International Society of
the Learning Sciences

Modeling Chance Processes in A Classroom's Ecological Investigation

Panchompoo Wisittanawat, Vanderbilt University, fai.wisittanawat@vanderbilt.edu
Amanda Dickes, Gulf of Maine Research Institute, adickes@gmri.org
Richard Lehrer, Vanderbilt University, rich.lehrer@vanderbilt.edu

Abstract: The work of ecologists entails structuring variability by parsing random and directed variability. Middle-grade students are often introduced to ideas about probability and statistics in mathematics, but these ideas are rarely employed in science investigations. This paper reports on a design study in one 7th-grade science classroom that participated in a citizen-science project investigating changes in invasive crab populations. Students surveyed crab abundance at one field site, contributing observations to a citizen-science database. Finding an unexpected ratio between male and female crabs in their sample, students compared the ratio obtained in the field to a simulated sampling distribution of ratios in light of an equiprobable assumption. Finding that their sample's sex ratio was improbable yet consistent with samples in the larger database instigated a search for ecological mechanism. Evidence of student thinking in classroom conversations point to seeds of distinguishing random from directed variability.

One fundamental task in ecological investigations is to understand the abundance and distribution of organisms across space and time. Abundance and distribution of organisms is not uniform across space and time as species vary, individuals within a species vary, and environmental gradients vary. Distribution-generating processes include random (e.g., genetic recombination or small fluctuations in temperature) and causal mechanisms (e.g., organism-environment interactions, organism-organism interactions, seasonal variation in temperature). Additionally, ecological sampling inevitably introduces chance variability: repeated samples vary just by chance. Sampling is key in ecological investigations as ecological processes occur across multiple spatial and temporal scales, making a census approach impractical. Teasing apart variability due to chance and variability due to other underlying patterns and processes is a messy but critical activity in ecological investigations (Albert et al, 2010). This critical activity is often hidden from students' view. Conceptual underpinnings for understanding variability and changes in ecology and evolution take time and coherent forms of support to develop. Students from younger ages need repeated opportunities to encounter and build on these foundational ideas (Lehrer & Schauble, 2012). This research addresses ways of supporting the development of foundational ideas about variability and change in ecosystems by positioning students as participants in classroom approximations of professional ecological practices, particularly those of visualizing, measuring and modeling variability in ecological investigations.

Research Context & Design

The work reported here takes place within a larger, multi-year design-based research study investigating how students' conceptions of and practices for representing variability inform their construction and revision of models of ecosystems, and how students come to view the function and utility of the ecosystem models they have constructed, as well as relations between different models. The context of this work is a statewide middle school citizen science curriculum focused on monitoring and mapping species of interest in the Gulf of Maine intertidal ecosystem with the particular goal of understanding how the relative abundance and distribution of native and invasive species of crabs is changing due to climate-driven factors. In addition to classroom work, students survey native and invasive crab abundance at local field sites and contribute their collected data to a larger database of survey data collected by students up and down the coast of Maine. In this paper we report on the first implementation of a data modeling trajectory co-designed with two middle-school science teachers. Our design focused on cultivating an appreciation of the role of chance in sampling so that as students collected and analyzed samples during their citizen science investigations, they had the means to account for uncertainty due to chance in their observed samples, and in this light, to anticipate when sample variability indicated a need to search for an ecological mechanism.

Instructional design

This instructional design builds on a data modeling perspective (Lehrer & Romberg, 1996). This perspective positions students as participants in approximations of statistical practices in professional science, especially those of visualizing, measuring, and modeling variability (Jones et al., 2017; Wild & Pfannkuch, 1999). Conceptual and representational cornerstones of these practices include mathematical descriptions of distribution and mathematical models of chance. These cornerstones in turn support model-based inference in light of uncertainty due to chance variability.

School curricula often allocate topics in probability and statistics to mathematics. Ideally, students come to science classrooms equipped with mathematical tools that they can utilize in science investigations. However, probability and statistics are often treated separately, and students may not meaningfully connect them in a way that this otherwise connected network of concepts and practices can readily be employed in service of scientific inquiry (Konold & Kazak, 2008; Moore, 1990). In this work, our instructional sequence aimed at both supporting students to develop mathematical means for visualizing, measuring and modeling variability while simultaneously advancing their scientific investigations (Lehrer, Schauble & Wisittanawat, 2020).

Prior instructional sequence: Visualizing and measuring variability

Prior to the instructional sequence that is the focus of this paper, we first supported students in developing ways of visualizing and measuring variability by challenging them to invent displays of data that they had generated. Students as novices to visualizing and analyzing data tend to focus on individual cases (e.g., the biggest crab or the crab they personally found) rather than on the aggregate (Konold et al., 2015). To support students to extend the view of data as mere cases to also encompass the view of data as an aggregate of distributed cases, we asked students to invent data displays that highlight characteristics (e.g., center and spread) of the data aggregate. Constructing data displays provided students opportunities to consider how to order, group and organize data, leading to structuring data as a distribution.

We searched for a candidate task (Lehrer & Schauble, 2021) in the invasive-crab investigation. That is, we searched for a task that provided opportunities to capitalize on children's conceptual resources in ways that would potentially support both the development of statistical practices and students' current scientific investigation. In this citizen-science investigation, students would record data of sizes of crabs found during their field survey. Prior to the field visit, students practiced measuring the size of a classroom's crab in residence (from their classroom's saltwater tanks). We posed the question: *How big is this green crab?* Each student measured the same crab and recorded their measurement. We then combined class measurements and provided groups of students with a copy of the class's batch of data. We asked students to develop displays of their class's measurement data to help other people quickly see what the crab's size really was. Consistent with prior work (e.g., Lehrer, Kim & Schauble, 2007), students generated different types of displays. The displays differed in how cases were grouped (e.g., no groupings, groups of same/similar values, groups of same size), how cases or groups of cases were ordered (e.g., by value or by frequency) and how cases were organized in relation to each other (e.g., on a consistent scale or not).

Next, we supported students to invent or use measures of center by posing the question: *What is our best guess of the crab's size given our class's data?* Students had already learned different measures of center, namely mean, median and mode, and they made use of these measures as estimates of the actual crab's size. To support students to invent or use measures of spread, we posed the question: *How precise were our class's measurements?* Here, students used range and interquartile range (and other similar invented measures that described the size of the middle or the biggest clump). Students also measured the same crab a second time, this time following a standard measurement protocol for measuring crab sizes—measuring the widest part of the carapace. This second data set invited comparisons between two data distributions and provided a situation for a test for robustness of measures of center and spread (e.g., whether changes in measures reflected changes in distributions). In this instructional sequence, students first invented both data displays and measures on paper, and we later introduced a digital tool, the Common Online Data Analysis Platform (CODAP), to facilitate future work with data.

Focal instructional sequence: Modeling sampling variability

The remainder of the paper focuses on this next phase of instruction. In this section we describe the tasks, tools, and rationale comprising the anticipated instructional sequence. In the Findings & Analysis section, we describe the enacted instructional sequence.

With ways to visualize and measure variability in hand, we aimed to extend students' conception of variability to include the role of chance. Again, we searched for a candidate task that served as an entrée into modeling chance processes within the context of the invasive-crab investigation. From the existing large citizen-science data set, we learned that school groups often found sharply unequal ratios between the number of male and female crabs in their quadrat surveys, with the majority finding more males than females. Across four years (sample sizes between 127 to 311 crabs), the sex ratios ranged from 68%-75% males. Across 19 field sites (with more than 10 crabs found in quadrats), the majority of them (17) sites reported more males, ranging from 56%-92% males. The other (2) sites reported more females (42% males and 45% males). We thus anticipated that in their upcoming field trip students would likely find an unbalanced sex ratio of the crabs in their quadrats, and most likely vastly more males than females. As in any ecological sampling, students' observed sample outcome would need to be considered in light of sampling variability. The key question we worked to support students to consider was as follows: *How likely could our sample's observed sex ratio occur just by chance?*

To consider this question, students first needed to develop a grasp for the structure of random variability. Individual outcomes of a random process are uncertain, but outcomes of long-run, repeated random process exhibit structure in the aggregate. We supported students to develop ideas about the structure of chance variability by engaging them in examining behaviors and outcomes of random devices (here, a physical and later digital spinner). We asked pairs/groups of students to conduct an experiment by spinning an equal-split "male-female" spinner 10 times. Students recorded the number of males per 10 spins and repeated the experiment to collect a total of 10 samples of 10 spins. We planned to have each pair represent the results of these 10 samples (each of which was comprised of 10 repetitions of the equiprobable process) as a dot plot of a sample statistic (the number of "males."). Our conjecture was that this would help students start to consider sample to sample variability that could be attributed solely to chance. Next, we planned to ask students to predict shape, and statistics of center and spread of a distribution constituted by growing the number of samples to include all those generated by every pair. The aim was to promote a conception of potential structure in a sampling distribution corresponding to elements of the variability-generating process (i.e., a center statistic representing the signal of its equiprobable structure and a statistic of variability that could be attributed to the randomness of the process and the sample size). Students would then organize the sampling distribution, examine its shape, and reason about the relations between the structure of the spinner and the shape of the sampling distribution. We planned then to introduced a digital spinner (in CODAP) as a way for students to explore effects of sample sizes on sampling distributions.

Next, we planned to ask students to use a (digital) spinner to construct models that represented a situation where there was roughly the same number of male and female crabs in the ocean. Students would then use the models to simulate repeated sampling of a sampleof crabs equal in number to that collected in the field. Based on the model-generated sampling distribution, students would consider how likely their observed sample's sex ratio could occur just by chance. Asking students to make inference about the sex ratio of crabs in the ocean, we anticipated that students would suggest collecting more data, especially if their observed outcome was unlikely to occur by chance. This would then motivate further exploration of the larger citizen-science data set to consider whether other school groups also found similar outcomes as theirs. As mentioned earlier, the sex ratio in the large data set skewed towards males. We would engage students in the discussion on whether and how they wanted to revise their chance models given new information they gained from exploring the larger data set.

Analysis & Findings

We reviewed fieldnotes and video recordings of classroom conversations with an eye towards how students constructed and revised models of chance to represent random variability in the crab sampling process. We attended to key moments in classroom conversations when students modeled random variability, and how students' emerging conceptions of and practices for representing variability supported their reasoning about the ecological system under investigation.

Exploring local data

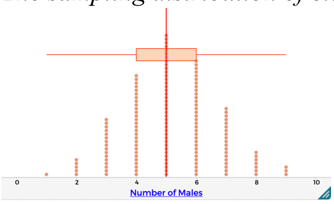
Students surveyed invasive crab abundance at a state park near the school and catalogued twenty-six crabs found in 16 1-m^2 quadrats sampled along a 60-meter transect. Field protocols specified identifying species of all surveyed crabs, measuring carapace length, recording the number of claws, and identifying sex. Students' exploration of their own sample revealed, as anticipated, a curious outcome: twenty-three of the twenty-six crabs were male, a "wild ratio" of 88% male to 12% female, as noted by the students. Although the outcome was surprising to students, they readily inferred characteristics of the entire population from their single sample. For example, students suggested that the male crabs were "overpopulating" the female crabs, that female crabs were "not surviving very well," and that the overall population of invasive crabs was actually *declining* due to a shortage of female crabs. The classroom teacher, Ms. S., pivoted student attention away from the sample at hand and invited them to imagine other possible samples. She asked, for example, whether it was possible to collect a sample with 88% males and 12% females just by chance? And, if we collected a second sample, could the sex ratio be 50/50? Students considered the "odds" of finding the same ratio of male crabs to female crabs in a second sample unlikely but also suggested several sources of natural variability which might have influenced their outcome including the time of day, the water temperature, and the protocols themselves (sampling the intertidal zone only versus sampling both the intertidal and the subtidal zones).

Examining chance variability with random devices

To consider the structure of variability in a repeated random process, students conducted experiments with random devices. Student groups conducted an experiment, collecting 10 samples of 10 spins on an equipartitioned Male-Female spinner as a way to simulate outcomes of repeated samples of 10 crabs. Each group made a distribution of their sample outcomes, and the teacher collected outcomes from all student groups to make a sampling distribution of 150 samples of 10 spins. Before examining the distribution, the teacher asked students to predict its shape. One student predicted a "hill in the middle," suggesting there could be a few outliers, but most outcomes would be "compacted to the middle." Another student thought the distribution would clump around 5, because the outcomes of his groups' 10 samples also clustered around 5 and included no outcomes of 1, 2, 8, or 9. A researcher asked if they expected more extreme outcomes in the 150-sample sampling distribution. The students thought there would be more extreme outcomes, but concluded most outcomes would still be in the middle. When Ms. S. showed students their aggregate sampling distribution, students noticed that their predictions bore out. Students noted that, as they predicted, "you can see...a clear hill and that five is the median and also the tallest." Additionally, students observed that extreme outcomes, such as 9 males, only happened 3 out of 150 times.

Figure 1

The sampling distribution of class's 150 samples of 10 spins on a 50M/50F spinner



Ms. S then introduced a digital spinner in CODAP, a resource she now framed as a "statistical tool" that expedited explorations of chance processes. Ms. S asked students to predict what happened if they ran the same experiment again, collecting 150 samples of 10 spins. Students predicted that the simulated sampling distribution would look similar to the one produced using hand-held spinners. Ms. S. ran the model while students watched, with much anticipation, in real time as the sampling distribution took shape and saw that the general shape of the distribution was similar to the previous one generated by physical devices. Ms. S then invited a student volunteer to draw on the whiteboard what the shape of the distribution might be if they were to continue the experiment and collect 150 more samples. The student drew a shape that traced the current distribution with a central clump around 5 (or 50%) males, and similar spread. Again, Ms. S. ran the model while students watched in real time as the sampling distribution took shape and saw that the general shape of the distribution did not change.

Using random devices to model chance variability in our crab sampling

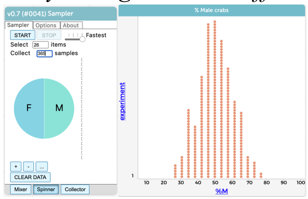
Having explored chance variability, the class now turned to consider their own observed sample in light of chance variability under the assumption of equiprobable sex ratio. Ms. S focused her students to the question at hand: *How likely could our sample's observed sex ratio occur just by chance?* She then asked students to return to the digital spinner and to use the spinner to help them answer the question.

Running their models, students noticed that more extreme outcomes, such as 80% or 90% males, were even less likely with this larger sample size of 26 crabs. For example, during small group work, Anna's model collected 150 samples of 26 crabs using a 50M/50F spinner, and she noted that her model produced no outcomes "over 80 or under 20 [percent male]." A researcher asked Anna if her model's outcomes made her think differently about how likely the class' observed outcome of 88% male could occur just by chance. Anna responded, "It makes me think that's it's even less likely. Like, I was thinking before that it probably was just by chance when we were [modeling] with 10 [crabs], but with 26 [crabs] it shows that it's even less likely because I didn't even get one point that was above 80 [percent male]."

In the subsequent whole class discussion, other students shared their models and similar observations as Anna's. Lacey, Morgan, and Griffin built a model they described as simulating the collection of "26 crabs a day every day of the year," that is, they used a spinner of 50M/50F and ran 365 samples of 26 spins (Figure 2). Ms. S asked the group about their outcomes:

Figure 2

Lacey, Morgan, and Griffin's model for collecting 26 crabs a day every day of the year



1. Ms. S: On any of those 365 days would we have had a day that we found 90% male crabs just by chance?
2. Lacey: Nope.
3. Ms. S: Really, not a single time in your model?
4. Lacey: Yes.

Having established with the class that outcomes like their observed outcome of 88% male rarely occurred in their simulated sampling distribution, Ms. S asked the class to consider:

5. Ms. S: What does this tell us about the likelihood of getting 88% male at [our field site]? Do we still think that could be attributed because of random chance or are we more convinced there might be something happening?

Ms. S's wondering invited a few students to propose ideas for revising the structure of the spinner to reflect a different sex ratio. Ben wondered if the "actual real" ratio was something different than 50/50, and suggested a ratio of 60 males to 40 females. Denver agreed, excitedly interjecting that he had been "trying to [share] *every single day,*" that there could be more male crabs than female crabs in the ocean. Some students wanted to investigate Denver's claim by going back to the field and collecting more data, with Caleb suggesting, "collecting every crab in the ocean." Other students thought they could first investigate by revising the chance model to reflect class members' guess that there were more male crabs. One student suggested making the spinners 4 slots males, and 3 slots females.

Exploring citizen-science data and revising models

In the following class, Ms. S. revisited students' earlier suggestion to revise the structure of the spinner and solicited the class for proposals for how to revise. Recalling Denver's exasperated statement at the conclusion of the last lesson, Ms. S. invited Denver to restate his proposal. Denver had consistently proposed that "there could just be more male crabs," and he now guessed the ratio to be "65 male, 35 female." As an alternative hypothesis, a researcher proposed maintaining the 50/50 spinner because that was the case for most animals and it was possible that crabs were incorrectly sexed during the field survey. Ms. S., on the other hand, proposed revising the spinner to resemble their observed sex ratio, citing "We practiced the protocol, we did a good job in the field, it's definitely 88% males." The teacher drew each of the proposed spinners—65/35, 50/50 and 88/12—on the whiteboard, and asked the class to discuss which one they wanted to put their votes behind. A handful of students thought it could still be 50/50. Most other students disagreed, however, but were unsure what the ratio should be:

6. Ben: I agree with Denver because I think the [50/50] spinner was inaccurate. I think what we should have is we should have 65 percent of the [spinner] be covered with males, and then 35 covered with females. I think that might be more accurate to our data. We found 88% males, it would just be crazy to find that on our first try if it was really 50/50. I don't think it's 88% males all the time because they would have trouble reproducing.
7. Denver: I agree with Ben, like, when we did the 50/50 spinner, we almost didn't have it up towards 90%, so I think that there's just going to be more males than females out in the ocean.

8. Anna: I agree, but I don't [know] if it's like the whole ocean has more male than female crabs, it might just be at [our field site]. I definitely think the number of male crabs to female crabs is more than 50/50.

In this exchange, students recalled the outcomes from the equiprobable model (e.g., "almost didn't have it up towards 90%") to argue the equiprobable assumption was inaccurate. At the same time, students were not convinced that the overall sex ratio could really be 88/12 as the crabs "would have trouble reproducing." Other students started to wonder whether their field site was exceptional in some way. As the class was leaning away from 50/50, but also away from 88/12, the researcher made connections to other students' ideas about other ways to find out what the sex ratio might be:

9. Researcher: I've also heard [another] idea, that we could just measure every crab in the ocean. But I'm wondering, is there another place we could look for the kind of data you're craving, Denver?
10. Denver: Well, there's all those other schools that went to those other beaches and we could see their data.
11. Researcher: And what could we do with their data?
12. Denver: We could see if it's close to ours because we found 88% males. If they found 85% males, we know that it's going to be close to more males than females.

As it was not feasible to "measure every crab in the ocean," the class followed Denver's suggestion and consulted "all the data from the other schools" by exploring the large citizen science crab survey data. Here, students also explored trends in sex ratio in relation to sites, time of year, shell conditions, etc. Through their explorations, students also determined the overall sex ratio (of 902 green grabs collected in the intertidal along the coast of Maine) was 71% males and 29% females.

Ms. S posed a challenge: "How could you revise your model to be a better representation of what we think is going on in the intertidal?" Students sketched their ideas for a revised spinner on paper, mostly drawing either a 71/29 spinner or an approximation of it (e.g., 70/30). The class agreed on revising the spinner to 71/29, and before running the model, several students predicted that, now that the structure of the spinner was 71/29, it would be more likely to get a sample of 88% males compared to the previous run with a 50/50 spinner. Because this model was difficult to accomplish on the spinner in CODAP, the class's revised model had to be run on a central computer using TinkerPlots (Konold & Miller, 2011). The 71/29-spinner model was projected on the board and the first sample of 26 spins was collected:

13. Ms. S: We got 66% [males], what should we do now?
14. Griffin: Run it again.
15. Ms. S: Run it again! How many times?

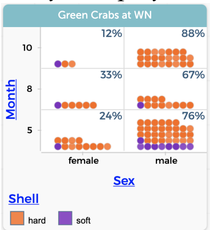
Students suggested running the model more times. This suggested a shift from viewing a single outcome as isolated to viewing an outcome as part of a long-run repeated random process, a shift foundational to understanding the structure of chance variability (Konold, 1989). Students were still developing a sense for how "long" the "run" needed to be, however, and only collected a relatively small number of samples. In response to the teacher's question, one student suggested collecting as many samples as the number of times middle-school students across the state went to collect samples in the field. The class took up that suggestion, and collected 70 samples, i.e., the number of field trips in the data set. As the model was running, the teacher asked the class to consider what the sampling distribution might look like, and students predicted the middle clump would be in the 70s. The resulting sampling distribution had a clump around 70% males, as students had predicted, and it included outcomes in the 80's and 90's, which rarely (or never) occurred in the sampling distribution generated under the equiprobable assumption. This led the class to conclude that their sample was consistent with the overall sex ratio in the larger data set.

Exploring mechanisms for sex ratio differences

Having established that their observed sex ratio was consistent with the overall sex ratio in the larger citizen science data set, the class now searched for possible mechanisms to account for this disparity between the sexes. Earlier, students had explored different data sets: their own field data, data from their site, and the multi-site citizen science data. In their explorations, one student, Emily, discovered an interesting pattern in the number of soft-shell crabs by sex at their field site. Ms. S. shared Emily's CODAP display (Figure 3) with the class, and students noticed that while there were more male crabs across the months, the percentages of male and female crabs surveyed were also changing. In particular, students noticed there were more female crabs found in May and August compared to in October (when they surveyed the site).

Figure 3

Emily's display showing shell hardness by sex and time of year



Ms. S revoiced many students' noticing that the sex ratios appeared to be changing from month to month and asks the class why the ratios might be changing. She notes, "We saw this in our data. … Why would this pattern exist *over the course of the year*?" Students considered variables such as temperature, as well as how molting cycles might influence foraging and refuge-seeking strategies:

16. Denver: We were thinking that it might be because of temperature but we were also thinking that if they're soft shells, they might want to hide and not be on the beaches.
17. Ms. S: Totally, what have we seen in our own tanks, do soft shells last very long? Unfortunately, no.
18. Bella: We thought that maybe it's because of the season, because the most male crabs were in colder … Like [August] and there were very few [male crabs] and that's like summer. And in October, it's colder.

Anticipating connecting students' ideas to scientists' account of the mechanism driving seasonal sex ratios, Ms. S. introduced a video clip of a conversation among the partner scientists on the citizen science project as they interpreted data on crab's sex ratios. Ms. S asked students to listen for explanations for why they see more males than females, and why those ratios "vary so much throughout the year." In the video, the scientists are considering a graph showing the percent male and percent female by year and county.

19. Ms. S: What have we heard so far?
20. Madison: It's mostly because of the molting that we don't see many females. We checked only in the Fall and that's when females go and hide and males are hard shells and don't need to hide.

Ms. S. called attention back to the question the class had been considering: Are there really more males or are the females just harder to find? She notes that the molting cycles, as Madison has suggested and that Emily investigated in their site data, have helped them think about "one part of our question: maybe the females are just hiding more and we can't find them." But she immediately followed this statement by asking "Or maybe there are just fewer of them?" In the video, the partner scientists consider this same question: Are males more easily observed in the Fall or are there more males in general during that time of year? The scientists discuss both movement and habitat preference as possible causes. However, the scientists conclude that this question is still not well understood, something Ms. S. calls attention to to make explicit the epistemic game they, and the scientists, are playing as well as the broader purpose of the citizen science project itself.

21. Ms. S: What did we hear there? Do we know for sure? Are there actually fewer females or are they harder to find?
22. Griffin: They're just harder to find. The males are molting in the spring and summer, and the females molt in the Fall and are harder to find in the Fall.
23. Ms.S.: And even [our scientist] isn't totally sure because they're harder to find or because there are actually fewer of them. So, we don't have a definitive answer and that's the point of all of this [citizen science], so that we can help [our scientist] get to the bottom of this.

Discussion

Ecology is both a challenging and fruitful context for supporting distinctions between random and directed variability. It is challenging because variability-generating processes in ecology occur across temporal and spatial scales that are not immediately tangible to students. Yet, when students have a hand in construction of samples, their direct experience in a sampling process supports their image of long-run repeated process (e.g., running the model to simulate visiting the field every day), and how their decisions about sampling can influence their outcomes.

In this iteration of our design study, we supported students in visualizing, representing, and modeling variability in the context of a citizen science ecological investigation. At the beginning of the focal instructional sequence, students did not consider their observed outcome in light of sampling variability. This was evident in how students immediately made inferences about crabs in the ocean from their single sample before they considered whether such a sample could be explained by chance. We supported students to consider how likely their observed sample could occur just by chance. Students' developing ways of representing and modeling sampling variability supported them to reject the earlier equiprobable assumption. Students then revised their model to consider whether their observed outcome was consistent with data collected by other schools in the citizen-science data set. Finding that their outcome was consistent with the larger data in turn motivated their search for plausible mechanisms that explain the observed difference in sex ratio.

The current study suggests two potentially fruitful areas for future iterations. First, we can further investigate students' connections to and use of publicly available data. The learning sciences community seeks a vision of Data Science education that is "grounded in consequential investigations in which learners pose questions, obtain data, and communicate findings within meaningful disciplinary contexts" (Wilkerson & Polman, 2020, p. 3). Grounded in consequential investigations, students can use and develop data science tools and methods to employ public data in purposeful ways. Here, we saw evidence of students making connections from their local investigation to publicly available data to further their investigation. Second, we elaborate and extend our instruction design to address relations across different forms of models, and we saw evidence of this practice emerging in this classroom. For example, when students noticed patterns on molting in the large-scale data, the teacher made the connection to their classroom microcosms which had emerged as a site for studying predation avoidance in relation to foraging and refuge seeking. We continue to consult our partner scientists, who employ field study and microcosms to craft models of climate impacts on intertidal ecosystems, to consider how students' inquiry can be supported to coordinate field and microcosm study to get a grip on sources of variability and change in these systems.

Acknowledgements

We are grateful for the enthusiastic support of our partner teachers and students. This material is based on work supported by the National Science Foundation under grant DRL- 2010119.

References

Albert, C. H., Yoccoz, N. G., Edwards Jr, T. C., Graham, C. H., Zimmermann, N. E., & Thuiller, W. (2010). Sampling in ecology and evolution – bridging the gap between theory and practice. *Ecography*, 33(6), 1028–1037.

Common Online Data Analysis Platform [Computer software]. (2014). Concord, MA: The Concord Consortium.

Jones, R. S., Lehrer, R., & Kim, M.-J. (2017). Critiquing statistics in student and professional worlds. *Cognition and Instruction*, 35(4), 317–336.

Konold, C., Higgins, T., Russell, S. J., & Khalil, K. (2015). Data seen through different lenses. *Educational Studies in Mathematics*, 88(3), 305–325.

Konold, C., & Kazak, S. (2008). Reconnecting data and chance. *Technology Innovations in Statistics Education*, 2(1), 1–37.

Konold, C., & Miller, C. D. (2011). *TinkerPlots: Dynamic Data Exploration*. Key Curriculum Press.

Moore, D. S. (1990). Uncertainty. In L. A. Steen (Ed.), *On the shoulders of giants: New approaches to numeracy* (pp. 95–137). National Academy Press.

Lehrer, R., Kim, M.-J., & Schauble, L. (2007). Supporting the development of conceptions of statistics by engaging students in modeling and measuring variability. *International Journal of Computers for Mathematics Learning, 12,* 195-216.

Lehrer, R., & Romberg, T. (1996). Exploring Children's Data Modeling. *Cognition and Instruction*, 14(1), 69–108.

Lehrer, R., & Schauble, L. (2021). Stepping carefully: Thinking through the potential pitfalls of integrated STEM. *Journal for STEM Education Research*, 4(1), 1–26.

Lehrer, R., Schauble, L, & Wisittanawat, P. (2020). Getting a grip on variability. *Bulletin of Mathematical Biology*, *82*:106.

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical inquiry. *International Statistical Review*, 67(3), 223–248.

Wilkerson, M. H., & Polman, J. L. (2020). Situating Data Science: Exploring How Relationships to Data Shape Learning. *Journal of the Learning Sciences*, *29*(1), 1–10.

