

## **Advancing Social Justice Through Data Literacy**

**Josephine Louie, Emily Fagan, Jennifer Stiles, Soma Roy, and Beth Chance**

**When students gain “critical data literacy skills,” they can analyze data in ways that help them understand power dynamics and ongoing challenges in society.**

Reflecting on the purposes of education, authors Heather McGhee and Victor Ray (2022) recently wrote in *The New York Times* that schools play a critical role in preparing citizens for democracy. They argue that to build a true democracy, young people must gain an unvarnished understanding of persistent social challenges--and the tools to confront them. We agree, and we believe critical data literacy is one set of tools that individuals must acquire to achieve social justice within a healthy democracy.

### **What Is “Critical Data Literacy”—and How Can We Teach It?**

Today, our world is inundated with data – on topics ranging from inflation and unemployment, patterns in extreme weather, the spread of viral diseases, and what students are learning (or not learning) in school. It’s left up to individuals to make sense of what’s fact or fiction and how to act in response. Unfortunately, up to 80% of individuals in business and public surveys express little confidence in their abilities to understand and work with data (H+K Global, 2016; Qlik, 2018). In our view, data literacy involves the ability to *read and write with data*—to extract meaning from diverse types of data and communicate what those data reveal, particularly in numeric, statistical, and graphical forms. Drawing on the work of Gutstein (2006), we consider “critical data literacy” to involve the ability to *read and write the world with data*. Critical data literacy builds on data literacy and incorporates ethical, social, and political perspectives. People who have critical data literacy routinely consider data ethics and how data can be used or

misused. Just as important, they examine data to understand power relations in society.

Recognizing and working to remedy the inequitable access various groups have to social, economic, and political resources is a hallmark of social justice, after all.

Schools have a big role to play in helping students develop critical data literacy to become informed citizens and create a more just world. As a team of researchers, statistics educators, and curriculum developers, we've worked to infuse skills for critical data literacy into schools. In 2018, we launched a project funded by the National Science Foundation titled *Strengthening Data Literacy across the Curriculum* (or SDLC) to create a set of curriculum materials to support high schools in promoting students' critical data literacy. Collaborating with over a dozen high school mathematics and social studies teachers in Massachusetts, our team tested and refined these materials in schools with high proportions of students from Black, Latinx, and low-income communities. We ultimately produced two curriculum modules, "Investigating Income Inequality in the U.S." and "Investigating Immigration to the U.S.," which have been implemented in high school non-Advanced Placement mathematics classes and in a small number of U.S. history classes. Each module (available for free at <https://go.edc.org/ussdata>) involves approximately 15 hours of instructional time to complete in full. In our study, teachers used our materials successfully after approximately 10 hours of professional development prior to implementing each module.

Both our modules are built on three key components:

**They center on topics and questions that engage students with social justice issues—** issues we have found to be interesting and culturally relevant for students involved. In the module about income inequality, for example, students examine large-scale data from the U.S.

decennial census and the American Community Survey (ACS), which is collected annually by the U.S. Census Bureau, to address questions such as: What is income inequality? How have higher- and lower-income earners been faring over time? How much income inequality exists between males and females, and does education help to explain the gender wage gap? In the module exploring immigration data, students encounter myths or claims about immigrants in the U.S. (adapted from Learning for Justice Staff, 2017, and Denhart, 2017). They use census and ACS data to investigate questions like: Are there more immigrants in the U.S. than ever before? or Are immigrants less likely than U.S.-born individuals to participate in the labor force?

**They include standards-based interdisciplinary content.** Curriculum activities focus on content aligned with multiple sets of high school learning standards. For example, students are asked to conduct data analyses that strengthen their conceptual understandings of measures of center and variability, sampling variability and margins of error, comparing quantitative distributions, and using conditional proportions to compare categorical data--central statistical ideas within the Common Core State Standards for Mathematics for secondary grades (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). Students also learn to anchor data work using a four-step data investigation process that's fundamental to data literacy: (1) formulating questions; (2) assembling data; (3) using statistical and other tools to analyze data; and (4) interpreting results (Bargagliotti et al., 2020). This process echoes the four-step Inquiry Arc that leaders of social studies education say should undergird students' social studies and civics learning (National Council for the Social Studies, 2013). Investigations of U.S. income inequality and immigration also align with recommended U.S. history frameworks nationally (National Center for History in the Schools, 1996) and with frameworks for teaching history in Massachusetts, where we piloted our new modules.

**They follow research-based learning approaches.** Following recommendations from statistics education research, our modules require students to learn to analyze large-scale datasets with accessible digital tools (Chance et al., 2007). In lessons in our modules, students used the Common Online Data Analysis Platform, or CODAP (at <https://codap.concord.org/>). This free, browser-based data analysis and visualization tool helps students develop a conceptual understanding of statistical ideas rather than focus on calculations. We built a Microdata Portal (a free plug-in within CODAP) that provides access to anonymized person-level data from multiple decades of the decennial census and the 2017 American Community Survey (see Figure 1). Using interactive tables and graphs, students assemble, manipulate, and analyze datasets containing information on thousands of individuals. They work in pairs or in small groups, and teachers support student thinking by facilitating group discussions and providing feedback. Each module's investigations also have a project-based orientation (Lenz & Larmer, 2020), providing scaffolding and a culminating investigation in which students explore one of the module's driving questions, using variables of their choice.

### **Data Literacy in Action: Opening Their Eyes to Inequities**

From spring 2019 through spring 2021, 15 high school mathematics teachers in 12 schools implemented iterations of these two modules in mathematics classrooms, and during the peak of the Covid-19 pandemic in fall 2020, several mathematics teachers taught the immigration module remotely. All these schools were located primarily in the metropolitan Boston region and had disproportionately large student populations from Black, Latinx, and low-income communities. To get a sense of how these curriculum elements strengthen students' critical data

literacy, let's look at some specific learning activities involved in our modules.

Generally, students begin each module by examining the census and ACS questionnaires and discussing the strengths and limitations of different data sources. To investigate how income inequality in the U.S. has changed over time, for instance, a lesson has students create and interpret graphs of wages reported by employed individuals in 1940, 1980, and 2017, displayed in constant 2017 dollars. Students can examine visual displays of wage distributions that draw from these data, including mean and median values and statistical outliers, to see how the wages of higher- and lower-income earners have diverged over time. These graphs also shed light on the degree to which wages in the United States differ for males and females and whether the wage gap can be explained by education levels (Figure 2 shows this data visualization). In the module's final investigation, students discover that even after controlling for other variables of their own choosing (such as occupation, marital status, and hours worked), the wage gap between male and female workers persists—which prompts discussion of possible sex discrimination in wages.

Students analyze data similarly to examine common claims about immigrants in the U.S. In one lesson, students create CODAP graphs using data from the census and ACS to investigate a claim that immigrants are less likely than U.S.-born individuals to be employed. Looking at 2017 ACS data, they discover that, contrary to this claim, overall rates of labor force participation were *higher* for immigrants that year than for people born in the U.S. Further investigations with additional variables continue to reveal the weakness of the original claim while revealing more complex patterns. For instance, for people without a bachelor's degree, employment rates among immigrants (63%) are higher than among U.S.-born people (55%), while for those with a bachelor's degree, employment rates among immigrants (73%) and the U.S. born (72%) are

similar.

We're seeing good results in these classes in terms of stronger understanding of core data concepts and how to interpret data related to social issues. In classrooms that used the income inequality module in fall of 2019, we found that students' statistics and data literacy outcomes—what we call their abilities to read and write with data—grew significantly based on pre- and post-module assessments that we adapted from validated instruments. Specifically, we measured significant growth in students' abilities to make sense of different data representations, reason about measures of center, and engage in multivariable thinking (Louie et al., 2022). Students' interests in data analysis also grew significantly, based on established pre- and post-module interest scales. Even during the difficult period when schools went remote due to the pandemic, students who learned through the immigration module showed significantly higher levels of interest in data and a stronger perception of the importance of working with data as measured by pre- and post-module surveys.

The modules also appeared to increase students' awareness of social justice issues. They strengthened students' ability to *read and write the world with data*. After completing the modules, some students expressed greater awareness and deeper understanding of specific social and political conditions in the world. One student wrote in a survey, "I enjoyed the lessons that discussed current issues such as income inequality between men and women and people with different levels of education. It connected the lesson with the real world and helped me understand both statistics and society in more in-depth ways." Another shared:

Our generation has a lack of understanding of real-world issues, and I feel like we're pretty influenced by each other and our parents. You can't really talk politics in school, but this is the closest you can probably get, where the kids can

form their own opinion. I bet you a lot of kids' opinions about this topic changed...I and a few other students were on a certain side or leaning towards one end of the issue, but in the end [the data showed what we assumed] was not the case.

Two other comments indicated that some teens felt empowered to work with data, potentially for social good:

I liked the income inequality data more than the statistics lessons we usually do because . . . it felt like we were our own statisticians analyzing data.

You always hear about it, like there's income inequality between different genders, and when we actually did the data, it was like—oh wow, this is real. People aren't just making things up. This is a real problem—and hopefully we can figure something out.

### **Teachers Can Tackle This—with the Right Support**

We've found teachers can get students working successfully with data—if they have help. To prepare teachers to try data-focused materials in their classes, our team provided 10 hours of professional development for each module. We shared an overview of module content and, casting teachers as learners, modeled how to teach the lessons and scaffold students in data-analyzing activities and projects. Through guiding teachers in using these materials, we've realized many teachers need additional learning or practice in five domains to successfully support students' critical data literacy.

**1. Data concepts.** We worked with both high school mathematics and social studies teachers to develop our materials, but focused module testing in mathematics classrooms. Consistent with prior research (Franklin et al., 2015), our collaborating math teachers were experienced in teaching mathematics, but less so in teaching data concepts and practices. Teachers need professional learning opportunities where they can apply core statistics and data concepts to real-world data before they will feel comfortable facilitating these activities among students.

**2. Datasets and sources.** Almost all our collaborating teachers had *not* worked with large-scale datasets previously. To ensure that teachers feel comfortable with the datasets they introduce to students, teachers need opportunities to become familiar with the origins, structures, strengths, and limitations of these datasets. We provided teachers with these types of opportunities during our PD sessions, when teachers could explore decennial census and ACS data from the Microdata Portal themselves, practice cleaning messy data, and discuss how different variables have been defined and coded. Organizations like ours that are part of the [Data Science for Everyone](#) coalition such as [youcubed.org](#), [instepwithdata.org](#), and [skewthewscript.org](#) offer teachers professional development and/or lesson materials to teach with real-world datasets.

**3. Data analysis tools.** Although our collaborating teachers had experience using spreadsheets, working with data visualization and analysis tools was new. The CODAP platform has an intuitive interface and is easy for students in secondary grades to learn, but teachers need time to play with tools like CODAP before they will feel confident in introducing these tools to students.

**4. Background on social issues.** Teachers don't need to be experts in social justice topics that data lessons focus on. But students' data investigations will benefit if teachers have a basic understanding of these issues—and key concepts or debates connected to them—and can point

students to reputable information sources to compare and reflect on their data findings. Many organizations study topics like income inequality and immigration and make their findings widely available; teachers should seek out such information.

**5. Ways to facilitate challenging conversations.** To promote critical data literacy, teachers should know how to guide students through conversations on why different groups in society may experience unequal social, economic, or political outcomes. These conversations can be difficult, particularly when issues such as racism, sexism, and other forms of group discrimination arise. Teachers need guidance on how to facilitate such discussions to avoid hurt or harm among participants. Although not a comprehensive list or a full solution, here are a few resources that educators might find helpful.

### **Tackling Data Head-on**

Democratic populations need critical data literacy to make sense of the data flooding our everyday lives and to understand and confront systemic social challenges. Schools should be at the forefront of fostering critical data literacy. Teachers need not only curriculum materials, but also support for giving students the understanding and skills to “read” data in various disciplines—including chances to collaborate with colleagues across departments. Our experiences reveal that teachers and school leaders *want* to provide such learning opportunities for students, and our research suggests such efforts may be well received. In the words of one student, “I think we need to do more of [these modules] in school, to be honest with you. A lot of people are afraid to talk about [these] issues... If more kids tackle it head-on, they’ll have a better understanding of how our world really works.”

## 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.

Chance, B., Ben-Zvi, D., Garfield, J., & Medina, E. (2007). The role of technology in improving student learning of statistics. *Technology Innovations in Statistics Education*, 1(1).

Denhart, M. (2017). *America's advantage: A handbook on immigration and economic growth*. George W. Bush Institute.

Franklin, C., Bargagliotti, A., Case, C., Kader, G., Scheaffer, R., & Spangler, D. A. (2015). *The statistical education of teachers*. American Statistical Association.

Gutstein, E. (2006). *Reading and writing the world with mathematics*. Routledge.

H+K Global (2016, November). Audiences love data, but what if they don't understand it? <https://www.hkstrategies.com/en/magnify-data-literacy-gap/>

International Forum for Social Development. (2006). *Social justice in an open world: The role of the United Nations*. Department of Economic and Social Affairs of the United Nations Secretariat. <https://digitallibrary.un.org/record/567990?ln=en>

Learning for Justice Staff. (2017, March 14). Ten myths about immigration. Learning for Justice. <https://www.learningforjustice.org/magazine/ten-myths-about-immigrationupdated>

Lenz, B. & Larmer, J. (2020). Project-based learning that makes a difference. *Educational Leadership*, 77(6), 66-70.

Louie, J., Stiles, J., Fagan, E., Chance, B., & Roy, S. (2022). Building toward critical data literacy with investigations of income inequality. *Educational Technology & Society*, 25(4), 142-163.

McGhee, H. C. & Ray, V. (2022, September 1). School is for making citizens. *New York Times*.

National Council for the Social Studies. (2013). *The college, career, and civic life (C3) framework for social studies state standards: Guidance for enhancing the rigor of K–12 civics, economics, geography, and history*.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common core state standards for mathematics*.

Qlik (2018), Lead with Data: How to Drive Data Literacy in the Enterprise, Qlik, PA.

[https://www.qlik.com/us/bi/-/media/08F37D711A58406E83BA8418EB1D58C9.ashx?galink=datlitreport\\_resource-library](https://www.qlik.com/us/bi/-/media/08F37D711A58406E83BA8418EB1D58C9.ashx?galink=datlitreport_resource-library)

**Josephine Louie** is a senior research scientist at Education Development Center (EDC), where she leads programs to improve learning in K-16 mathematics, science, and data literacy. **Emily Fagan** is a project director and senior curriculum developer at EDC, a former teacher and a co-author of books on formative assessment in mathematics. **Jennifer Stiles** is a research associate at EDC and co-leads a mentoring program for scholars in STEM education. **Soma Roy** is a professor of statistics at California Polytechnic State University and helps lead projects that promote the use of real-world data and simulation-based methods in education. **Beth Chance** is a professor of statistics at California Polytechnic State University who has co-authored textbooks for introductory statistics.

*Note.* This work was supported by the National Science Foundation (NSF) under Grant No. DRL 1813956. Opinions, findings, conclusions, or recommendations are those of the authors and do not necessarily reflect the views of the NSF.