Cognition and Instruction

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hcgi20

Remembering What Produced the Data: Individual and Social Reconstruction in the Context of a *Quantified Self* Elementary Data and Statistics Unit

Victor R. Lee, Joel Drake, Ryan Cain & Jeffrey Thayne

To cite this article: Victor R. Lee, Joel Drake, Ryan Cain & Jeffrey Thayne (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, DOI: 10.1080/07370008.2021.1936529

To link to this article: https://doi.org/10.1080/07370008.2021.1936529

Published online: 10 Jul 2021.

COGNITION AND INSTRUCTION
https://doi.org/10.1080/07370008.2021.1936529

Remembering What Produced the Data: Individual and Social Reconstruction in the Context of a *Quantified Self* Elementary Data and Statistics Unit

Victor R. Lee[a], Joel Drake[b], Ryan Cain[c], and Jeffrey Thayne[d]

[a]Graduate School of Education, Stanford University, Stanford, CA, USA; [b]Institutional Research & Assessment, Brigham Young University – Idaho, Rexburg, ID, USA; [c]Weber State University, Ogden, UT, USA; [d]Department of Psychology, Brigham Young University – Idaho, Rexburg, ID, USA

ABSTRACT

Given growing interest in K-12 data and data science education, new approaches are needed to help students develop robust understandings of and familiarity with data. The model of the *quantified self*—in which data about one's own activities are collected and made into objects of study—provides inspiration for one such approach. By drawing on what one already knows about their self and their prior experiences, it may be possible to bootstrap students' abilities to interpret and make sense of data. Taking that possibility seriously, this article describes some of the gains observed in students' statistical reasoning following a quantified self, wearables-based elementary statistics unit and provides a theoretical framework drawing from cognitive psychology, embodiment, and situative perspectives to characterize how prior experience is used as a resource in data sense-making when the data are about students' own physical experiences. This framework centralizes and interrogates the work of "remembering" prior experiences and articulates how remembering is involved in interpreting quantified self data. Specifically, the framework emphasizes that remembering in service of data interpretation is a reconstructive act that draws from both general and specific embodied resources and that the work of reconstructive remembering in the classroom is both individual and multi-participant work. To demonstrate measured learning gains and illustrate the framework, written assessment results and descriptive cases of student and teacher discussions about quantified self data from two sixth-grade classes participating in a classroom design experiment are provided. Both a discussion of and recommendations for ethical considerations related to quantified self data in education are also provided.

Introduction

By virtue of their work, those who conduct empirical research develop a heightened familiarity with their data. This includes an awareness of where measurements were taken, what instruments were used, how data were recorded, the decisions that informed these activities, and what data-driven inferences were ultimately made. Studies of professional scientists in action demonstrate this. For example, Hall, Stevens, and Torralba (2002) observed how the "substantive link between fieldwork and laboratory analysis" (p. 184) for research entomologists enabled them to immediately recognize the species represented in charts of their data. Roth (2014) has also documented how fisheries scientists' continual reconsideration of the relationship between their data collection procedures and the resulting representations led the science team toward conceptual change. These realizations took time and collaborative effort for the professional scientists to make. Yet, knowing where and how data were collected aided in the subsequent work of data analysis and interpretation.

CONTACT Victor R. Lee vrlee@stanford.edu Graduate School of Education, Stanford University, 485 Lasuen Mall, Stanford, CA 94305, USA.

In the learning sciences, such accounts of professional data work serve as informative and aspirational targets for the design of inquiry-oriented learning environments. We seek, for instance, to provide students with appropriately scaled versions of these kinds of professional experiences. Yet, it is still uncommon for students to work with and reflect upon data from the beginning of data collection to the end of data analysis. The reasons for this range from inadequate curriculum materials (Cai, Lo, & Watanabe, 2002), underestimation of children's data-related competences (National Research Council, 2007), insufficient teacher preparation with respect to data (Roth, McGinn, & Bowen, 1998), or even the perception of insufficient time during the school day to collect and analyze data. To mitigate these challenges, learning scientists have designed scaffolds to support examination of large, existing data sets (e.g., Quintana et al., 2004), developed technologies to simplify record keeping tasks during students' fieldwork (Lo & Quintana, 2013), and provided access to real-time data sensors that can produce immediate data representations of commonly taught phenomena, such as a cooling curve (Tinker & Krajcik, 2001). In all, these studies show the promise of reducing the overhead of data use in classrooms, and they describe pathways for youth to work with actual data.

In recent years, another strategy for supporting student learning with and about quantitative data has emerged that could greatly reduce the amount of time and effort needed for students to collect data. This new strategy involves having youth use wearable sensor devices to collect numerical records about their own experiences. For instance, youth may wear or carry devices that record GPS coordinates from their journeys around their communities to help them critically reflect on official maps and available options for their own mobility (Taylor, 2017; Taylor & Hall, 2013). Youth could also use visualizations of their own heart rate data to learn more about how their bodies respond to different kinds of physical stress and exercise (Kang et al., 2016; Lee & Dumont, 2010). Researchers designing for such wearable-based inquiry in learning activities like these see this wearable technology approach as advantageous for several reasons, including 1) the embedded sensors allow for data to be passively obtained and 2) youth have subsequent opportunities to use the data to reflect upon where, how, and when these data were collected (Lee & Shapiro, 2019). With such sensors, a youth can then dedicate their mental resources to the tasks immediately at hand (e.g., riding a bike through the city), and the wearable device will do the work of recording data. Later, that youth should be able to apply their subjective experience and memory with the measured activity to the task of making sense of the resulting data representations. This is ultimately a bootstrapping hypothesis in which engagement in a lived experience should support reasoning about data from that experience later.

Yet, prior research on microcomputer-based labs (MBLs) and probeware raises questions about whether this bootstrapping hypothesis is correct. Research with probeware has shown significant student learning gains related to graph comprehension (e.g., Metcalf & Tinker, 2004; Mokros & Tinker, 1987; Redish, Saul, & Steinberg, 1997). However, those learning gains rely on simultaneous, linked visualizations of sensor data. A delay of even 20 s between when data were obtained and when data are visualized undermines data comprehension (Brasell, 1987). Students struggle to remember and track what within the data corresponded to what happened in the data collection event. As such, the hours or days of delay between data collection and subsequent data examination associated with wearable and portable sensors could be problematic. Furthermore, some reports have argued that the data visualizations typically given with this class of wearable sensing devices are inherently unfamiliar and confusing for students to interpret (Ching & Schaefer, 2015).[1] It is unclear if students can easily make sense of representations of automatically obtained data in meaningful ways.

[1]For a specific example of a wearable-based representation that could pose challenges for youth beyond what Ching and Schaefer (2015) provide, consider the data representations used by the Apple Watch. In the Apple Watch, physical activity records are stored as three concentric open circles that are completed at variable rates and are distinguished by color and

Thus, existing research indicates possibilities for wearable-based inquiry while also raising questions about the suitability of contemporary portable and wearable sensor technologies for educational settings. The current article provides evidence toward answering two of these questions: Can such technologies be successfully integrated into a range of classrooms in a manner that produces learning gains? And if so, how are the interpretive challenges associated with sense-making of visualizations of previously collected, person-specific wearable technology data negotiated within the classroom?

Regarding the former question, the current article will report on the quantifiable learning gains of a cohort of sixth grade students. We have previously produced evidence showing that wearable-based inquiry approaches can produce greater learning gains in some school settings than standard instruction of similar duration (Lee, 2019; Lee, Drake, & Thayne, 2016; Lee & Thomas, 2011). The current study continues in that work by looking at learning gains in a still yet further refined version of a wearable-based inquiry unit. It also more deliberately focuses on statistical reasoning as operationalized as part of a learning progression articulated by Lehrer, Kim, Ayers, and Wilson (2014).

Regarding the latter question about interpretive challenges, we provide three descriptive case studies from students in the two classrooms who enacted this version of a wearable-based inquiry unit to illustrate how interpretive work unfolds around the interpretation of data from students' own physical activities. With those cases, we rely upon, refine, and extend prior theory from cognitive psychology related to reconstructed memory and transactive memory. As we mentioned above, delays between when the data were collected and when they were viewed can lead to some information loss in what is recalled over time. However, in the context of the kinds of data that these students obtain, this loss is not inherently problematic. What is recalled may vary in its specificity and precision, but those features that are remembered can be effectively leveraged when students are tasked with thinking about data. Two of the cases we present below were selected because they illustrate this. One shows how students articulate their data interpretations when what they recall is fairly general with respect to remembered embodied experience. The second case shows how students articulate their data interpretations when what they recall is more specific. A third and final case shows both general and specific aspects of what is recalled and how those are mobilized and articulated by more than one participant in a discussion around a canonical data representation. Throughout all the cases, we also show how remembering and interpreting data in classrooms are not solitary activities taken on by individual students. Rather, in the complex environment of a classroom, we can see that thinking through and with these data involves multiple people in the work of remembering.

Theoretical background

Remembering reconsidered

We know from prior research that when probeware and sensor-based data are shown concurrently with the sensed phenomenon, data display (i.e., graph) comprehension tends to improve (Redish et al., 1997). When there was a time delay between the phenomenon and the data display and students must do the work of remembering the events that corresponded with the production of data, comprehension suffers (Brasell, 1987). To explain why this difference existed, Brasell made an appeal to working memory load. In this view, the co-occurring physical event and real-time display features would be stored as a single memory unit. On the other hand, delay requires that memory of the event being quantified and features of a data display each take a slot of

arrows that point in different directions. Moreover, once the circle is closed, any additional physical activity from the day is recorded over the existing circle.

limited working memory. In addition, some additional cognitive effort must be made to maintain those two items in working memory or to be able to correctly recover event-related memories later to do the work of graph comprehension.

This notion of working memory affecting comprehension is sensible in the context of a controlled research study where students are aware that they should attend to the immediate events. However, a presumed benefit of wearable sensors is that data will be obtained in the background of whatever a student is already doing. This means that students can and should engage with measured activities with the same level of care and attention they would if they were *not* carrying a sensor with them. By doing so, youth should have access to memory of those activities in order for those memories to have a bootstrapping effect. At first glance, this expectation appears to conflict with the working memory account. On the one hand, delay of data representation should make things more difficult for memory, and on the other, delaying data representation comprehension should make things easier.

This turns us toward consideration of how we conceive of "memory" and the work of "remembering." Without disputing that we do rely upon short-term and long-term memory, our proposal is to view remembering as more of a reconstructive act than one of retention and retrieval (Loftus & Loftus, 1980). One central tenet about this reconstructive view of memory is that what is remembered is not a veridical representation of what actually happened. When we do the work of remembering, we rely on a subset of what we actually had encountered and reconstruct an understanding *post hoc*. To illustrate, we borrow a distinction made by Kahneman (2011) in his discussion of how we remember experiences of discomfort or pain. A theme that emerged in Kahneman's research has been a distinction between an *experiencing self* and a *remembering self*. In his own words:

> The *experiencing self* is the one that answers the question: "Does it hurt now?" The *remembering self* is the one that answers the question: "How was it, on the whole?" Memories are all we get to keep from our experience of living, and the only perspective that we can adopt as we think about our lives is therefore that of the *remembering self*. (Kahneman, 2011, p. 381)

Throughout much of his career, Kahneman explored how the experiencing self and remembering self operate differently when reporting on concurrent versus retrospective judgments of discomfort or pain, such as the amount of immediately felt discomfort of a hand submerged in icy water (e.g., Kahneman, Fredrickson, Schreiber, & Redelmeier, 1993) or the discomfort of a colonoscopy (e.g., Redelmeier & Kahneman, 1996). He and his colleagues found that when the remembering self was operating, people do not make rational judgments. Rather, peak discomfort and the final level of discomfort had the most predictive power for how an uncomfortable experience was judged, as opposed to the total amount of discomfort experienced. This finding demonstrates that reconstruction is at work. In our view, this begins to point the way toward a productive, alternative view toward how data displays of past personal experience are comprehended.

For instance, imagine that we go on a 30 min run. During individual seconds and minutes of the run we experience a range of feelings and sensations, such as fatigue or exertion (e.g., when we increase our speed or begin running uphill) or relief (e.g., when reaching flat terrain). In addition, we may come to a sudden stop because a car is racing through an intersection that we are about to cross, or we may run more quickly because a vicious looking dog is barking at us. These moments (and the feelings that accompany them) would all be noted by our experiencing self as they happen. However, *after* the run, we may associate the run with generalities of other accumulated memories of running, and describe it as being more or less typical, despite the unique experiences along the way. If asked for detail, we may recount a few specific events from that particular run (i.e., almost getting hit by a car or running from a dog), but our remembering self will not resurrect all recently experienced aspects of the run. When new or atypical experiences are repeatedly remembered and eventually considered to be regular and predictable, then those aspects of experience will become generalized memory and future unusual encounters from the experiencing self can become candidates for possible generalization or remain as unusual standout moments that we can access retrospectively. This is a similar mechanism to our modification of and inference from enduring knowledge structures, such as scripts (Schank & Abelson, 1977) and cases (Kolodner, 1993). It is also similarly documented in cognitive psychology literature related to individuals' event memory formation (Zacks & Tversky, 2001; Zacks, Tversky, & Iyer, 2001).

Recognizing that this process of reconstructive remembering is a common one, we propose the following extension of Kahneman's distinction between the *experiencing* and remembering self to the ability to understand what events are depicted in a delayed display of personal data. We may have continuous engagements with the world, but we only have access to a fraction of our experiences when we subsequently reflect on those engagements. For this reason, we will base our "remembering" at least partly on our memory and knowledge of what *typically* happens and is *likely* to be experienced from that kind of activity. The remembering self engages in a process of reconstruction that overlays a few incomplete memories of specific, vivid experiences onto a more general familiarity of the activity in question. The reconstructed nature of experience is not necessarily problematic—a core argument of this paper is that these reconstructions of prior experience, though incomplete, can still serve as productive resources for students when they work with delayed-display wearable device data. As will be discussed and illustrated through the descriptive cases later, general remembrances, specific remembrances, and the combination of them can all be productive when comprehending data that reflects past experiences. This holds especially true for the context of the classroom.

Socially-distributed memory

The above account of general and specific memory that the remembering self does to access and reconstruct what had been encountered by the experiencing self is individualistic in nature. If the data interpretation situation we are considering is a lone person looking at their personal data dashboard, that individualistic account can provide some useful labels for what is happening in moments of interpretation and sense-making. However, the situations that we are presently trying to design and support take place in schools and classrooms with multiple people as participants. Given that, some additions to this framework are in order. What is remembered need not be restricted to a single person. By virtue of multiple people being present and involved, it can be distributed and shared (Sutton, Harris, Keil, & Barnier, 2010). Thus, in multi-person learning environments (such as classrooms), remembering may be expanded so that it is viewed as not just an individual activity but also as a collaborative one (Bietti, 2012). While tempting to see this as a replacement, our intention in recognizing the social nature of remembering is intended to be more pluralistic. That is, we think that social and individualistic perspectives on remembering can be maintained in parallel and serve as adjuncts with one another.

To illustrate, let's return to the example of a remembered run above. Instead of it being about the experience of one person, imagine if the 30 min run was done by running partners. We would appropriately expect that, while they followed the same route at roughly the same pace, they may each remember some things differently from their respective viewpoints and given their own expected sensitivities to the environment. One runner might have fondness for dogs and thought of the barking dog as adorable and only reacting loudly to being scared, whereas our first runner had remembered that dog simply as vicious and frightening. One runner might have been hungry just before the run and remembered that shortly after the dog incident, they passed a sidewalk food vendor with delicious looking food. The other runner might distinctly attend to part of the running route because they have a friend who lives down the street that they ran past. If given the opportunity to remember and verbalize their reconstructed memories of the run together, we would expect that something different would result compared to what each runner would have reported alone. Discussing what each remembers may lead to elaborations and new information being surfaced. Indeed, this is a known phenomenon in group memory research. When more than one individual is involved in a memory task involving a shared experience, the group tends to remember more than any individual would alone (Harris, Paterson, & Kemp, 2008; Rajaram & Pereira-Pasarin, 2010).

Intuitively, this may not sound controversial. However, much of the extant psychologically oriented research on memory has focused on what individuals are capable of doing. In the space of cognitively oriented memory research, there are only a few contingents that focus on what groups of individuals remember, and potentially remember better, in dialogue with one another. One way multi-person remembering is viewed is through what is referred to as transactive memory theory (Wegner, 1987). In transactive memory theory, the work of remembering is expected to be distributed among multiple individuals. Through discursive transactions, this information is surfaced for the group through contributions from specific individuals who are known to have relevant remembered information to share. Part of the ability of a transactive memory system to work well involves individuals knowing that others know and remember relevant things and that those can and should be accessed during a memory task. In what some consider optimal cases, having categories and items to be remembered delegated to different individuals represents one especially effective transactive memory strategy (Harris et al., 2008).

Lab-based research has suggested that groups of people who have some prior relationship with one another will have often developed their own effective distributed memory strategies that do not exist for ad hoc groupings (Hollingshead, 1998; Wegner, Erber, & Raymond, 1991). Long-married and long-partnered couples are often examined to help researchers understand some of these distributed memory strategies and how discourse within these couplings is used to prompt one another in ways that yield more accurate or more detailed memories (e.g., Browning, Harris, & Van Bergen, 2019; Grysman et al., 2020). Given the tendency for research in this area to be based in psychology laboratories, pre-collegiate classroom students have not been a focal population for this research perspective. However, many of us already recognize that students in a classroom have had an abundance of shared experiences together that can span hours, days, months, or even years. Given that recognition, we could reasonably hypothesize that when tasked with remembering previous shared experiences, students in a classroom can also leverage what each other individually remembers. This could take the form of corrections, and it could take the form of elaborations through additional testimonies. While new or more memories may not always be cued in these exchanges (Meudell, Hitch, & Boyle, 1995), the potential for the discursive interactions between multiple parties to enhance memories of prior events remains a compelling possibility and worthy of consideration (e.g., Harris, Keil, Sutton, Barnier, & McIlwain, 2011; Huebner, 2016; Sutton et al., 2010). It is one way in which we intend to extend beyond points we made above regarding remembering as general and specific reconstruction.

Embodiment within the school day

Having discussed some of how experiences are remembered, individually and with the support of others, we now orient toward thinking about what aspects of experience are being remembered. School day activities, and specifically recess activities, are the main objects for youth data analysis in our line of design-based research. Consequently, what happens at recess is what we are interested in students remembering. In some of our prior work, we have discussed how recess has unique qualities that make it especially appealing for providing students with opportunities to think and learn with data (Lee & Drake, 2013). Among those are the more overtly embodied aspects of recess. Much of the traditional school day experience for students is, for good or ill, largely extended times spent sitting at desks or tables. Free movement is often restricted to transitions between classes or activities, physical education, and designated breaks that are typically formalized in the elementary school day as recess. While a strong take of an embodied perspective may suggest that attending to bodies is a requirement for understanding any human activity, for our current work, recess stands out because it is one of the times when active and highly discretionary bodily movement is privileged for youth.

The recognition of embodiment has several implications for the design of educational experiences (Abrahamson & Lindgren, 2014; Hall & Nemirovsky, 2012; Lindgren & Johnson-Glenberg, 2013). At the same time, it is important to acknowledge that embodiment as a research perspective has multiple facets (see Lee [2015] for a summary). For the current work, we touch upon two facets that connect to our focus on student wearable activity tracker data, and specifically those data that come from the active movement time of recess. One of those facets emphasizes mental representation and looks at how our physical bodies provide bases for spatial and sensory-based schemata that shape conceptualization (Barsalou, 1999). For instance, conceptual metaphor theory suggests that sensory and movement-based metaphors are invoked in our conceptualizations of topics such as emotion and lifespan (as hot/cold or as a journey, respectively; Lakoff, 1980). Our bodies shape our understanding of certain topics to the degree that transcends any specific situation in which those ideas are invoked. Proponents of this view of embodiment have seen these body-originating ways of thinking as so fundamental that some have gone so far as to argue that all cognition related to specific disciplines, such as mathematics, is based in conceptual metaphor (Lakoff & Nunez, 2000). In learning sciences research, specific metaphors may or may not be called out, but drawing from associated intuitive schema developed from body-based interaction with the world has had appeal (Lindgren & Johnson-Glenberg, 2013). There have been intentional design efforts to invoke the underlying schema that are the bases for these metaphors, often instantiated or elicited through gestural or body-based movements, to support mathematical and scientific reasoning (Howison, Trninic, Reinholz, & Abrahamson, 2011; Lindgren, Tscholl, Wang, & Johnson, 2016).

In the current article, our expectation is that the underlying schemas comprise much of the general remembrance that is involved in memory reconstruction. Interactionally, these schemas will be instantiated through taken-as-shared articulations and mutual understanding within a classroom of known physical activities and how those connect to bodily experience. For instance, a student complaining that they have to walk uphill to get home from school would not be challenged with the idea that that is easier than a condition where the path they must walk to get home is more level. Uphill movement would be intuitively and mutually understood as involving more physical effort. To understand this, we would not need to know specifics about the hill that must be walked, the exact angle of incline, or what roads and buildings surrounds that hill. We simply need to recognize that there is an activity we already know about generally that is “walking home” and that in that student’s mentioning of it, it has an uphill component. For those that are familiar with the complex systems view of intuitive knowledge (Smith, diSessa, & Roschelle, 1994), this could reasonably be understood as activation of a network of prior knowledge that has as part of its core intuitive body-based understandings related to effort and motion (e.g., diSessa, 1993).

The second facet of embodiment literature relevant to the current work is through the lens that acknowledges all action is inherently situated. In this view, bodies in space and place that encounter a dynamic physical and social world are examined as encompassing and complex configurations that yield what we call familiarity and experience. This favors a *person plus context* unit of analysis that takes seriously the work of social and social-material coordination as embodied activities. Learning sciences work that represents this tradition well includes studies that deeply consider space and mobility as embodied resources (Leander, Phillips, & Taylor, 2010). For instance, this can involve mobility and mapping of lived environments (Taylor & Hall, 2013), movements and body positions of multiple individuals in a designed learning setting (Shapiro et al., 2017), or coordinated multiparty mathematical activity (Hall, Ma, & Nemirovsky, 2015). With respect to the latter study by Hall and colleagues, we also can see embodiment as socially distributed and collaboratively orchestrated. Thus, a situated (or situative—see Greeno, [1998]) account lends itself well to thinking about something historically treated as based in an individual as being multi-person in nature. In our current study, embodied experiences and their situativity figure most prominently by foregrounding how students' remembrances of school day activity are tied to specific situations that involved specific activities in specific places with specific *other people*. Note that while we emphasize here the situated facet of embodiment literature, viewing learning activity through this lens does not, in our view, preclude maintaining perspectives from the first facet of embodiment, which is more individual based.

Together, these two ways of looking at embodiment help make more salient certain features of how students can remember recess activity. Those features include an intuitive sense of more and less exertion and generally what school day and recess activities are associated with those. It also includes the specific social and physical experiences that students have and are bringing to these learning activities. This intentionally corresponds to the distinction of general and specific remembrances discussed above where some developed ideas about what are common activities and what is more or less exertion intersect with specific moments of prior experience. It also connects to the distinction between individual and socially distributed memory we have discussed above.

Quantified self and student experiences with familiar data

The opportunity for students to work with (and learn from) data obtained from their routine activities dovetails with what some have labeled as the *Quantified Self* movement (Lee, 2013). This has also been called personal analytics (Lee, 2018) or personal informatics (Li, Dey, & Forlizzi, 2010). The central assumption of this movement is that we can gain new insight or knowledge from reviewing quantitative data about ourselves and our experiences. By doing so, we may find unexpected regularities in our behavior, or be able to track change over time with respect to a certain goal. It is worth noting that in quantified-self communities, these are not typically academic goals but rather ones related to health, wellness, or routine habits (Lee, 2014). The hobbyist community that has been most prominently associated with quantified self is largely adults located in major metropolitan areas around the world. These individuals are loosely coordinated with one another through online forums (e.g., social media, discussion boards, 'meetup' platforms) and publicly shared video recordings of individual quantified self project presentations (Choe, Lee, Lee, Pratt, & Kientz, 2014; Lee, 2014). Therefore, the types of activities and experiences that are typically quantified are those that represent adult routines and interests.

It is also worth noting that there are some similarities with other learning sciences educational experiences that centralize students' sense-making of data, such as the work of Enyedy and Mukhopadhyay, (2007), Philip, Olivares-Pasillas, and Rocha (2016), and Rubel, Hall-Wieckert, and Lim (2017). However, there is a difference in emphasis in that the current work is a focus on data from the self rather than local neighborhood- or city-scaled geography or the surrounding community. Explicit focus on the self, as a physically embodied entity that engages with the world in ways that can produce traces of individual activity that can be retained for subsequent quantifiable analysis as a means for learning, constitutes a quantified self paradigm for K-12 data science instruction. Presently, we adhere to that paradigm, and we also intentionally focus on and design instruction around disciplinary content––specifically, statistical ideas related to center, variation, and distribution. We also "stay on school grounds" and work with classrooms to unpack what may seem routine and find new things to notice or understand about the "ordinary" activities around them.[2]

Overall, self-quantification, as understood here to be highly personal and about one's own day-to-day activities and experiences, has become a more pervasive practice (Lupton, 2016). The overall growth in self-quantification has been enabled by the proliferation of mobile handheld devices (and the person-specific "data exhaust" that these and other devices produce), as well as the rise of commercial wearable technologies such as fitness and activity trackers (Lee & Dumont, 2010). These are the same kinds of technologies that learning scientists seek to leverage through wearable-based inquiry in order to reduce the educational overhead of data collection (e.g., Ching & Schaefer, 2015; Kang et al., 2016). There has been thus far only a few explicitly named efforts to connect quantified self to education and the design of learning activities (e.g., Lee et al., 2016; Potapov, Vasalou, Lee, & Marshall, 2021; Rivera-Pelayo, Zacharias, Müller, & Braun, 2012).

Our present interest is in the overlap between the reconstructive work of remembering, the embodied nature of what is generally and specifically remembered, the data that are situated within the quantified self paradigm, and the opportunities for disciplinary (e.g., statistical) learning. For an individual, we would consider this as involving the remembering self overlaying specific embodied experiences on top of general expectations to construct a remembered experience. These are driven by or mapped onto quantified self data. Both directions for data interpretation and remembering are valid and are likely to happen. Much will depend on the way in which a data interpretation activity is designed and which interpretive sequence it promotes. What is key here is that introducing quantified self data representations could provide opportunities for students to develop greater fluency with data by providing numerical values to which the remembering self can link salient experiences. For instance, a student may not immediately be proficient at reading time-ordered displays of activity data. However, if a wearable-device equipped student recalls an unusual moment when they had to run back to their classroom during lunch, the recall of that embodied experience can help students to locate that event as a visual landmark in the data display (e.g., as an unusually high set of heart-rate values; we have previously referred to these as *eventmarks*, see Moher et al., 2014). Similarly, if students recall that activities in their daily language arts lesson are fairly routine, they may develop general expectations for how movement data from language arts *should* look, and how much variability is to be expected across a set of days, when seeing language arts represented across several days as their own physical activity data.

We believed that similar remembrance activities could be built into elementary grade learning experiences to support the development of statistical knowledge. This would involve opportunistically appropriating common routines within a school setting as quantifiable experiences and focusing on those as objects for discussion and analysis (Lee, Drake, Cain, et al., 2015). It would also leverage the presence of multiple people comprising the classroom. Quantified self is often treated as a practice of more deliberately understanding a single individual. This solipsistic approach is reflected in the quantified self hobbyists' community tendency to favor an "N of 1" design (Heyen, 2020). Classrooms, as discussed above, have multiple people who can collect data, and the unit of analysis can be more than just a single person. This creates new opportunities for thinking about *quantified selves* in addition to students thinking about their individual quantified self (Lee et al., 2016).

Thinking about these configurations and resources for the purposes of instruction, we recognize a need to consider what is a reasonable target for students to realize in terms of their statistical thinking. Existence proofs already show that elementary school students can learn important

[2]We do depart from school grounds in some of our other quantified self studies, such as using a university gymnasium or giving youth time to examine data they produce while at home. See Lee and Dumont (2010) and Lee and Briggs (2014) for examples.

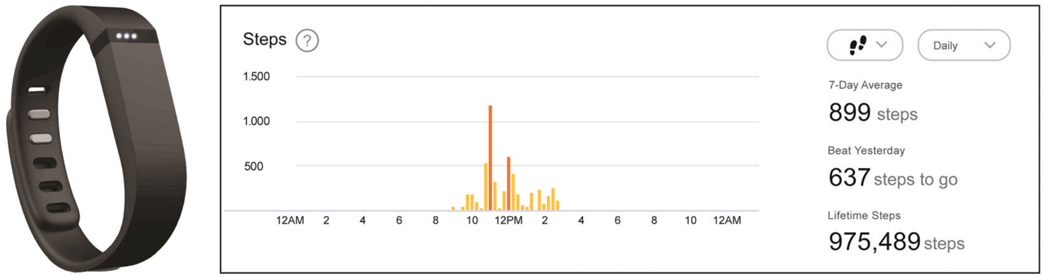

Figure 1. Fitbit Flex Wrist-Based Wearable Device (left) and a Screen Capture of the Dashboard Showing How Many Steps Have Been Taken (right). Note. Rather than show numbers of steps taken as some wearable activity trackers do, the display shows only LED dots that increment over a known interval, such as 2000 steps. In the dashboard, the steps are shown in 15 min increments even though the data are stored on Fitbit servers in minute intervals.

statistical topics often reserved for later grades, such as variation and distribution of data (Lehrer & Schauble, 2004; Petrosino, Lehrer, & Schauble, 2003), when given adequate opportunity to reflect on real data. Other notable examples within research have shown that children also can learn distributional shape and how it emerges, particularly as more data are collected, with thoughtfully designed activities or technologies that have built-in display capabilities to make key features and patterns in a set of data visible (Abrahamson, 2009; Ben-Zvi, 2000; Konold et al., 2002). Our own work has focused on helping students to realize those accomplishments and also to develop more robust understandings of averages and other measures of center through new activity design that involves quantified self/selves approaches and also relies on the use of visualization technology. The emphasis on averages and measures of center is in part because those are topics taught repeatedly at multiple grade levels, yet research has repeatedly shown that even adults still struggle with the underlying ideas (Mokros & Russell, 1995; Strauss & Bichler, 1988; Watson & Moritz, 2000). It was also content matter that was expected to be covered based on local standards that our partnering school were following. Yet this strong orientation toward using quantified self data for developing these statistical intuitions distinguishes our approach from these cited studies. More details about activities in our approach are shared below.

Design of the learning activities

This article reports on data from the fourth, and most refined, iteration of a multiyear design research project (The Design-Based Research Collective, 2003), in which we strategically repurposed wearable technologies that were originally used for athletic performance and personal fitness tracking to generate data for statistical inquiry. By "most refined," we mean that during this fourth iteration of the design project, the curriculum materials, tools, and designed learning interactions were cohesive and complete enough that, when implemented, they required very little unexpected sequence or activity modification, and appeared to bring about the intended outcomes. Throughout the project, our high-level design conjecture (Sandoval, 2014) was that deployment of passively obtained data about students' own activities could serve as a vehicle for producing both specific and general components of remembered embodied experience that would support statistical understandings. To realize this conjecture, we appropriated wearable Fitbit activity trackers (Figure 1) and data visualization software for new uses in the classroom.

Wearable step tracking devices and data visualization software

As we described above, we chose to use wearable devices through various iterations of this project because, unlike previous probeware-style sensors, wearables have the ability to record data from ongoing activities and routines without dedicated user attention to data collection (Lee &

Shapiro, 2019). Many of these devices record large amounts of data, often at minute or multi-second intervals. Currently, physical activity trackers on the market include the ability to track steps taken, heart rate, geographic position, gain in altitude, calories burned, breaths taken, and number of times one moves from a seated position to a standing one. Most fitness and activity trackers work through the use of small three-axis accelerometers that produce data that are processed through proprietary algorithms. Some produce data that, when tested independently by exercise scientists under laboratory conditions, can be quite accurate (Diaz et al., 2015; Takacs et al., 2014). Furthermore, others in the field have been taking similar approaches to considering the learning potential associated with wearables (Ching & Schaefer, 2015; Kang et al., 2016; Lee & Shapiro, 2019; Moher et al., 2014; Taylor, 2017).

Ultimately, we opted to focus on *steps taken* as a focal measure for elementary students. Steps are familiar, countable entities that are measured by a variety of tracking devices. Further, we found that wrist-based activity bracelets, and specifically the Fitbit Flex (Figure 1), were promising for classroom use for at least two reasons: First, the visibility of the device on one's arm reduced device loss (when compared to clip-on devices that attach to clothing); second, the Fitbit Flex does not display the number of steps taken on the device itself—rather, it simply displays more lights when certain values are reached (i.e., a new light comes on after every 2,000 steps taken). Keeping data out of the immediate purview of the experiencing self helped to keep students oriented toward their daily activities, rather than on their immediate judgments of the numbers shown on the device. We have previously reported on challenges encountered when students receive immediate numerical feedback (Lee, Drake, & Williamson, 2015). These include students becoming preoccupied with processing delays as data are being updated on the device and the added work of needing to mathematically compute how many steps were taken during specific moments.

Typically, physical activity data from these devices are presented to consumers through an online "dashboard." Some early research with youth has suggested that these dashboards are counterintuitive for students to read and interpret (Ching & Schaefer, 2015). Considering these devices and tools are marketed toward adults, it is unsurprising that they are not youth friendly. However, there has been, within the educational research community, some research into tools that support data visualization and interpretation that could render data more approachable.

One of the most noted early tools was *Tabletop*, a data visualization tool for youth that linked icons with records in a database and allowed creation of graphically organized sets based on constraints and inclusion criteria (Hancock, Kaput, & Goldsmith, 1992). *TinkerPlots* (Konold & Miller, 2005) and *Fathom* (Finzer, 2005) are comparable tools developed later that have many of the same capabilities, support additional visualization types, and have features that help link probability with statistics (Lehrer et al., 2014). Most recently, a new online platform, *CODAP* (codap.concord.org), has been under development that also provides these capabilities through a web interface. Commercial options using similar paradigms exist as well (e.g., *Tuva*). In addition, statistical visualization tools appear in some mathematics education research, such as the computer mini-tools appearing in teaching experiments run by Cobb and colleagues (e.g., Cobb & Tzou, 2009), and homegrown visualization tools that can show distributions and variation have been developed for use in statistics education research (delMas & Liu, 2005).

Given the targeted age group and the kinds of data that could be obtained from Fitbit Flex devices, we selected *TinkerPlots* as the visualization tool for our designed intervention. It allowed students to easily see data over time, examine individual data points to see which student in the class produced that record, and examine sets of data points for overall distributional shape. We had to develop an additional custom tool, previously described elsewhere (Lee et al., 2016), to extract Fitbit step data at minute-long intervals for specific times of day. Once the tool was developed, the data could be imported into *TinkerPlots* and manipulated through its drag-and-drop interface.

The enacted elementary data and statistics unit

In consultation and collaboration over multiple years with practicing fifth and sixth grade teachers, we developed a 4-week unit to focus on Common Core-based state standards related to data and statistics, including understanding statistical variability, describing distributions of data in terms of center and spread, understanding how measures of center are produced from a larger set of numbers, and displaying data in canonical data representations such as histograms and box-and-whiskers plots. These are considered both within the purview of statistics education and as foundational ideas for data science education. This unit was iteratively created and refined over 4 years, with more formally written activity plans and support materials developed and refined each year of the project.

Daily lesson organization

The unit consists of roughly 20 h long daily lessons, corresponding to the amount of time the teachers in our collaborating site would normally have allocated to mathematics and these topics. Each day, the students checked out an assigned Fitbit Flex at the start of school and checked it in again before leaving school at the end of the day. As a result, students were continuously collecting data throughout the school day for the duration of the unit. The intention in collecting all these data was to create a relatively large data set (numbering in the 100,000 s) that would enable retroactive data queries from students. That way, students could later pursue projects where they compared all data from one class and compare those to another class or look for class trends over time.

The school day proceeded normally until the morning math period where the teachers enacted our designed unit. The designed structure of each math period followed a similar, two-part approach that involved a “Daily Starter” portion and a “Main Lesson” portion.

Daily starters

In order to facilitate frequent encounters with and opportunities to reflect on students’ activity data, the first fifteen minutes of each class were dedicated whole-class activity data explorations, which we called “Daily Starters.” During a Daily Starter, Fitbit activity data from one or more students was displayed in *TinkerPlots* and projected in front of the class. At the beginning of the unit, the data set examined included every minute of the school day for a given student. After a few days, the class would notice that there were common contours in the data and would then choose to transition to examining a more focused period of time (i.e., morning recess). This reduced the number of data points the students examined at a time and focused their efforts on recess times, which had more variability in step counts than the more structured portions of their school days.

The classroom teacher facilitated discussions each day and would use the Daily Starter as a time to not only have students discuss current data but also revisit content from the Main Lessons on days prior. For instance, following the instructional model used by Lehrer and Schauble (2004), groups of students invented different data displays and converged on histograms as one way to organize data to see distributional shape. During the next school day’s Daily Starter, the teacher included a task where students organized their activity data into a histogram and discussed what that representation captured and what it did not.

We intentionally designed the Daily Starter class discussions to have students engage their remembering selves in thinking about their previous day’s data and make their strategies for interpreting these activity data plots open to public inspection. Although we allotted 15 min to each Daily Starter, the teachers were welcome to let the activities extend longer if they felt the conversation was especially productive. On a few occasions, the teacher and students found the

Daily Starter discussion fruitful enough that they extended it for the entire class period. However, the two classrooms were still careful to keep pace with each other so that they generally covered the same content in the Main Lesson each day.

As the unit progressed, the Daily Starters adapted so that students were not simply examining their data but were generating hypotheses about their school day activities, identifying activities based on the shape of the data, and attempting to control distributional characteristics by changing the nature of their activities. For example, during these Daily Starter discussions, some students would speculate that certain playground activities generally require more steps than others over the same period of time. Other times, students would suggest ways in which one could alter their movement throughout recess to create smaller or larger ranges in their data.

Main lessons

After the Daily Starter, a teacher-facilitated "Main Lesson" followed. Over the four years of this design study, we designed and refined these lessons with two main goals: 1) to build on existing research for how to present students with statistical ideas and 2) to complement and extend students' explorations of activity data in the Daily Starters. For instance, previous research (e.g., Lehrer & Schauble, 2004) has suggested that inventing representations of data and critiquing them can improve students' understanding for why some forms of data representation are more useful for some tasks than others. Thus, for 2 days, the main lesson involved groups of students designing their own representations of data that they had obtained from measuring common objects in their school and then critiquing the various features of those representations. This motivated the use of putting data into bins and producing representations that showed and ordered distributions in order to see features of center and spread as would be expected and associated with canonical data representations (e.g., an ordered dot plot or a box-and-whiskers plot). Another example involved making inferences from their data by comparing distributions of all students who did two different recess activities (Watson & Moritz, 1999). Through a combination of whole-class discussions, small-group work, and individual activities, the students learned to produce canonical displays of data (specifically, histograms and box-and-whiskers plots, as were required by their state's standards) and calculate and apply measures of center and variability. At the end, they looked at a subset of pooled class data to make a comparison between two different conditions under which wearable device data were collected. As mentioned before, these topics were regularly connected back to their activity data on subsequent Daily Starter activities.

Participants

The students in this study came from a Title I K-8 school located in the Rocky Mountain region of the United States. For most of the school day, the students were in one of four mixed fifth/sixth-grade classrooms. However, the students were separated by grade level during math time. Two of the combination grade teachers, Ms. Bryson[3] and Ms. Hayley, taught sixth-grade mathematics. These sixth-grade math classes were our focus because of their need to cover related elementary statistical content according to local standards documents. Only the sixth-graders wore the Fitbit Flex devices. Ms. Bryson had over 30 years of elementary teaching experience. Ms. Hayley was in her 3rd year of classroom teaching but had extensive experience in early childhood education prior to getting her teaching credential.

At the beginning of every academic year, the students took a pre-assessment for math that the school used to sort them into one of two math classes. Ms. Bryson received the students who had higher scores on the pre-assessment (hereafter "Class A"), and Ms. Hayley received the students

[3]All names of teachers and students are pseudonyms.

who had lower scores ("Class B"). This sorting was reportedly a school-wide strategy intended to allow each teacher to tailor their instruction to the needs of their respective class. We did observe that = students in Class B expresssed more frustration with mathematics in class, and Ms. Hayley often found it necessary to review much earlier mathematics content, such as multi-digit subtraction, with the students, in the middle of lessons. We did not get an official count, but we were told by the two teachers that more students in Ms. Hayley's class had individualized education plans and were a more neurodiverse group. Ms. Hayley had an aide who provided assistance with during math instruction, whereas Ms. Bryson did not.

Class A had 27 students, of whom 26 were consented to participate in our research. Class B class had 29 students, all of whom were consented. As is common with students at rural schools in the United States, the population was predominantly but not exclusively White. Specific racial demographic data were not obtained from the students or the school. However, according to census data, the town in which the school was located was 95% White, less than 5% Hispanic or Latino,[4] and less than 1% population each of other named census categorized races and ethnicities, and less than 2% for the categories of other or two or more races. From our discernment, Class A was entirely made up of White students. There were a pair of twin girls of East Asian descent and one Latina girl in Ms. Hayley's math class, but otherwise the rest of the Class B's students were discernably White.

Data collection and data analysis

Written pre- and post-assessments

To assess student progress in the unit and help us to understand the degree to which our designed activities supported students in reaching intended learning goals, we administered written pre- and post-assessments using an instrument developed by Lehrer et al. (2014). This instrument was developed explicitly to assess how students were performing along a statistical thinking learning progression informed by a body of previous classroom-based design research (e.g., Lehrer & Schauble, 2004; Petrosino et al., 2003). As described in the extant literature, learning progressions represent potential trajectories for student development and learning of practices and content in a specific disciplinary area (Duncan & Hmelo-Silver, 2009). They are intended to be informative for both design and assessment. Through an iterative instrument validation process, Lehrer et al. (2014) identified a set of constructs that their written instrument measured. These included:

- Conceptions of Statistics (CoS)—ability to describe and characterize distributions and measures of center through computation and estimation;
- Data Display (DaD)—ability to create and read canonical data representations such as density plots and histograms;
- Metarepresentational Competence (MRC)—ability to critically evaluate alternate forms of data representation relative to what is to be communicated and who is to be reading a data representation (this is based on existing research related to children's ability to do complex work with creating and evaluating representations [diSessa, 2004]);
- Informal Inference (InI)—ability to draw inferences from multiple sets of distributional data based on overall appearance in data displays and assuming inherent variability;
- Modeling Variability (MoV)—ability to understand how changes in measurement systems will produce variability and the inherent variability associated with more and less precise measurement approaches.

[4]We use the US census terminology here.

Emma is a fan of the Telluride Prospectors, a minor league baseball team in her hometown. Emma tracked the number of points they scored in their 30 home games. The list below shows how many runs they scored in each of their games.

1, 2, 3, 6, 3, 4, 3, 6, 8, 4, 5, 5, 2,
6, 5, 5, 4, 3, 3, 11, 4, 2, 0, 9, 4, 4, 1, 0

Given these scores, make a display that helps you think about how many runs yoiu expect the Prospectors to usually score during a home game.

Figure 2. Example of an assessment question.

Table 1. Summary of pre- and post-assessment question types.

Assessment	Question sets	Total responses	Multiple choice response type	Open response type
Pre	7	20	5	16
Post	8	25	7	21

We made slight localization modifications to the items (Figure 2) from this instrument and created pre and post-tests that posed intentionally similar questions in different orders. For the post-test, some surface features of the items were changed (e.g., the context described in the problem and numerical values), but the difficulty and performance expectations of the items were not. The tests each had seven or eight problem sets that each had several scorable sub-questions. In total, the pre- and post-tests each had 20–25 items (scorable sub-questions) that could be assessed, with a mix of multiple choice and open response questions among them. The post-test was slightly longer in following with the original instruments we were provided. The exact breakdown of multiple choice and open response questions is summarized in Table 1. The tests were administered during the students' mathematics period the day before the unit began and the day after the unit ended. A sample item with our specific contextual adaptations appears in Figure 2.

Assessment scoring process

Lehrer et al. (2014) had developed assessment rubrics for each item in the written tests and provided them to us. These rubrics identified specific, assessed constructs (e.g., *Data Display* or *Informal Inference*); a stated performance for each identified level; and examples of student responses that matched each performance level. It is important to note that each rubric item included as many as eight performance levels (inclusive of half levels when a performance was characterized as in transition but distinct enough to merit its own row in the rubric), resulting in a complex scoring system designed to accommodate many degrees of nuance in student responses. A summary of the constructs and low, intermediate, and high performances as applicable to our population is shown in Table 2.

To enable quantitative analysis, we converted these performances into numerical values using a linear conversion scale. Threshold performances were indicated by an integer. For example, a Data Display performance of DaD2(c) would have a threshold value of 2. Subthreshold performances were denoted by letters from "a" to "d." We mapped these to decimal values on a 0.2 increment. So, a performance scored as DaD2(c) would correspond to a numerical value of 2.6. The 0.2 increment allowed us to account for transitional values between subthresholds, denoted by a minus (-). This would be considered a 0.1 deduction on the unmodified score—for example, DaD3(c-) would convert to 3.5.

Table 2. Highest, intermediate, and lowest performance levels on constructs of interest for our statistics unit.

Construct	Lower performance	Intermediate performance	Highest performance
Data Display (DaD)	DaD1(a): Create or interpret displays without relating to the goals of the inquiry.	DaD4(a): Display data in ways that use its continuous scale (when appropriate) to see holes and clumps in data.	DaD5(b): Describes aggregate properties of displays using quantities such as proportions.
Metarepresentational Competence (MRC)	MRC1(a): Recognize that displays represent data, but misinterpret one or more elements of the display.	MRC3(a): Compare displays by indicating what each shows about the structure of the data.	MRC4(b): Coordinate qualities of multiple displays with what those displays show and hide about the data.
Conceptions of Statistics (CoS)	CoS1(a): Use visual qualities of the data to summarize the distribution.	CoS2(a): Calculate the statistic for central tendency.	CoS3(d): Predict how a statistic is affected by changes in its components or otherwise demonstrate knowledge of relations among components.
Modeling Variability (MoV)	MoV1(a): Attribute variability to specific sources or causes.	MoV2(a): Informally estimate the magnitude of variation due to one or more causes.	MoV2(b): Describe how a process or change in the process affects variability.
Informal Inference (InI)	InI1(a): Make a judgment or prediction according to personal experience or beliefs.	InI3(b): Compare two distributions based on one value such as the maximum, mean, or the number of cases above a cut-point.	InI5(a): Compare two distributions based on proportions within defined regions, such as above/below a cut-point or IQR.

Note. Because the Lehrer group's learning progression accounts for a broad spectrum performance across many age groups, we only show the highest performance that were likely attainable through participation in our unit.

Two members of the research team scored the open responses on the written tests, with each researcher scoring exams from one class. To maintain consistency with each other and with previous assessments, reliably scored tests from a previous year were used as a training set.

Assessment scoring reliability

A subset of five students from each class were randomly selected and cross-coded to verify final scoring had remained reliable given possible thresholds, subthresholds, and transitional performances. Reliability was high based on testing with a weighted Cohen's Kappa ($\kappa = 0.905$), a reliability test we used because of the ordinal nature of the scores.

Classroom video

The second major data source was video obtained of each day of the unit in both Class A and Class B. Two high-definition video cameras were placed in each classroom with one operated by a researcher and the other remaining stationary in the back of the room. Many of the procedures discussed in Derry et al. (2010) were followed in the use of video for observational records. Specifically, content logs were produced of all recorded lessons and group viewing took place over a period of several months for selected videos. The goal in group viewing was to identify instances in which there was multi-party sense-making around projected student activity data. Within these instances, we focused on articulating what resources were being used by students, including explicit mention and discussion of previous physical activities, as those were indicative of the various kinds of work being done by a remembering self. These occasions took place most

Table 3. Mean performance for Ms. Bryson's class (Class A) on the pre- and post-assessments.

	Pre-Assessment		Post Assessment	
Construct	Mean	SD	Mean	SD
CoS	0.231	0.375	1.984	0.567
DaD	1.444	1.077	3.020	0.813
MRC	2.700	1.169	2.696	1.452
InI	2.394	1.086	3.110	0.919
MoV	0.890	0.809	1.675	0.635

often during Daily Starter activities, which focused on discussing and analyzing segments of activity data.

For the current article, we selected three cases from our review of relevant logged video data (over 3 weeks from each class of Daily Starter activities, or roughly 30 days equivalent of video footage). The rationale for these cases is discussed when the cases themselves are presented below.

Quantitative results

Because the two classes were organized based on their performance on a school-designed assessment at the beginning of the year, we report on each class' pre- and post-performance separately.

We first report on the scores of students in Class A, which consisted of students with higher performance on the school's sorting assessment. We performed a repeated samples *t*-test ($N = 24$), which suggested that the overall performance on the post-test was significantly higher on the post-test than on the pretest ($t = -11.281$, $df = 23$, $p < 0.001$, $d = 2.303$). In addition, there appeared to be improvement across four out of the five constructs (Table 3). The performance was significantly greater in Conceptions of Statistics (CoS; $t = -13.59$, $df = 23$, $p < 0.001$, $d = 2.774$), Data Display (DaD; $t = -6.094$, $df = 23$, $p < 0.001$, $d = 1.244$), Informal Inference (InI; $t = -3.305$, $df = 23$, $p < 0.01$, $d = 0.675$), and Measures of Variability (MoV; $t = -4.332$, $df = 23$, $p < 0.001$, $d = 0.884$). The one construct that did not show significant improvement was Metarepresentational Competence (MRC; $t = 0.012$, $df = 23$, $p = 0.99$, $d = 0.002$). A visual appraisal (Figure 3), suggests that there was little net change in Ms. Bryson's class on Metarepresentational Competence.

Class B ($N = 26$) also showed a statistically significant overall improvement from pretest to post-test ($t = -10.131$, $df = 25$, $p < 0.001$, $d = 1.987$; Table 4). When *t*-tests were performed on each construct, unlike Class A, there was significant improvement in Metarepresentational Competence (MRC; $t = -4.76$, $df = 25$, $p < 0.001$, $d = 0.934$; Figure 4). There was also significant improvement across Conceptions of Statistics (CoS; $t = -10.32$, $df = 25$, $p < 0.001$, $d = 2.024$), Data Display (DaD; $t = -7.133$, $df = 25$, $p < 0.001$, $d = 1.399$), Informal Inference (InI; $t = -7.105$, $df = 25$, $p < 0.001$, $d = 1.393$), and Measures of Variability ($t = -5.776$, $df = 25$, $p < 0.001$, $d = 1.133$).

While these suggest gains for students in Class B across all measured constructs, those students did not reach the same performance levels as their peers in Class A. The students in Class A scored significantly higher on the post-test than students in Class B on each construct ($p < 0.05$). However, the overall gain scores of students in Class A were *not* significantly higher than the overall gain scores of students in Class B ($t = 0.330$, $df = 47.44$, $p = 0.743$, $d = 0.092$). Students in Class A gained significantly more than students in Class B on Conception of Statistics ($t = -3.810$, $df = 45.796$, $p < 0.001$, $d = 1.084$), and students in Class B gained significantly more than students in Class A on Metarepresentational Competence ($t = 3.229$, $df = 47.454$, $p = .002$, $d = 0.907$) and Informal Inference ($t = -3.810$, $df = 45.796$, $p < 0.001$, $d = 1.084$). In short, even though this was intentionally an ambitious unit for students, we noted that those who were sorted

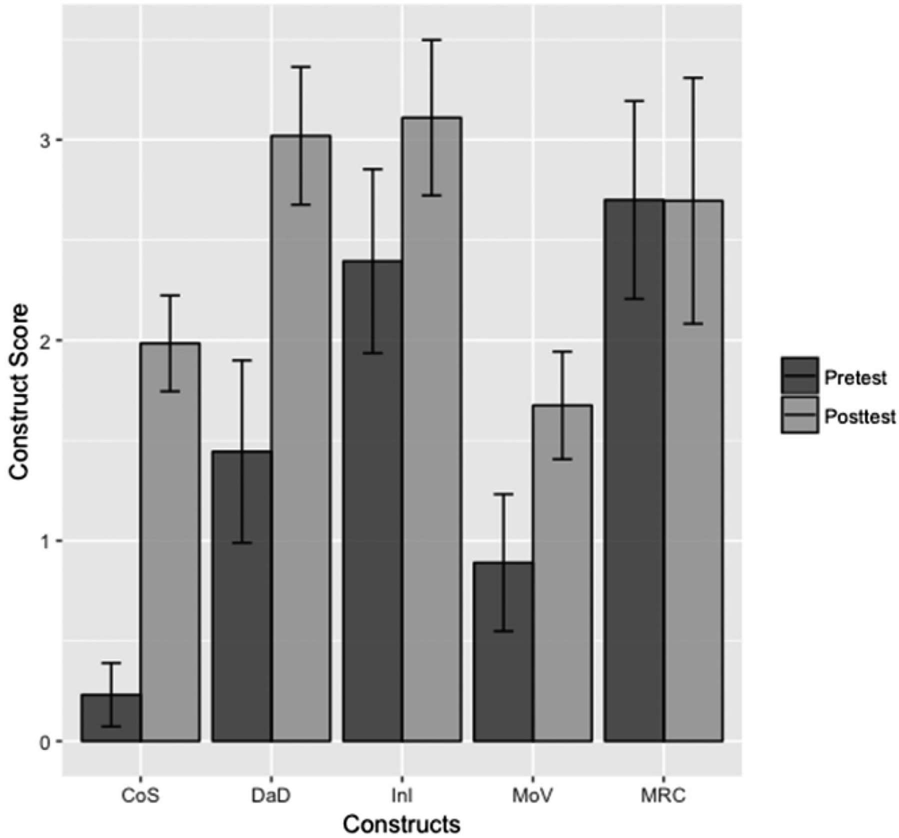

Figure 3. Mean pre- and post-test performance for class A (Ms. Bryson) across five measured constructs. Note. Conceptions of Statistics—CoS, Data Display—DaD, Informal Inference —InI, Modeling Variability—MoV, and Metarepresentational Competence—MRC. Pretest scores are the darker columns. Error bars depict the 95% confidence interval. Class A showed significant gains across all constructs except MRC.

Table 4. Mean performance for Ms. Hayley's class (Class B) on the pre- and post-assessments.

	Pre-Assessment		Post-Assessment	
Construct	Mean	SD	Mean	SD
CoS	0.142	0.148	1.182	0.580
DaD	0.270	0.478	1.954	1.202
MRC	0.646	1.095	2.296	1.544
InI	0.571	0.651	2.248	1.175
MoV	0.585	0.495	1.249	0.660

by the school as lower performing were able to demonstrate measurable progress on all constructs.

Discussion of quantitative results

From these pre- and post-test scores, we infer that this version of a wearable-based inquiry unit we had developed could produce significant improvements for a range of sixth-grade students. Previous studies which were able to have same-school staggered control groups on earlier versions of the unit have shown significant differences with traditional units covering similar content (Lee & Thomas, 2011; Lee et al., 2016). It seems plausible that this could have been true here if

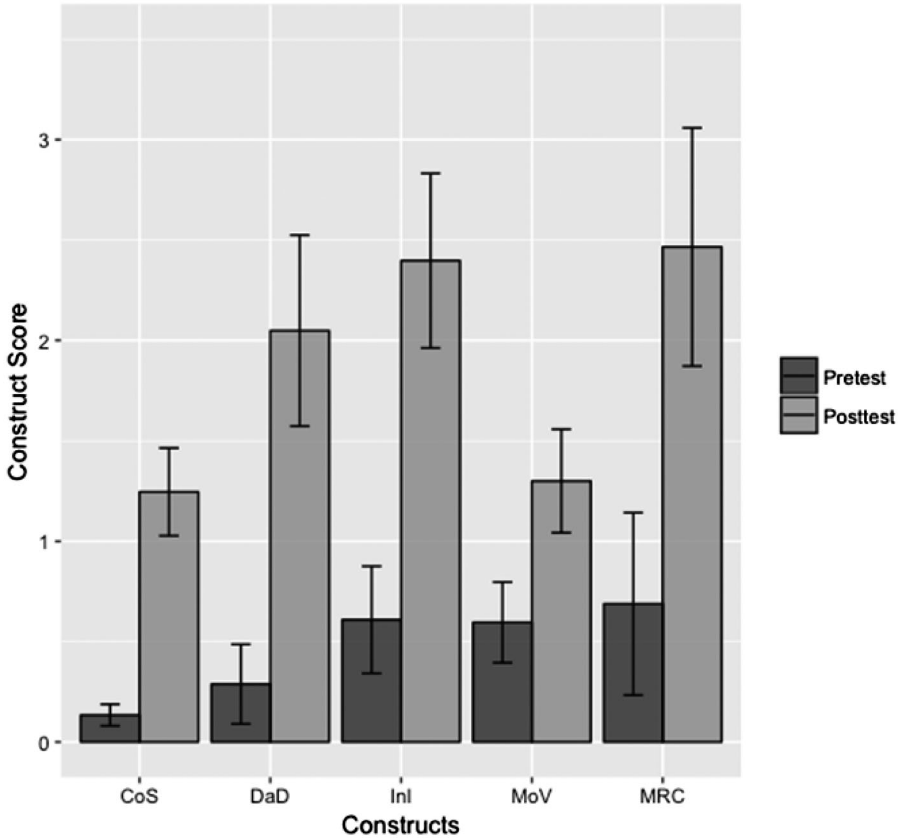

Figure 4. Mean pre- and post-test performance for class B (Ms. Hayley) across five measured constructs. Note. Conceptions of Statistics—CoS, Data Display—DaD, Informal Inference—InI, Modeling Variability—MoV, and Metarepresentational Competence—MRC. Pretest scores are the darker columns. Error bars depict the 95% confidence interval. Class B showed significant gains across all constructs.

we were able to arrange for an acceptable control group. However, as part of our negotiations with this particular school, we agreed to help both sixth-grade classes follow the same unit in parallel and could not establish a staggered controlled comparison. Therefore, claims from these specific quantitative results about any differences from a possible control condition must remain speculative.

There were, however, differences in baselines and gains between the two classes. Given the school's test-based splitting of the classes, that could be anticipated. One class may have students who test well, and the other had students who do not produce as high of scores on written forms of assessment. Each teacher also customized and adjusted instruction for their respective students. For instance, based on our video records, the modifications in Class B that Ms. Hayley initiated often involved remedial computational work. This may have limited opportunities for Class B students to do the same kinds of activities as Class A. We do note that the classification and sorting of the students was a choice made by the school and may have involved tendencies to characterize certain performances as less valued than others in school-designed conditions (McDermott, 1996).

With those caveats known, it is worth noting that the largest gains for Class B were in the areas of Data Display, Informal Inference, and Metarepresentational Competence, constructs which place more emphasis on representing and interpreting data than on computational ability. This suggests that the overarching support for interpreting data—the continual push to remember and reflect on what experiences and activities were represented in a set of data—may have played an important role. This was a key part of each day's lesson by design, and that did happen regularly across both classes.

Class A showed the highest post-test performance on all these constructs. However, despite this high performance, Class A showed no significant improvement in Metarepresentational Competence. The same amount of time had been spent on the associated lessons in both classes, and the quality of instruction seemed generally comparable, although future analyses may uncover some undetected difference in lesson enactment. Beyond that, there could still be other reasons for the lack of significant change in Class A. For one, the assessment instrument only had one item related to metarepresentational competence on each of the assessments.

Another potential reason for the limited change in metarepresentational competence in Class A could be the students. We may have observed a kind of ceiling effect in that the students in Class A already demonstrated higher levels of normative performance on metarepresentational competences when the pretest was given that were developmentally appropriate. In the original study of metarepresentational competence (diSessa, Hammer, Sherin, & Kolpakowski, 1991), the population was identified as an advanced group of sixth grade students who were able to "invent graphing" over the course of a week. It may be that by sixth grade, the broad suite of abilities related to use and understanding of representations has already matured to this point for some students of that age range and specifically those students had been presorted into Class A.

Finally, while the expectation for scoring to be based on a written rubric and for that to be verified with reliability checks, we also note also that raters were not blind to which were pre- and which were post-tests. It is possible that the raters could have skewed toward being more generous in the post-test scoring given the pretests and post-tests were labeled as such and had correspondingly labeled rubrics for distinct question sets.

Qualitative cases of data discussions

The quantitative data described above can inform the extent to which the designed unit was able to achieve some of its intended learning goals with the range of students who participated in this unit, but those data do not necessarily illustrate whether nor how working with their own quantified self data could be supportive of student learning. The cases presented in this article, taken together, were selected to provide some initial evidence that students can and will productively reconstruct memories of their own wearable device data when given the opportunity. The third case in particular demonstrates the proficiency that students have developed in thinking about canonical data representations (box plots) in terms of their reconstructed memory and routine familiarity with the measured physical activities.

A second reason for selecting these cases was to illustrate the collective aspects of memory reconstruction in service of activity data inspection. Our particular design and research context was in a classroom. When discussing our theoretical perspective above, we stated that while remembering is often presented as an individual reconstructive activity, remembering in a classroom can actively involve multiple individuals. Different students will make contributions that can add to or otherwise redirect others' data comprehension processes by making different aspects of a remembered activity more or less salient. Each of the three cases exemplifies the kinds of interactions that took place around activity data that could be supportive of any learning gains shown in the pre- and post-test measures.

Finally, we selected these specific cases to emphasize implications of leveraging general and specific remembrances when interpreting activity data displays. These map onto the different emphases related to embodiment perspectives. General remembrances were for activities that are familiar and understood, through personal experience, as involving different levels of exertion and demand. These were not bound to the details of specific episodes. Specific remembrances go to specific moments and situations that were previously encountered.

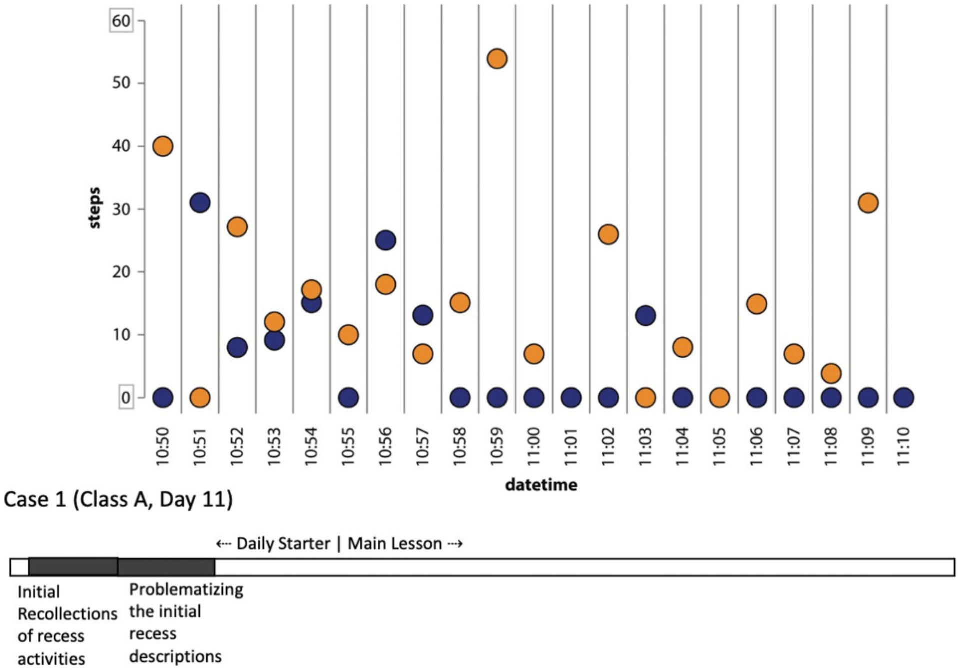

Figure 5. Julius' and Sophia's recess data. Note. Orange appears as the lighter dots, blue appears as the darker ones. The bar at the bottom shows when the portions of the associated case study took place in the class period.

In the first case, the emphasis was on general remembrances of different routine activities, as the students involved in that conversation talked about features or tendencies associated with people or with particular physical activities but not situated in precise times nor sequences. The second case situates data interpretation based on remembrance of specific events. There, the alignment of what is remembered with what data are shown suggests a very direct and specific interpretation of data where unusual moments and sequences of events for a particular day are important for guiding data interpretation. The third and final case involves some of the same aspects of both general and specific remembrances demonstrated in the prior cases. However, its purpose is to demonstrate the proficiency that students have developed in thinking about canonical data representations (box plots) in terms of their physical activity.

Case 1: Interpreting data from general remembrances

The first case illustrates how general remembrances are instantiated in classroom discussion. In a classroom discussion, it is important to consider that many students have developed a generalized sense of what happens during various physical activities that are discussed in the classroom, and this is mobilized through different testimonies of what produced certain patterns of activity data. These different accounts can be debated and synthesized to develop a group consensus that has more detail, and potentially more validity, than what could be accomplished by a single student alone.

The case below comes from the 2nd week of the unit as enacted in Ms. Bryson's class (Class A). At this point in the unit, the students had determined that the morning break (20 min of outdoor recess) was the most interesting part of the day for them to examine. Recess data from two students, Julius and Sophia, had been projected for the class to see (Figure 5). Each minute of recess was represented with a circle in *TinkerPlots*, and the students were represented by the colors orange and blue. All minutes were shown sequentially.

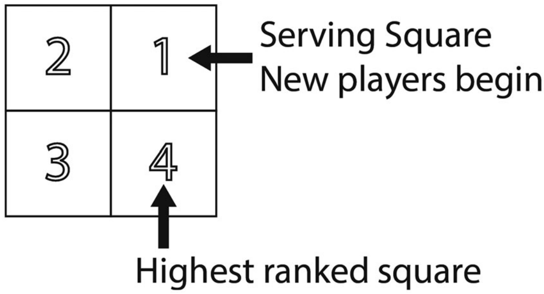

Figure 6. A typical four square court and the main player positions. Note. Players enter the game in the serving square and advance counter-clockwise to the next open square by getting players in higher ranked squares out. If the player in the highest ranked square is knocked out, they return to the serving square rather than going to the end of the line of waiting players.

The class was told that the data came from Julius and Sophia, but they were tasked with figuring out who corresponded with what color of data. This episode began after the teacher asked Julius to tell the class what he did at recess on the previous day. During these excerpts, all students remained seated and deictic references were made from afar. The excerpts below were almost all continuous but are broken into "scenes" to aid the reader in following the class discussions. In interpretive summary sections, specific line numbers of transcript as presented in this manuscript are referenced.

Initial recollections of recess activities

1. Julius:I think I, yeah, I played Four Square and I- I wasn't very good at it but I always play in the other squares, other than Mason and Caden's squares.
2. Ms. Bryson:I don't know what that means.
3. Julius:It means-
4. Mia:In the nicer squares.
5. Julius:The nicer squares, where you don't get out as fast.
6. Ms. Bryson:Okay. But what effect does that have on your data?
7. Julius:I don't know. Maybe it doesn't have as big of spikes.
8. Ms. Bryson:Interesting observation. Sophia, tell us what you did yesterday at break.
9. Sophia:We just walked around and talked most of break.
10. Ms. Bryson:You just walked around and talked most of the break?
11. Jed:Somebody was sitting a lot

Julius began with the statement saying he believed he had played Four Square (Line 1), a playground game that involves four players standing in each of four roughly 8′ x 8′ connected and ranked squares. In the game of Four Square (Figure 6), each player must pass a playground ball that bounces in their square into another player's square such that it either bounces in that player's square or touches the other player. Failure to successfully pass the ball leads to the player leaving the game (known as getting "out") and standing at the end of the line of people waiting to play. The players ranked lower than the out player advance to higher-ranked squares. The person at the head of the line enters the game in the lowest ranked square, serves the ball, and play continues.

Julius was uncertain about specifics related to his previous day's Four Square games. He had not been told beforehand that his data were going to be examined the next day, and he began with remembering what activity he had done for his recess (Four Square). He was quick to add that he did not play in Four Square courts frequented by two other students, Mason and Caden (Line 1). At that point, Julius's account of recess had surfaced general knowledge related to recess and knowledge about his own activities during typical recess. This included: 1) Julius was not a very skilled Four Square player; 2) there were differences in difficulty across Four Square courts; and 3) two students in the class, Mason and Caden, were associated with the more difficult and competitive courts. This was conveyed through Julius's use of "always" (Line 1) to imply some regularity to his behavior and court selection. No specific part of a particular Four Square game was being introduced.

This description of recess activity, however, was unfamiliar to Ms. Bryson (Line 2, where she asked for clarification), who then was told that the squares Julius played in are considered to be the "nicer" squares (Lines 4–5) where students play less aggressively. Ms. Bryson's unfamiliarity with this designation of nicer squares was noteworthy in that it was a stark instance of how talk of students' routine physical activities positioned the students as experts on what things happen during the morning break. Here, Ms. Bryson needed to ask a question about a routine bit of playground knowledge that she did not already know. With respect to our account of remembered experience as it unfolds in the classroom, general remembrances need not be shared by everyone in the group. Ms. Bryson had no direct experience with playing Four Square with these students, so she was already limited in what she could contribute. However, what was most important was that those who were involved in reconstructing remembered activity could treat some of the surfaced information as shared. In a classroom setting, where many experiences are shared by students, it is likely that general knowledge about recess activities will be shared by some subset of individuals. In this interaction, Mia had some of this prior knowledge and could contribute (Line 4). While not necessarily flattering to Julius, Mia's contribution was not challenged but was, in fact, validated and taken up by Julius (Line 5).

Ms. Bryson then returned focus to the data and data displays by asking what effect Four Square play would have on the data (Line 6). Julius responded with some hesitation and uncertainty, eventually suggesting that the data display would not "have as big of spikes" (Line 7), implying that his recess could not be represented by the blue data (the lack of objection by any students, and Julius specifically, following Line 13 below is taken as evidence he had been referring to the blue line). At a minimum, this was evidence for us that Julius had been considering some of the display's features and how they would correspond, after some initial prompting from Ms. Bryson, with his gameplay. As our perspective on remembered experience is tuned toward how students make sense of data representations depicting their own prior activity, this exchange is at least a baseline from which potential features of the representation begin to get mapped to what is remembered even if clarity had not been established on what precisely a given data display was showing.

Sophia was then asked to publicly share what she did during recess, and she responded that she was part of a group that "walked and talked most of break" (Line 9). We characterize this type of general description as a *gloss*. In this example, walking and talking are routine and general bodily activities that take place at recess. No additional detail is necessarily needed and that level of generality is deemed sufficient and reasonable as an information contribution. It would be comparable to when someone is asked how their day went, and the response is vague and general, such as "fine, I just went to work and then came home." In our review of data, we consider glosses from students that are not immediately self-amended to be highly suggestive of moments of general remembrance rather than specific ones.

Following the teacher's confirmation of Sophia's gloss, Jed chimed in, without being called on, with "somebody was sitting a lot" (Line 11). Although this was just a fleeting comment, it was an interesting one from our theoretical perspective. Due to the lack of distinct inflection or accenting of "somebody," this appeared to be a genuinely new observation Jed had made based on the data rather than a remark to implicate a specific person in the class as having sat a lot. If that interpretation is correct, then from looking at the features of the projected data, Jed had made some determination that the activity that was being depicted must have been sitting. This was because there were several minutes in the blue data with zero steps taken in each of those minutes. At this point, Jed was making a connection between what was shown in the data display and a familiar embodied (albeit stationary) activity that could generate such data. However, it was unclear to us if Ms. Bryson heard that outburst, and the class discussion proceeded without a formal acknowledgment of Jed's statement.

Problematizing the initial recess descriptions

The utterances below continue immediately after the previous utterances.

12. Ms. Bryson:Who do you think is who? [half of the class raises a hand] Abby?
13. Abby:Sophia is blue and Julius is orange.
14. Ms. Bryson:What are you basing that on?
15. Abby:Because me and Sophia were just sitting around talking, not walking, not running around or anything. Just lollygagging around.
16. Ms. Bryson:Were you sitting?
17. Sophia:Kind of.
18. Ms. Bryson:What does that mean?
19. Sophia:I guess we would sit for 1 minute and then we'd stand and walk around.
20. Ms. Bryson:Okay. Is there anyone who disagrees with Abby?
21. Avery:I don't know, but I would think that Julius would go and then get out [of Four Square] and stand in line and not really move, and then play and then get out again.
22. Julius:Four Squares don't have very big lines. The ones I go in don't have very big lines.
23. Avery:That's why I'm kind of questioning myself
24. Julius:So they have a one to five-person line.
25. Avery:Abby is probably right. That would mean they only stand in line for a minute or something.

In this excerpt, Abby posited that Sophia must have been represented by the blue dots (Line 13) because she (Abby) spent recess with Sophia. According to her, they were "just sitting around … lollygagging … " (Line 15). This was again a gloss of remembered activity, although it is worth noting that this came from a different student. Sophia still maintained that there was some walking (Line 19), although the doubt introduced by Abby was enough to prompt Ms. Bryson to seek additional information (Line 20). Through class discussion, glosses can become opportunities for other speakers to push for more precision and more specifics in what students can remember about their prior experiences. Based on our familiarity with the broader video corpus, it appeared that Ms. Bryson did that often. This is an opportunity to actively encourage collective remembering.

However, rather than continuing to discuss Sophia's walking, Avery refocused the discussion on Julius' activities and how they could correspond to the displays. She speculated Julius would be in line a lot, which would not involve a lot of movement. This implied Julius' recess was represented by the blue data. She added that his play would be fairly brief when he was active in the game (Line 21), hinting that Julius was known by others in the class to be mediocre at Four Square. There were no specifics of when he would be relegated to standing line nor how often this would happen, suggesting a speculation based on general awareness of how Julius played in the "nicer" squares. Julius objected to Avery's suggestion and tried to make the case that even if he was in line a lot, the lines for the squares he was in were relatively short ones (Lines 22 and 24). Again, that was a general feature of the games and the court where he played rather than recall of a specific moment standing in line.

At this point, two claims derived from general knowledge related to recess were being considered. One shared, general claim was that Julius was mediocre at Four Square. The implication for understanding the data was that he should have several moments of little or no movement (e.g., while he was standing in line). The other claim was that the Four Square courts where Julius played had lines that moved quickly, so Julius would actually have had more opportunities to be moving. Here, general knowledge about Four Square, Julius' play, and how play progresses in certain courts were invoked and became the basis for some conflicting speculations (Avery's claim first, then Julius' counter claim second). In this moment, we can see general remembrance as being tied to the class' shared understanding. That is, knowing about Four Square and how students play are general in that it is understood as known by multiple people and not connected to a specific remembered moment.

Continuing with the excerpt, we see that Avery accepted the validity of the second claim, in which Julius played on courts that generally had short and fast-moving lines (Line 23). This was made more evident when Avery stated that she agreed with Abby's assertion that Sophia and Julius were represented by the blue and orange dots, respectively (Line 25). She backed this up by noting that the shorter lines "mean they only stand in line for a minute," which would preclude the long periods of zero-step minutes present in the blue data.

Considering how general remembrance and knowledge were involved in interpreting the data, we note that 1) multiple parties played a role in reconstructing memories of recess and 2) even without recall of specific events or moments, consideration of different aspects of Four Square as typically played led to some convergence in interpretation of the related data. In verbal contributions, one way that such knowledge would be expressed was as a gloss. Another way was through the use of modifier words that do not situate the activity in a specific time, such as "always" (Line 1). General knowledge is locally shared, although it may not be shared by all parties involved in the conversation. Indeed, Ms. Bryson's lack of shared knowledge may have been beneficial in terms of positioning the students as having more expertise in this area than her. This could have contributed to the extended talk about what Sophia and Julius were doing in this episode.

With respect to statistical ideas, this was an early activity in the first few days of the unit. It was not directly tied to any topic in the assessments. However, it was still important to establish ways of thinking about and describing data in the unit. Students were just getting familiar with the idea that their quantified self data records were being segmented by steps per minute and that these values varied but also came from specific moments. The students were demonstrating an understanding that each case (i.e., each data point or data record) had magnitude and different attributes that were associated such as time of day and individual who was being measured. They were also seeing, informally, variation even though it was ordered by time. These were important ideas for thinking about distribution later in the unit. They would later review those same data points collapsed over their quantities to see variation and distribution. Furthermore, these conversations were inviting identifications of overall differences between distributions of data. There were times later when the data from Sophia and Julius were represented so that difference between them was more visible. This links to informal inference, which was one area that students showed improvement on through this unit.

Case 2: Interpreting data from specific remembrances

As the first case was illustrative of heavy reliance on general remembrance, the second case emphasizes specific remembrances. Here, specific remembrances are tied to specific events or activities. They can be expressed as, but are not necessarily, mentions of unusual events or specific narrative sequences of what was experienced at a given time or place. For the interpretation of activity data, we were interested in how specific remembrance supported students' data interpretation.

The second case also comes from Class A. It took place the day after the first case and again during the Daily Starter activity, which was similar in its presented classroom objective: the students were to again identify whose morning recess was represented in a set of data, although this time there were three students' data projected instead of two. The students were Keri, Caden, and James. None of these three students knew their data were going to be projected when it was produced the previous day. At the beginning of the Daily Starter, the class knew the data were from those three students, but did not know which color (Red, Blue, and Yellow) corresponded to which student. James was unexpectedly absent from school the day of this activity, and thus was not discussed. Like the first case, this case is broken into scenes to facilitate reading.

Keri asserted the yellow data were her data

The class session began with Ms. Bryson stating whose data were projected, observing that James was absent, and asking the class to identify which student corresponded with which data. She stressed the importance of justifying claims, but Ms. Bryson offered no examples or criteria for justifications. The case analysis below begins after Keri raised her hand and Ms. Bryson called on her to speak.

26. Keri: I was playing jump rope and I am yellow.
27. Mrs. Bryson: Why do you say that?
28. Keri: Because I was waiting and walking around while Ms. Wilson [another teacher] was getting the jump ropes. Then I was jumping. Then I swung the rope. And I don't think that would be very high because we didn't move.
29. Ms. Bryson: So you don't think that would be very high.
30. Keri: Not as high.
31. Ms. Bryson: So you are going to say that you are yellow?
32. Keri: Yes
33. Nick: So wait, when you were waiting for the jump ropes to be cut, did you get, like when you were actually jump roping, was it the last half of recess? [Keri nods] OK, I think you're right.

Initially, Keri stated that she had been playing jump rope and that she believed her activity was depicted with the yellow dots (Line 26; Figure 7), in a declaration that is akin to the actor-perspective phenomenon described in Roberts and Lyons (2020). When asked for justification (Line 27), she recalled her experience of waiting around while Ms. Wilson got the jump ropes (Line 28). That is a specific remembrance in that it was more temporally based and involved a specific individual doing something that was situated in a time and place. The temporality was partially indicated by her subsequent use of "then" for the next parts of recess (jumping, followed by swinging the rope). It also directly implicated Ms. Wilson as a specific actor at a specific time (i.e., the previous day, at the beginning of recess). Following that sequence, Keri stated that the latter set of values, when she was swinging the jump rope, would not be "very high" because she did not change locations. In making this comment, she appeared to be referring to the observation that in the last columns (Figure 7), the yellow circle was positioned below the blue and red ones.

Ms. Bryson made an explicit statement to make clear the inference (Line 30), and that was followed by some confirmation that Keri was indeed claiming that her data were depicted with the yellow circles (Lines 29–32). Following this, another student in the class, Nick, interjected (Line 33). Of note, he first stated the specific remembrance of waiting for the rope to be obtained. However, he interrupted himself and instead asked if the time spent "actually jump roping" was in the last half of recess. Keri confirmed that was the sequence, and that was enough for him to

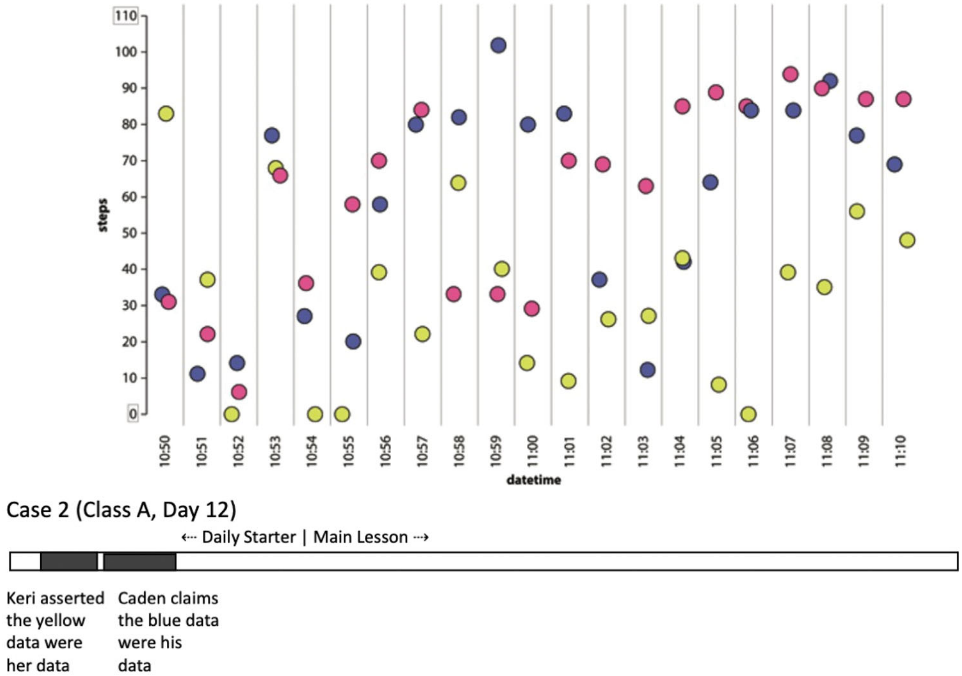

Figure 7. Keri's, Caden's, and James' Data. Note. Keri, Caden, and James were represented by the yellow, blue, and red dots, respectively. In the leftmost bin, yellow appears with the highest value, then blue, and then red is the lowest value dot. The bar at the bottom shows when the portions of the associated case study took place in the class period.

offer his agreement. This transaction was a clarification in what was remembered and surfaced for the class.

What we interpret as having taken place was that ostensibly, the lower values in the data plot that were in the earlier part of recess were connected to the specific event of waiting for the jump ropes to be obtained. Then, a process of segmenting the specific sequence of events was offered as a way to think about the data as depicted, with more active values (higher values) being in the latter portions of the representation. Keri's rationale seemed to also involve the observation that the yellow circles were below the other circles, corresponding to her activity levels not being "very high" since she did not move. When Nick participated in this, he accepted the specific remembrance for what Keri did the day before during recess. He tacitly accepted the sequence of events as presented, but he segmented recess into halves, with actual jumping taking place in the second half. This implied that waiting and cutting jump ropes took place in the first half. Those activities aligned adequately with how Nick was reading the plot.

Caden claims the blue data were his data

Caden, another student whose data were projected, had wanted to share his thoughts with the class as well. Recall that Caden was one of the two boys mentioned during the first case for playing Four Square in the challenging courts. This exchange took place immediately after the preceding excerpt, except it had one speaking turn where Ms. Bryson called on Caden to speak. Once Ms. Bryson called on him, he offered the following:

34. Caden:Well, I think I would be blue because at the start no one was really there, so I wasn't doing very much stuff to move around. But toward the end I started moving around in my

square a lot and, uh, when it, like, got to being between 100 and 110, the highest part I think that would be like the part where the ball went way far away and I went and got it or something

35. Ms. Bryson:OK, do you remember that happening in the middle of recess?
36. Caden:Yeah, that happened. It happened a couple times.
37. Ms. Bryson:Did it?
38. Caden:Yeah.It happens every day.

In this exchange, Caden proposed that his color was blue (Line 34). He began talking about specific remembrances, such as the beginning of recess and "no one was really there" and moving around in his own square a lot at the end. Caden also identified a specific point in Figure 7 that had the highest value "when it got to being between 100 and 110" and tied that to a specific incident, "where the ball went way far away" and he went to retrieve it. This again suggests a specific remembrance of having to travel far at a particular time. The specific timing of this was questioned by Ms. Bryson (Line 35), and Caden confirmed that this indeed happened at the middle of recess (Line 36). He also added that in that recess, the ball going far away happened "a couple times," which would seem plausible given the blue data. That specific information was checked by Ms. Bryson (Line 37) and verified by Caden (Line 38). Interestingly, he generalized the experience of a ball getting away from the game by saying "it happens every day."

The identification of specific moments was not requested, and thus this excerpt was fortuitous for illustrating specific remembrances. The specific remembrances that Caden relied upon were the beginning and end of recess, the time he had to run and get the ball in the middle of recess, and that the act of retrieving the ball happened two times. As this has likely happened with some regularity, the statement of "it happens every day" is unlikely to be the moment at which a specific remembrance becomes a general one in how he is storing that information. However, in classroom discourse, its generality appeared to serve as boosting his credibility in stating that those incidents happened during the previous day's recess.

Ultimately, one of the distinguishing qualities of this transaction compared to the first scene in Case 2 was articulation of certain embodied events that stood out in each youth's memory. Another was the presentation of time and sequence, such as what happened "at the start" (Line 34). The interpretation process involved some segmenting of the projected data representation and connecting certain registered features of a representation with a particular event from recess. The teacher played a key role in pushing for elaboration and verification from each student who identified a set of points with their own experience. Had Ms. Bryson not asked for why Keri thought her jump rope data were the yellow set or if running after the ball happened to Caden the day before, further inspection and alignment of specific recalled activities to the data representations may not have happened. To a lesser extent, Nick also served as a verifier, although it may be he was just trying to follow along rather than press Keri for more detail and support. Regardless, these pushes enabled more alignment of time portions to remembered events, making chronology more important to how these data were interpreted.

A challenge that Caden addressed was explaining causes of variability, which he did with recounting a time he had to run after the ball. With respect to other statistical ideas, much of what took place in the first case also took place for the second one. The students were gaining more familiarity with the data points. Multiple attributes were being seen, and new comparisons were getting made. Presumably, maximum values were being recognized, which are implicated in computing ranges. They also can represent outlier values, and other learning activities with student data explored the effect of outliers on measures of center.

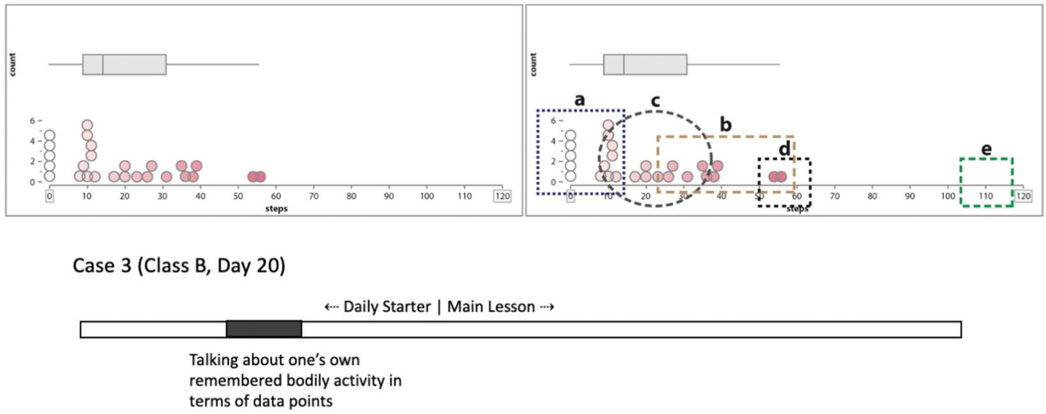

Figure 8. Density plot of Lisa's recess data (L) and Density plot of Lisa's recess data (R). Note. Regions are marked based on where students were pointing and motioning while talking about it (r). The bar at the bottom shows when the portions of the associated case study took place in the class period.

Case 3: Connecting remembered activities to a canonical data representation

The final case came from Class B. Whereas the first two cases primarily isolate and emphasize general and specific remembrances, this final case draws on both general and specific aspects of what is remembered and extends those ideas into a new imagined situation. It is in transactions like the one presented here where we believe additional productive and exploratory student interaction around canonical statistical ideas and representations takes place in the context of students' quantified self data.

This case comes from the 20th day of the unit, meaning it was almost 2 weeks after the two cases above in addition to being from a different class. The day of the selected episode, a set of five students had been identified to show their preceding day's recess data in front of the class. However, rather than view a time-ordered data representation as had been done in the first two cases, the activity data were shown in a form of density plot (the unit had been designed to transition from reading activity data ordered by time to reading data in different forms of aggregate and density displays). In the current data display format, the placement of individual dots along the horizontal was determined by number of steps taken. Dots with the same number of steps as another already plotted were stacked vertically. As the morning recess break was only 20 min, the density plots were not heavily populated. Still, this was a vehicle for looking at and discussing ways of representing variability, including the box-and-whiskers plot convention. *TinkerPlots* was capable of automatically rendering a box-and-whiskers plot, so the students saw both the density plots and the box-and-whisker plots simultaneously (Figure 8, left). In a box-and-whiskers plot, the outer lines of the box correspond to the 25th and 75th percentiles and the line splitting the box corresponds to the median. The outer whiskers correspond to the remaining values and extend to maximum and minimum values in the density plot.

Talking about one's own remembered bodily activity in terms of data points

For the first 10 min of class, the five students whose recess data were being shared each came to the front of the class and described how they would compute their interquartile range (IQR, i.e., the difference between the 25th and 75th percentile values) and then quickly returned to their seats. Mrs. Hayley then asked who had the smallest and the largest IQRs of the five students who presented, with many students calling out Lisa as having the smallest. Mrs. Hayley asked the class why it was smaller, and Lisa promptly returned to the board and stated the following:

Figure 9. Density plots of Adrian's recess data. Note. Density plot of Adiran's recess data (l) and density plot of Adrian's recess data with region marked based on where Lisa (see Lines 27–29) was pointing and motioning while talking about it (r). The bar at the bottom shows when the portions of the associated case study took place in the class period.

26. Lisa: Actually, what we really did, for the first beginning part [of break] like we started hanging up posters for art and then we kind of stopped and started talking about other things and then we started hanging more posters up. So, I think when we were hanging up posters was more around here [motions hand over Region "a" in Figure 8, right], and then we stopped and were walking around and talking was over here more [motions hand over region "b" in Figure 8, right].

Lisa offered a specific remembrance that she and a friend had spent their break largely indoors, helping to hang up posters inside the school building. That was followed by talking with her friend, and then later hanging up more posters. While these are seemingly uneventful activities for one's recess, they stood out for Lisa. Beyond just sharing what she remembered as that particular recess' events, she then proceeded to identify how the various dots on the density plot likely corresponded to those activities ("hanging up posters was more around here … "). It is possible she was referring to the plot as being time-ordered. However, it could also be that she had parsed the data display and the specific remembered activities into low values and high values. Regardless, if it was one interpretation or another or even a hybrid of both,[5] this action was noteworthy both because it was unsolicited and because these circles could be seen as just a set of points on the screen but were made into a meaningful representation of her recess experience. That suggested a usefulness in her remembrance of the embodied activity that was being mobilized into how she thought through and with a display of the data.

Comparing data from self to data from another person

Following Lisa's speaking turn, she returned to her seat and the teacher asked the class who had the largest IQR of the five students that day. That student was Adrian. Adrian was then invited to the front of the class and asked to say what happened with his recess to produce such a large IQR. With some prodding from the teacher, Adrian provided a short gloss of his recess (Figure 9, left),

[5]Based on what unfolds starting at utterance number 29, we believe that she was interpreting the display correctly and thinking of hanging up pictures as a minimal exertion activity compared to walking around in the later part of recess. However, given the nature of talk in classroom discussions, we acknowledge it is not possible to know for sure.

stating that he and his friends were "just running around" and that "sometimes [he] stopped a little bit to get a drink." Adrian did not indicate the specific moments when these events happened.

Following Adrian's brief time speaking in front of the room, the class was again posed with the question as to what made the IQR for Lisa and Adrian so different. A few students offered their thoughts about the data. Then Lisa raised her hand and returned to the front of the class. Lisa's data display was again projected, about five minutes after utterance 26 (Figure 9).

27. Lisa: Since mine's like more over here, like to the 10s and 20s [motions over Region "c" in Figure 8, right], I took less steps because we were just hanging posters and not very active. And can you go to Adrian's?
28. Ms. H: Sure. [changes the projected data from Figure 8 to Figure 9]
29. Lisa: I think he was like, playing more with his friends [motions over Region "a" in Figure 9, right], like whatever they were doing at break. More running more often, and not walking like we were. He was just like running some more.

Lisa's main point was that the activities that involved more minutes between her and Adrian had different steps. Lisa had been largely hanging posters, and that region where she had motioned over earlier was not as far to the right side of the data display (Line 27). In Adrian's case, he had been running and thus produced higher values that led his distribution to have a much larger IQR (Line 29). In considering her exact words, Lisa's use of "whatever" in describing what Adrian and his friends were doing at break suggests that she had some distance from what Adrian had been doing. She did not know specifically what he had done and relied instead on his brief summary.

These excerpts show how knowing what produced the data can be used to make sense of and make informal inferences related to distributions of data. The next segment, which came three minutes after, demonstrates how thinking about data in terms of recess activity was involved in thinking about data representation and measures of center.

Imagining new activity and how it would affect the box-and-whiskers plot

With Lisa's data projected, Samantha asked to go to the front of the board to share something she had realized amidst this class discussion and comparison of recesses and IQR. At this moment, the teacher asked how one would change the IQR in the data.

30. Samantha: If you have, these are the same [motions over Region "d" in Figure 8, right] but you get like five extra minutes of recess and for those five extra minutes Lisa went outside and ran like really fast, like going downhill, and ended up somewhere here [motions over Region "e" in Figure 8, right]. And if like then four dots were over here [Region "e"] and one dot was over here [on the left half of the plot, Region "a" in Figure 8, right], it would shift over like four dots. So this [the median line] is like right here [motions just a little to the right of the existing median line], and these would most likely be a little more stretched out [motions her hand from the 75th percentile line rightward slightly].

Samantha's utterance and gesture essentially demonstrated how the median had some sensitivity to outliers. Her point was that if new, much higher valued data points were added—as a result of running, like what Adrian had done during his recess—to Lisa's data set, then the median and the 75th percentile line would be affected slightly. Her count of how many dots would be involved was not exactly correct, but her gestures and use of language (e.g., "little more stretched out") both communicate a modest effect from the outliers. This bears resemblance to the gesture-speech mismatch phenomenon where correct gestural reference can precede the articulation of correct declarative verbal statements (Church & Goldin-Meadow, 1986). More importantly, she was reasoning about how the data (the dots) were connected to a canonical data display (a box-and-whiskers plot) by way of considering what she had already understood about the physical activities that produced these kinds of data.

This was noteworthy in that Samantha was not reading existing activity depicted in Lisa's projected data (Figure 8). Rather, she proposed a hypothetical situation with five more minutes of recess. That additional recess time would involve running really fast, which she described as being downhill (their playground had a small hill on it) to create data points that would be on the far right of the distribution. Characterizing this as running downhill also reflects that Samantha had harnessed both an embodied and quantitative intuition of what kind of recess activity could produce data in the territory of 110 steps per minute. (Interestingly, she included one point on the left part of the plot, perhaps to reflect the minute needed to get situated or to walk from the art room to the playground. This was not elaborated.)

Following this, the students engaged in an expanded discussion about how different box-and-whisker plots could be produced from different hypothetical recess activities. This discussion led to suggestions from students to try to produce those hypothetical plots using that day's recess activities. In all, what this suggested to us was that the students were connecting their prior remembrances of what produced the data they were examining with how it could be represented canonically. With respect to statistical thinking, students were engaging in discussions and speculations about measures of distribution and in the box-and-whiskers plots, anchored in descriptions of their own bodily activities. Box plot comprehension is noted as being difficult for students in the statistics education literature, and some suggest reserving that representation for high school (Bakker, Biehler, & Konold, 2005) or using other transitional representations with students (Konold, 2007). In this case, younger students were successful in interpreting and comparing box plots and in predicting how they would change when more data were added.

Discussion

This article has been based on the observation that experience in producing data appears to support one's ability to interpret and make sense of data later. The increased availability of new portable and wearable sensor technologies seems to support that kind of work. We have sought to leverage those technologies and show that in the context of a designed unit, students can engage in substantive sense-making and data interpretation with wearable device data. They can demonstrate progress in areas such as elementary statistical reasoning, which is foundational for a prospective K-12 data science education. We have borrowed a distinction between experiencing selves and remembering selves and observed that remembering selves may not capture everything encountered by an experiencing self. However, enough can still be reconstructed within a classroom such that students can work with data that come from the world of the quantified self.

We believe the consideration given to these three selves (experiencing, remembering, and quantified) echoes and adds to ongoing concerns within the learning sciences. First, with respect the experiencing self, there is the obvious point that students already have a multitude of experiences in their daily lives that, if taken seriously by educators, can provide fodder for reflection and inquiry. In the cases we have presented above, we predominantly discussed recess and playground games. In focusing on recess, our goal was to privilege the routine and daily activities of students as resources where youth are already knowledgeable and have pools of shared understanding. Playing Four-Square or eating school lunch are routine activities that, given appropriate design decisions, can allow for scientific or statistical practices to surface. In essence, we are advocating for a revaluation of routine youth experience as a resource for the design of learning activities. This need not be confined to school settings. Families have recurring and routine experiences that could make for productive inspection (e.g., Lee & Dubovi, 2020), as do afterschool programs and other spaces where we recognize learning takes place. In those spaces, when appropriate steps are taken to respect privacy and suit the goals of individuals in those settings, we see new possible self-reflective learning opportunities.

At the same time, we acknowledge that what is recalled from such experiences relative to what was encountered is inherently limited. This is the constraint associated with the remembering self—we will not remember everything we experience. Instead, we build and rely upon general knowledge about our experiences and tend to be most vivid with just a few detailed specifics. While this limitation might be seen as an impediment, we argued by way of example that this constraint of memory can be used in productive ways. Through the cases presented above, we have argued that recalling specific, unusual, embodied activities can aid a student in memory reconstruction for sense-making and data interpretation. The participation of other students and the teacher who were also inspecting the data contributed to the reconstruction as well.

In this paper, we believe that specific remembrances can also serve other important roles. For instance, students seeing an unusual feature in a representation of a routine activity that is tied to a specific memory can create a space for generating new hypotheses that could be tested, as in Case 3. In another paper (Drake, Cain, & Lee, 2017), we have examined how a specific remembrance that a student had about a recess period involving her jump rope activity led to discrepancy in how a set of data were being interpreted by students in her class. This then led to the entire class devising, implementing, and then critiquing a series of tests over multiple days to determine whether jumping was registered by the wearable device as a step. The need to further experiment and refine the procedures of experimentation was driven in large part by students recognizing the limitations of their recollections. This led to the realization that they needed to be more deliberate and planful in their collection of new data once uncertainty had been established.

That kind of problematization around data and measurement (i.e., wondering if jumps get counted as steps) was facilitated by the acceptance and strategic use of quantified self data. Used in this way, these accrued quantified self data have the potential to help us develop new insights not just about our behaviors but also about the nature of data and quantification. For instance, we may distinctly and confidently remember we did something that was inappropriately counted by a wearable sensor, such as mountain biking on a bumpy trail and being credited with thousands of steps that were not actually taken during that time. In that situation, our awareness of what we had actually physically done can provide an opportunity for circumspection about what the wearable device is actually measuring. This critical frame—which invites a form of ground-truthing of data (Taylor & Hall, 2013)—can then be extended to other instruments and their outputs. Referencing the above jump rope example again (Drake et al., 2017), the question of whether jumps counted as steps raised precisely that issue. The students ultimately came to recognize that the wearable devices actually detected large, high-impact lower body movements. Jumps were counted as steps by the device even though the students would not consider a jump to be ontologically identical to taking a step. Having the opportunity to see those discrepancies in what is encoded as data and what we believe we are measuring can be valuable if managed properly in a classroom setting as it could prevent students from viewing measurements from a stance of naïve realism. Stated differently, our measurement records in research are taken as proxies for some object, feature, or action we have named in the world. The measurements themselves are ultimately just referents to those entities but are not the entities themselves. This is a fundamental epistemic concern that can be realized and encountered in the face of abundant and automatically collected but also already familiar data. In using wearable technology that records data passively, this is one way to also productively trouble what are already given data ontologies (e.g., a step) when other research designs for student data collection would suggest students engage in the creation of those ontologies (e.g., determining what is a worthwhile measurement unit).

Ethics, privacy, and data responsibility

One might take our overall position as advocating strongly for the quantified self in education. We would consider our position to be more nuanced. First, we recognize that quantification of our lives can be a cause for concern and consternation due to as-yet unresolved issues of privacy, data ownership, and implications of identity (Pangrazio & Sefton-Green, 2019). For this particular project, we took measures to keep data protected and obtain informed consent for these learning activities that would involve student data. When a student is uncomfortable with sharing their data, we advocate prioritizing student autonomy and choice. If a student does not want data to be collected or for collected data to be shared, our feeling is that those data should not be collected and that very clear and understood measures are taken to honor and respect the wishes of the individual of whom the data are about. Prioritizing the preferences and desires of the individual has been a guiding orientation in our work. For example, in Ms. Hayley's class, one student was concerned that from this data corpus of steps taken, their weight might be inferable to their peers. That would not be possible. The student did, however, make explicit that he wanted to collect step data using the Fitbit device and to use their data for the planned learning activities. Given that this student seemed apprehensive enough about the idea that weight could be inferable led us to make a decision to exclude their step data for certain activities—such as the daily starters—so that they would not need to even consider it as being a risk. We intentionally also took additional protective measures such as creating project specific accounts that could not be linked by anyone else to the students who participated in this project.

Since this work has been undertaken, our relationships to data have become even more complicated (see Pangrazio & Sefton-Green [2019] for further exploration and examination of how we could conceptualize those relationships). We are not currently positioned to offer a resolution to the various societal and individual concerns that are legitimately emerging with respect to personal data, but we do advocate for an intentional focus on helping students learn more about data so that both the risks and consequences of constant personal "data-fication" can be better understood. With respect to conducting educational research or design with data being collected passively about student activity (physical or otherwise), we offer two ethical recommendations that should be highlighted as part of or beyond what an ethical review board would consider. The first recommendation is to devise ways to ease awareness that data collection is likely in progress. The students should be able to, without an unreasonable amount of work on their end, infer that immediate continuous data collection is likely. This does not mean that students need to be interrupted and explicitly reminded and asked to give new statements of consent each day. That would be needlessly burdensome and intrusive for both the students and the researchers and provide little, if any, benefit to anyone. Rather, if a student were to wonder if data were being collected, they could find some indicator or clue that is likely happening. In the case of a wearable data generating device, that could be simply indicated by the presence of the wearable. We believe that the students understood the device-collected data about their activities when worn and did not when it was removed. Additionally, they were regularly working with data each day during this unit and discussing what they did to collect data with the wearables. Those were all indicators that we felt were present and unobtrusive. However, in other situations, like when video recording, that could be signaled by a red light on the camera, by the presence of a researcher standing by or maneuvering the camera, or the lone camera being plainly visible and positioned in a way that looks like it is there to record what is happening. Where ease of awareness is problematic is when there is no disclosure and an excessive amount of steps involved in finding some indicator that data collection may be happening.

The second recommendation we have is to be considerate of agreements with respect to scope of subsequent data use. In our case, students knew we were a university-based research team and that they were part of a study and part of a curricular unit involving their own data. It was understood that the data were to be used by the class for the students to learn certain ideas and practices, but beyond that, the researchers were going to do whatever it is that researchers do with the data. It was taken as given that we had no intention of using the activity data that were collected in order to sell students new products, evaluate their health, or rank their school. Even though one student had concerns that might imply something related to their health (weight) that were not intended, we opted to rely on the student's perception of scope to guide our decision-making.

As we have discussed elsewhere (Lee, 2013), the idealized quantified self model loops the data about the person right back to the person. Some of the concerning data scenarios that are part of our current societal discourse are ones where the data do not ever return, as data, to the person who generated it. Instead, a third party takes the data and uses it in service of their own objectives.[6] We do not claim that such work by third parties is inherently bad. Many experiences that we value and enjoy are enabled by this third-party data arrangement. However, the question that must be considered is whether that scope of data use was part of the explicit or implicit agreement that was made when the decision to allow collection of the data was made.

While we encourage vigilance and attentiveness to risks and possible means of mitigation, we also recognize that quantified self and other genres of data-rich learning, like wearable-based inquiry, are already appearing and growing in prominence. We believe that these should be thoughtfully considered in the learning sciences. When viewed from the side of the advocates and enthusiasts, these developments are exciting because they represent a new palette for us to explore as researchers and designers. Yet we need to develop new ways of thinking about how we comprehend and relate to quantified self data. That motivated us to elevate a view of reconstructive memory as being an important part of data sense-making.

We would also like to note that we do not see streams of quantified self data as inherently superior scaffolds for students to learn statistics or data literacy. However, such data can be incorporated as supports or scaffolds in ways that produce new interactions that we can then better understand and engineer into learning tools and experiences. We designed and examined a situation where the quantified self could serve to mediate between students' remembering selves and experiencing selves in service of learning. It could be, instead, that the quantified self serves to mediate between students and behavioral goals (e.g., getting more exercise throughout the school day), which has been the tendency elsewhere in the quantified self literature (Li et al., 2010). However, that need not be the only mediational role that quantified self can play, and the current study sought to demonstrate and explore that.

Reflections on theoretical contributions

In terms of theoretical contributions, we believe that the work presented here has accomplished a few things. First, it has highlighted some features surrounding memory that have been well established and accepted in cognitive psychology. These include the reconstructive and partial nature of memory. Within that framing, we also cast general and specific components of an embodied experience as being worthy of consideration and applicable to the reconstructive work of remembering. We anchored these reconstructive acts of remembering in the reading and interpreting of specific data visualizations. Because of the increased pervasiveness of data, we are now creating records of all manner of experiences and will be, among other activities, reflecting and reconstructing what events and activities produced those data. Those records are often connected to our embodied experiences, creating new questions and new opportunities for us to explore.

The work here is a mobilization and modification of prior work on remembering given that new reality with a specific focus on how remembering interacts with data representations. In the learning sciences literature, instances of such interpretive and sense-making work around

[6]Sometimes a third party gets the data and retains it but provides access to the data to the person involved.

representations and inscriptions, have been areas of growing interest. Related lines of research preceding the current work or that are otherwise contemporaries include Nemirovsky's et al.'s work on fusion (1998), Roberts & Lyons's work on actor-perspective taking (2020), and Enyedy et al.'s work on liminal blends (2015). The current work differs from Nemirovsky, Tierney, and Wright (1998) in that that work was based in a situation involving a near simultaneity of data representation and experience (whereas we have a case of a day's worth of delay), and from Roberts and Lyons (2020) in that it was tied to the viewers' enacted experiences rather than imagined ones. With respect to the Enyedy, Danish, and DeLiema (2015) work, there are some overlaps in that records of students' own bodily activity are subsequently jointly examined by multiple students. However, their use of play behaviors was one that involved specific content-themed roles, and the play behaviors examined here were ones that took form within the existing role of "student at school."

Our account relies heavily, although not exclusively, on the construct of memory. Others may place stronger emphasis on discourse and semiosis, which we also consider to be important. Regardless of differences in emphasis, there is a common kinship amongst these different projects and analytical endeavors. We all recognize and appreciate that digital traces of human activity have great potential, that bodies play instrumental roles in our thinking and learning, and that interpreting and reading various forms of representation and inscription are quite complex and are more than acts of translation.

While remembering is an anchor in the current work, we do wish to state that this process of remembering is not only good for retrospective inscriptional interpretation. These processes and their products can provide a basis for imagining potential variations of to-be-lived experience. The students in the third case showed how they could reason prospectively in flexible ways such that they could devise expectations for how data representations would change and look different in imagined situations. Samantha, for instance, was able to predict and justify how the data representation would look different if five additional minutes were added.

In addition, we have added to this account a social perspective on remembering. This is appropriate in that many learning activities, including the ones we designed for these classrooms, are social spaces. The many people in those spaces can contribute to a distributed remembering. Expanding the act of remembering beyond the individual is not unfamiliar territory for learning scientists. Some have offered the step of identifying objects that support and augment memory (Norman, 1993), which immediately changes the unit of analysis to person plus a cognitive artifact. Hutchins (1995) has been invoked as another take on distributed memory by arguing that a complex sociotechnical system with multiple actors and multiple artifacts—such as an airline cockpit—may be the appropriate unit to consider as doing the work of remembering. We do not see these instances of distributed memory as being in conflict with what we presently advocate. However, we would situate our version of distributed remembering as being one that speaks predominantly to multiple people having partaken in an ephemeral experience and working together around a sort of experiential residue—the data that are generated—to reconstruct what was jointly experienced or observed. We are also situated in what many would recognize as a canonical learning environment. Given this focus on recess activities as the source of data for this study, the students were the experts and made the strongest contributions to the group process of remembering. However, when one takes this more distributed view of remembering, it is also reasonable to expect that even in activities where teachers are more directly knowledgeable, that students are still empowered to shape what is collectively known and agreed upon.

Digital technology is a prominent feature of this work and others that we cited. This is in part due to growing attention to digital data. However, we believe strongly that analog approaches are valuable and analyzable in the terms we offer as well. For example, the *Dear Data* project (Lupi & Posavec, 2016) has served as inspiration for the first author in their teaching and for others in their educational designs for K-12 data and data science education experiences (Rubin, 2020).

Dear Data involves tracking a week's worth of personal data and then creating a novel hand-drawn visualization of those data to share such as adaptations of the *Dear Data* project. In its original form, those visualizations were inscribed on postcards that were mailed between the two *Dear Data* project participants over the course of a year. For others, those visualizations could become imprints on a t-shirt (Stornaiuolo, 2020). Regardless of the final visualization product, this general model of relying on self-tracked data serves as an effective vehicle for learners to consider how data are made, what are the challenges of representing it, and how data can be relevant and consequential to their own lives. While it has not been referred to in this way, we would maintain that *Dear Data* is at its core, a subtype of quantified self activity that leans more heavily on the esthetic dimension of data representation. It is quantified self because it is made up of data that come from everyday activities and experiences. Ultimately, we believe that the saturation of data in our day-to-day experiences is only increasing. Learning designs like these are going to become more common and ways of understanding the thinking and learning that can take place within them will be needed.

Acknowledgements

Our appreciation and thanks go out to Doug Clark and Mike Horn for enthusiastic early support and advocacy for this paper. Special thanks to Joshua Danish and Noel Enyedy for their editorial work, to the Cognition and Instruction editorial assistants, and to the anonymous reviewers who expressed strong interest in this paper, offered thoughtful suggestions, and encouraged us to include some more bold elaborations and assertions that now appear. Finally, our deepest thanks go out to the participating teachers and students, who are kept anonymous for confidentiality purposes.

Funding

This work was supported by funding from the National Science Foundation (US) under Grant No. DRL-1054280. The opinions expressed herein are those of the authors and do not necessarily reflect those of the National Science Foundation.

ORCID

Victor R. Lee http://orcid.org/0000-0001-6434-7589
Joel Drake http://orcid.org/0000-0003-4480-9154
Ryan Cain http://orcid.org/0000-0003-2472-7872
Jeffrey Thayne http://orcid.org/0000-0002-2777-8969

References

Abrahamson, D. (2009). Orchestrating semiotic leaps from tacit to cultural quantitative reasoning—The case of anticipating experimental outcomes of a quasi-binomial random generator. *Cognition and Instruction, 27*(3), 175–224.

Abrahamson, D., & Lindgren, R. (2014). Embodiment and embodied design. In R. K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (2nd ed., pp. 358–376). Cambridge, UK: Cambridge University Press.

Bakker, A., Biehler, R., & Konold, C. (2005). Should young students learn about box plots? In M. C. Gail Burrill (Ed.), *Curricular development in statistics education* (pp. 163–173). Voorburg: International Statistical Institute.

Barsalou, L. W. (1999). Perceptual symbol systems. *The Behavioral and Brain Sciences*, *22*(4), 577–660. doi:10.1017/s0140525x99002149

Ben-Zvi, D. (2000). Toward understanding the role of technological tools in statistical learning. *Mathematical Thinking and Learning*, *2*(1–2), 127–155. doi:10.1207/S15327833MTL0202_6

Bietti, L. M. (2012). Joint remembering: Cognition, communication and interaction in processes of memory-making. *Memory Studies*, *5*(2), 182–205. doi:10.1177/1750698011404986

Brasell, H. (1987). The effect of real-time laboratory graphing on learning graphic representations of distance and velocity. *Journal of Research in Science Teaching*, *24*(4), 385–395. doi:10.1002/tea.3660240409

Browning, C. A., Harris, C. B., & Van Bergen, P. (2019). Successful and unsuccessful collaborative processes in strangers and couples performing prospective memory tasks. *Discourse Processes*, *56*(8), 655–674. doi:10.1080/0163853X.2018.1541398

Cai, J., Lo, J. J., & Watanabe, T. (2002). Intended treatments of arithmetic average in U.S. and Asian school mathematics textbooks. *School Science and Mathematics*, *102*(8), 391–404. doi:10.1111/j.1949-8594.2002.tb17891.x

Ching, C. C., & Schaefer, S. (2015). Identities in motion, identities at rest: Engaging bodies and minds in fitness gaming research and design. In *Learning technologies and the body: Integration and implementation in formal and informal learning environments* (pp. 201–219). New York, NY: Routledge.

Choe, E. K., Lee, N. B., Lee, B., Pratt, W., & Kientz, J. A. (2014). Understanding quantified-selfers' practices in collecting and exploring personal data. Paper presented at The Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Toronto, Ontario, Canada.

Church, R. B., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition*, *23*(1), 43–71. doi:10.1016/0010-0277(86)90053-3

Cobb, P., & Tzou, C. (2009). Learning about data analysis. In W. Roth (Ed.), *Mathematical representation at the interface of body and culture* (pp. 135–170). Charlotte, NC: IAP.

delMas, R., & Liu, Y. (2005). Exploring students' conceptions of the standard deviation. *Statistics Education Research Journal*, *4*(1), 55–82.

Derry, S. J., Pea, R. D., Barron, B., Engle, R. A., Erickson, F., Goldman, R., … Sherin, B. L. (2010). Conducting video research in the learning sciences: Guidance on selection, analysis, technology, and ethics. *Journal of the Learning Sciences*, *19*(1), 3–53. doi:10.1080/10508400903452884

Diaz, K. M., Krupka, D. J., Chang, M. J., Peacock, J., Ma, Y., Goldsmith, J., … Davidson, K. W. (2015). Fitbit?: An accurate and reliable device for wireless physical activity tracking. *International Journal of Cardiology*, *185*, 138–140. doi:10.1016/j.ijcard.2015.03.038

diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction*, *10*(2–3), 105–225. doi:10.1080/07370008.1985.9649008

diSessa, A. A. (2004). Metarepresentation: Native competence and targets for instruction. *Cognition and Instruction*, *22*(3), 293–331. doi:10.1207/s1532690xci2203_2

diSessa, A. A., Hammer, D., Sherin, B. L., & Kolpakowski, T. (1991). Inventing graphing: Meta-representational expertise in children. *Journal of Mathematical Behavior*, *10*(2), 117–160.

Drake, J., Cain, R., & Lee, V. R. (2017). From wearing to wondering: Treating wearable activity trackers as tools for inquiry. In I. Levin & D. Tsybulsky (Eds.), *Optimizing STEM education with advanced ICTs and simulations* (pp. 1–29). Hershey, PA: IGI Global.

Duncan, R. G., & Hmelo-Silver, C. E. (2009). Learning progressions: Aligning curriculum, instruction, and assessment. *Journal of Research in Science Teaching*, *46*(6), 606–609. doi:10.1002/tea.20316

Enyedy, N., Danish, J. A., & DeLiema, D. (2015). Constructing liminal blends in a collaborative augmented-reality learning environment. *International Journal of Computer-Supported Collaborative Learning*, *10*(1), 7–34. doi:10.1007/s11412-015-9207-1

Enyedy, N., & Mukhopadhyay, S. (2007). They don't show nothing I didn't know: Emergent tensions between culturally relevant pedagogy and mathematics pedagogy. *Journal of the Learning Sciences*, *16*(2), 139–174. doi:10.1080/10508400701193671

Finzer, W. (2005). *Fathom dynamic data software*. Emeryville, CA: Key Curriculum Press.

Greeno, J. G. (1998). The situativity of knowing. *American Psychologist*, *53*(1), 5–26. doi:10.1037/0003-066X.53.1.5

Grysman, A., Harris, C. B., Barnier, A. J., & Savage, G. (2020). Long-married couples recall their wedding day: the influence of collaboration and gender on autobiographical memory recall. *Memory, 28*(1), 18–33. doi:10.1080/09658211.2019.1673428

Hall, R., Ma, J. Y., & Nemirovsky, R. (2015). Re-scaling bodies in/as representational instruments in GPS drawing. In V. R. Lee (Ed.), *Learning technologies and the body: Integration and implementation in formal and informal learning environments* (pp. 112–131). New York, NY: Routledge.

Hall, R., & Nemirovsky, R. (2012). Introduction to the special issue: Modalities of body engagement in mathematical activity and learning. *Journal of the Learning Sciences*, *21*(2), 207–215. doi:10.1080/10508406.2011.611447

Hall, R., Stevens, R., & Torralba, T. (2002). Disrupting representational infrastructure in conversations across disciplines. *Mind, Culture, and Activity*, *9*(3), 179–210. doi:10.1207/S15327884MCA0903_03

Hancock, C., Kaput, J. J., & Goldsmith, L. T. (1992). Authentic inquiry with data: Critical barriers to classroom implementation. *Educational Psychologist*, *27*(3), 337–364. doi:10.1207/s15326985ep2703_5

Harris, C. B., Keil, P. G., Sutton, J., Barnier, A. J., & McIlwain, D. J. F. (2011). We remember, we forget: Collaborative remembering in older couples. *Discourse Processes*, *48*(4), 267–303. doi:10.1080/0163853X.2010.541854

Harris, C. B., Paterson, H. M., & Kemp, R. I. (2008). Collaborative recall and collective memory: What happens when we remember together? *Memory (Hove, England)*, *16*(3), 213–230. doi:10.1080/09658210701811862
Heyen, N. B. (2020). From self-tracking to self-expertise: The production of self-related knowledge by doing personal science. *Public Understanding of Science*, *29*(2), 124–138. doi:10.1177/0963662519888757
Hollingshead, A. B. (1998). Retrieval processes in transactive memory systems. *Journal of Personality and Social Psychology, 74*(3), 659–671. doi:10.1037/0022-3514.74.3.65
Howison, M., Trninic, D., Reinholz, D., & Abrahamson, D. (2011). The mathematical imagery trainer: From embodied interaction to conceptual learning. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1989–1998). Vancouver, BC: ACM.
Huebner, B. (2016). Transactive memory reconstructed: Rethinking Wegner's research program. *The Southern Journal of Philosophy*, *54*(1), 48–69. doi:10.1111/sjp.12160
Hutchins, E. (1995). How a cockpit remembers its speeds. *Cognitive Science*, *19*(3), 265–288. doi:10.1207/s15516709cog1903_1
Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Strauss, and Giroux.
Kahneman, D., Fredrickson, B. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When more pain is preferred to less: Adding a better end. *Psychological Science*, *4*(6), 401–405. doi:10.1111/j.1467-9280.1993.tb00589.x
Kang, S., Norooz, L., Oguamanam, V., Plane, A. C., Clegg, T. L., & Froehlich, J. E. (2016). SharedPhys: Live physiological sensing, whole-body interaction, and large-screen visualizations to support shared inquiry experiences. In *Proceedings of the the 15th international conference on interaction design and children* (pp. 275–287). New York, NY: ACM Press.
Kolodner, J. L. (1993). *Case-based reasoning*. San Mateo, CA: Morgan Kaufmann.
Konold, C. (2007). Designing a data analysis tool for learners. In M. C. Lovett & P. Shah (Eds.), *Thinking with data* (pp. 267–292). New York, NY: Lawrence Erlbaum.
Konold, C., & Miller, C. A. (2005). *TinkerPlots*. Emeryville, CA: Key Curriculum Press.
Konold, C., Robinson, A., Khalil, K., Pollatsek, A., Well, A., Wing, R., & Mayr, S. (2002). *Students' use of modal clumps to summarize data*. Paper presented at the Sixth International Conference on Teaching Statistics. Cape Town, South Africa.
Lakoff, G. (1980). *Metaphors we live by*. Chicago, IL: University of Chicago Press.
Lakoff, G., & Nunez, R. E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York, NY: Basic Books.
Leander, K. M., Phillips, N. C., & Taylor, K. H. (2010). The changing social spaces of learning: Mapping new mobilities. *Review of Research in Education*, *34*(1), 329–394. doi:10.3102/0091732X09358129
Lee, V. R. (2013). The Quantified Self (QS) movement and some emerging opportunities for the educational technology field. *Educational Technology*, *53*(6), 39–42.
Lee, V. R. (2014). What's happening in the quantified self movement? In J. L. Polman, Kyza, E. A., O'Neill, D. K., Tabak, I., Penuel, W. R., Jurow, A. S., O'Connor, K., Lee, T., D'Amico, L. (Ed.), *Learning and becoming in practice: The International Conference of the Learning Sciences (ICLS) 2014* (Vol. 2, pp. 1032–1036). Boulder, CO: ISLS.
Lee, V. R. (2015). Technology meets body, body meets technology. In V. R. Lee (Ed.), *Learning Technologies and the Body: Integration and Implementation in Formal and Informal Learning Environments* (pp. 1–20). New York, NY: Routledge.
Lee, V. R. (2018). Personal analytics explorations to support youth learning. In R. Zheng (Ed.), *Digital technologies and instructional design for personalized learning* (pp. 145–163). Hershey, PA: IGI Global.
Lee, V. R. (2019). On researching activity tracking to support learning: A retrospective. *Information and Learning Sciences*, *120*(1/2), 133–154. doi:10.1108/ILS-06-2018-0048
Lee, V. R., & Briggs, M. (2014). Lessons learned from an initial effort to bring a quantified self "meetup" experience to a new demographic. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication* (pp. 707–710). ACM. doi:10.1145/2638728.2641321
Lee, V. R., & Drake, J. (2013). Quantified recess: Design of an activity for elementary students involving analyses of their own movement data. In J. P. Hourcade, E. A. Miller, & A. Egeland (Eds.), *Proceedings of the 12th international conference on interaction design and children 2013* (pp. 273–276). New York, NY: ACM.
Lee, V. R., Drake, J., Cain, R., & Thayne, J. (2015). Opportunistic uses of the traditional school day through student examination of fitbit activity tracker data. In M. U. Bers & G. Revelle (Eds.), *Proceedings of the 2015 ACM SIGCHI interaction design and children conference* (pp. 209–218). Boston, MA: ACM.
Lee, V. R., Drake, J., & Williamson, K. (2015). Let's get physical: K-12 students using wearable devices to obtain and learn about data from physical activities. *TechTrends*, *59*(4), 46–53. doi:10.1007/s11528-015-0870-x
Lee, V. R., Drake, J. R., & Thayne, J. L. (2016). Appropriating quantified self technologies to improve elementary statistical teaching and learning. *IEEE Transactions on Learning Technologies*, *9*(4), 354–365. doi:10.1109/TLT.2016.2597142

Lee, V. R., & Dubovi, I. (2020). At home with data: Family engagements with data involved in Type 1 Diabetes management. *Journal of the Learning Sciences*, *29*(1), 11–31. doi:10.1080/10508406.2019.1666011
Lee, V. R., & Dumont, M. (2010). An exploration into how physical activity data-recording devices could be used in computer-supported data investigations. *International Journal of Computers for Mathematical Learning*, *15*(3), 167–189. doi:10.1007/s10758-010-9172-8
Lee, V. R., & Shapiro, R. B. (2019). A broad view of wearables as learning technologies: Current and emerging applications. In P. Diaz, A. Ioannou, K. K. Bhagat, & J. M. Spector (Eds.), *Learning in a digital world - Perspectives on interactive technologies for formal and informal education* (pp. 113–133). Springer. doi:10.1007/978-981-13-8265-9_6
Lee, V. R., & Thomas, J. M. (2011). Integrating physical activity data technologies into elementary school classrooms. *Educational Technology Research and Development*, *59*(6), 865–884. doi:10.1007/s11423-011-9210-9
Lehrer, R., Kim, M.-J., Ayers, E., & Wilson, M. (2014). Toward establishing a learning progression to support the development of statistical reasoning. In *Learning over time: Learning trajectories in mathematics education* (pp. 1–49). Charlotte, NC: Information Age Publishers.
Lehrer, R., Kim, M.-J., & Schauble, L. (2007). Supporting the development of conceptions of statistics by engaging students in measuring and modeling variability. *International Journal of Computers for Mathematical Learning*, *12*(3), 195–216. doi:10.1007/s10758-007-9122-2
Lehrer, R., & Schauble, L. (2004). Modeling natural variation through distribution. *American Educational Research Journal*, *41*(3), 635–679. doi:10.3102/00028312041003635
Li, I., Dey, A., & Forlizzi, J. (2010). A stage-based model of personal informatics systems. In *Proceedings of the 28th international conference on Human factors in computing systems* (pp. 557–566). New York, NY: ACM. doi:10.1145/1753326.1753409
Lindgren, R., & Johnson-Glenberg, M. (2013). Emboldened by embodiment: Six precepts for research on embodied learning and mixed reality. *Educational Researcher*, *42*(8), 445–452. doi:10.3102/0013189X13511661
Lindgren, R., Tscholl, M., Wang, S., & Johnson, E. (2016). Enhancing learning and engagement through embodied interaction within a mixed reality simulation. *Computers & Education*, *95*, 174–187. doi:10.1016/j.compedu.2016.01.001
Lo, W.-T., & Quintana, C. (2013). Students' use of mobile technology to collect data in guided inquiry on field trips. In *Proceedings of the 12th International conference on interaction design and children* (pp. 297–300). New York, NY: ACM Press.
Loftus, E. F., & Loftus, G. R. (1980). On the permanence of stored information in the human brain. *American Psychologist*, *35*(5), 409–420. doi:10.1037/0003-066X.35.5.409
Lupi, G., & Posavec, S. (2016). *Dear data*. New York, NY: Princeton Architectural Press.
Lupton, D. (2016). *The quantified self*. Cambridge: John Wiley & Sons.
McDermott, R. P. (1996). The acquisition of a child by a learning disability. In S. Chaiklin & J. Lave (Eds.), *Understanding practice: Perspectives on activity and context* (pp. 269–305). Cambridge, UK: Cambridge Unviersity Press.
Metcalf, S. J., & Tinker, R. F. (2004). Probeware and handhelds in elementary and middle school science. *Journal of Science Education and Technology*, *13*(1), 43–49. doi:10.1023/B:JOST.0000019637.22473.02
Meudell, P. R., Hitch, G. J., & Boyle, M. M. (1995). Collaboration in recall: Do Pairs of people cross-cue each other to produce new memories? *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *48*(1), 141–152. doi:10.1080/14640749508401381
Moher, T., Ching, C. C., Schaefer, S., Lee, V. R., Enyedy, N., Danish, J. A., … Fong, C. (2014). Becoming reflective: Designing for reflection on physical performances. In J. L. Polman, E. A. Kyza, D. K. O'Neill, I. Tabak, W. R. Penuel, A. S. Jurow, K. O'Connor, T. Lee & L. D'Amico (Ed.), *Learning and Becoming in Practice: The International Conference of the Learning Sciences (ICLS)* 2014 (Vol. 3, pp. 1273–1282). ISLS.
Mokros, J. R., & Russell, S. J. (1995). Children's concepts of average and representativeness. *Journal for Research in Mathematics Education*, *26*(1), 20–39. doi:10.2307/749226
Mokros, J. R., & Tinker, R. F. (1987). The impact of microcomputer-based labs on children's ability to interpret graphs. *Journal of Research in Science Teaching*, *24*(4), 369–383. doi:10.1002/tea.3660240408
National Research Council. (2007). *Taking science to school*. Washington, D.C: National Academies Press.
Nemirovsky, R., Tierney, C., & Wright, T. (1998). Body motion and graphing. *Cognition and Instruction*, *16*(2), 119–172. doi:10.1207/s1532690xci1602_1
Norman, D. A. (1993). *Things that make us smart: Defending human attributes in the age of the machine*. Reading, MA: Addison-Wesley Pub. Co.
Pangrazio, L., & Sefton-Green, J. (2020). The social utility of 'data literacy'. *Learning, Media and Technology, 45*(2), 208–220. doi:10.1080/17439884.2020.1707223
Petrosino, A., Lehrer, R., & Schauble, L. (2003). Structuring error and experimental variation as distribution in the fourth grade. *Mathematical Thinking and Learning*, *5*(2–3), 131–156. doi:10.1080/10986065.2003.9679997

Potapov, K., Vasalou, A., Lee, V. R., & Marshall, P. (2021). *What do teens make of personal informatics? Youth responses to self-tracking practices within a classroom setting*. Paper presented at the ACM CHI Conference on Human Factors in Computing Systems. ACM. doi:10.1145/3411764.3445239

Philip, T. M., Olivares-Pasillas, M. C., & Rocha, J. (2016). Becoming racially literate about data and data-literate about race: Data visualizations in the classroom as a site of racial-ideological micro-contestations. *Cognition and Instruction*, *34*(4), 361–388. doi:10.1080/07370008.2016.1210418

Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., … Soloway, E. (2004). A scaffolding design framework for software to support science inquiry. *Journal of the Learning Sciences*, *13*(3), 337–386. doi:10.1207/s15327809jls1303_4

Rajaram, S., & Pereira-Pasarin, L. P. (2010). Collaborative memory: Cognitive research and theory. *Perspectives on Psychological Science*, *5*(6), 649–663. doi:10.1177/1745691610388763

Redelmeier, D. A., & Kahneman, D. (1996). Patients' memories of painful medical treatments: Real-time and retrospective evaluations of two minimally invasive procedures. *Pain*, *66*(1), 3–8. doi:10.1016/0304-3959(96)02994-6

Redish, E. F., Saul, J. M., & Steinberg, R. N. (1997). On the effectiveness of active-engagement microcomputer-based laboratories. *American Journal of Physics*, *65*(1), 45–54. doi:10.1119/1.18498

Rivera-Pelayo, V., Zacharias, V., Müller, L., & Braun, S. (2012). Applying Quantified Self approaches to support reflective learning. In S. Dawson & C. Hathornwaite (Eds.), *LAK '12 Proceedings of the 2nd International conference on learning analytics and knowledge* (pp. 111–114). New York, NY: ACM. doi:10.1145/2330601.2330631

Roberts, J., & Lyons, L. (2020). Examining spontaneous perspective taking and fluid self-to-data relationships in informal open-ended data exploration. *Journal of the Learning Sciences*, *29*(1), 32–56. doi:10.1080/10508406.2019.1651317

Roth, W.-M. (2014). Learning in the discovery sciences: The history of a "radical" conceptual change, or the scientific revolution that was not. *Journal of the Learning Sciences*, *23*(2), 177–215. doi:10.1080/10508406.2014.893435

Roth, W.-M., McGinn, M. K., & Bowen, G. M. (1998). How prepared are preservice teachers to teach scientific inquiry? Levels of performance in scientific representation practices. *Journal of Science Teacher Education*, *9*(1), 25–48. doi:10.1023/A:1009465505918

Rubel, L. H., Hall-Wieckert, M., & Lim, V. Y. (2017). Making space for place: Mapping tools and practices to teach for spatial justice. *Journal of the Learning Sciences*, *26*(4), 643–687. doi:10.1080/10508406.2017.1336440

Rubin, A. (2020). Learning to reason with data: How did we get here and what do we know? *Journal of the Learning Sciences*, *29*(1), 154–164. doi:10.1080/10508406.2019.1705665

Sandoval, W. (2014). Conjecture mapping: An approach to systematic educational design research. *Journal of the Learning Sciences*, *23*(1), 18–36. doi:10.1080/10508406.2013.778204

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals and understanding: An inquiry into human knowledge structures*. Oxford, UK: Lawrence Erlbaum.

Shapiro, B. R., Hall, R. P., & Owens, D. A. (2017). Developing & using interaction geography in a museum. *International Journal of Computer-Supported Collaborative Learning, 12*(4), 377–399. doi:10.1007/s11412-017-9264-8

Smith, J. P., diSessa, A. A., & Roschelle, J. (1994). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *Journal of the Learning Sciences*, *3*(2), 115–163. [Database] doi:10.1207/s15327809jls0302_1

Stornaiuolo, A. (2020). Authoring data stories in a media makerspace: Adolescents developing critical data literacies. *Journal of the Learning Sciences*, *29*(1), 81–103. doi:10.1080/10508406.2019.1689365

Strauss, S., & Bichler, E. (1988). The development of children's concepts of the arithmetic average. *Journal for Research in Mathematics Education*, *19*(1), 64. doi:10.2307/749111

Sutton, J., Harris, C. B., Keil, P. G., & Barnier, A. J. (2010). The psychology of memory, extended cognition, and socially distributed remembering. *Phenomenology and the Cognitive Sciences*, *9*(4), 521–560. doi:10.1007/s11097-010-9182-y

Takacs, J., Pollock, C. L., Guenther, J. R., Bahar, M., Napier, C., & Hunt, M. A. (2014). Validation of the Fitbit One activity monitor device during treadmill walking. *Journal of Science and Medicine in Sport*, *17*(5), 496–500. doi:10.1016/j.jsams.2013.10.241

Taylor, K. H. (2017). Learning along lines: Locative literacies for reading and writing the city. *Journal of the Learning Sciences*, *26*(4), 553–574.

Taylor, K. H., & Hall, R. (2013). Counter-mapping the neighborhood on bicycles: Mobilizing youth to reimagine the city. *Technology, Knowledge and Learning*, *18*(1–2), 65–29. doi:10.1007/s10758-013-9201-5

The Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, *32*(1), 5–8.

Tinker, R., & Krajcik, J. S. (2001). *Portable technologies*. New York, NY: Kluwer Academic.

Watson, J. M., & Moritz, J. B. (1999). The Beginning of Statistical Inference: Comparing two Data Sets. *Educational Studies in Mathematics, 37*(2), 145–168. doi:10.1023/A:1003594832397

Watson, J. M., & Moritz, J. B. (2000). The longitudinal development of understanding of average. *Mathematical Thinking and Learning*, *2*(1–2), 11–50. doi:10.1207/S15327833MTL0202_2

Wegner, D. M. (1987). Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G. R. Goethals (Eds.), *Theories of group behavior* (pp. 185–208). New York, NY: Springer New York.

Wegner, D. M., Erber, R., & Raymond, P. (1991). Transactive memory in close relationships. *Journal of Personality and Social Psychology*, *61*(6), 923–929. doi:10.1037/0022-3514.61.6.923

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, *127*(1), 3–21. doi:10.1037/0033-2909.127.1.3

Zacks, J. M., Tversky, B., & Iyer, G. (2001). Perceiving, remembering, and communicating structure in events. *Journal of Experimental Psychology: General*, *130*(1), 29–58. doi:10.1037/0096-3445.130.1.29