THE IMPACT OF SELF-COLLECTED DATA ON STUDENTS' STATISTICAL ANALYSIS

Karoline Smucker
The Ohio State University
smucker.37@osu.edu

Azita Manouchehri
The Ohio State University
manouchehri.1@osu.edu

Context plays a crucial role when students engage in analysis and interpretation of data, and is broadly recognized as one of the main elements that separate the studies of mathematics and statistics (Cobb & Moore, 1997). The ways that students engage with the environment surrounding data within a statistical task are of interest to researchers because of this. In this analysis of a teaching experiment with third grade students, interactions with self-collected wingspan data led to insights about graphs as students attended to individual values, but students varied in their willingness to use the data to describe/infer more broadly.

Keywords: Data analysis and Statistics, Measurement, Elementary School Education

When working with data, the context surrounding it plays an essential role in developing students' understanding. Context is one feature that can separate the study of statistics from the study of mathematics (Cobb & Moore, 1997). Watson (2009) posited that one of the goals of statistical literacy is for "students to be able to tell a story from a context with a distribution" (p. 33). With the potential for students' statistical investigations to involve the collection and use of survey or measurement data (Lovett & Lee, 2016; Makar, 2014), it is of particular importance to carefully unpack the impact of self-collected data (which may include the students themselves) on their statistical analysis. In this work, we investigated the ways that students in a third grade classroom engaged with context as they completed an activity involving wingspan data. The guiding research question included: how would the use of self-collected measurement data impact students' descriptions and their ability to infer based on their data collection and display?

Conceptual Framework

According to Watson (2007), a learner's ability to engage productively with context is one of the hallmarks of statistical literacy. When "meaningless or nonexistent" context is provided for data, students will often apply incorrect or inappropriate procedures, while when the context is "interesting, and relevant to students' worlds" they are more likely to use statistical ideas (Doerr & English, 2003, p. 111). However, when students engage with the context too personally, they may also make assumptions that are not supported by statistical evidence (Watson, 2007). Students may focus so closely on the story behind data that they prefer making conclusions based on their informal knowledge of the data's context over patterns in the data itself (Ben-Zvi et al., 2012; Biehler et al., 2018; Pfannkuch, 2011).

Within statistics, there are two broadly defined categories; descriptive and inferential. Descriptive statistics focuses on the "organization, summarization, and presentation" of data (Paparistodemou & Meletiou-Mavrotheris, 2008, p. 83). Inferential statistics looks at how patterns for a particular group could be viewed more broadly. Recently, research has focused on the importance of exposure to tasks which build inferential reasoning for students as young as the elementary grades (Makar, 2014; Pfannkuch, 2011; Watson, 2001). Graphical displays are an important part of both descriptive and inferential statistics. They can help students see the distribution of data more clearly, with distribution being defined as the overall picture of data based on expectation (center), variation, and shape (Watson, 2009). Watson (2007) concludes based on her longitudinal work with students in both the elementary and secondary grades that learners can benefit from opportunities to create their own displays and compare/contrast them with displays made by others to investigate their "success in telling the story of data" (p. 61). Despite this much remains unknown about the intricacies and

In: Sacristán, A.I., Cortés-Zavala, J.C. & Ruiz-Arias, P.M. (Eds.). (2020). *Mathematics Education Across Cultures: Proceedings of the 42nd Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*, Mexico. Cinvestav / AMIUTEM / PME-NA. https:/doi.org/10.51272/pmena.42.2020

impact of such an approach. This work looks specifically at the context of self-collected measurement data and how the nature of the data may have impacted elementary students' display creation and descriptions/inferences.

Methods

This research aimed to investigate the ways that a group of third grade students from a Midwestern State engaged in a task related to class wingspans. Using characteristics of a teaching experiment methodology (Steffe & Thompson, 2000), a focus group of five students were recruited to participate in the study. The wingspan task unfolded in three stages. First, the focus group were asked to measure each other's wingspans, and to consider reasons for observed differences. Second, the focus group collected measurements from the rest of their class and used these measurements to create displays individually or in pairs. Students had previously been introduced to bar graphs and pictographs involving categorical data but had not used them to represent quantitative data. Finally, participants presented their displays to the focus group. The purpose of this discussion was to consider the ways that different displays conveyed similar information, examine the overall distribution of class wingspans, and hypothesize about the distribution of wingspans for other classes or grade level groups. Video recordings of discussions, transcriptions of the activity, and student work served as data sources for analysis.

Data analysis consisted of three phases. First, video-data were segmented to isolate instances during which students seemingly contemplated the self-collected or personal nature of the data. These were broadly defined to consist of occasions where they referenced specific data values or general characteristics of specific groups (i.e. their class, other classes, other grades). The second phase of analysis focused on identifying ways that the context shaped students' interactions with the data towards description or inference. The third phase concentrated on analyzing the type of descriptions/inferences identified during the second phase of analysis.

Findings

The context of the physical measurement of wingspans impacted the ways that students engaged with the data in a descriptive sense. It led to a focus on individual students' measurements which both helped and hindered students' perceptions as they attempted to describe the data. Including self-collected data based on their own classroom seemingly led to an appreciation for how individual values fit within the data, but at times this focus on individuals appeared to inhibit students from looking at their displays more holistically.

Personalization of data

Locating self-based measurements amongst the data set was critical to enhancing students' understanding of the characteristics of displays produced. This approach to data reading seemed to have motivated them to both create and understand their graphs (Lovett & Lee, 2016). Finding one's own measurements in the display anchored their interpretations and descriptions of the overall data as depicted in this vignette (all names are pseudonyms):

Julie: Everybody's was in the 50s but mine – but that kind of makes sense because I'm shorter than everybody.
Researcher: Oh, so you think that maybe it's connected to your height?
Katelyn: Yeah, because people say that if you're really tall then you have a big wingspan
Julie: Yeah like Katelyn and Bob are the tallest people and they have the longest – um
Bob: Wingspan

In both this and other instances, students referenced the data values of both themselves and others in the class. One wingspan, which was lower than all the others, is given an interpretation when Julie recognizes that the value represents her and makes a potential connection with height. This is pursued further by others who recognized that the higher wingspans also represent those in the group who are taller. In this way students were able to give some justification for the variation in the data and make a potential connection with height. Their prior experiences may have impacted this interpretation, as evidenced by Katelyn's statement about what "people say" about height and wingspan.

Interpreting based on individual measurements or personal experiences led to varying approaches to inference. When asked what the graph for another, unspecified third grade class might look like, students first believed that the graphs might be similar, as when Bob and Julie referenced their displays (Figure 1) and said, "It would look a lot like this/ours", or when Jeff said, "Most of them would be in the 50s". This provides evidence that some students viewed their data as representative of third graders in general. Others believed the graphs would look different, but for varying reasons. Abbie based her analysis on another specific class, saying, "If I were to measure Ms. Johnson's class, she has some really tall kids," leading her to conclude that they would have more values on the right side of the graph. Katelyn was unwilling to describe what a graph for another class would look like because they have "different people – one class could have all the tall people". In this case, envisioning specific classes led to inferences regarding a single case (Abbie) or an unwillingness to make any conclusion (Katelyn), characteristics that have been seen previously in young students with regards to informal inference (Ben-Zvi et al., 2012; Makar, 2014). Abbie's and Katelyn's observations also provide further evidence that students inferred a more general relationship between height and wingspan; for them, referencing height appears to be sufficient to justify potential variation in wingspan.

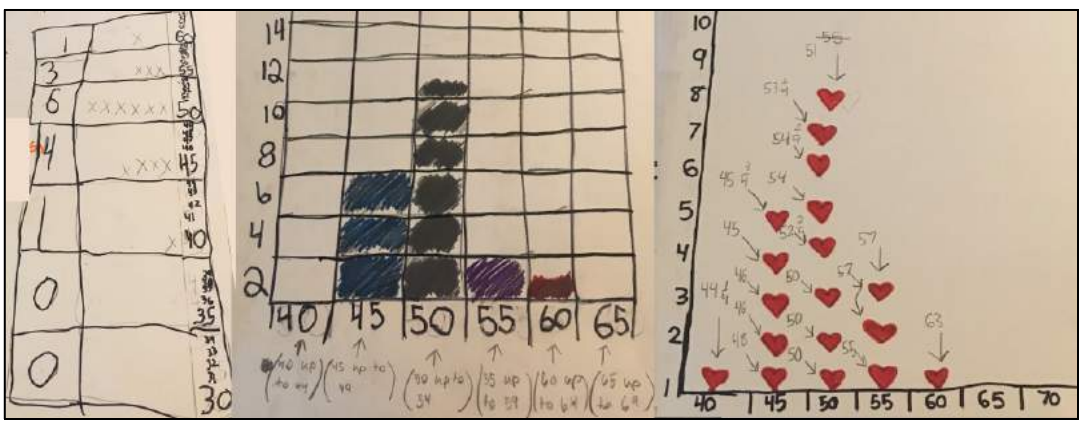

Figure 1: Jeff/Bob's display (left), Julie's display (middle), Katelyn's display (right)

Differences in Displays

The self-collected data were helpful for engaging students in interpreting their wingspans at an individual level, but the data collection and display creation process also led to some issues which kept students from seeking structural characteristics of the data. Students were intent on critiquing the format of others' displays (things like labels and titles), and characterized them as either correct or incorrect. This tendency can plausibly relate to the procedural and standalone nature of instruction on graph creation that students often experience (Friel et al., 2001). This led to a focus on minor discrepancies in students' data lists. Some values were left out accidentally by one or more of the participants, while other values differed slightly based on measurement. While this led to a discussion of measurement variability, students were also intent on tracking and noting these differences. Questions like, "Did you get John?" or "Which number is Maggie?" were common. When asked to compare their graphs, two students said, "They're all different". The following captures the difficulty students experienced in viewing and interpreting their displays holistically:

Researcher: Look at these side by side and look at the shape. What do you notice?
Abbie: Oh, it's the same
Jeff: Oh they're the same
Bob: It's the same thing
Jeff: It's the same thing except these are pictures and those are like bars (pointing)
Abbie: Wait wait wait!
Researcher: What do you mean it's the same thing?
Abbie: It's not the same thing – it's not the same thing
Katelyn: It's not the same thing because Julie is missing a 40

Note that while students' graphs had similar structures (Figure 1), Julie's graph does not include a value in the 40-44 inch category, whilst the other two graphs do. Thus, an exchange that initially seemed to be moving students towards a discussion of the more general characteristics of their displays transitioned to a focus on an individual value. Discrepancies related to the self-collected nature of the data seemed to inhibit their ability to summarize the wingspan characteristics in this instance. Despite this, students did note some of the general characteristics of the data based on their displays. A consensus was reached that the "average" for the class was "50-55 inches" because it was "the highest" value amongst the data set. Thus, students' conception of "average" in this case appeared more closely aligned with ideas related to the formal ideas of "median" and "mode", a phenomenon previously reported by researchers who had considered students of the same age group (Makar, 2014; Mokros & Russell, 1995). Katelyn also noted during the discussion how the data "all go up and down", a reference to the unimodal shape of the distribution. Students determined that potential values could be anywhere between 40 and 65 inches, and believed this to apply to third graders in general. While some students envisioned that the graph for another class might look different in terms of "average", students also believed that the wingspans of another class would fall within this range, demonstrating awareness of reasonable variation between samples.

Discussion/Limitations

This research investigated how the use of self-collected measurement data would impact students' statistical analysis. The results suggest that positioning themselves within the data at times helps students distinguish relationships between values. However, the personal nature of the data, along with slight differences in representation due to measurement/sampling variability, at times kept students' attention on the differences between their graphs and the individuals represented by each element as opposed to aggregate characteristics of the distribution. However, the exploratory/preliminary nature of this study limits the scope of its conclusions. A future study comparing students engaged in tasks utilizing self-collected versus teacher-provided (not self-collected) data would be essential to gain further evidence of whether the characteristics observed in this study are truly related to the use of self-collected data.

References

Ben-Zvi, D., Aridor, K., Makar, K., & Bakker, A. (2012). Students' emergent articulations of uncertainty while making informal statistical inferences. *ZDM*, *44*(7), 913–925. https://doi.org/10.1007/s11858-012-0420-3

Biehler, R., Frischemeier, D., Reading, C., & Shaughnessy, J. M. (2018). Reasoning About Data. In D. Ben-Zvi, J. Garfield, & K. Makar (Eds.), *International Handbook of Research in Statistics Education* (pp. 139–193). Springer International Publishing.

Cobb, G. W., & Moore, D. S. (1997). Mathematics, statistics, and teaching. *The American Mathematical Monthly*, *104*(9), 801–823. https://doi.org/10.1080/00029890.1997.11990723

Doerr, H. M., & English, L. D. (2003). A Modeling Perspective on Students' Mathematical Reasoning about Data. *Journal for Research in Mathematics Education*, *34*(2), 110–136.

Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education*, *32*(2), 124–158. https://doi.org/10.2307/749671

Lovett, J. N., & Lee, H. S. (2016). Making sense of data: Context matters. *Mathematics Teaching in the Middle School*, *21*(6), 338. https://doi.org/10.5951/mathteacmiddscho.21.6.0338

Makar, K. (2014). Young children's explorations of average through informal inferential reasoning. *Educational Studies in Mathematics*, *86*(1), 61–78. https://doi.org/10.1007/s10649-013-9526-y

Mokros, J., & Russell, S. J. (1995). Children's concepts of average and representativeness. *Journal for Research in Mathematics Education*, *26*(1), 20. https://doi.org/10.2307/749226

Paparistodemou, E., & Meletiou-Mavrotheris, M. (2008). Developing young students' informal inference skills in data analysis. *Statistics Education Research Journal*, *7*(2), 83–106.

Pfannkuch, M. (2011). The role of context in developing informal statistical inferential reasoning: A classroom study. *Mathematical Thinking and Learning*, *13*(1–2), 27–46. https://doi.org/10.1080/10986065.2011.538302

Steffe, L. P., & Thompson, P. W. (2000). Teaching experiment methodology: Underlying principles and essential elements. In R. Lesh & A. E. Kelly (Eds.), *Research design in mathematics and science education* (pp. 267–307).

Watson, J. M. (2001). Longitudinal development of inferential reasoning by school students. *Educational Studies in Mathematics*, *47*, 337–372.

Watson, J. M. (2007). *Statistical Literacy at School: Growth and Goals*. Routledge.

Watson, J. M. (2009). The influence of variation and expectation on the developing awareness of distribution. *Statistics Education Research Journal*, *8*(1), 32–61.