Making Sensor Data Dear

Accessible sensing platforms as conduits to localized, personally meaningful data science

Leah F. Rosenbaum[1], and Paulo Blikstein[1, 2]

[1] Teachers College Columbia University, USA
[2] Instituto Tecnologico de Monterrey, Mexico

Abstract
As data has become an increasing part of our lives, data science education has sought to support youth in becoming critical consumers of data and authors of data-based claims. Relatively abstract sensor data can pose challenges for young people's sense-making. We share youth work from a week-long, museum-based data science camp to illustrate the ways youth used portable, low-cost environmental sensors to situate explorations of sensor data in hyper-local contexts in their homes and communities. We also share challenges that some youth faced in translating their locally grounded curiosities into questions appropriate for sensor data research. Even as topics and technologies evolve, we position our findings as underscoring the continued relevance of constructionism's commitment to personally meaningful projects that authentically reflect learners' interests.

Keywords and Phrases: Student-centered learning, place-based data science, low-cost sensor technology

1. Introduction & Background
As data increasingly shape our daily lives, data science education takes up the challenge of equipping learners with the skills and practices necessary to be critical consumers and authors of data-based claims. Resonant with constructionism's commitment to learning within personally meaningful projects (Papert, 1980), a number of diverse approaches aim to situate data and data science perspectives as relevant to learners' lives. Place-based data science learning grounds data-based narratives in spaces and locations important to its learners (e.g., Lanouette et al., 2024). Quantified self approaches position learners' own biological data (e.g., daily steps, heart rate, galvanic skin response) as a resource for learning about data science and their own bodies (e.g., Lee et al., 2021). Web-based data mining offers one avenue for youth to create their own big data sets on which to pursue data science investigations rather than analyzing data created or curated by others (e.g., Acquah et al., 2024). Though outside formal research, the Dear Data Project aims to convey the aesthetic beauty of data that many find compelling (Lupi & Posavec, 2016).

Rosenbaum, L. F., & Blikstein, P. (2025). Making Sensor Data Dear. Accessible sensing platforms as conduits to localized, personally meaningful data science. *Constructionism Conference Proceedings, 8/2025*, 437–444. https://doi.org/10.21240/constr/2025/98.X

We contribute to this conversation by pursuing the following research question: *What are the personally meaningful places, spaces, and phenomena in which youth conduct sensor-based data science projects of their own design?* We share how museum-based programming that used a low-cost, learner-focused, sensor-based data collection platform enabled middle school youth to collect and use sensor data to explore phenomena rooted in their personal experiences and motivations. We share both thematic analysis of youths' research questions and contexts and two cases that illustrate how youth brought lived experience to bear in formulating and making sense of their research projects. Despite some challenges that we document, we advocate for both technological and curricular developments that support more youth in establishing such a personal relationship with data. We argue that constructionism's tenet of personally meaningful projects applies even as content areas and learning technologies evolve and can continue to inform efforts to position data as a tool youth can use to explore and understand their worlds.

2. Methods

2.1 Setting & Participants

We draw on data from three iterations of a week-long data science camp ("Data Innovation Lab") hosted by a science museum in a major U.S. city between April and August, 2024. Researchers collaborated with museum educators to create a community-based data science curriculum (Cohn, 2008; Phillips et al., 2019) that was both playful and learner-centered (Honey, 2013). Within this approach, facilitators encouraged youth to pose questions about personally meaningful places in their homes and communities. Facilitators supported youth in posing and refining their research questions, selecting and configuring sensors to collect relevant data, and creating visualizations to present their findings. The 5-day camp concluded in a presentation session in which youth shared posters about their data explorations with their peers and family members.

Participants were 11- and 12-year-old youth who participated in the Data Innovation Lab. Youth were recruited for the camp by museum staff in collaboration with the public schools these youth attend. Participant demographics are representative of the neighborhoods surrounding the museum in which they live.

2.2 Materials

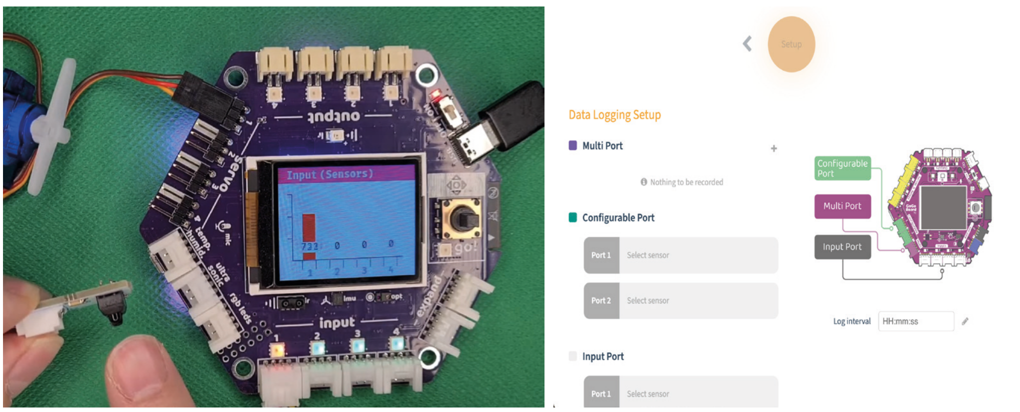


Figure 1: The "GoGo Board" available to campers as a make-and-take sensor platform. The above board (left) is equipped with a proximity sensor. Youth used the web interface (right) to configure their boards and sensors.

The camp curriculum used the GoGo Board hardware and software platforms (Figure 1) for sensor data collection and analysis. The GoGo Board includes a large, customizable display screen that obviates the need for computers and mobile phones in the field, making it ideal for use on-the-go and in the varied spaces youth occupy in their daily lives. Each participating youth received their own GoGo Board that they took home each afternoon of camp and which they kept when the camp concluded. As the GoGo Board supported a range of digital and analog sensors - e.g., temperature, humidity, light, air quality, water quality, and heart rate - it allowed participants to collect a variety of data types to characterize their chosen phenomena of interest.

2.3 Data Sources & Analysis

In total, 47 campers completed data science projects and consented to participate in research. Of those, 12 campers worked in pairs of two, resulting in 41 unique projects. We analyzed the final posters that youth created to share their week-long data projects. Specifically, we analyzed the research questions youth posed, the motivations and contexts they shared for those questions, and the data visualizations they created to share their collected sensor data. As we are primarily interested in the ways in which youth situated data explorations in personally meaningful contexts, we categorized each project by the primary location stated in each research question. The resulting categories were: bedroom, home, community, and museum. Bedroom was identified as a special space within the home given its role for many youth as a space of solitude and personal expression. Museum was identified as a special place within the community given that it hosted the data science camp.

We also selected two participants as focal cases to illustrate how youth positioned data as a tool for exploring personally meaningful questions. We selected focal participants based on the following criteria:

- Each collected data in a different focal space;
- Each used a different focal sensor; and
- They were not analyzed in prior writing about this project.

Focal youth are referred to by pseudonyms.

3. Findings

We first characterize the places in which youth situated their data collection projects and then share two focal cases to illustrate how youth used data to explore and understand personally meaningful questions.

3.1 Youth-selected spaces for data collection projects

Using their portable sensor kits, youth collected and explored data in a range of familiar places and spaces. Community-based projects were the most common (Figure 2), with youth exploring topics including transit (bus, car, and subway), retail spaces (crowd size and noise levels), and air and water quality. In the next most common setting, their homes, youth posed questions and collected data about a similarly wide range of topics, including recreation (gaming), pet behavior, bathroom occupancy, air and water quality, and temperature. Youth also explored topics of temperature and air and water quality in the museum, often when they encountered technical issues with collecting data in their homes or communities. Though less common, youth also explored topics specific to their bedrooms (dust and insect prevalence). When given the tools and freedom to do so, youth chose a rich array of locations and topics in which to root their data science projects.

Bedroom (2)
"Where are the spiders in my room coming from?"
"How much dust is on my bed?"

Home (12)
Temperature: "Why is the living room so hot and the basement so cold?"
Recreation: "How hot does my iPad get?"
Bathroom: "How many people use the bathroom at my home? … It is always occupied."
Water quality: "How safe is the water in my house?"
Air quality: "What is the air quality of the 4th floor garbage shoot and how much trash is in it?"
Pets: "Are my birds louder [than the outside birds]?"

Community (18)
Transit: "How loud can it be on top of [a] car?" "Is the <route number> bus too crowded when I'm going home?" "How is the train construction causing change to the human body?"
Retail: "How many times do people enter a store and how loud will it get?"
Air quality: "How to factories affect our air quality?"
Water quality: "Does the river get dirtier as the ferry drives by?"

Museum (9)
Temperature: "How does the temperature change as more people walk into the museum?"
Air quality: "What is the air like in <museum>?"
Water quality: "Does <museum> have good quality water for visitors?"

Figure 2: The spaces where youth used their low-cost sensor kits to collect data about questions they posed.

3.2 Focal youth: Using data to explore meaningful questions about transit and garbage

We next share two example projects to illustrate in greater detail how youth expressed personal interests and brought data to bear on questions that mattered to their everyday lives.

Hope wanted to study how crowded it got on the bus she took to and from the museum each day. She mentioned that "this is important because it helps me and others know if we can get a seat or a good spot at this specific time (3:15-4:10)" (poster). She drew a map of the route (Figure 3, left), contextualizing her bus route within landmarks that included shops, residences, parks, and streets. Some places she named specifically (e.g., McDonald's, CVS, Dunkin Donuts) while others she described more vaguely ("I don't know Ave"), giving a richly personal view of the bus route in question.

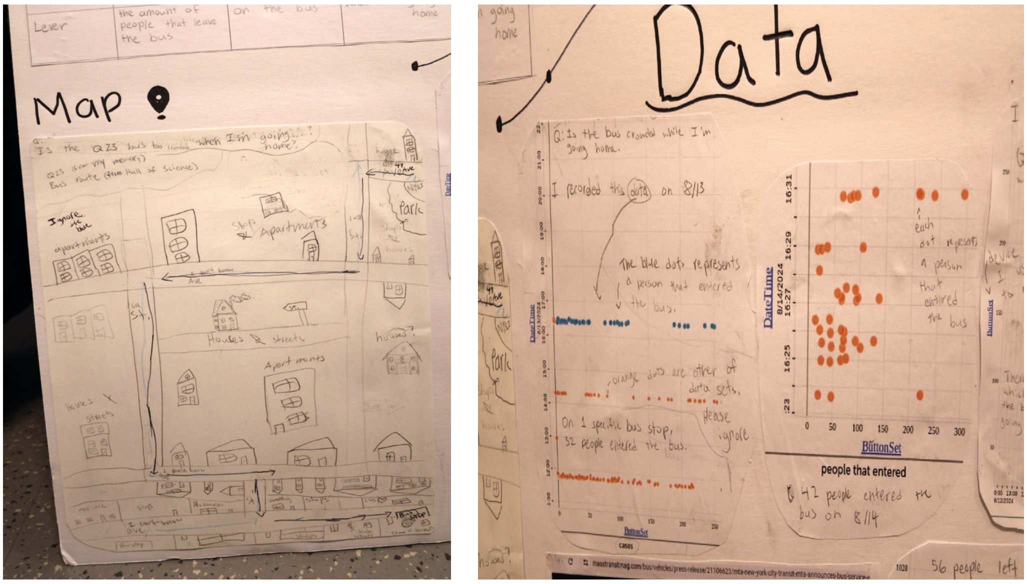
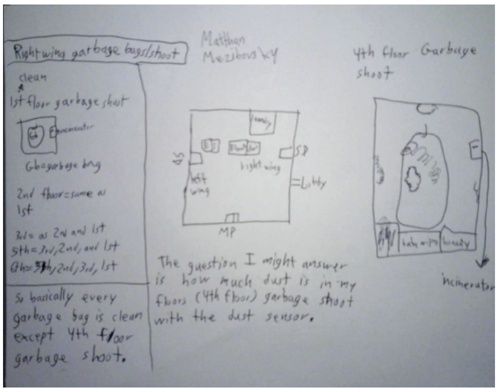
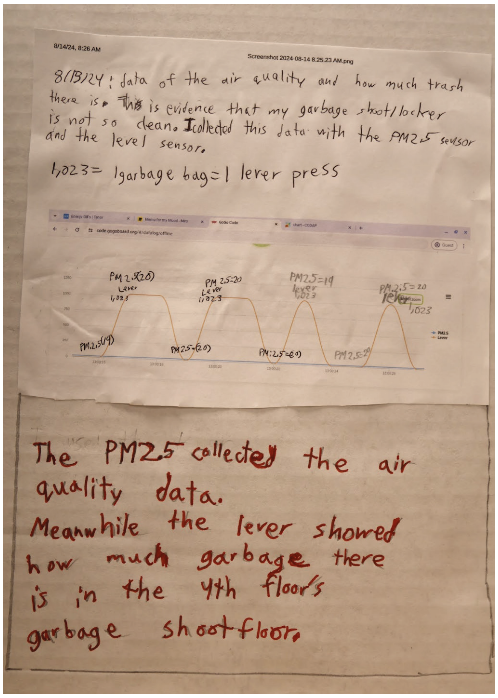


Figure 3: Hope's map of her bus route (left) and her data of how many people entered the bus while she rode it (right).

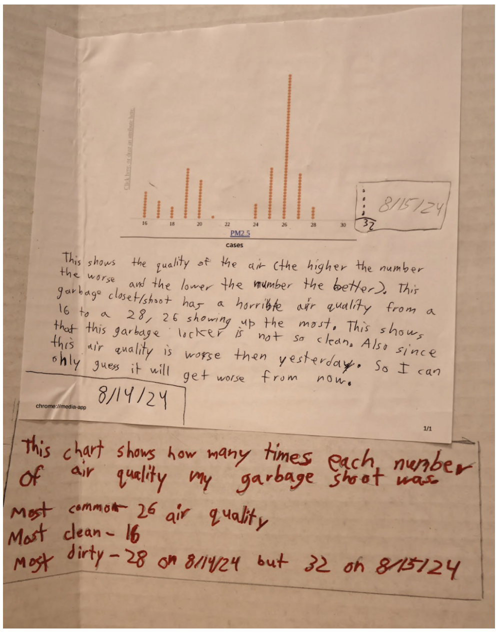
Using a lever and button set, Hope collected data over two days of her commute home from the museum, pressing a button or the lever every time someone entered the bus. While Hope could have collected these same counts by means of simple tally marks, the sensors she used associated each count with a timestamp. Her resulting plots (one in Figure 3, right) clustered her data points - she explained “each dot represents a person that entered the bus” (poster) - in roughly 2-minute increments corresponding to the time between bus stops. Hope’s sensor data both addresses her original question about the total number of people on the bus and also includes this spatial information reflecting the detailed map she provided to situate her project. Notably, Hope’s conclusion from her project was that the bus crowdedness “depends on the day, time, and bus stop/schedule 😐” (poster). We interpret her comment to suggest that, as much as her project reflected a question and context that mattered to her, it also helped her learn about the challenges and contextuality of data science investigations.

Matteo similarly situated his project within a phenomenon important to his daily life: the amount of trash in his floor’s garbage shoot. In drawing a map of his apartment building’s garbage situation, Matteo described that “basically every garbage [shoot] is clean except 4th floor garbage shoot” (poster). In his sketch of the garbage room in question (Figure 4, left), Matteo included details such as the incinerator as well as the garbage bags, baby wipes, and laundry that were contributing to cleanliness issues.





"8/13/24: data of the air quality and how much trash there is. This is evidence that my garbage shoot/locker is not so clean."



"8/14/24: This shows the quality of the air ... This garbage closet/shoot has a horrible air quality from a 16 to a 28, 26 showing up the most. ... Also, since this air quality is worse than yesterday. So I can only guess it will get worse from now."

Figure 4: Left, Matteo's sketch of his apartment's 4th floor garbage room. Right, his plots of air quality and lever (proxy for garbage bag) quantifying the garbage issue and its impacts.

To quantify his floor's garbage problem and its impact on air quality, Matteo used a lever sensor and a PM2.5 dust sensor to collect data across three days. On the first day, he counted four trash bags (the four orange peaks in Figure 4, top right) and noted air quality measurements in the range of 19-20 ppm of PM2.5 (Figure 4, top right). The next day, he again collected air quality measurements, noting the range of 19-26 micrograms per cubic meter of PM2.5 (the orange histogram in Figure 4, bottom right). He returned again the next day and gathered four air quality readings of 32 micrograms per cubic meter of PM2.5 (Figure 4, bottom right). From this trend in air quality measurements, Matteo inferred that the air quality of the garbage room would "only get worse from now" (poster), presumably until the garbage bags were removed.

4. Discussion & Conclusions

In sharing youth work from a 5-day data science camp, we identified a range of places and phenomena in which middle schoolers rooted data science explorations when given the technology and agency to do so. From their bedrooms to their bus routes, youth posed questions that reflect their everyday curiosities and frustrations. And they gathered sensor data to make sense of and corroborate their experiential knowledge of those phenomena. In some cases, sensor data enhanced observations that youth could have made with their senses alone (e.g., Hope's button set added time to her count of people on her bus, clustering her data according to bus stops). In other cases, the sensors youth used gave them access to a level of information about their surroundings that they could not have achieved without the use of technology (e.g., Matteo's use of the air quality sensor), offering a powerful new layer of perception on the issues they identified. Akin to Papert's gears (1980), we argue that the personal meaning youth can attribute to sensor data of places and spaces in their lives offers an unusual affective benefit to their data science learning, a benefit we see as echoing findings of other place-based data science education (e.g., Lanouette, 2022).

We also acknowledge limitations of this work. As some of the research questions quoted above indicate (e.g., "Where are the spiders in my room coming from?" "Why is the living room so hot and the basement so cold?"), not all youth framed their explorations in ways that identified variables and sources of variance key to data science investigations. Resonant with other findings in science and statistics education (e.g., Arnold & Franklin, 2021; Manz et al., 2024), youth may need particular support in the question formation and evolution process to help recognize when and where sensor data is appropriate and where other sorts of data are more appropriate for exploring the phenomena they care about. Relatedly, youth may need support formulating data collection plans that accurately represent their research question (e.g., Hope counting the number of people entering the bus but not the number existing, complicating her exploration of the number of available seats). Finally, youths' interpretation of their collected data did not always align with scientific norms. For example, Matteo saw the increase in PM2.5 in his garbage room from 16 to 32 and concluded that "this garbage closet/shoot has a horrible air quality" when both values fall within the acceptable 24-hour standard for air quality. Similar to support for strengthening research questions, curricular materials could provide reference frames for sensor data when youth collect data within a narrow range. Despite these challenges, we see the degree of voice reflected in youths' projects as evidence for the continued value of constructionism's commitment to personally meaningful projects even within newly emerging fields like data science education.

Acknowledgements

We appreciate the work of the youth participants in the Data Innovation Lab as well as of Anthony Negron, Kate Maschak, and Ray Ferrer in facilitating the week-long camp. This material is based upon work supported by the National Science Foundation under Grant No. 2314089. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

Acquah, A., Barany, A., Johnson, M. A., Scarola, A. S., Rivera, C., & Walker, J. T. (2024). *Sandboxing in Data Science: An Exploration of Youth Learning Using Open-Inquiry Approaches for Computing-based Data Mining*. Proceedings of the 18th International Conference of the Learning Sciences-ICLS 2024, pp. 2433-2434.

Arnold, P., & Franklin, C. (2021). What makes a good statistical question? *Journal of Statistics and Data Science Education*, *29*(1), 122–130.

Cohn, J. P. (2008). Citizen Science: Can Volunteers Do Real Research? *BioScience*, *58*(3), 192–197. https://doi.org/10.1641/B580303

Honey, M. (2013). *Design, make, play: Growing the next generation of STEM innovators*. Routledge.

Lanouette, K. (2022). Emotion, place, and practice: Exploring the interplay in children's engagement in ecologists' sampling practices. *Science Education*.

Lanouette, K., Cortes, K. L., Lopez, L., Bakal, M., & Wilkerson, M. H. (2024). Exploring Climate Change Through Students' Place Connections and Public Data Sets. *Science Scope*, *47*(3), 18–25.

Lee, V. R., Drake, J., Cain, R., & Thayne, J. (2021). Remembering what produced the data: Individual and social reconstruction in the context of a quantified self elementary data and statistics unit. *Cognition and Instruction*, *39*(4), 367–408.

Lupi, G., & Posavec, S. (2016). *Dear data*. Chronicle books.

Manz, E., Stoler, A., Federico, L., Patton, S., Weaver, L., Diaz Silveira, G., & Nassar, S. (2024). What Should We Investigate? Using a classroom decomposition chamber to support the development of investigation questions. *Science and Children*, *61*(5), 33–40.

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. Basic Books.

Phillips, T. B., Ballard, H. L., Lewenstein, B. V., & Bonney, R. (2019). Engagement in science through citizen science: Moving beyond data collection. *Science Education*, *103*(3), 665–690.