REVEALING CONCEPTUAL DIFFICULTIES WHEN INTERPRETING HISTOGRAMS: AN EYE-TRACKING STUDY

Lonneke Boels[1,3], Rutmer Ebbes[1], Arthur Bakker[1], Wim van Dooren[2], Paul Drijvers[1]
[1] Utrecht University, The Netherlands
[2] KU Leuven, Belgium
[3]Christelijk Lyceum Delft, The Netherlands
L.B.M.M.Boels@uu.nl

Many people misinterpret histograms. The conjecture is that some of these misinterpretations emerge from the application of interpretation strategies associated with case-value plots. To investigate this, eye-movement data were collected from six university students solving questions on histograms and case-value plots. Analysis of gaze data and cued retrospective verbal reports showed that participants seemed to use a histogram interpretation strategy, a case-value plot interpretation strategy or an elimination strategy. Several participants appeared to use a single preferred strategy without distinguishing between the type of graphs at stake. As conjectured, some participants applied a case-value plot strategy also to histograms. In addition, analysis of gaze data and verbal reports suggest that more experienced participants abandoned their initial strategy when necessary.

BACKGROUND

Histograms are widely used to represent data. The literature on histograms documents many misinterpretations when people use histograms (e.g., Boels, Bakker, Drijvers, & Van Dooren, 2017; Cooper & Shore, 2010; Whitaker & Jacobbe, 2017). These misinterpretations persist despite various interventions (Garfield & Ben-Zvi, 2007). The aim of the current explorative case study is to reveal, with help of eye-tracking and cued retrospective verbal reports, possible causes of specific misinterpretations.

In line with the literature (e.g., Cooper & Shore, 2010), a histogram is defined as a graph with bars displaying continuous data of one variable, measured at interval or ratio measurement level and its density or—when bins are equal—(relative) frequency on the vertical axis. A systematic review of the literature revealed that two big ideas of statistics—*centre* and *variability*—play an important role in numerous misinterpretations students have when interpreting histograms (Boels, Bakker, Van Dooren, & Drijvers, 2018). For example, students are often not aware of the difference between a case-value plot or value-bar chart (a bar graph with one bar for each single case) on the one hand and a histogram on the other, despite the fact that centre and variability are depicted differently in both types of graphs (e.g., Cooper & Shore, 2010). For this explorative case study we focus on students' conceptual difficulties with measures of centre, because the target audience of our larger project—secondary school students—learns more about centre than about variability.

Data can be represented in a histogram, but also in a case-value plot, distribution bar graph, dot plot, or box plot. The conjecture for our study is that students' misinterpretations of histograms stem from the application of interpretation strategies associated with other types of graphs, in particular case-value plots. This conjecture is in line with findings from, for example, Cooper and Shore (2010). This leads to the research question: What are the differences between the interpretation strategies underlying students' answers when interpreting histograms and case-value plots?

METHOD

Participants and procedure

The sample included four bachelor and two master students (two male and four female) aged 19–27 years. The students participated voluntarily and a five euro reward was provided for participation. Participants were selected using convenience sampling. They were tested in a one-on-one setting in a quiet room to minimize distractions. Before starting the task, consent was signed. The aim of the study was explained as well as how to operate the software. The eye-tracker was calibrated using five points on the screen. After completion of all tasks, participants were asked for a verbal report using their own gazes as a cue to improve the quality of the report (Van Gog & Jarodzka, 2013).

Materials: construction of the items

The items were constructed in such a way that students would either be estimating the arithmetic mean, or comparing the arithmetic means of two graphs. The multiple choice answers were constructed so that a confusion of the arithmetic mean with other measures of centre would become visible (mode, midrange), as well as confusion of the frequency with the measured value. These confusions are in line with the misinterpretations found in the literature (Boels, Bakker, Van Dooren, & Drijvers, 2018).

For six of the questions with histograms, a case-value plot was constructed with the same salient features such as number of bars, “shape,” range and variable (weight) resulting in a total of twelve questions, see Figure 1 for an example of the graphs for the first and fourth item. In line with the recommendations of Orquin and Holmqvist (2017) the items (stimuli) were constructed so that the “stimuli […] differ systematically on one or more features” (p. 6). These features were carefully chosen so that a student applying the same interpretation strategy on both the histogram and the case-value plot could be expected to answer the question for the case-value plot correctly and for the histogram incorrectly, or vice versa.

The arrangement of the items was such that there were never more than two graphs of the same type (histogram or case-value plot) in succession. The first two items were single histograms as in Figure 1a, followed by an item in which participants had to compare two case-value plots. Graphs with the same salient features never followed one another directly. The question for the histogram in Figure 1a was: What is approximately the average weight of the parcels delivered by Thijs? The question for the single case-value plot in Figure 1b was: What is approximately the average weight that has been collected per person?

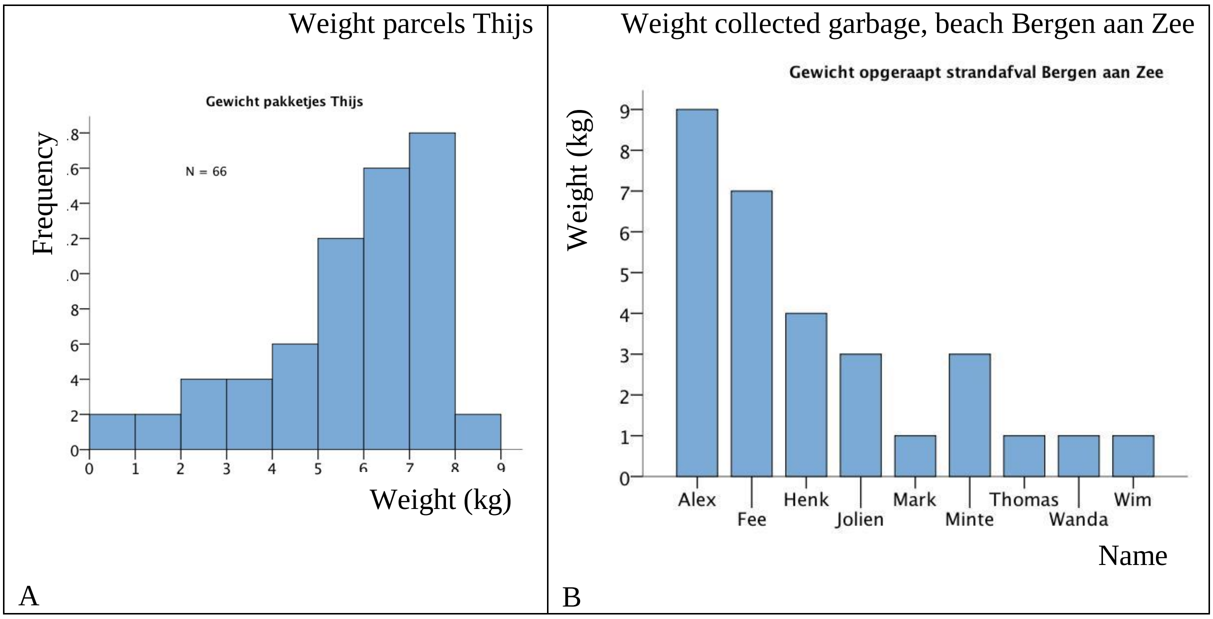

Figure 1. An example of a histogram (a) and a case-value plot (b)

Equipment and data analysis

The eye-tracker Tobii Pro X2-60, allowing for a sampling rate of 60Hz, was mounted on a laptop with a 13-inch display using adhesive mounting brackets. The Tobii Pro Studio 3.4.7 software recorded in real time where people were looking on the screen. Online survey software Qualtrics was used to present the tasks and record the answers. An USB audio-recorder was used to record the verbal reports. Gaze data and the verbal reports were coded qualitatively by the first two authors using open coding, axial coding and selective coding (Corbin & Strauss, 1990) to infer the strategies that participants use. Gaze data were also displayed in heat maps, showing in gradients from green via yellow to red where a participant looked most often (see e.g., Figure 2).

PRELIMINARY RESULTS

Participants used various strategies that were classified as either a *histogram interpretation strategy*, a *case-value plot interpretation strategy* or an *elimination strategy*. An example of a histogram interpretation strategy is referring to the arithmetic mean as a point on the horizontal axis where the distribution seems in balance. This becomes apparent from gaze data through a focus mostly on the horizontal axis and the height of the bars, and from the verbal data through referring to a balance. For instance, a participant mentioned: “Here I also was busy with that balancing.”

An example of a case-value plot interpretation strategy is compensating high bars with low bars in order to obtain bars of equal length, corresponding to the arithmetic mean. This becomes apparent from gaze data through students’ focus on the tops of the bars and the corresponding vertical axis values, and from verbal data through students mentioning “redistributing” or “compensating”.

A more generally applicable strategy is eliminating multiple choice possibilities that seem impossible answers. This becomes apparent from gaze data through a focus mostly on answers, and from verbal data through mentioning eliminating answer options. For instance, a participant mentioned: “I knew from the start that neither one nor nine could be an option.”

Most participants seem to make use of a single strategy for both case-value plots and histograms. In Figure 2 the number of fixations of one participant is shown from green (less), via yellow, to red (most). For instance, this participant interpreting a histogram looked mostly at the frequency axis (yellow spot around value 7 on the vertical axis, arrow in Figure 2a) and the same height in the graph area (red spot), indicating a horizontal looking pattern. This horizontal looking pattern is more in line with the strategy that leads to the correct answer for the case-value plot (see Figure 2b), but seems to be incorrectly applied to histograms. Applying inappropriate strategies also becomes apparent in the verbal data. For instance when a participant mentioned: “in this [a histogram other than below] it is harder to redistribute [i.e. compensate high and low bars; a case-value plot interpretation strategy], because there is only one really high bar”.

Another result from the gaze data is that some participants do not seem to read the title of several graphs. For example in Figure 2a and 2b no green spots are located on the title at the top of the graph.

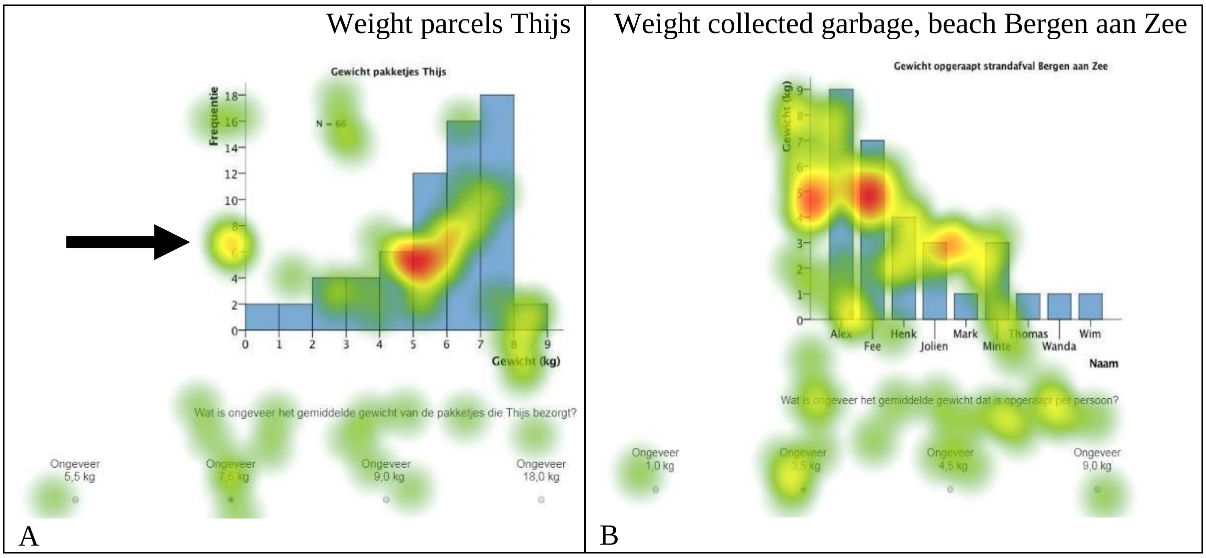

Figure 2. Histogram (a) and case-value plot (b) with heat map overlay, corrected for systematic error. For the labels on the axis, see Figure 1. The arrow (a) points at the fixations around frequency 7.

CONCLUSIONS AND DISCUSSION

The first tentative conclusion is that participants used an initial preferred strategy when interpreting histograms and case-value plots, without distinguishing the type of graph at stake. This strategy can be classified as a histogram interpretation strategy, a case-value plot interpretation strategy or an elimination strategy. The second conclusion is that our findings are consistent with the conjecture that students often fail to distinguish between case-value plots and histograms (Cooper & Shore, 2010).

The use of eye-tracking technology is still quite new in statistics education research, so its relevance is worth discussing. Based on our findings, eye-tracking data triangulated with cued retrospective verbal reports seem to have the potential to confirm or refute conjectures on the interpretation strategies that cause people to misinterpret histograms.

Eye-tracking data also appear to have some advantage over purely verbal reports in revealing interpretation strategies that participants are unaware of—including abandoning the initial strategy before arriving at the final answer—as these were not always reported in the verbal data. This is in line with findings from other research, for example from Epelboim and Suppes (2001) who state that "Spoken or written protocols have limited value because only the inferences that reach 'awareness' can be reported" (p. 1561). They also indicate that these inferences may be the result of several smaller steps that cannot be verbalised by the participant without interrupting the thinking process.

Furthermore, when students could not explain their strategy, gaze data nevertheless seemed to indicate a strategy. When dealing with participants having difficulty in verbalizing their thoughts, gaze data may offer an opportunity to reveal their strategies. This is also relevant for the next phase of our project in which we will expand the research to a larger number of secondary school students. We anticipate that some of these students will find it difficult to express their thoughts, so we expect an advantage of using eye-tracking.

Our research implicates that in education it might be necessary to give more explicit attention to case-value plots in addition to histograms (see also Cooper & Shore, 2010). We further expect that eye-tracking research has the potential to give more insight into the causes of students' difficulties with other big ideas in statistics, such as the concept of variability, or the concept of distribution. We therefore believe eye-tracking can support solving the question why some misinterpretations in statistics remain so persistent.

ACKNOWLEDGEMENT

The research is funded with a Doctoral Grant for Teachers from The Netherlands Organisation for Scientific Research, number 023.007.023 awarded to Lonneke Boels.

REFERENCES

Boels, L., Bakker, A., Drijvers, P., & Van Dooren, W. (2017). Conceptual difficulties with histograms: A review. In B. Kaur, W. K. Ho, T. L. Toh & B. H. Choy (Eds.), *Proceedings of the 41st PME Conference, Vol. 1* (p. 172). Singapore: PME.

Boels, L., Bakker, A., Van Dooren, W., Drijvers, P. (2018). *Conceptual difficulties when interpreting histograms: A review.* Manuscript in preparation.

Cooper, L. L., & Shore, F. S. (2010). The effects of data and graph type on concepts and visualizations of variability. *Journal of Statistics Education, 18*(2).

Corbin, J., & Strauss, A. (1990). Grounded theory research: Procedures, canons and evaluative criteria. *Zeitschrift Für Soziologie, 19*, 418–427.

Epelboim, J., & Suppes, P. (2001). A model of eye movements and visual working memory during problem solving in geometry. *Vision Research, 41*, 1561–1574.

Garfield, J., & Ben-Zvi, D. (2007). How students learn statistics revisited: A current review of research on teaching and learning statistics. *International Statistical Review, 75*, 372–396.

Orquin, J. L., & Holmqvist, K. (2017). Threats to the validity of eye-movement research in psychology. *Behavior Research Methods.* doi.10.3758/s13428-017-0998-z

Van Gog, T., & Jarodzka, H. (2013). Eye tracking as a tool to study and enhance cognitive and metacognitive processes in computer-based learning environments. In: R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 143–156). New York, NY: Springer.

Whitaker, D., & Jacobbe, T. (2017). Students' understanding of bar graphs and histograms: Results from the locus assessments. *Journal of Statistics Education, 25*(2), 90–102.