Journal of the Learning Sciences

ISSN: 1050-8406 (Print) 1532-7809 (Online) Journal homepage: https://www.tandfonline.com/loi/hlns20

From Data Collectors to Data Producers: Shifting Students' Relationship to Data

Lisa Hardy, Colin Dixon & Sherry Hsi

JOURNAL OF THE LEARNING SCIENCES, 00: 1–23, 2019

ISSN: 1050-8406 print / 1532-7809 online
DOI: https://doi.org/10.1080/10508406.2019.1678164

From Data Collectors to Data Producers: Shifting Students' Relationship to Data

Lisa Hardy, Colin Dixon, and Sherry Hsi
The Concord Consortium

This paper contributes a theoretical framework informed by historical, philosophical and ethnographic studies of science practice to argue that data should be considered to be actively produced, rather than passively collected. We further argue that traditional school science laboratory investigations misconstrue the nature of data and overly constrain student agency in their production. We use our "Data Production" framework to analyze activity of and interviews with high school students who created data using sensors and software in a ninth-grade integrated science class. To understand the opportunities for students to develop act with and perceive agency in data production, we analyze in detail the case of one student as she came to use unfamiliar technologies to produce data for her own personally relevant purposes. We find that her purposes for producing data emerged as she worked, and that resistances to her purposes were opportunities to act with and perceive her own agency, and to see data in new ways. We discuss implications for designing science learning experiences in which students act as agents in producing and using data.

Sensor-based laboratory investigations should be a setting in which students develop an understanding of the origins and uses of data, and in particular an understanding of the ways in which data are produced as artifacts of a human, technical and material activity. Due to conceptual learning goals and sensitivities of sensors as devices, educators often carefully prescribe methods to ensure that the data are amenable to a target conceptual interpretation (Lehrer, Schauble, & Lucas, 2008). In these investigations, the role of the students is to carry out procedures rather than act as agents in producing and using data. School science

Correspondence should be addressed to Lisa Hardy, The Concord Consortium, 6550 Vallejo Street, Suite 101C, Emeryville, CA 94608, USA. E-mail: lhardy@concord.org

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hlns.

laboratories thus obscure the ways that data are produced for human purposes, in interaction between designed technologies and the material world. However, if learners are to work with and critically evaluate data, as students or as citizens, they must be able to interrogate how and why the data came to exist in the first place.

This paper contributes a framework for understanding data production that foregrounds (a) data as produced and (b) student agency in their production. We use this framework to analyze activity of and interviews with high school science students using sensors and software to produce data in inquiry projects. We present the case of one student who came to align her use of unfamiliar sensors and software with her own personally relevant goals. We discuss how those goals emerged through use of the tools, recognition of their affordances as well as constraints on their use, and through resistance to those purposes from the material world.

THEORETICAL BACKGROUND

Philosophical, historical and ethnographic studies of scientific practice problematize the relation among data, natural phenomena and scientific knowledge (Hacking, 1983; Latour & Woolgar, 1986; Pickering, 1995). Data do not exist as prima facie evidence of natural phenomena, to be passively *collected* independent of scientists' intent to and means of collecting them. Rather they are *produced* as artifacts of activity in which humans, technologies and the material world each play an active role.

Data and the Dance of Agencies

Pickering's (1995) analysis of scientific practice focuses on the work of the early particle physicist, Donald Glaser. Glaser's goal was to capture traces of cosmic particles using a device called a *bubble chamber*. As particles passed through the chamber, they would create tracks of bubbles in the chamber liquid and trigger an imaging system. Glaser's early chamber designs failed—it was only after re-thinking the particle's interaction with the triggering mechanisms, and revising the chamber design that Glaser would successfully capture particle tracks.

Pickering theorized the particle's apparent resistance to capture as an exertion of ***material agency***. In response to these material resistances, Glaser exerted his own ***conceptual agency:*** "mangling" conceptual models and practices, designing new forms of devices and new accounts of their functioning, and shifting his goals. Pickering likened this dialectic of material resistance and scientists' accommodation to a "dance of agencies" (Pickering, 1995). Importantly, the agency of the scientist is not located only in their mangling of concepts and practices, but in their directing this dance in pursuit of broader goals and purposes (Pickering, 1995).

Scientific understanding and technologies co-evolve as scientists re-engineer devices to better "capture" the phenomena of interest (Ihde, 1979; Pickering, 1995; Tala, 2009). These technologies come to form the material culture of science, embodying both the intent to pursue particular lines of inquiry, as well as knowledge about how to do it (Baird, 2004; Pickering, 1995). The tools thus embody a form of ***disciplinary agency*** with which scientists can align in pursuit of their goals.

On one hand, the acceptable uses of technologies for producing, transforming and representing data imbue data with persuasive, rhetorical power within a community of practice—enough so that scientific technologies have been likened to "rhetorical devices" (Latour, 1987; Tala, 2015). On the other hand, the discipline also places constraints on acceptable uses of the tools it provides, and which may be in tension with scientists' goals, purposes and particular orientations to practice. Pickering (1995) describes how Glaser's practice was motivated by a desire to align new work with his existing areas of expertise, and with an orientation toward, or preference for, "small science"—science that could take place in a "peaceful environment," responsive to the individual scientist and local funding mechanisms. Situated within the rise of "big science," Glaser's *goals* had psychological, social and moral dimensions. They also had direct consequences for his data: though his laboratory-scale chamber could produce data sufficient to do "interesting physics," the data required significant manipulation to make them useful to questions of elementary particle physics (Pickering, 1995, p. 43).

Below, we outline three implications of the "dance of agencies" for the nature of data.

Data are Part Fact, Part Artifact

Data always bear traces of their technical and material origins (Ihde, 1979; Kroes, 2003). This is particularly apparent in modern scientific practice, as the objects of inquiry (e.g. trends in global temperature, plate tectonics, bacterial resistance) are inaccessible to unaided perception due to their abstractness, or temporal or spatial scale. To observe such phenomena requires the careful design, development and orchestration of specialized technical instruments (Hacking, 1983; Pickering, 1995). For instance, in the early science of photography, on any developed photographic plate might appear bright areas due to stray light, dark spots where chemicals sat for too long, or abnormalities due to aberrations of the camera lenses. These photographic data were a convolution of fact and artifact—a blend of observed phenomena with sensitivities of the device (Geimer, 2000). Scientific work requires teasing apart the features of data attributable to an underlying phenomenon from those that are artifacts of its manner of production (Geimer, 2000; Harré, 2003; Kroes, 2003). Understanding the peculiarities of the material context of data production can be critical. For example, controversy emerged when scientists revised claims about trends in global temperatures after realizing that ship water intake systems had artificially inflated measured temperatures (Karl et al., 2015). Here, the shift was not due to a new analytical technique or theory, but to a refined understanding of the contexts in which the data were produced.

Data are Theory-laden

The tools used to produce and transform data are themselves neither natural nor neutral, but are "theory-laden" technologies with an "in-built intention" to reduce complex phenomena to useful forms of data (Ihde, 1979; Tala, 2009). A temperature sensor, for example, creates a stable and reliable transformation between a complex physical behavior and a quantitative value interpretable as a temperature. Such tools embody disciplinary, historical knowledge both about the nature of phenomena and how they can be captured as data. The data not only function as evidence, but also as a social contribution to establishing and maintaining lines of inquiry (Latour, 1987). Consideration of the "phenomenological profile" (Ihde, 1979) of technologies may be of particular importance for learners' abilities to critically evaluate data, as even seemingly low-tech survey instruments have a consequential profile (e.g. measurement of gender as binary or ethnicity as mutually exclusive categories).

Data are Created in Pursuit of Human Goals and Purposes

Data exist because someone intended to create them. Phenomena worthy of study, as well as what constitutes valuable and persuasive data, are negotiated not only within disciplinary, but also economic and political boundaries (Latour, 1987; Machamer & Osbeck, 2004; Pinch & Bijker, 1984). This is especially evident in the case of sensor networks for monitoring environmental toxins. Such networks are funded, installed and maintained by political or industrial powers, and may exclude particular locations or kinds of impacts (Ottinger, 2010).

At the same time, data also embody smaller-scale personal, psychological and professional purposes (Machamer & Osbeck, 2004; Pickering, 1995). For example, monitoring a stream or species can involve remote sensors, specialized modeling software or extended fieldwork. As such, an ecological study might offer the chance to tinker with hardware, gain experience with new methods or to spend hours (or weeks) outdoors. A scientist's decisions in pursuing particular lines of inquiry might be responsive to topics being investigated, but also to the financial resources available, to professional or family obligations or to preferences for working in particular sociotechnical configurations (Azevedo, 2013).

Data and Agencies in School Science

Sensor-based science laboratories should be a natural context in which students come to understand the nature of data. However, school science laboratories rarely resemble the dance of material, disciplinary and human agencies characteristic of science practice.

Seldom are students given the opportunity to join or lead this dance, as they rarely have the *epistemic agency to* shape knowledge building work in their classroom (Miller, Manz, Russ, Stroupe, & Berland, 2018). Student epistemic agency is often in tension with (or, asked to *surrender* to) the agency of the discipline—a tension reflected in the dual goals of (1) building on students' prior knowledge by engaging them in scientific sense-making and figuring-it-out, and (2) arriving at normative disciplinary conceptual understandings via specified scientific practices (Miller et al., 2018; Stroupe, 2015).

School science investigations, in particular, are often characterized by a strong in-built disciplinary agency (Greeno, 2006), in which educators carefully prescribe acceptable uses of scientific tools and methods to ensure that the data produced are consistent and amenable to a target conceptual interpretation (Lehrer et al., 2008). Below we outline three problems that this poses, both for student agency and their understandings of data.

Problem 1: The material world plays no role in the emergence of practice, or technical and conceptual knowledge.

To make data useable for conceptual learning goals, educators intentionally *limit* the possibility that students encounter material resistances. Whether by giving detailed instructions, or time to plan experiments in detail in advance to later be "carried out," we often attempt to prevent anything that might resist a straightforward conceptual interpretation of the data. As a result, the material context of an investigation is rarely called into question, and revision of an experiment or design is rarely motivated. Yet as Stroupe (2015) argues, scientific practice should evolve with a growing awareness of the material context of an investigation. An investigator may learn that some detail of their experimental setup matters to their data, or become aware of the need to control for a confounding variable. Just as resistances prompt the "mangling" of conceptual models and practices, they can also spur technological innovation and new interpretive accounts of how data was produced (Pickering, 1995).

Manz (2015) specifically argues for designing material resistances into science instruction, as contending with uncertainties related to the material context of investigation creates opportunities for students to re-negotiate local epistemic practices, including how to collect, represent and evaluate data. Encountering resistances should thus provide opportunities for students to act with agency. By limiting encounters with material resistances, school science investigations also limit opportunities for students to tease apart *fact* and *artifact*, to see data as contingent on material and technological origins.

Problem 2: Disciplinary tools and methods are prescribed in advance of a recognized need to use them.

Often, the tools of science are not resources for students' goal-directed activity, but are themselves the goal of the activity: students should use such tools in particular, disciplinarily appropriate ways (Russ & Berland, 2018). For instance, we might ask students to perform a *control of variables strategy*, though this is a solution to a problem students may not have identified a need to solve. When educators prescribe particular tools or practices for producing data—whether or not they are relevant to students' own objectives—the "slippage" that results poses problems for students' agency and sustained sense-making (Manz, 2015; Russ & Berland, 2018).

In contrast, when students engage in the goal-oriented activity of figuring-it-out, they can adopt disciplinary tools as resources and become progressively aware of the constraints on their use that are imposed by the discipline (Russ & Berland, 2018). This approach is rarely extended to students' laboratory experiences, despite its strong theoretical resemblance to the *dance of agencies* in science practice.

Problem 3: Purposes for producing data are fixed and unnecessarily narrow.

Educators' purposes are often to produce data that supports particular *conceptual* learning goals (e.g. to reproduce a particular phenomenon in the classroom, or to ask a question of a particular conceptual model). Data are produced primarily for use as evidence in evaluating models or scientific claims (Duncan, Chinn, & Barzilai, 2018).

Little attention has been given to students' emergent purposes for producing data, despite accounts of scientific practice that highlight emergence and evolution of scientists' goals (Pickering, 1995). Just as scientists' goals and purposes are not narrowly conceptual, but to orient their work toward a "desired future state of culture" (Pickering, 1995), we must broaden the sense of *epistemic* *agency* in science education to include students' engagement with aspects of activity relevant to prior knowledge, identities, areas of expertise and desired learning trajectories (Azevedo, 2013; Calabrese Barton & Tan, 2010; Holland, Lachicotte, Skinner, & Cain, 2001).

STUDENTS AS DATA PRODUCERS

Both to support and investigate students' developing agency in data production, we present a framework that foregrounds (1) data as *produced*, and (2) student agency in data production (Table 1).

In our research, we use Data Production as a "sensitizing" framework (Stevens, O'Connor, Garrison, Jocuns, & Amos, 2008) to guide both design and analysis. In part due to a theoretical focus on *emergence* and in light of the broad variety of ways in which students act with agency in data production, our framework does not yet allow us to generate specific predictions or testable hypotheses about the relation between student agency and their understanding of data. Rather, we use the framework to direct attention toward opportunities for students both to deepen understandings of data, and to develop, perceive and act with agency in data production.

METHODS

Context of Research

The InSPECT project aims to integrate computational thinking into high school Biology classes through programming for data acquisition and control in the context of laboratory investigations. To that end, we are (1) building low-cost hardware kits including sensors and actuators (Figure 1), (2) developing an online programming environment, Dataflow, that interfaces with these kits (Figure 2) and (3) designing laboratory activities in which students learn both about Biology and about the tools that produce, transform and store data. The labs use everyday materials that are familiar and easy to modify, to support student agency and bring "transparency and aesthetics back to scientific investigation." (Resnick, Berg, & Eisenberg, 2000).

Our research team began working with Mr. B at a summer teacher workshop attended by teachers from Northern California who learned about the hardware kits and develop ways they might be used in their classrooms. Mr. B. is a mid-career science teacher with a background in physics and engineering. In Spring 2018, he taught three classes of 9th grade Integrated Science at an urban, tuition-free, public charter school. Of its students, 83% quality for Free or Reduced Price lunch, 78% are English Language Learners, and 90% of graduates would be the first in their

TABLE 1
A Data Production Framework

Characteristic of data	*Implications for pedagogy and design*
Data are produced in interaction between instruments and the material world.	**Students engage with the material world**, • responding to material resistances • designing and re-designing tools • constructing interpretive accounts of data that include material setups and designs, in addition to scientific concepts and phenomena • drawing on practical knowledge of materials and tools
Data is produced by tools and practices that embody disciplinary knowledge and values.	**Students use disciplinary tools and practices**, • as resources in a goal-directed activity • with recognition and understanding of their affordances for particular purposes • encountering constraints on their appropriate use over time
Data is produced for human purposes.	**Students' goals and purposes in data production** • need not be conceptual, but can include other purposes such as being playful, technical, design-oriented, personally relevant or others • shift and emerge in activity, entangled with the use of disciplinary tools, concepts and practices • when resisted, prompt re-structuring of the activity or development of new purposes

families to go to college. Mr. B.'s goals for using InSPECT kits and activities with his students were that his students develop their own questions and conceptual models through designing and collecting data with their own experimental setups.

Our research in Mr. B's class spanned two weeks. The first two days were an introduction to hardware and software. As part of these initial activities, students used the sensors and actuators to "build something useful"—for example, a system that turns on a heater when the temperature drops below a set threshold. We hypothesized that a short time to tinker might be an engaging introduction to the

FIGURE 1 An InSPECT hardware kit including Raspberry Pi computer, sensors and a relay.

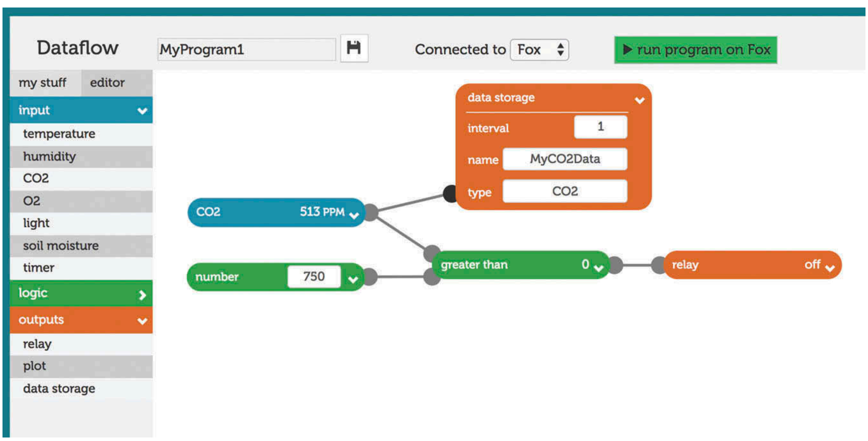

FIGURE 2 A screenshot of the dataflow programming environment.

tools, particularly for students with interests in computers and programming, and which might support increased agency later in investigations.

In the third session, Mr. B showed each class a setup for collecting data about CO2 levels in a closed container with spinach leaves. He asked students to build their own setups and create two 10-minute duration datasets: one with spinach under a bright lamp, and another in the dark. In the next class session, Mr. B led a discussion about the students' data in which they developed a "whole class conceptual model"—two graphs on the board conveying consensus about what the groups' datasets looked like. Led by Mr. B, the class discussed what might have caused the CO2 levels to decrease in the light and increase in the dark. Students contributed a variety of ideas: that the plants might be decomposing in the dark, re-releasing CO2 they had just taken in, or that plants are *always* releasing CO2 and are taking it faster than it is released only when under a bright light. Mr. B. did not direct the students to the accepted scientific interpretation, but rather left open some uncertainty in the discussion, and had students brainstorm a list of variables that might have affected their results (e.g. light intensity, temperature).

Students then had the next week to conduct independent investigations using the sensor kit as well as additional materials and equipment from the school's workshop. Students that did not have a project in mind were invited to investigate one of the variables listed in the previous class. Throughout the independent projects, Mr. B visited groups of students to offer guidance and support. At the end of this second week, students presented their data and project summaries.

Data Sources

Researchers took field notes of overall classroom activity and recorded video of focal groups of consented students, as well as photographs of student work and lab notebooks. Before lab activities, students took background surveys about their experiences with and interest in science and technology. Following the students' inquiry projects, three interviewers conducted semi-structured interviews, lasting an average of 45 minutes, with 18 students about their projects and experiences. The interview protocol included prompts about what the students did and how they felt in class, with the interviewer asking follow-up questions, for instance about what they thought was happening, or when they felt *surprise* or *frustration*.

Data Analysis and Case Selection

After transcribing interviews and organizing field notes from the classroom sessions, research team members read through all interviews, and began to code using a scheme derived from our Data Production framework (Table 2).

TABLE 2
Coding for the Development of Student Agency Interacting with Data

	Codes and code descriptions
Material	• *Resistance*: narratives about problems they encountered: things breaking, things failing, things surprising them, unexpected or surprising occurrences • *Description*: students describe or note aspects of the material setting, such as the setup they built, sensor placement • *Interpretive account*: students give interpretive accounts of their data that incorporate aspects of the material setting or setup
Disciplinary	• *Prior-use*: students talk generally about sensors and computational tools, outside of class or in other contexts • *Methods*: students describe methods, procedures, instructions, practices or routines for producing data • *Tool behavior*: students discuss sensors or software, including sensor limitations, fluctuations, noise, non-ideal sensor behavior • *Science*: students talk about science or science class, apart from these experiences
Human	• *Intentions (mine)*: students express goals/intentions/purposes/efforts to produce data or particular kinds of data • *Intentions (others')*: students talk about scientists' or other students' goals/intentions/purposes/efforts to create data • *actions*—students relate features of data to their own decisions/actions/choices • *Self*—students describe their own state (e.g. surprised, confused)

We also included a separate code for descriptions or evaluations of data (e.g. accurate, weird) that could, depending on context, point as much to a purpose to produce a particular sort of data, as to a resistance to that purpose (e.g. "the data was *messed up*"). The goals of our coding at this stage were to draw attention to particular experiences, or moments within each students' narratives, in which they made connections between their purposes, tool use and the materiality of their investigations, and that may have been important to either their development of agency or understanding of data. The initial round of coding flagged what students each made relevant in their accounts and identified moments that may have been important to their own trajectories, yet due to a variability across student projects and types of problems encountered did not lead to robust themes across the interviews. We began to sketch out some of the ways students engaged in data production, and oriented toward different aspects of the activities. We characterized these aspects or orientations as *conceptual, playful, design, technical* (*programming/technology), data* or *personal*.

We drew observational data (field notes, audio, video and student artifacts) into our analysis, while focusing on a subset of seven students who represented these differing ways of approaching the activities. These students pursued a variety of projects (e.g. designing a container to burn spinach and measure CO2 released, or to create *flat data* by gradually adjusting the light intensity), yet each encountered some form of resistance that prompted a rethinking of their approach, or of their data. At this point in our analysis, we sought a detailed understanding of how those resistances may have been opportunities for students to act with agency, as well as how they may have influenced students' understandings of their data. In order to provide an understanding of student experiences in these activities, and to illustrate an examination of learning through a lens of data production, we focused our analysis to a single in-depth case, Hope.

The case we present is a case of an orientation we had characterized as *personal*, involving (a) connection of data production activities to prior learner's interests and experiences across settings, and (b) deeper reflection on one's learning. Hope first came to our attention when we learned during classroom observations that she enjoyed art, and did not like programming. On the background survey, she indicated that she never used electronics kits or programmed computers, though enjoyed learning about technology. Further, in her interview, Hope articulated unique insights about the technologies in relation to her own learning trajectory. While she described herself in her interview as "bad at technology," she came to use the sensor kit to pursue a project that stood out not only in its personal relevance, but also the extent to which she related her findings to multiple domains of prior knowledge.

Results: Hope's Agency in Data Production

Learning the Basics

Though initially reluctant to engage with unfamiliar and "confusing" technologies, during the first class sessions Hope became interested in the data the sensor kit was generating:

> It collected a lot of data really fast and it was crazy … I knew, like, technology is going to get more advanced and so I'm like, you need to learn the basics because I know that I don't want to grow up and I'm 30 and I'm like, what is this?

Using sensors to measure carbon dioxide levels during the initial investigations with spinach leaves, Hope carefully reproduced the setup demonstrated by Mr. B, including a cardboard box that would block out surrounding light. She analogized her work to produce data to a "recipe," saying, "you can't make a good chocolate cake unless you have a recipe … if you didn't have the CO2 levels go up and down, something must have went wrong."

She noted that this use of technology in science class was surprising to her, and contrasted these activities with her prior experiences in school science, which she described as feeling "old," or "in the past," producing knowledge or replicating experiences that everyone else already had:

> Because you don't really see people play with sensors or all that, and it was like a science class, you know, science like, 'oh, do this volcano.' 'Okay, okay, yeah.' So it wasn't really any much computer in it. And I'm really bad with technology. I'm like, even with my phone: I know nothing about it … so I felt like this was interesting, was different. It was kind of this generation thing. It wasn't like, 'do this, just put soda and mint together.' 'Cause honestly, like no … It just made us feel like you're present. Like right now, not in the past. Because some things you just feel like are so old, everybody knows it. Maybe put dye in the water, let the plants grow and like you kind of know what's going to go, like—it's just so basic and now it just felt like it was more like our level, our generation.

Asked to elaborate, Hope spoke about the realization that the sensors and software had a vast amount of knowledge embedded in them:

> Everything that you're using had so much information. Like the sensor, like it's small but it's impacted with so much … like you can't just put your finger in [the container] like, 'oh I know how much CO2 is coming out.' Like no, people had to take time, knowledge, just to make this together.

With Hope's insight into the sensors as powerful, designed technologies came the feeling that she herself knew very little:

> I feel like I knew so little about the world … like something so simple I thought I knew so much about, I just knew nothing about it. And just playing with technology just made it like I was like, 'come on, you need to get yourself in science more,' like you need to know at least a little.

Hope would again feel she knew little about the world after her initial investigation with spinach leaves. Hope thought that the leaves must have been dying in the dark, releasing the CO2 they took in while in the light (see Figure 3). Coming to see that plant death might not be as simple as she had thought, Hope decided that in her independent investigation she would need to "learn the basics":

> Before I did this spinach at all, like before I was just thinking of like I want it big, I want to do something crazy. And after [I] did the spinach thing and I was like … let me learn the basics. And then that kinda just changed what I was thinking, I was like, I need to go back to small.

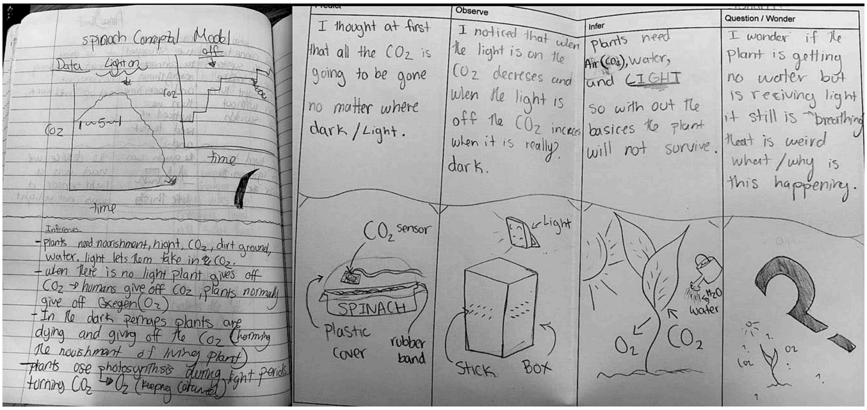

FIGURE 3 (a) Hope's lab notebook, with sketches of the graphs from the Spinach Lab. (b) Hope's worksheet with columns for predict, observe, infer, question/wonder.

Hope became curious about what she saw as a "really basic" question: "just to see when a plant dies … how would we know what's dead and why is it doing it?" Hope gardened at home, and saw her project as a chance to learn more about why her plants often died, when for others they thrived. She investigated plant death by comparing CO2 data for lettuce that she kept either in a closed or open container overnight—two conditions she explained as parallel to *refrigerated* and *unrefrigerated* lettuce.

Interpretation: Connecting Disciplinary Tools and Emergent Purposes

Just as Pickering (1995) was careful to note that scientists' *goals* and *purposes* aren't independent of, or prior to, the tools that implement them, so we see that in Hope's case, her goals emerged alongside her use of technologies. Recognizing the in-built knowledge of the sensors and their affordances for "see[ing] CO2," Hope was able to align with the agency of the discipline in forming and pursuing her own emergent goals. She did not have a singular or pre-determined goal to develop conceptual knowledge about plants, and likewise we see Hope's agency not just in her development of that knowledge. Rather, we locate her developing agency in her formation and pursuit of several, broader goals—goals partially *informed* by her prior knowledge about plant death, but also by her interests in gardening, desired learning trajectories to learn a little about science and technology, an emergent orientation to investigate "small" or "basic" questions, and a preference for doing work that felt new and "in the present." And rather than constrained, in this case Hope's own agency was extended by the affordances of disciplinary tools. Her subsequent personal and purposeful use of the tools represents an alignment with the agency of the discipline.

Getting "The Data I Need"

In conducting her independent investigation, Hope created two datasets by duplicating two setups, one for each experimental condition. She carefully worked around the sensitivities of the sensors, attempting to isolate them from their surroundings.

> Because I didn't want [the sensor] to touch plants and I didn't want to touch the plastic. I don't want anything to interfere with it and mess up our data … like it's going to, if you put your hand on it, it's going to affect it, if you breathe on it it's going to affect it. So I'm like, I really wanted to keep it isolated but still I wanted to get the results I needed.

Asked whether her data might have turned out differently if she had constructed her setups differently, Hope expressed her concern that other students' breathing in the classroom could have interfered with her results:

> I used a box and newspaper—obviously that's not gonna work … sometimes like if you want to have absolute data you really want, like I would say, I keep thinking of having this empty room just for like the plants. Like I don't want anything to interfere. I don't want some people breathing oxygen and breathing out CO2 and affecting the plant's data and making gas all like—no, when you always think about scientists, you always think this white room. They just cover it up and then the windows—people just seeing it … And so I felt like I'm pretty sure I've got some people's breath into those data, and I'm like, I'm not trying to have that, um, but I don't think it's really going to affect it a lot because I kind of covered it with plastic wrap even though it could go under, and I know it can. Um, but I really try to keep myself away from the data … I'm like, this is the data I need.

Interpretation: Encountering Resistances to Purpose

Hope shifted from seeing her classroom activity as following a "recipe" for data with features known in advance, to an activity in which the sensors, the material world, and Hope herself were active agents. Though the sensors were useable to her to investigate plant death, Hope became aware that they came with sensitivities to the material context, and corresponding constraints on their use—that they had to be isolated from the material environment to produce "absolute" data. Hope's data were then part fact, part artifact—perhaps attributable to the lettuce dying in her container, or to her classmates' breath. Evaluation was no longer dependent on how well she had followed a recipe, but on the degree to which she had managed to isolate the sensors from the material world. Her narrative of her data production took on the intentional structure characteristic of science practice (Pickering, 1995): foregrounding herself and her own intentions, she tried to avoid touching or breathing on the sensors, "not trying to have" people breathing and affecting the data, so she could get the results she "needed." While admittedly not isolating things to the same degree as scientists with their "white rooms," she saw herself as analogously engaged in combatting the resistance of the material world.

Death of a Lettuce

When she began analyzing her data, Hope was unable to see any apparent differences across her two conditions. This was an artifact of the way that the Dataflow software auto-scales the view of any single graph (see Figure 4).

Hope showed the graphs to the researcher/first author, who then offered to show Hope how to import the data into another data analysis tool, CODAP, so that they could be viewed simultaneously. Though Hope found value in the way Dataflow initially made her graphs appear similar (arguing that she must have done a good job setting up her two conditions similarly), by using new tools to graph the datasets together, she was able to see that they were quite different:

> It was so different. It was crazy to feel like when you look at them they look exactly, the curve that is exactly the same … [but] when you look close to the numbers, there were different numbers … I didn't even pay attention to the numbers until I put them together.

Hope found that the CO2 in her closed condition rose more quickly, which she attributed to the lettuce dying more as it was closed off from things it needed to live. Seeing the power of this new tool to control how her data looked, Hope pulled up a chair for another research team member and asked for guidance on how to label and change the colors of each graph, and select the points at the beginning of her graph, where difference were most apparent (Figure 5).

Hope then said she felt proud of her data, "because it gave me kind of an answer, like, gave me an answer to my question. That's not really usual. [Usually] you go search it up. And so like I'm like, 'wow, I figured something [out] by myself.'" Over the course of her investigations, Hope began to think of *plant death* not as plants just "turning brown," but ceasing photosynthesis. She connected this to food that may be producing CO2 in her refrigerator ("every time I open it … a bunch of CO2 comes out"), to ideas about "the circle of life" ("I always felt like trees … like get in CO2, and it just disappears … but nothing is disappearing in this world, it's always coming back") and to global warming

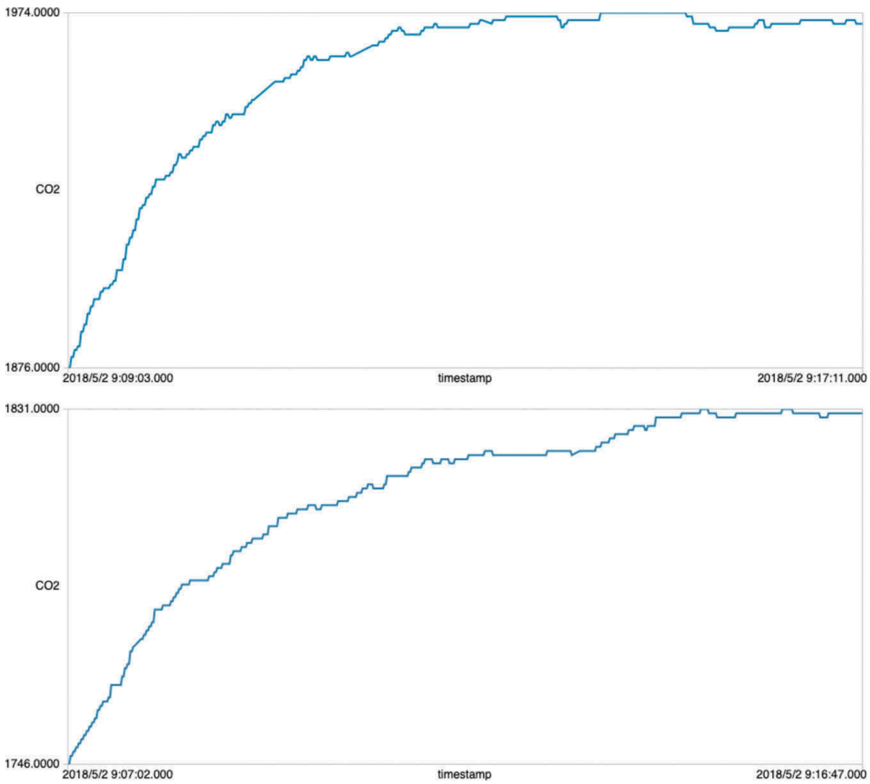

FIGURE 4 Hope's "box" and "open" graphs, as shown by the Dataflow interface.

("there's too much carbon dioxide out there, it's not getting stored? Is that the thing that's happening? I'm pretty sure that is.")

After Hope had related these insights, we asked if she ever had a "hmm … " moment, when she felt that something might not be right:

> Like I felt like I knew everything, like I *got* this. And then when I was done with the project [Mr. B] kept asking me all this question and [I] was like, 'I wasn't even thinking this … like, what?' And it made me feel like I could've done better. Like, you always start reckless and small, and then after that you keep making progress to make it better and more efficient.

Beginning to question if she really knew that the lettuce in her *closed* condition was really *more dead*, Hope thought about other approaches she could have tried: creating additional experimental conditions exposed to light so she might

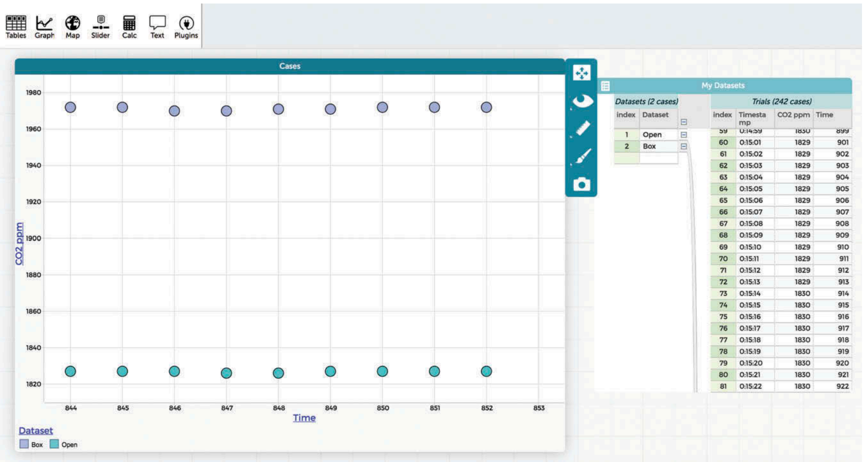

FIGURE 5 Hope's view of her graph and data table within CODAP, recreated from classroom photographs and dataflow server logs.

be able to see a more "complete difference," or including an oxygen sensor to see if the lettuce was still producing oxygen even as the CO2 rose. At the end of the work, Hope shared her graphs with her classmates. At Mr. B,'s suggestion to "call it something cool," instead of *effects of refrigeration on plant CO2*, she titled her presentation *Death of a Lettuce*.

Interpretation: Connecting Resistances to Purposeful Use of Disciplinary Tools

That the Dataflow software did not initially reveal differences between Hope's datasets posed a *resistance* to her goal of understanding such differences. This resistance became an opportunity for Hope to see the affordances and limitations of disciplinary tools both for sense-making and for communication, and to see her data in new ways. At the same time, a new purpose emerged through her use of the new software: to *emphasize* the differences between her datasets. She pursued that purpose in ways that were at once personal and esthetic (making things look like she wanted, including picking particular colors), but also critical to her inquiry (revealing differences in her two conditions, and allowing her to communicate those results most clearly). Once again, Hope's agency was extended through her use of disciplinary tools as resources for her own emergent, goal-driven activity. In this case, it was their affordances for responding to a particular resistance she had encountered that made them relevant.

While we view Mr. B's suggestion to "call it something cool" as an invitation into broader ways of engaging in and talking about science, his questioning of her findings represented a form of push-back by the agency of the discipline: an encounter with the normative constraints on a satisfactory scientific explanation of empirical results. Instead of being discouraged, Hope saw this push-back as a natural part of a scientific practice that may have started "reckless and small," but would grow in time.

DISCUSSION

Student Agency in Data Production

Hope's case provides an account of student agency in engaging in the "dance" of human, material and disciplinary agencies of science practice. Scientists' agency is not solely in their development of scientific concepts and practices, but in their pursuit of broader goals (Pickering, 1995). Likewise, student goals for engaging in data production need not center around particular conceptual models or questions. To summarize Hope's goal as to *measure the effects of refrigeration on plant output of CO2* would seem narrowly scientific, but would be an incomplete and uncharitable view of her work. Rather, Hope had multiple, intersecting goals: to pursue questions relevant to her life experiences, to not be left behind by technological advances, to ask questions she felt were fundamental or "basic," and to work in ways that felt "in the present." While the strong disciplinary agency that characterizes school science investigations overly constrains student agency in several ways (Greeno, 2006; Miller et al., 2018; Russ & Berland, 2018), in particular it may be in tension with students' agency to pursue particular lines of practice (Azevedo, 2013).

Relationship to the Discipline

However, while disciplinary agency might *limit* student agency, we also saw ways in which it can *extend* student agency. As the usefulness of disciplinary tools arose through Hope's activity, so too did new, previously unformed purposes. We also note that even encounters with the constraints of the discipline played a role in the development of Hope's agency, as responding to push-back led her to envision revisions to her investigation. That is, becoming attuned to the constraints of the discipline can lead to an awareness of ways to work around them. Thus, rather than being defined or developing separate from the agency of the discipline, student agency develops as they encounter both constraints of and affordances of the tools provided by the discipline. To develop *agency in science* is not only to shape classroom knowledge-building practices, or to engage with material agency in the way scientists do (Manz, 2015; Miller et al., 2018), but to also become progressively and reciprocally engaged with the agency of the discipline.

Designing for Resistances

Future research should consider ways to invite students to join in and lead a dance of human, material and disciplinary agencies. While disciplinary agency has historically been overly strong in school science, Hope's case suggests that making room for material agency in science classrooms may be one way to shift some power from the discipline to the student. That is, *material resistances* may provide opportunities for students to act with and perceive their own agency (Manz, 2015). In particular, we saw in Hope's case that *resistances to purpose* may be opportune moments for students to develop an awareness of the affordances and constraints of the tools of the discipline for particular purposes (Russ & Berland, 2018). We agree with Manz (2015) that material resistances should be designed into science instruction, and in subsequent iterations of the InSPECT activities we have designed in active roles both for the sensors and the material world. For instance, in a laboratory in which students measure carbon dioxide levels in their breath before and after exercise, students are now asked to use available materials to design a way to capture these CO2 levels when a direct measurement of undiluted breath is outside of the range of the available sensors. These activities open up space for engaging in design and iteration and allow for divergent data production techniques. We conjecture that this may broaden opportunities for students to take an agentive role in data production, and perceive ownership of the data and their practice.

Understanding Data

The argument we have made is that agency in data production is likely necessary for students to develop an understanding of data as *produced*: through interaction between instruments and the material world, using technology that embody knowledge and values, in pursuit of human goals and purposes. We saw the beginnings of such an understanding in Hope, evidenced by shifts in how she talked about her data, and herself as actively combatting interference from the material world. To design science labs to support and build on these understandings, in future work it will be necessary to make more explicit connections between data in the science classroom, to data in science, society, and our everyday lives. Across many encounters with data, we must encourage reflection on the material context, the tools that produce data (including the knowledge and values embedded in them), and the human purposes for which data are created.

CONCLUSION

Our title, “From data collectors to data producers,” is intended to convey the ways in which data should be thought of as *produced*, and also the shifts we believe are necessary in science laboratory investigations for students to act as *agents* in data production. We propose that educators stop talking about *data collection*, as if data were sand dollars waiting on the beach. Instead, students and scientists engage in *data production*, building technologies to create and capture traces of natural phenomena relevant to their particular purposes. Though slight, we believe that this shift will help combat idealized and reductive ways of conceptualizing scientific practice that misconstrue the nature of data, and instead support an understanding of data that will be necessary for young people’s ability to participate productively and critically in science and society.

FUNDING

This work was supported by the Division of Research on Learning in Formal and Informal Settings DRL-1640054.

ORCID

Lisa Hardy http://orcid.org/0000-0001-6185-8893
Sherry Hsi http://orcid.org/0000-0001-8356-3279

REFERENCES

Azevedo, F. S. (2013). The tailored practice of hobbies and its implication for the design of interest-driven learning environments. *Journal of the Learning Sciences*, *22*(3), 462–510. doi:10.1080/10508406.2012.730082

Baird, D. (2004). *Thing Knowledge: A philosophy of scientific instruments*. Berkeley: University of California Press. Retrieved from https://www.ucpress.edu/book/9780520232495/thing-knowledge

Calabrese Barton, A., & Tan, E. (2010). We be burnin: Agency, identity and learning in a green energy program. *Journal of the Learning Sciences*, *19*(2), 187–229. doi:10.1080/10508400903530044

Duncan, R. G., Chinn, C. A., & Barzilai, S. (2018). Grasp of evidence: Problematizing and expanding the next generation science standards’ conceptualization of evidence. *Journal of Research in Science Teaching*, *55*(7), 907–937. doi:10.1002/tea.21468

Geimer, P. (2000). Noise or Nature? Photography of the Invisible around 1900. In H. Nowotny & M. Weiss (Eds.), *Shifting boundaries of the real: Making the invisible visible* (pp. 51–63). Salt Lake City, UT: Robert Smithson, Spiral Jetty.

Greeno, J. G. (2006). Authoritative, accountable positioning and connected, general knowing: Progressive themes in understanding transfer. *Journal of the Learning Sciences*, *15*(4), 537–547. doi:10.1207/s15327809jls1504_4

Hacking, I. (1983). *Representing and intervening : Introductory topics in the philosophy of natural science*. Cambridge: Cambridge University Press.

Harré, R. (2003). The materiality of instruments in a metaphysics for experiments. In H. Radder (Ed.), *The Philosophy of scientific experimentation* (pp. 19–38). Pittsburgh, PA: University of Pittsburgh Press.

Holland, D. C., Lachicotte, W. S. J., Skinner, D., & Cain, C. (2001). *Identity and agency in cultural worlds*. Harvard University Press. Retrieved from http://www.hup.harvard.edu/catalog.php?isbn=9780674005624

Ihde, D. (1979). Heidegger's philosophy of technology. In *Technics and Praxis. Boston studies in the philosophy of science* (Vol. 24, pp. 103–129). Dordrecht: Springer. doi:10.1007/978-94-009-9900-8_9

Karl, T. R., Arguez, A., Huang, B., Lawrimore, J. H., McMahon, J. R., Menne, M. J., … Zhang, H.-M. (2015). Possible artifacts of data biases in the recent global surface warming hiatus. *Science*, *348*(6242), 1469–1472. doi:10.1126/science.aaa5632

Kroes, P. (2003). Physics, experiments, and the concept of nature. In H. Radder (Ed.), *The Philosophy of scientific experimentation* (pp. 68–86). Pittsburgh, PA: University of Pittsburgh Press.

Latour, B. (1987). *Science in action : How to follow scientists and engineers through society*. Harvard University Press. Retrieved from http://www.hup.harvard.edu/catalog.php?isbn=9780674792913

Latour, B., & Woolgar, S. (1986). *Laboratory life: The construction of scientific facts*. Princeton, NJ: Princeton University Press. Retrieved from https://press.princeton.edu/titles/2417.html

Lehrer, R., Schauble, L., & Lucas, D. (2008). Supporting development of the epistemology of inquiry. *Cognitive Development*, *23*(4), 512–529. doi:10.1016/j.cogdev.2008.09.001

Machamer, P., & Osbeck, L. (2004). The social in the epistemic. In P. Machamer & G. Wolters (Eds.), *Science values and objectivity* (pp. 78–89). University of Pittsburgh Press. doi:10.2307/j.ctt5vkg7t.8

Manz, E. (2015). Resistance and the development of scientific practice: Designing the mangle into science instruction. *Cognition and Instruction*, *33*(2), 89–124. doi:10.1080/07370008.2014.1000490

Miller, E., Manz, E., Russ, R., Stroupe, D., & Berland, L. (2018). Addressing the epistemic elephant in the room: Epistemic agency and the next generation science standards. *Journal of Research in Science Teaching*, *55*(7), 1053–1075. doi:10.1002/tea.21459

Ottinger, G. (2010). Buckets of resistance: Standards and the effectiveness of citizen science. *Science, Technology, & Human Values*, *35*(2), 244–270. doi:10.1177/0162243909337121

Pickering, A. (1995). *The mangle of practice: Time, agency, and science*. University of Chicago Press.

Pinch, T. J., & Bijker, W. E. (1984). The social construction of facts and artefacts: Or How the sociology of science and the sociology of technology might benefit each other. *Social Studies of Science*, *14*(3), 399–441. doi:10.1177/030631284014003004

Resnick, M., Berg, R., & Eisenberg, M. (2000). Beyond black boxes: Bringing transparency and aesthetics back to scientific investigation. *Journal of the Learning Sciences*, *9*(1), 7–30. doi:10.1207/s15327809jls0901_3

Russ, R. S., & Berland, L. K. (2018). Invented science: A framework for discussing a persistent problem of practice. *Journal of the Learning Sciences*, 1–23. doi:10.1080/10508406.2018.1517354

Stevens, R., O'Connor, K., Garrison, L., Jocuns, A., & Amos, D. M. (2008, July). Becoming an Engineer: Toward a three dimensional view of engineering learning. *Journal of Engineering Education*, *2008*, 355–368. doi:10.1002/j.2168-9830.2008.tb00984.x

Stroupe, D. (2015). Describing "Science Practice" in learning settings. *Science Education*, *99*(6), 1033–1040. doi:10.1002/sce.21191

Tala, S. (2009). Unified view of science and technology for education: Technoscience and technoscience education. *Science & Education*, *18*(3–4), 275–298. doi:10.1007/s11191-008-9145-7

Tala, S. (2015). *Physics as technoscience - From research labs to educational labs*. University of Helsinki, Report Series in Physics HU-P-D224. Retrieved from http://hdl.handle.net/10138/153055