datasets—and indeed, of all natural and social phenomena—and how issues such as omitted variables or variable confounding can affect data interpretations. Drawing on interviews with present-day data scientists as well as other sources (e.g., Donoho, 2017; EDC, 2016), H.S. Lee et al. (2022) propose a six-part data investigation process that draws out the need to process or prepare data for analysis; use data visualization and other statistical tools to explore and analyze data; consider the models that are used to analyze large datasets (as well as their limitations); and communicate clear data stories and implications to varied audiences. Contributing further to these ideas, D'Ignazio and Bhargava (2015) argue that data literacy in the era of Big Data should include the ability to identify when and where data are being collected about *us*; a general understanding of what algorithms are and how they draw conclusions about individuals; and the ability to weigh “the real and potential ethical impacts of data-driven decisions for individuals and for society” (p. 3).

The need for critical data literacy. These last recommendations highlight growing concerns about how the data revolution may affect fundamental aspects of our lives and social relations. Scholars have argued that throughout history, groups in society have held different levels of social, economic, and political power. Without vigilance and countermeasures, those who hold power are likely to use data to maintain existing power hierarchies (e.g., Bhargava, 2015; D'Ignazio & Klein, 2017; Philip et al., 2013). Some have noted that the ability of actors to collect vast amounts of data from us—with or without our acquiescence—leads to loss of personal autonomy (e.g., Pangrazio & Selwyn, 2019; Raffaghelli, 2020). Others have noted that data are never neutral and are collected to serve specific interests (e.g., Bhargava, 2015; Skovsmose, 1994). Complex and biased algorithms, which governments and corporations deploy on large-scale data to categorize and predict people's characteristics and behaviors, can worsen social and economic inequality (e.g., O'Neill, 2016; Vakil & Higgs, 2019) and aggravate racism and sexism (e.g., Buolamwini & Selwyn, 2018; Noble, 2018). Data “illiterates” risk exploitation by those who control today's data and data structures (Tygel & Kirsch, 2016). A dystopian picture that emerges is one where social and technological elites control not just our labor but also our data, degrading individual dignity, collective decision-making, and democracy itself (Bhargava, 2015; Raffaghelli, 2020; Skovsmose, 1994).

To counter these threats, scholars have begun to advocate for greater *critical* data literacy. This type of data literacy goes beyond the constant questioning of data sources, analytical approaches, and conclusions that statistics educators say are essential to the statistical thinking process. Critical data literacy also includes ethical consideration of whether and how to collect and use data, to avoid harming the dignity and privacy of others (e.g., Baumer, 2022; NASEM, 2018). It includes a keen awareness of unequal power structures in society and an explicit attention to how power dynamics affect the purposes and uses for which data are deployed. People with critical data literacy are alert to the personal and social harms that powerful interests can inflict with data. They are also equipped with the knowledge and motivation to act in ways that can achieve more just outcomes for themselves and society. Examples of such actions may include adjusting one's privacy settings in online websites or even deploying online obfuscation tactics to “mitigate, evade or perhaps sabotage dominant structures of data reuse and recirculation” (Pangrazio & Selwin, 2019, p. 431). They may include collecting video data of a racist social encounter through a mobile phone, to author one's own telling of the encounter (Gutierrez et al.,

2019). They may also include conducting analyses of social and economic data to uncover systemic social injustices (Weiland, 2017). Almost 30 years ago, Tate (1995) provided an illustrative case where a class of middle school students gathered and analyzed data about the problems associated with liquor stores located near their school, and then confronted authorities with their data to push for change.

Possible social futures through critical data literacy. This last example highlights how ideas at the heart of critical data literacy are not new. These ideas involve examining data to understand unequal sociopolitical arrangements and to create a more just world. Freire (1970/2012) characterized the world as a place in which powerful groups dominate others for the former's material benefit. He advocated a form of education in which educators work as co-investigators with individuals of oppressed groups to examine their life situations, build a critical awareness and understanding of the power structures that oppress them, and transform the world to a place where all individuals have the "freedom to create and construct, to wonder and venture" (p. 68) and to pursue their own goals without dehumanizing limits.

Scholars such as Frankenstein (1983, 2009), Gutstein (2003, 2006), and Skovsmose (1994) have drawn from the writings of Freire to describe how mathematics educators can advance these ideals. These scholars have discussed efforts to work with learners to examine real-life problems using mathematics, statistics, and data to raise people's critical consciousness of injustices in society and to fight forces of social oppression. For example, in a middle school serving students with low-income and Latinx backgrounds, Gutstein (2003, 2006) led mathematics classes in which students analyzed data to explore racial discrimination in housing prices and police stops. He describes his approach as teaching mathematics for social justice, where students use mathematics and data to develop social and political consciousness of systemic inequalities, social agency to work toward more just social arrangements, and, among non-dominant groups, a sense of pride in their own cultural and social identities.

Building on this tradition, educators and researchers have continued to engage in teaching mathematics and statistics for social justice and transformation (e.g., Bartell, 2011; Berry et al., 2020; Gregson, 2013; Gutstein & Peterson, 2013; Kokka, 2020; Lesser, 2007; Raygoza, 2016; Ridgway, in press). By bringing real-life socioeconomic problems and data into classrooms, these efforts seek to raise students' awareness of social injustices and to equip students to advocate for change. Scholars outside of mathematics education have also articulated visions of what a society empowered with critical data literacy could look like. Bhargava et al. (2015) envision a future "where citizens demand to have a voice in how and by whom data is used, what it is used for, and use data to fulfill their goals in an ethical and equitable manner," to achieve "a more inclusive society" (p. 19). Philip et al. (2013) describe a world where students see themselves "as doers and creators of data science," as "people who can engage with and use data for their own purposes and goals," and who recognize data "as a powerful way to understand and address societal issues," attending to those with "perspectives that are marginalized" (pp. 114–115).

What are examples of efforts to promote critical data literacy in K–12 settings?

Recognizing the need for greater critical data literacy in the population, scholars from varied disciplines have been developing and studying strategies to promote such literacy in both formal and informal K-12 education environments. V. R. Lee et al. (2022) recently completed a systematic review of data science interventions in K-12 education viewed through the principles of data feminism as articulated by D'Ignazio and Klein (2016). Raffaghelli (2020) conducted a recent examination of efforts to promote data literacy for social justice in higher education. There has also been a growing body of innovative research examining how students make sense of data when using new technologies and data platforms (e.g., Gould et al., 2016; Hautea et al., 2017; Lee et al., 2015; Podworny et al., 2022), when drawing on varied types of data (e.g., Higgins et al., 2021; Wilkerson and Laina, 2018), and in interdisciplinary settings (e.g., Matuk et al., 2022; Radinsky et al., 2014). In this brief paper, I do not try to duplicate prior reviews, nor do I

examine studies where the focus has been on exploring tools and contexts for supporting data literacy. Instead, I focus on a set of eight learning interventions that have sought to advance students' critical data literacy in service of ethical or social justice goals. I selected this set after searching the learning sciences and mathematics and statistics education literatures for recent studies (i.e., published primarily within the last several years) where authors studied learning interventions with explicit critical or social justice learning objectives with or through data. I do not claim that these eight examples represent the full body of work that currently exists in this domain. Instead, I selected these examples to illustrate a range of recent approaches, and I examine them with analytic lenses drawn from Gutstein's framing of social justice pedagogy and Bargagliotti et al.'s framing of the data inquiry cycle. My goal is to help connect perspectives and spark conversations across the learning sciences, social justice mathematics/statistics, and other education research communities about where prior efforts to promote critical data literacy have brought us and where we might go next.

Eight examples: Learning contexts and participants. Appendix Table 1 summarizes basic features of eight studies that have aimed to promote critical data literacy among learners in K-12 education.

- Six studies occurred in formal school settings – three in high school mathematics classes (Kokka, 2020; Louie et al., 2021a, 2021b; Rubel et al., 2016), and three in high school science, social studies, or media arts classes (Stornaiuolo, 2020; Taylor et al., 2020; Van Wart et al., 2020).¹ The remaining two studies took place in a summer workshop for middle-grades participants at a public library (Kahn, 2020) and at a high school community center program (Vakil et al., 2020).
- All but one of the interventions were designed by a team of researchers and educators. The one exception was in Kokka's² study, where the teacher designed the intervention. Teachers delivered the interventions in Kokka's, Louie's, and Rubel's studies; intervention developers and researchers helped to facilitate learning activities with local educators in the remaining studies.
- Six studies directed their interventions toward communities with high proportions of youth from historically marginalized or non-dominant groups. Kahn's study focused on two focal pairs of siblings, where one pair identified as African American and the other as White. Kokka's study deliberately focused on students who were primarily from White and affluent backgrounds.

Critical literacy goals: Different emphases and strategies. The interventions emphasized different critical literacy goals involving ethical or social justice learning objectives with data. Set within a frame borrowed from Gutstein (2003, 2006), the four studies described below leaned toward helping students to “**read the world**” with data—that is, to learn from data about social or human conditions, and to gain social and political consciousness through such learning.

- ***Uncover social inequality with data.*** In studies by Kokka, Louie, and Rubel, students actively worked with data from the U.S. Census or other sources to uncover stark inequalities in society. Students in Kokka's study examined housing inequality by comparing housing prices and household incomes in their own town and the U.S. overall. In Louie's study, students investigated income inequality in the U.S. by comparing the wages of males and females (both before and after controlling for a third variable) and of higher- and lower-income earners over time. Students in Rubel's study grappled with the social ramifications of the lottery by comparing median household incomes and lottery revenues by local neighborhood. By awakening students'

¹ The intervention studied by Taylor, et al. (2020) was also conducted in an undergraduate course on learning theory.

² To improve readability, I refer to each study from this point onward using only the name of the study's first author.

consciousness of persistent and quantifiable group disparities, the interventions aimed to trigger questions about the larger social forces at play as well as the fairness or justice of these outcomes.

- ***Confront the ethics of Big Data.*** In Vakil’s study, high school participants viewed a demonstration of the programming and inside workings of an online application that can track and display all Twitter posts issued from any geographic address. By helping learners literally “read” other people’s tweets and linked personal data, and to experience what it is like to be both surveyor and surveilled, this intervention aimed to raise awareness and spark ethical questions about the threats to individual privacy that current data-gathering technologies present.

Borrowing again from Gutstein’s framing, the other studies leaned toward efforts to help students “**write the world**” with data. The four studies outlined below included a strong emphasis on helping learners develop agency with data, to feel that they can tell their own data-based stories.

- ***Envision new public spaces with data.*** Van Wart and Taylor both studied interventions that equipped students with different types of tools (e.g., air-quality sensors, wearable cameras, GPS devices) to collect data about the conditions of local parks and other public spaces. Each intervention was geared to help students recognize that a variety of information (e.g., photos, drawings, audio recordings, expressed hopes and dreams) can serve as valid forms of data. Students compiled data into presentations about how local places can or should be improved, and they shared their findings and recommendations with family and local town authorities. A primary goal was to empower students with data to reimagine future environmental and city planning outcomes, and in the process, build a “third space” in which the ideas of students could help transform the thinking of dominant others and create social change.
- ***Craft one’s family history with data.*** In Kahn’s study, learners in a summer library-based workshop assembled data from their own recollections, family members, and large-scale datasets to craft a story of their family’s geobiography, detailing where their ancestors settled over time and the reasons for their migrations. To create their family’s story, participants explored both the question of “What moves families?” and “What moved my family?” These questions were designed to help participants recognize the variability in data when individual family stories do not match aggregate social patterns. Family members who shared alternate or more nuanced information about past family experiences also provided participants with opportunities to recognize how data can be contested and can shape different stories.
- ***Express one’s self with data.*** In Stornaiuolo’s study, students in a high school media arts class engaged in a project to identify a topic of personal interest, collect and analyze data about the topic, and design a T-shirt to convey their own personal data story. Drawing on the example of *Dear Data* (Lupi & Posavic, 2016), the intervention aimed to help students express themselves with data, and to recognize their own power to define and produce data. By choosing what topics to investigate, what data to collect, and what data to display, students had opportunities to become “authors and architects” of data and to learn that one can control one’s data story by deciding what data to share or highlight.

Although each intervention may have emphasized either “reading” or “writing” the world with data, most of the interventions also made efforts to advance the other critical literacy goal. For example, students in Rubel’s study not only analyzed map data to identify geographic disparities in lottery participation; they also collected interview data from community members to shape explanations about why people from different communities played the lottery. Using data from their own simulations, students also discovered the low probability of winning the lottery and created posters with their findings

to counter official state lottery advertising. In Kahn’s study, participants not only collected data from their own past to create their family story; they also learned about population migration patterns (and the difficult social conditions that often prompted them) to contrast aggregate trends with their own family’s journey. In these studies, the interventions created opportunities for learners to encounter and consider unequal social conditions as revealed through large-scale social and economic databases (“reading the world with data”), and to create their own data-based personal, family, or neighborhood stories that may counter existing narratives (“writing the world with data”).

Data literacy goals: Different highlights and strategies. Just as the eight interventions emphasized different critical literacy goals, they also highlighted different data literacy goals (framed here as aspects of the data inquiry cycle). Data literacy goals often corresponded with or supported critical literacy goals. The settings for each intervention, as well as the types of data and tools that each intervention used, may have also played a role in influencing data literacy goals.

- **Analyze data quantitatively.** In Rubel’s, Kokka’s, and Louie’s studies, the interventions took place in mathematics classrooms. Students in each of these interventions engaged heavily in analyzing data (step 3 of the data inquiry cycle, as articulated by Bargagliotti et al., 2020). Students employed mathematical and statistical problem-solving with data to uncover discrepancies between actual levels of risk in winning the lottery and the messages conveyed in state advertising (Rubel), or to surface quantitative disparities in socioeconomic outcomes by demographic group or geographic location (all three studies). In the process, students worked actively with and deepened their understanding of concepts such as measures of center, variability, probability, and proportions.
- **Define and collect data.** In contrast, the other five studies occurred in learning environments outside of mathematics classrooms. Studies by Taylor, Van Wart, Kahn, Stornaiuolo, and Vakil (as well as Rubel) highlight the learning possibilities that arise with a strong focus on data collection (step 2). Students in these studies collected varied types of data directly or indirectly from the local environment, people in the community, or online sources, and in the process confronted how people define, measure, and construct data. In Vakil’s study, students also had opportunities to peer into the “black box” of computational algorithms (D’Ignazio and Bhargava, 2015), and in the process learned how vast amounts of information can be collected from people and put to uses of which they are not aware.
- **Communicate with data—for different audiences and purposes.** Students in each study were tasked with interpreting data and communicating findings (step 4), but the types of audiences and the goals of communication differed. In Kokka’s and Louie’s studies, students examined data both to uncover and explore possible explanations behind unequal social and economic outcomes. Final classroom discussions or presentations were therefore designed to *describe* and *explain* larger social phenomena with data. In Taylor’s, Van Wart’s, Rubel’s, and Vakil’s studies, students used data to try to *persuade* community members about existing problems in the community or society at large, and to consider adopting alternative social actions or visions. In Kahn’s and Stornaiuolo’s studies, students used data to *share stories* about their family’s past or to *express* aspects of their personal identity. Working in a makerspace to design T-shirts, the students in Stornaiuolo’s study had particularly strong opportunities to explore not only the constructed nature of data, but also its visual and aesthetic possibilities.
- **Examine questions with data.** In line with its expressive orientation, the intervention in Stornaiuolo’s study gave students wide latitude to choose their own personal issue or question to explore with data (step 1). The other studies presented or constrained the primary questions for

student investigation to help students reach critical or data literacy learning goals. The types of choices that students could make in the other studies included selecting which side of the family to explore when constructing a family geobiography (Kahn), which community space to target for investigation and improvement (Taylor), and which demographic variables to add to one's analysis to try to explain the male-female wage gap (Louie).

What have been outcomes and challenges of these efforts?

Each study describes successes in advancing critical data literacy outcomes among students. All studies provide qualitative evidence from learner interviews, intervention observations, or learners' artifacts to support claims about growth in students' critical understandings of the world and of data. Using pre- and post-intervention assessments, Louie and Rubel show quantitative gains in students' understandings of data concepts, while the remaining studies describe students' engagement in the data inquiry cycle to suggest learning of data literacy goals. Successes that authors describe include:

- ***More critical perspectives of society.*** Rubel, Kokka, and Louie share comments from students indicating that by learning to "read" patterns in large-scale socioeconomic data, they became more aware of pervasive disparities in economic outcomes across geographic areas and demographic groups. The data analysis and discussion activities within each intervention opened the eyes of many students to these disparities by revealing recurrent and measurable data patterns that were difficult to dispute. In classroom conversations and final posters, some students in Rubel's study came to voice critical opinions of the lottery. In group discussions and interviews, a majority of the privileged students in Kokka's study shared concerns over the levels of housing and economic inequality they saw in their data and displayed signs of "civic empathy." In interviews, some students in Louie's study said they previously did not believe that income inequality was a significant issue but changed their minds after viewing the data. Through both words and body language, students in Vakil's study demonstrated shock and concern over the power of new technologies to access people's data and to violate their privacy.
- ***Stronger agency in authoring with data.*** The four studies in which interventions focused heavily on student data collection appeared to help students recognize data's constructed nature and to "write" with data to convey their own community, family, or personal stories. When students in Taylor's study discovered through their data collection activities that there were no tours for tourists in their multicultural neighborhood in Queens (unlike many such tours in neighboring Manhattan), they designed their own bus tour to showcase the assets in their neighborhood. When students in Van Wart's study did not find a park inventory exercise meaningful, they were able to help change the focus of the activity to create a vision of how to redesign the park to better meet their community's needs. In Kahn's and Stornaiuolo's studies, the participants who completed their final projects literally authored their own family or personal stories with data that they had gathered, shaped, curated, and assembled.

Along with successes, the eight studies also describe challenges that the interventions faced in promoting critical data literacy. Challenges include:

- ***Critical sociopolitical views may not arise naturally or easily.*** Rubel describes how a student in her study did not develop a critical view of the lottery as intervention designers had intended. Like what Brantlinger (2013) found in prior research, this student resisted the sociopolitical perspective that designers had hoped all students would adopt. Kokka relays how during a research interview, one student in her study drew upon negative stereotypes of low-income groups when making sense of existing housing policies. Like Enyedy and Mukhopadhyay (2007),

Kokka found that exposure to new data may not change existing beliefs about social arrangements and outcomes. Stornaiuolo found that under her intervention in a media makerspace environment, students gravitated toward personal topics to explore with data (e.g., their favorite activities) but did not extend their explorations to larger cultural or social issues. These authors suggest that learning designers and facilitators need to anticipate and prepare for the many types of headwinds that can hinder development of sociopolitical awareness of the world and of data.

- ***Discussing social and political inequality can be fraught.*** Two mathematics teachers in Louie’s study shared concerns that some students from non-dominant groups may have felt despondent and disempowered by data investigations revealing persistent trends in income inequality. These teachers were aware of unionization efforts or cases of collective action that may offer hope for change, but they did not have the data or time to spell out and discuss such examples. And although Louie, Kokka, and Rubel did not report situations in which students voiced views that denigrated students from other demographic groups, Philip et al. (2016) point out the subtle ways in which students may express such views during data discussions of social inequality, and the harm these views can cause. Again, intentional preparation and skilled facilitators may be needed to protect against these harms.

Reflections and future directions

The eight studies leaned toward supporting learners’ abilities to either “read” or “write” the world with data, emphasizing different components of the data inquiry cycle. Thinking of critical data literacy as reading and writing the world with data may be helpful in varied K-12 education settings, where educators may draw parallels with reading and writing about the world with text. As suggested by Tygel and Kirsch (2016) and as illustrated by the different examples, a Freireian approach toward promoting critical data literacy would bring educators and learners together to examine data about issues in learners’ immediate worlds, uncovering challenges in these worlds, and “co-authoring” (p. 119) alternative visions that are supported and communicated with data. This approach could complement the data feminism perspective described by D’Ignazio & Klein (2016), which calls on people to follow principles such as “examine power,” “challenge power,” and “embrace pluralism” when working with data. In addition, in a world where control over large data infrastructures now provides access to systemic power, efforts to raise critical consciousness of social and economic inequality and exclusion need to turn a lens toward the oppressive dangers associated with large data infrastructures (as Vakil’s study illustrates). Some may argue that there has always been a need for a critical literacy education that builds social awareness and agency to create a more just world. The emergence of Big Data and the associated threats it poses to everyone provide an opportunity to clarify and commit to the idea of critical data literacy for all.

Below I share a few additional reflections on the eight interventions’ approaches, findings, and implications for future work.

- ***Promoting critical data literacy is an interdisciplinary task.*** As described at the beginning of this paper, critical data literacy encompasses ways of thinking, skills, and understandings that cross traditional disciplines. The interdisciplinary nature of critical data literacy is reflected in the interdisciplinary teams that designed all but one of the interventions examined in this paper. To foster critical data literacy among K–12 learners, it seems important to create spaces where learning designers and educators from disciplines such as mathematics, statistics, computer science, social studies/history, city planning, arts and communication can collaborate, and where educators can deepen their understandings of critical data literacy and how to support it.

- ***Finding ways to scale promising approaches is a needed next step.*** All but two of the eight interventions involved large teams of learning designers and small numbers of students or participants. If we believe that it is important to promote critical data literacy widely within K–12 education, then we need to identify the key components of promising strategies from small-scale exploratory studies and begin to test these components in larger-scale efforts. Louie’s study has taken a step in this direction, and Taylor’s study has begun to adapt its curriculum to multiple sites, but these efforts are still relatively small in scope and more work is needed to design interventions for scalability.
- ***We need frameworks and tools for assessing critical data literacy.*** Louie and Rubel assembled pre- and post-intervention instruments to measure students’ learning of specific data concepts, but they did not measure students’ growth in critical data literacy because no established measure of this construct exists. The eight studies discussed in this paper drew upon qualitative sources of data to describe changes in students’ critical perspectives toward society and data. If we want to identify interventions with the potential to improve students’ critical data literacy at a larger scale, then we need to develop frameworks and tools to measure this construct.
- ***Can students demonstrate critical data literacy without quantitative reasoning?*** The studies by Taylor, Van Wart, Kahn, and Stornaiuolo describe interventions that were conducted outside of mathematics classrooms, in settings where students may have learned deep lessons about the constructed nature of data and how they can be the authors of their own data stories. It is unclear, however, what skills students learned or brought to bear to analyze data quantitatively. More discussion is needed about the types of quantitative analysis practices that individuals should demonstrate to be considered critically data literate. If it is not possible to help learners develop these practices within a single intervention or disciplinary setting, then we need frameworks and coordination across time and learning contexts to support these practices.
- ***We need to learn more about promoting critical data literacy among different groups.*** Seven of the eight studies focused on interventions that served learners primarily from non-dominant populations. If critical data literacy is a goal for all learners, then we need more research on whether and how strategies developed and tested with learners from certain groups (e.g., Black or Latinx) may work with those from other groups (e.g., White). Effective strategies for different groups will require sensitivity to their historical positions of power in society, and strategies to face the different headwinds that history and context may stir up.

References

- Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L., & Spangler, D. A. (2020). *Pre-K-12 guidelines for assessment and instruction in statistics education II (GAISE II)* (2nd ed.). American Statistical Association.
- Bartell, T. G. (2011). Learning to teach mathematics for social justice: Negotiating social justice and mathematical goals. *Journal for Research in Mathematics Education*, 44(1), 129-163.
- Baumer, B. S., Garcia, R. L., Kim, A. Y., Kinnaird, K. M., & Ott, M. Q. (2022). Integrating data science ethics into an undergraduate major: A case study. *Journal of Statistics and Data Science Education*, 30(1), 15-28.
- Berry, R. Q., Conway, B. M., Lawler, B. R., & Staley, J. W. (2020). *High school mathematics lessons to explore, understand, and respond to social injustice*. Corwin Press.
- Bhargava, R., Deahl, E., Letouze, E., Noonan, A., Sangokoya, D., & Shoup, N. (2015). *Beyond data literacy: Reinventing community engagement and empowerment in the age of data*. Data-Pop Alliance (Harvard Humanitarian Initiative, MIT Media Lab and Overseas Development Institute) and Internews.
- Brantlinger, A. (2013). Between politics and equations: Teaching critical mathematics in a remedial secondary classroom. *American Educational Research Journal*, 50(5), 1050-1080.
- Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of the 1st Conference on Fairness, Accountability, and Transparency*, PMLR, 81: 77-91.
- D'Ignazio, C., & Bhargava, R. (2015). Approaches to building big data literacy. *Bloomberg Data for Good Exchange Conference*. New York City.
- D'Ignazio, C., & Klein, L. F. (2020). *Data feminism*. MIT Press.
- Donoho, D. (2017). 50 years of data science. *Journal of Computational and Graphical Statistics*, 26(4), 745-766.
- Education Development Center [EDC]. (2016). *Profile of the data practitioner*. Oceans of Data Institute.
- Engel, J. (2016). Open data at the interface of mathematics and civics education: Challenges of the data revolution for the statistics curriculum. *Journal of Mathematics and Statistical Science*, 264-273.
- Enyedy, N., Mukhopadhyay, S. (2007). They don't show nothing I didn't know: Emergent tensions between culturally relevant pedagogy and mathematics pedagogy. *Journal of the Learning Sciences*, 16(2), 139-174.
- Frankenstein, M. (1983). Critical mathematics education: An application of Paulo Freire's epistemology. *Journal of Education*, 165(4), 315-339.
- Frankenstein, M. (2009). Developing a criticalmathematical numeracy through *real* real-life word problems. In L. Verschaffel et al. (Eds.), *Words and Worlds: Modelling Verbal Descriptions of Situations* (pp. 111-130). Sense Publishers.
- Freire, P. (1970/2012). *Pedagogy of the oppressed* (30th ed.). Continuum International.
- Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, 7(1), 1-51.

- Gould, R., Machado, S., Ong, C., Johnson, T., Molyneux, J., Nolen, S., Tangmumrunkit, H., Trusela, L., & Zanontian, L. (2016). Teaching data science to secondary students: The Mobilize introduction to data science curriculum. *Proceedings of the International Association for Statistical Education (IASE) Roundtable Conference*, Berlin, Germany.
- Gregson, S. A. (2013). Negotiating social justice teaching: One full-time teacher's practice viewed from the trenches. *Journal for Research in Mathematics Education*, 44(1), 164–198.
- Gutstein, E. (2003). Teaching and learning mathematics for social justice in an urban, Latino school. *Journal for Research in Mathematics Education*, 34(1), 37–73.
- Gutstein, E. (2006). *Reading and writing the world with mathematics*. Routledge.
- Gutstein, E. & Peterson, B. (Eds.) (2013). *Rethinking mathematics: Teaching social justice by the numbers*. Rethinking Schools.
- Hautea, S., Dasgupta, S., & Hill, B. M. (2017). Youth perspectives on critical data literacies. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 919–930.
- Higgins, T., Rubin, A., Mokros, J., Sagrans, J., & Ren-Mitchell, A. (2021). When the data drive the learning. *Proceedings of the International Association for Statistical Education (IASE) 2021 Satellite Conference*, Online Conference.
- Kahn, J. (2020). Learning at the intersection of self and society: The family geobiography as a context for data science education. *Journal of the Learning Sciences*, 29(1), 57–80.
- Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures & their consequences*. Sage Publications.
- Kokka, K. (2020). Social justice pedagogy for whom? Developing privileged students' critical mathematics consciousness. *The Urban Review*, 52(2), 778–803.
- Lee, H. S., Mojica, G. F., Thrasher, E. P., & Baumgartner, P. (2022). Investigating data like a data scientist: Key practices and processes. *Statistics Education Research Journal*, 21(2), Article 3.
- Lee, V. R., Drake, J., & Williamson, K. (2015). Let's get physical: K-12 students using wearable devices to obtain and learn about data from physical activities. *Tech Trends*, 59(4), 46–53.
- Lee, V. R., Pimentel, D. R., Bhargava, R., D'Ignazio, C. (2022). Taking data feminism to school: A synthesis and review of pre-collegiate data science education projects. *British Journal of Educational Technology*, 53, 1096–1113.
- Lesser, L. M. (2007). Critical values and transforming data: Teaching statistics with social justice. *Journal of Statistics Education*, 15(1), 1.
- Louie, J., Stiles, J., Fagan, E., Roy, S., Chance, B. (2021a). Data investigations to further social justice inside and outside of STEM. *Connected Science Learning*, 3(1).
- Louie, J., Roy, S., Chance, B., Stiles, J., Fagan, E. (2021b). Promoting interest and skills in statistical and multivariable thinking with social justice data investigations. *Proceedings of the International Association for Statistical Education (IASE) 2021 Satellite Conference*, Online Conference.
- Lupi, G., & Posavec, S. (2016). *Dear Data*. Princeton Architectural Press.

- Matuk, C., DesPortes, K., Amato, A., Vacca, R., Silander, M., Woods, P. J., & Tes, M. (2022). Tensions and synergies in arts-integrated data literacy instruction: Reflections on four classroom implementations. *British Journal of Educational Technology*, 53(5), 1159–1178.
- National Academies of Sciences, Engineering, and Medicine [NASEM]. (2018), *Data science for undergraduates: Opportunities and options*. National Academies Press.
- Noble, S. U. (2018), *Algorithms of oppression: How search engines reinforce racism*. NYU Press.
- O’Neil, C. (2016), *Weapons of math destruction: How Big Data increases inequality and threatens democracy*. Crown.
- Pangrazio, L., & Selwyn, N. (2019). ‘Personal data literacies’: A critical literacies approach to enhancing understandings of personal digital data. *New Media and Society*, 21(2), 419–437.
- Philip, T. M., Schuler-Brown, S., & Way, W. (2013). A framework for learning about big data with mobile technologies for democratic participation: Possibilities, limitations, and unanticipated obstacles. *Technology, Knowledge and Learning*, 18(3), 103–120.
- Philip, T. M., Olivares-Pasillas, M. C., & Rocha, J. (2016). Becoming racially literate about data and data-literate about race: Data visualizations in the classroom as a site of racial-ideological micro-contestations. *Cognition and Instruction*, 34(4), 361-388.
- Podworny, S., Husing, S., Schulte, C. (2022). A place for a data science introduction in school: Between statistics and programming. *Statistics Education Research Journal*, 21(2), Article 6.
- Radinsky, J., Hospelhorn, E., Melendez, J. W., Riel, J., & Washington, S. (2014). Teaching American migrations with GOS census webmaps: A modified “backwards design” approach in middle-school and college classrooms. *Journal of Social Studies Research*, 38, 143-158.
- Raffaghelli, J. E. (2020). Is data literacy a catalyst of social justice? A response from nine data literacy initiatives in higher education. *Education Sciences*, 10, 233.
- Raygoza, M. C., (2016). Striving toward transformational resistance: Youth participatory action research in the mathematics classroom. *Journal of Urban Mathematics Education*, 9(2), 122–152.
- Ridgway, J. (2016). Implications of the data revolution for statistics education: The data revolution and statistics education. *International Statistical Review*, 84(3), 528–549.
- Ridgway, J. (Ed.) (in press). *Statistics for empowerment and social engagement: Teaching Civic Statistics to develop informed citizens*. Springer.
- Rubel, L. H., Lim, V. Y., Hall-Wieckert, M., & Sullivan, M. (2016). Teaching mathematics for spatial justice: An investigation of the lottery. *Cognition and Instruction*, 34(1), 1–26.
- Rubin, A. (2020). Learning to reason with data: How did we get here and what do we know? *Journal of the Learning Sciences*, 29(1), 154–164.
- Skovsmose, O. (1994). *Towards a philosophy of critical mathematics education*. Springer Science+Business Media.
- Stornaiuolo, A. (2020). Authoring data stories: Adolescents developing critical data literacies. *Journal of the Learning Sciences*, 29(1), 81–103.
- Tate, W. F. (1995). Returning to the root: A culturally relevant approach to mathematics pedagogy. *Theory into Practice*, 34(3), 166–173.

- Taylor, K. H., Silvis, D., Kalir, R., Negron, A., Cramer, C., Bell, A., & Riesland, E. (2019). Supporting public-facing education for youth: Spreading (not scaling) ways to learn data science with mobile and geospatial technologies. *Contemporary Issues in Technology and Teacher Education*, 19(3), 529–542.
- Tygel, A. F., & Kirsch, R. (2016). Contributions of Paulo Freire to a critical data literacy: A popular education approach. *The Journal of Community Informatics*, 12(3), 108–121.
- Vakil, S., & Higgs, J. (2019). Education: It's about power. *Communications of the ACM*, 62(3), 31–33.
- Vakil, S., Marshall, J., & Ibrahimovic, S. (2020). “That’s bogus as hell!”: Getting under the hood of surveillance technologies in an out of school STEM learning environment. *Proceedings of the 14th International Conference of the Learning Sciences (ICSL)*, Nashville, TN.
- Van Wart, S., Lanouette, K., & Parikh, T. (2020). Third spaces for data science education using participatory digital mapping. *Journal of the Learning Sciences*, 29(1), 127–153.
- Weiland, T. (2017). Problematizing statistical literacy: An intersection of critical and statistical literacies. *Educational Studies in Mathematics*, 96, 33–47.
- Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67(3), 223–265.
- Wilkerson, M. H., & Laina. (2018). Middle school students’ reasoning about data and context through storytelling with repurposed local data. *ZDM Mathematics Education*, 50(7), 1223–1235.

Appendix Table 1. Examples of studies of learning interventions to promote critical data literacy in K-12 education settings³

Study authors	Learning context and students' project focus	Learning designers and facilitators	Student participants	Critical literacy (i.e., ethical/social justice) goals	Data literacy goals	Students' data sources	Students' data analysis tools
Rubel et al., 2016	High school mathematics curriculum (15 lesson sequence); analyze the social and political geography of the local lottery	<i>Designers:</i> University researchers, teachers, mappers, informal educators <i>Facilitator:</i> Math teacher	Students from 4 sections of gr. 12 remedial math; public school is 100% Latinx/ Black and low-income, in Northeast city	Recognize and consider how the lottery and geographic distribution of resources can be unjust	Improve learning of probability, combinatorics; remediate concepts of median, average, proportion	Lottery simulations; data from people in neighborhoods; digital maps with project-curated layers and variables	GIS platform; web-based tool to house students' assembled data
Taylor et al., 2019	High school & undergraduate science/teacher ed curriculum; examine how spaces should be improved for youth learning and development	<i>Designers:</i> University & non-profit researchers & educators <i>Facilitators:</i> Developer team, teachers, graduate students	High school students and undergraduates in historically underserved neighborhoods in Chicago, NYC, Seattle	Promote civic participation and input into community development among youth	Develop understanding of collecting and interpreting spatial, real-time, dynamic data	Open-source map data; student-collected video, photos, text annotations, audio of community interviews	GPS devices, wearable cameras, GIS software, mobile mapping platform, paper and pen; mobile augmented reality tools
Van Wart et al., 2020	High-school A) summer science program (5 weeks); study air quality of regional transit system; B) social studies unit (16 weeks); park planning project	<i>Designers:</i> University researchers <i>Facilitators:</i> Researchers, university students	A) 11 students ages 15-18, all Latinx, Black, or Asian, from urban low-income schools; B) two 11th grade classes 99% BIPOC, 80% low income	Develop a "third space" that engages multiple perspectives toward collective ends and to shift what counts as knowledge	Develop students' data collection, analysis, and argumentation skills	A) Students collect air quality data in the community and the subway; B) students assemble observations of and visions for improving a local park	Air quality sensors; participatory mapping tool supporting multiple data types (e.g., photos, sensor data)

³ Studies are listed chronologically by date of publication and to juxtapose similarities and contrasts in critical and data literacy goals among projects.

Study authors	Learning context and students' project focus	Learning designers and facilitators	Student participants	Critical literacy (i.e., ethical/social justice) goals	Data literacy goals	Students' data sources	Students' data analysis tools
Kahn, 2020	Free summer library workshop for middle- and high-school youth (6 sessions, 3 weeks; create a family geobiography)	<i>Designers:</i> University researchers <i>Facilitators:</i> research team with library staff	17 participants primarily from African American backgrounds, focus on two pairs of siblings, in Nashville	Support students' critical perspectives toward data and the phenomena they represent	Develop skills in storytelling with large-scale data	U.S. Census and global demographic data; public records websites, mapping tools (e.g., Google Earth)	Social Explorer, Gapminder
Stornaiuolo, 2020	High school media arts class in a makerspace; examine personally meaningful topic for a week, create data visualizations for T-shirts	<i>Designers:</i> Media arts teacher, with support from university researchers <i>Facilitator:</i> Teacher and researchers (participant observers)	31 students in gr. 10-12 required arts credit, 92% Black or Latinx and 100% low income, in an urban public school in the Northeast	Recognize data's role in reproducing educational injustice; empower students to design and author their own futures	Understand data forms and collection methods, analysis tools, visualization approaches	Publicly available datasets	Online tools such as Google Trends and DataBasic
Vakil et al., 2020	High school summer workshop and after-school program; examine ethics of social media surveillance, cell-phone tracking, facial recognition technology	<i>Designers:</i> university researchers, non-profit leaders, tech activists, CS professional, civics teacher <i>Facilitators:</i> civics teacher and design team members	11 students: 6 Black, others biracial, Asian, Latinx, white; attend a community center on border of urban and suburban area in Midwest	Interrogate role of surveillance and other advanced technologies in local community	Understand how the technologies work "under the hood," explain social impacts to community members	Data collected by students from class demos, virtual reality exercise, walking tour of city, meetings with local activities and tech experts	GeoMedia, allows users to pull tweets within a certain geographic radius
Kokka, 2020	Curriculum units designed to teach mathematics for social justice in a	<i>Designer:</i> teacher <i>Facilitator:</i> teacher	10 focal students in an elite independent K-8 school with a	Understand sociopolitical conditions, develop civic	Support learning of measures of central	U.S. Census data (printed on posters)	Paper and pencil

Study authors	Learning context and students' project focus	Learning designers and facilitators	Student participants	Critical literacy (i.e., ethical/ social justice) goals	Data literacy goals	Students' data sources	Students' data analysis tools
	gr. 6 class; examine issues such as housing prices, household incomes, and cost of living in local town		social justice focus in CA; 6 Caucasians, 1 Indian American, 1 mixed race, 2 unidentified	empathy, promote taking action	tendency, bar and line graphs, percentages, rate, and proportion		
Louie et al., 2021a, 2021b	High school mathematics curriculum modules (3 weeks); analyze income inequality in the U.S., nationally and by individual attributes	Designers: university statistics faculty, math curriculum developer, researchers Facilitators: teachers at local schools	Over 180 gr. 12 students in non-AP 4 th -year math/ statistics classes, 55-91% Black and Latinx, 38-65% low income, from 6 urban schools in the Northeast	Build a deeper awareness of the scope and persistence of social and economic inequalities in the U.S.	Improve understanding of entire data investigation process, measures of center, variability, multivariable thinking	U.S. Census and American Community Survey microdata from IPUMS-USA	CODAP