

# Reflections on gaze data in statistics education

Lonneke Boels<sup>1,2</sup> 

<sup>1</sup>University of Applied Science Utrecht, Utrecht, the Netherlands

<sup>2</sup>Utrecht University, Utrecht, the Netherlands

## Correspondence

Lonneke Boels, University of Applied Science Utrecht, Utrecht, the Netherlands.  
Email: [lonneke.boels@hu.nl](mailto:lonneke.boels@hu.nl)

## Funding information

Nederlandse Organisatie voor Wetenschappelijk Onderzoek, Grant/Award Number: 023.007.023

## Abstract

Gaze data are still uncommon in statistics education despite their promise. Gaze data provide teachers and researchers with a new window into complex cognitive processes. This article discusses how gaze data can inform and be used by teachers both for their own teaching practice and with students. With our own eye-tracking research as an example, background information on eye-tracking and possible applications of eye-tracking in statistics education is provided. Teachers indicated that our eye-tracking research created awareness of the difficulties students have when interpreting histograms. Gaze data showed details of students' strategies that neither teachers nor students were aware of. With this discussion paper, we hope to contribute to the future usage and implementation of gaze data in statistics education by teachers, researchers, educational and textbook designers, and students.

## KEYWORDS

eye-tracking, histogram, machine learning algorithm, teaching statistics

## 1 | INTRODUCTION

Despite being promising, gaze data are still uncommon in statistics education [1,2]. In a review, Strohmaier et al. [3] found only four studies in statistics education using eye-tracking [4–7]. Strohmaier et al. described the added value of gaze data so that students' solution processes become visible—including approaches that students never articulated—and that these processes are neither disturbed nor interrupted, and that gaze data can make some complex cognitive processes visible (eg, for statistical thinking). Gaze data provide teachers and researchers with a new window into these processes.

The goal of the present discussion article is to reflect on what we have learned so far from our own studies in which we collected gaze data of students and teachers. In addition, we discuss possible future applications of eye-

tracking in statistics education. By discussing what and how teachers can learn from gaze data, we aim to contribute to the future usage and implementation of gaze data in statistics education by teachers, textbook designers, and eventually, students.

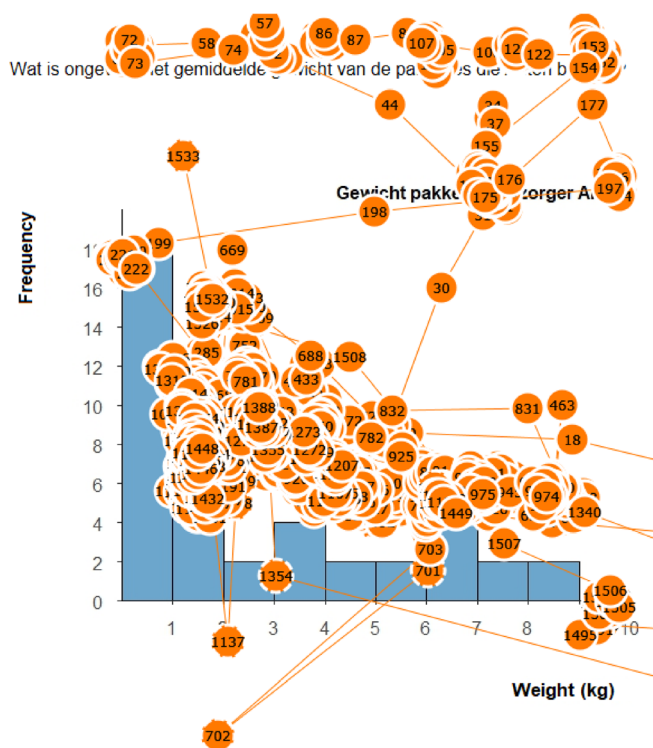
Gaze data have been around for some decades but are now becoming more widely available. The nature of these data is quantitative (eg, time, location coordinates). Furthermore, these data are often aggregated into qualitative (non-numerical) forms such as videos (dynamic gazeplots) and images (static gazeplots, heatmaps). Both qualitative and quantitative forms allow for various types of analysis. These are new forms of data as they are new to most statistics teachers—and researchers. We, therefore, first need to understand what gaze data are and how these can be analyzed before statistics teachers can use such data for their teaching or teach their students to work with them (Box 1).

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Author. *Teaching Statistics* published by John Wiley & Sons Ltd on behalf of Teaching Statistics Trust.

### BOX 1 Example from empirical research

In our study [8], students were asked to *estimate* the arithmetic mean weight from graphs with bars (histograms and case-value plots). We asked several students what strategy they used. We found three most common strategies, one correct strategy, and two incorrect strategies. In one of the incorrect strategies, students read off values from the vertical scale. They explained, for example, that they made all bars equally high by cutting the long bars and distributing the pieces over the other bars. The gaze pattern belonging to this strategy (Figure 1) roughly follows a horizontal line. Note that this particular student did not seem to have looked at the word frequency nor at the word weight along both axes even though they were not told that this graph was a histogram.



**FIGURE 1** Example of a student applying an incorrect strategy to a histogram to estimate the mean weight of the packages that a postal worker delivered. Note that places where students fixate on the screen are represented by circles (called fixations) and fast transitions between the two are represented by thin lines (called saccades). This student never seemed to have looked at the axes titles. Translated for readers' convenience. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

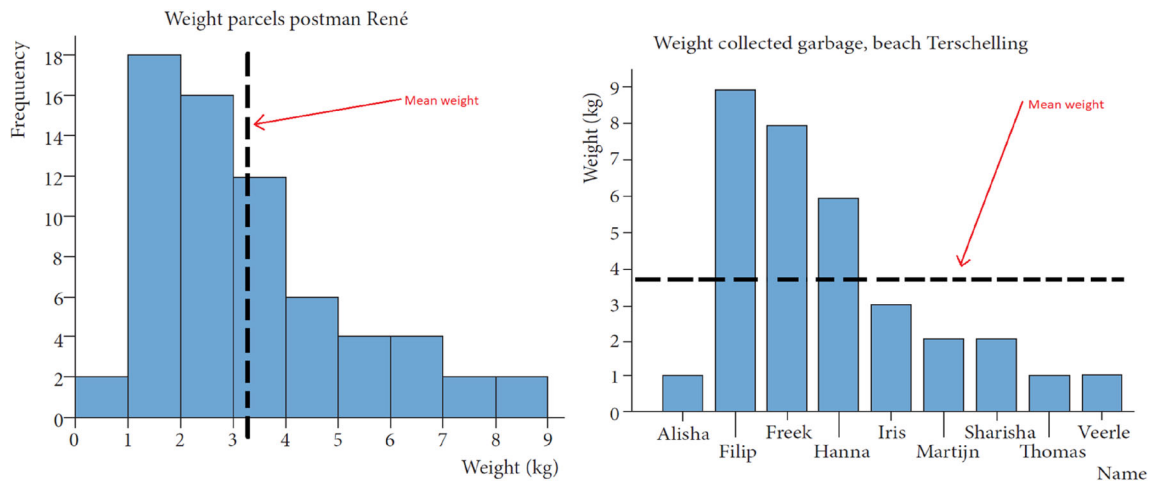
## 2 | BACKGROUND INFORMATION ON EYE-TRACKING

### 2.1 | Students' gaze patterns on statistical graph tasks

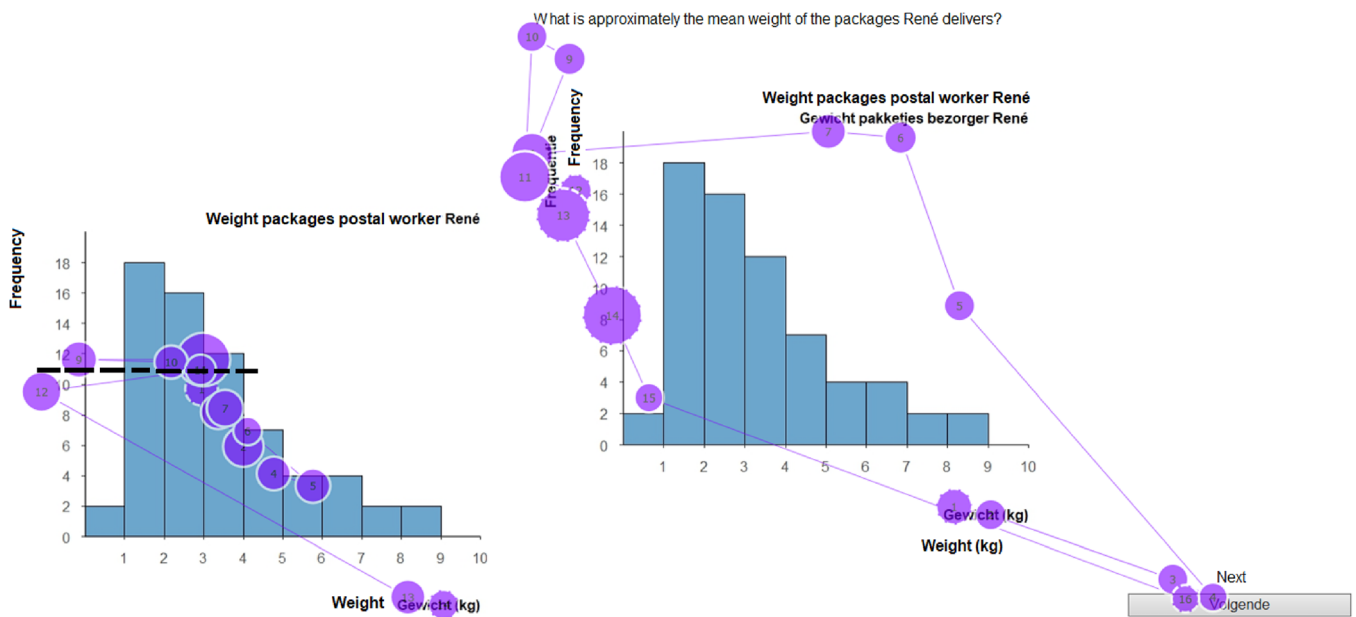
As an example, we discuss a study in which 50 Grades 10–12 pre-university track students (15–19 years old) solved several tasks. All students have had statistics in Grades 7–9, including reading off values from histograms. Some of the tasks in the study required them to estimate arithmetic means of statistical graphs (eg, histograms, case-value plots [8]). We used the mean, as it can be seen as a precursor for estimating variability from a graph. Moreover, Gal [9] already argued that asking students to estimate the mean from a graph can reveal gaps in students' knowledge. For an elaborate discussion on the use of the arithmetic mean and a review of the literature on the mean, we refer readers to the original article [8]. A possible strategy for estimating the arithmetic mean from a histogram is that students search for the balancing point of the graph [10,11], a strategy not familiar to these students. Although computational strategies are also possible, most students used a visual search strategy for finding this mean.

The gaze data were collected with a Tobii XII-60 eye-tracker mounted on a laptop, which tracks where a participant looks on the screen of a laptop through harmless infrared light. Participants were asked to estimate the mean weight from either a histogram or a case-value plot (see Figure 2 for examples of both graph types).

We first qualitatively analyzed students' gaze data by looking at the videos generated by the Tobii Studio Software version 3.4.5. After watching students' gaze behavior in more than 600 trials (with histograms or case-value plots) and analyzing students' retrospective verbal reports, we found out that the perceptual form of students' gaze patterns on the graph area was most relevant for students' strategies on the tasks [8] (indicated by a dotted line, Figures 2 and 3). For example, several students showed a horizontal gaze pattern when estimating the arithmetic mean from a histogram (Figure 3, left) which indicates a strategy for interpreting the graph at hand as if it was a case-value plot (Figure 2, right). Whether students did or did not read axes titles did not seem to influence students' strategies for solving the histogram tasks. For instance, a student used an incorrect strategy for finding the mean from a histogram (Figure 3, left). Next, this student checked the titles of the horizontal and vertical axes (Figure 3, right) but stuck to their incorrect answer for the mean weight ("ten," even though gazes seem to be around eleven). While a secondary school teacher may always emphasize that students should carefully read the graphs, graph titles,



**FIGURE 2** Possible correct strategies for finding the arithmetic mean from a histogram by estimating the balancing point (left) and from a case-value plot by making all bars equally high (right). Both strategies were found in research with high school and university students and teachers. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



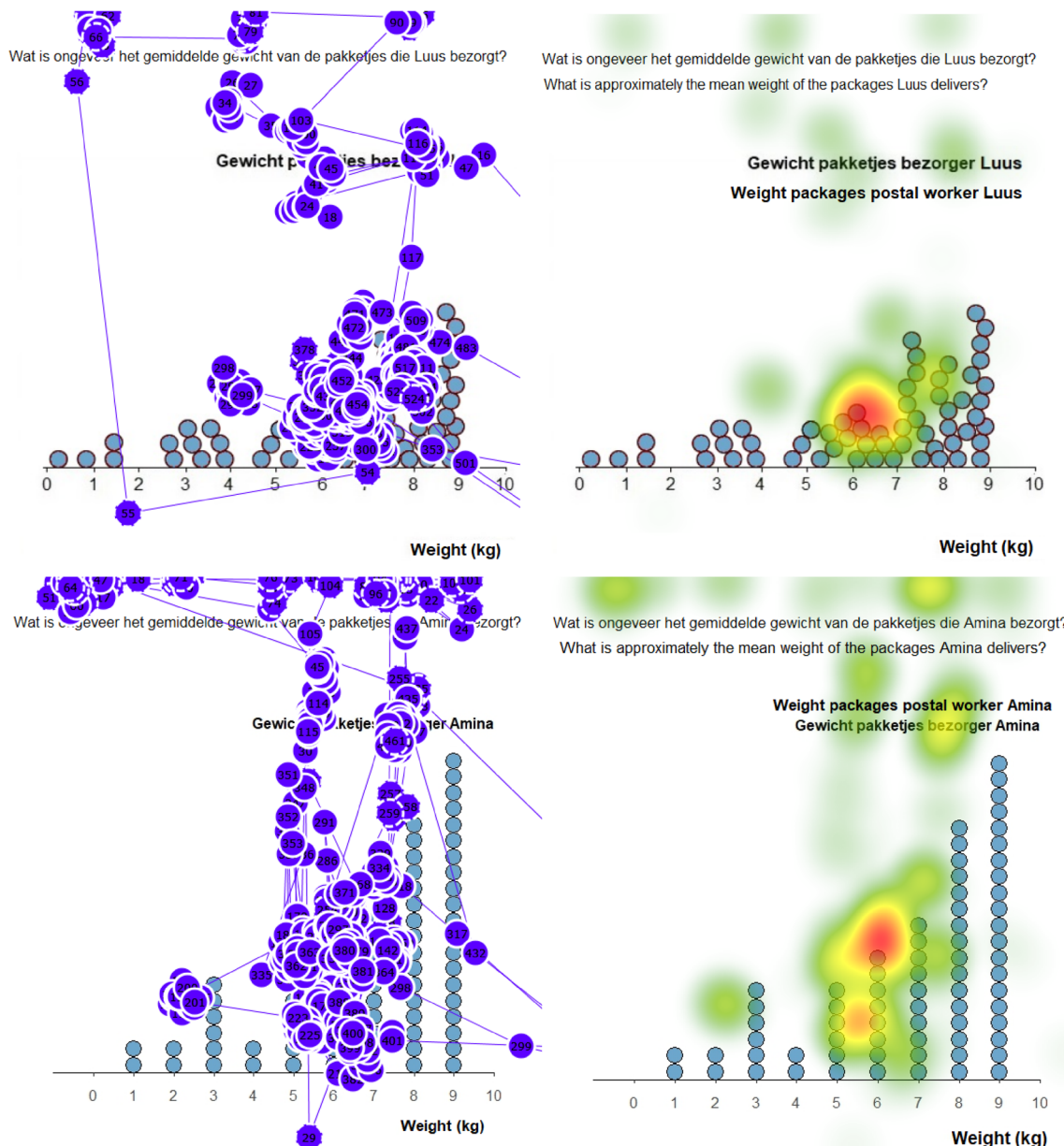
**FIGURE 3** Part of a gaze pattern of a student incorrectly estimating the arithmetic mean weight from a histogram (left) followed by checking the axes and graph titles. Translations and black line segments are superimposed for the convenience of the reader. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

and axes titles, this research suggests that for some students this may not be enough.

In a follow-up study with machine learning algorithms (MLAs), we used the fixations on the graph area [12]. In addition, we constructed a simple model, which decided whether the gaze pattern on the graph area was more horizontal or more vertical. From that study, we learned that both our simple model and the MLAs could quite accurately classify whether students used a correct (Figure 2, left) or incorrect strategy (Figure 3, left) for estimating the mean from a histogram. This not only confirmed the results of our previous, qualitative

study but also indicated that the process of strategy cognition can be automated. This opens up possibilities for future use of gaze data in statistics education, for example in an intelligent tutoring system or teacher dashboard [cf. 13].

Data were also collected on students' gazes on messy and stacked dotplots. Lyford [14] found that some students had difficulties interpreting stacked dotplots. As we expected students to have fewer difficulties with messy dotplots, we provided them with both (Figure 4). Although our students did not learn dotplots in previous grades, they were quite capable of estimating means from



**FIGURE 4** Gaze pattern of a student correctly interpreting a messy (top) and stacked (bottom) dotplot. Correct answers: 6 (both graphs; answers were accepted as correct when lying within the range 4.9–7.1). Translated for the readers' convenience. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1111/tesl.12340)]

single dotplots (see also [15]) and did slightly better with messy dotplots than stacked dotplots [eg, 16].

## 2.2 | Review of the literature on eye-tracking

A major advantage of eye-tracking is that it can make students' task-specific strategies visible at a great level of detail (eg, [17]). In addition, it can make strategies visible that participants are unaware of or are unable to articulate (eg, [18]). However, there is no simple relation between gaze patterns in general and students' strategies

(eg, [19]). Moreover, not every gaze is part of students' strategies (eg, [20]). Research is needed to reveal how specific gaze patterns relate to students' strategies for different topics (eg, [21]). In addition, data triangulation, for example through cued recall (retrospective reporting with students' own gazes as a cue), will be needed until clear patterns have been found for specific tasks and topics and in different communities. We found such patterns for estimating means from histograms and case-value plots for university students [5], teachers [22], and high school students [4,8] in the Netherlands. Future research is needed to find out if gaze patterns on those tasks are similar in different cultural settings and educational

systems around the world. In addition, future research is needed to reveal if and how gaze patterns are related to task-specific strategies for other tasks in statistics and mathematics education.

A question remains regarding what measures are useful for analyzing task-specific gaze data in statistics and mathematics education. Most eye-tracking studies use gaze data measures that are temporal (eg, total fixation duration or dwell time, reaction times, time to first fixation, total reading time), count (fixation count, number of saccades between relevant or irrelevant parts of the stimuli) or both (eg, [17,23,24]). The advantage of these measures is that they can be computed easily. The disadvantage, however, is that these computational measures often only provide a global insight into students' thinking processes—neglecting the level of detail Kaakinen [17] refers to—and that these measures do not provide task-specific guidance that can be used for learning or instruction. A question left for future research is: which of these metrics (if any) are relevant to statistics education researchers?

Reviews of eye-tracking studies in other fields (eg, communication, teacher education, medical education) provide several insights relevant to the use of eye-tracking research in statistics education: (1) data triangulation is highly recommended [19], (2) most studies that use eye-tracking to study learning contribute to general theories such as information processing or multimedia learning (eg, [23,25]), (3) eye tracking can be used for student learning, and (4) changes in gaze behavior occur during learning (eg, [26]).

### 2.3 | Eye-tracking in statistics education

Although eye-tracking has been around for some time, its use is still in its infancy in statistics education. Besides our conference papers [4,5,22], only two other studies were found in a literature review [3]. One study was on Bayesian reasoning strategies [6]. The other study was on statistical numeracy as a moderator of (pseudo) contingency effects on decision behavior [7]. Both used quantitative gaze measures, such as total time spent on an “Area of Interest” (AOI). The most common methods for handling gaze data are computational and stem from cognitive sciences that usually aim for more general strategies, such as self-regulated learning. However, quantitative measures, such as traditional time measures, can hide visual scanning patterns [27]. A similar argument can be made for count measures such as the percent of fixations on specific parts of the screen [28].

Spatial measures (eg, scanpath, fixation position, fixations sequence, gaze patterns) can disclose the kind of detailed information Kaakinen [17] refers to. Spatial

measures seem better suited for providing detailed information about students' thinking processes [29]. However, spatial measures are still quite uncommon in educational eye-tracking research [30]. Moreover, when people refer to scanpaths, they usually mean a sequence of transitions between areas of interest (eg, [1,31]).

Recently, two more studies were conducted. A study with dotplots in primary education investigated differences in students' strategies regarding local and global viewing of graphs [32]. In a study with boxplots, gaze data of university students including pre-service teachers were analyzed regarding their strategies. Strategies included comparing areas of the boxes and medians of boxplots [33].

### 2.4 | Points of attention: time investment, data triangulation, and ethical considerations

Before discussing possible future applications, it is important to highlight three points of attention in eye-tracking research. First, the substantial time investment it takes to initiate such research as an early researcher—in our case, roughly 1 year full-time for preparation and data collection, and then over half a year for qualitative data analysis. Second, the already mentioned necessity to triangulate data. Third, the ethical component of using students' gaze data. Already large amounts of data are collected every day when people click on a website or buy groceries in a supermarket. Although this can have several advantages, such as music websites offering music that you might like based on your previous choices, it has a downside too. Fry [34] provides several examples of improper data use to train machine learning algorithms that decide who is to be invited for a job interview or who is turned down for a loan. A similar ethical discussion about using such data needs to be started about gaze data. An advantage of gaze data is that they can provide insight into people's thinking processes at a level of detail that was not available before, or at least not without influencing the thinking process. One reassurance is that there is no general relationship between eye movements and thought processes and that this has to be figured out for specific tasks and situations each time. However, it is conceivable that in the future, faster methods will become available for this rather than the manual work that is still required today. It is, therefore, important that an ethical discussion be held now about who may collect and use students' gaze data. Some ethical questions could be: Are we going to hand this over to large tech companies—just as we did with earlier data—or will this remain reserved for non-commercial parties? Can students—and teachers—refuse to make their data available (something that currently seems

impossible when using Google Classroom)? Is the collection of gaze data fundamentally different from data that are already currently being collected about us and our students? It is important to think about these kinds of questions now and not after the technology has long been implemented in education. Therefore, this discussion can no longer wait. Currently, gaze data are already being used regularly in the gaming world, and it will likely be a matter of time before they are introduced into education.

In the next sections, the potential uses of gaze data in statistics education are discussed. In statistics education, gaze data are still an unusual or “non-traditional” source of gaze data. In addition, the way we analyzed gaze data—through spatial measures and MLAs—is also relatively new in education, even outside of statistics education. New developments happen quickly, but when we conducted the research reported here, we could not find a study within statistics education that took such an approach.

### 3 | POSSIBLE FUTURE APPLICATIONS OF EYE-TRACKING IN STATISTICS EDUCATION

#### 3.1 | Gaze data in a feedback tool

Gaze data can be used in—or as—a feedback tool [35]. A first possibility is to provide students with their own gaze data after solving one or more tasks and ask them to describe the strategy they used. From our analysis, we have indications that this may also help students reflect on their chosen strategy [12]. In our study, students had no difficulties interpreting their own gaze data during the cued recall. However, we used a way of showing their gaze data that differed from the visuals in previous Figures. Instead, we illuminated the location where students looked—through a kind of spotlight—and made the rest of the graph darker. This highlighted where a student looked instead of covering it with a red circle. Having students individually look back at their eye movements in this way is time-consuming and, therefore, not feasible for regular use in classrooms.

A second option is to provide students with the gaze data of other students. For example, students could be given a video or gazeplot of a gaze pattern from students who correctly interpreted the graph at hand. Next, these students could be asked to reflect on the other students' gaze patterns.

A third possibility is to provide students with immediate, personalized feedback based on their gaze data (eg, [35]). Automatic feedback will become possible when a number of conditions are fulfilled. First, there need to be

distinctive eye movement patterns that can be linked to specific strategies as we showed in our first study [8]. Second, after training, an MLA or an interpretable model needs to be able to extract these patterns from the gaze data of new students or for the same students on new, similar tasks. Both conditions are met for single histogram tasks as we showed in our second study [5]. A study investigating whether an MLA can be trained for dual histogram tasks is currently in progress. Third, there is a need for an inexpensive option to measure eye movements. Webcams seem to offer possibilities for this but with less accuracy (eg, [36]). This requires further research. Only when eye-tracking is inexpensive will large-scale application in the classroom, during homework or distance learning, and in MOOCs become feasible. Fourth, it is necessary to find out what form of feedback is useful. Does it work to let students see their previous eye movements or is another form of feedback needed?

A future possibility could also be to investigate how the process of interpreting their own gaze data can help students' reasoning with, for example, data, data representations, center, and variability. In addition, gaze data can be used to understand students' cognitive processes. Teachers often want to know what students are paying attention to. Posing questions during an intervention or experiment can shift students' attention from where they were at that moment to what they think the teacher is asking for. Viewing patterns can potentially provide similar insight into what students pay attention to without disrupting students' thought processes.

#### 3.2 | Gaze data as an informative tool for teachers

A first possibility is to provide teachers with *information* based on students' gaze data so that feedback can be given. Several questions then still remain open. Is it better if such a system reports back which students used a correct strategy, which students used an incorrect strategy, and for which students the strategy is unclear so that the teacher can intervene in a targeted way? Is it helpful—or necessary—to then provide the teacher with a record of students' eye movements, and if so, in what form?

The last question refers to a second possibility, namely, to provide students' gaze data directly to teachers. Gaze data can provide them with insight into students' reasoning. Many studies infer students' reasoning toward their answers from students' answers (eg, [37,38]). However, it could be that students use a productive strategy for solving the task at hand and still answer

incorrectly, or vice versa. As correct reasoning in statistics education is valued, this could provide teachers with a new tool to discover such correct reasoning. From our experience, it can sometimes even be possible to infer from the gaze data that a student started with a correct strategy that was abandoned for some reason. In addition, from previous research, we know that when students develop a sense of a topic, their gaze patterns change (eg, [39,40]). That research suggests that it is a combination of students' actions, perceptions, and feedback from the environment that results in a change in gaze patterns (cf. [41]). An example can be found in a study where students move the tops of two bars up on a tablet [42]. Students receive feedback on how well this is done according to an unknown—to the students—rule (a proportion) through the coloring of the screen or bars: green for correct and red for incorrect. A change in students' gaze patterns (from looking between the top of the bars to the middle of one bar and the top of the other bar) helps their coordination. A delay between such a change in gaze pattern and students' verbal reflections can also indicate readiness for learning [43] or that the task lies within the zone of proximal development [44]. In one study, we also found indications of changes in gaze patterns [12] that, triangulated with other data, suggest learning. Khalil et al. [2] found that the looking patterns of novices and experts differed. This raises the question of what the differences between the gaze patterns of novices and experts are in other statistical tasks.

The second possibility raises several other questions. Can teachers identify students' strategies—from students' gaze data—when students are interpreting a statistical graph? Do they need instruction for that and if so, how should such instruction look? Teachers could not only be asked whether they think a student had performed a correct strategy (or what strategy they think the student used) but also if the strategy was inappropriate, what kind of intervention they would do. Such questions were asked by bachelor students [unpublished thesis study] who (re)used the students' gaze data from our study [8]. Participants of their study were secondary school STEM teachers. To the best of our knowledge, their study is one of the first that provided STEM teachers with the opportunity to interpret and thus reason with this non-traditional kind of data. As their data collection was relatively small and hindered by the COVID-19 pandemic, we plan to investigate this further.

### 3.3 | Gaze data as a new form of data

In tertiary education, some students and (pre-service) teachers might also collect gaze data themselves. There

are already courses on eye-tracking. These are mostly taught in the neurosciences or psychology departments. Students collecting data in such courses are often interested in memory and cognitive load theory, self-regulated learning, and metacognitive skills or strategies. In marketing, gaze data are used for inferring decision behavior. Courses that focus on mathematical or statistical task-specific strategies inferred from gaze data seem to be rare.

### 3.4 | Gaze data to revise instructional design

Gaze data can also be used to revise the instructional design. Although we have not found studies that do so in statistics education, such studies do exist for mathematics education. Examples of using multimodal data for revising the design can, for example, be found in studies on proportions [45], equal areas [46], and trigonometry [47].

### 3.5 | Possible tools for using and analyzing gaze data: heatmap, raw data, videos, and static gazeplots

As this special issue is about “non-traditional data,” and because we were interested in spatial gaze patterns, our focus will now shift to how gaze *data* can be analyzed. The following will reflect on our choices and provide some guidance for making choices in future research and education.

In our studies, we analyzed the collected gaze data in two ways: a qualitative analysis of the videos of the gazes on graph tasks and quantitative analyses using the raw data of these tasks in an MLA. For these analyses, we used two types of data obtained through *data moves* [48] either by the Tobii Studio software [49] or by us: videos of the gazes (sometimes called dynamic gaze plots see Figure 5 middle left for an example), and ‘raw’ data that consist of  $x$ - and  $y$ -coordinates of the gazes on the screen for approximately every 17 milliseconds (Figure 5, left for a plot of such data). For the latter, we only used coordinates that were on the graph area.

Other data moves are possible as well. For example, instead of videos of the gazes, a static gazeplot can be used (middle right), here created through the Tobii software. Another possibility is using heatmaps (middle [5]), here created through the Tobii software (see also Ref. [21]). Heatmaps have the advantage that they aggregate the gaze data but a disadvantage is that time and spatial information (eg, the order of the fixations or saccades) are thrown away. Changing representations of gaze data

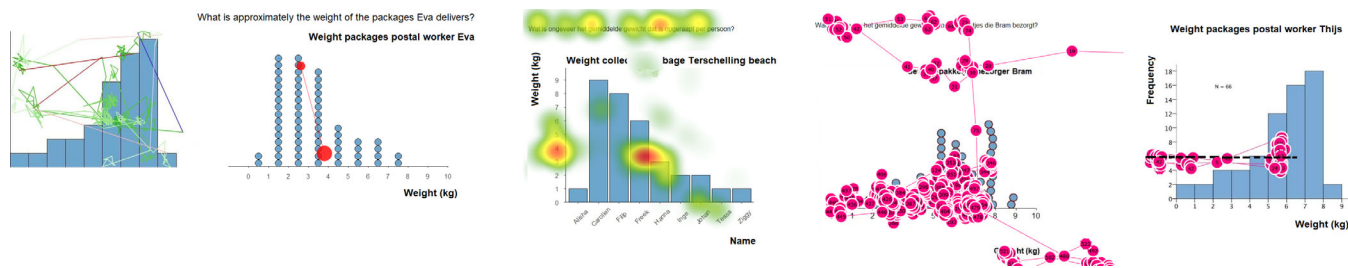


FIGURE 5 Examples of different ways to visualize and use gaze data. Raw scanpath (left), video still of smoothed gaze pattern (middle left), heat map (middle), gaze plot (middle right), and scanpath in gaze plot (right) indicating an imaginary horizontal line (superimposed for the reader). Permission for the use of the most left graph is kindly granted by Enrique Garcia Moreno-Esteva. Translated to English. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

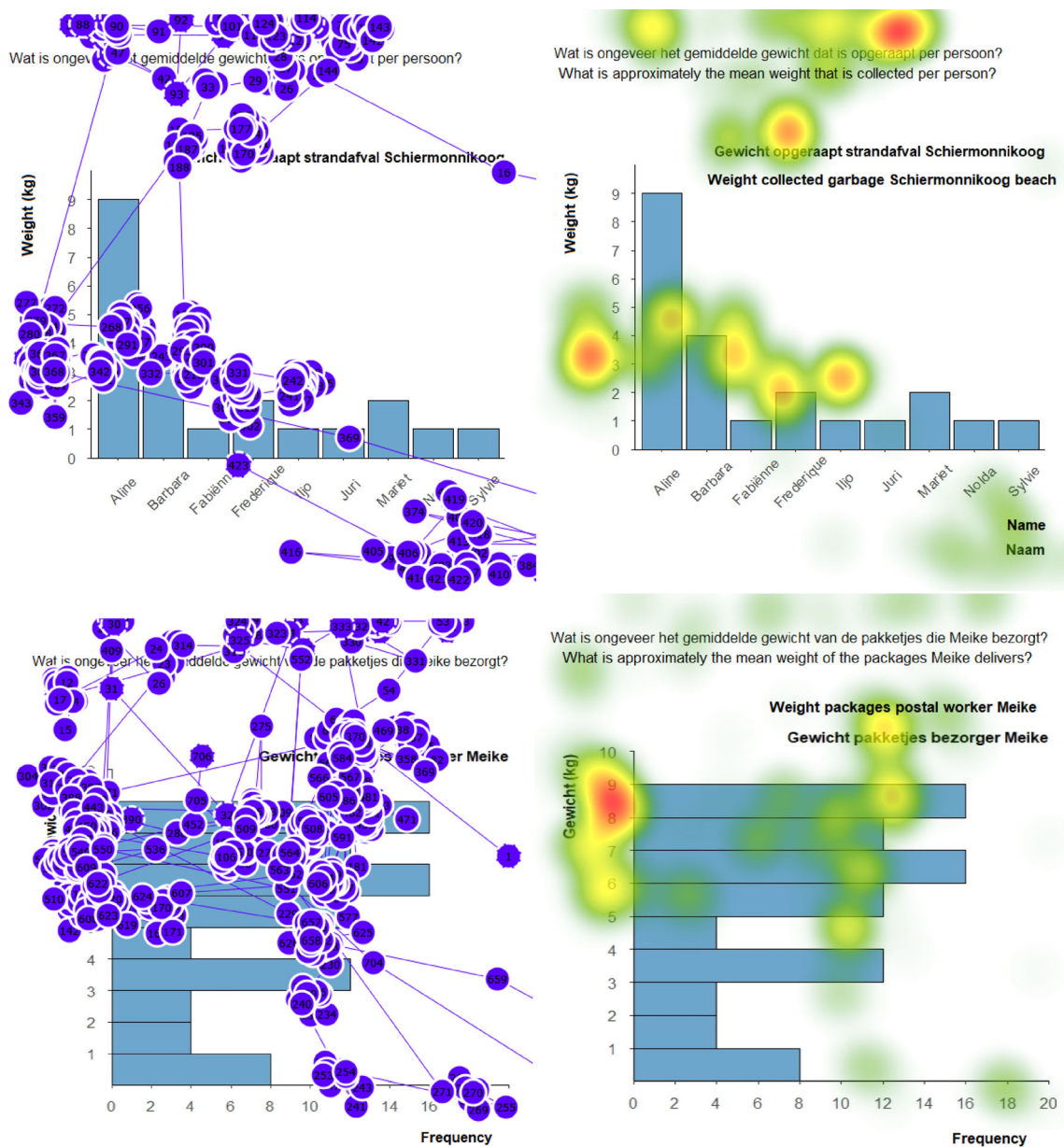


FIGURE 6 Gaze patterns (left) and heatmaps (right) of a student correctly interpreting a case-value plot (top) and horizontal histogram (bottom). Correct answers: 3 (top) and 7 (bottom). Translated for readers' convenience. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



from one form into another is part of transnumeration [50] and can be helpful in understanding what the gaze data can tell us.

For our qualitative analysis, we found the videos of the gazes to be the best approach. In this way, we could see the order of the gazes and what students attended to. Our attention was mostly on the saccades—fast transitions between fixations. Fixations are the positions on the graph where students looked. In the raw data depicted in Figure 5, fixations are indicated by short lines in a star-like form together. Saccades are long lines between those fixations. In the video still and statistic gazeplots, fixations are indicated by circles, and saccades are represented by the thin lines between them.

Another possibility for the qualitative analysis of spatial gaze data is heatmaps (Figure 5, middle). These have the advantage that attention is drawn to the fixations. Fixations on locations where the student spent little time are green and go to yellow and then red where more time is spent in total. Static gazeplots (Figure 5, middle right) have the advantage that all gaze data are shown and that the order is given (by numbers on the fixations). Unfortunately, for students spending a longer time on the graph area, the pattern, order of fixations and saccades, and graph can get hidden behind all the fixations. The most relevant part of the pattern can be isolated (Figure 5, right) but requires days of manual work as this needs to be done for every student and for every task separately, and the judgment of what belongs to this pattern—and what does not—is part of the qualitative analysis. Future research is needed to find out if heatmaps or static gazeplots can be used in similar ways for both qualitative analysis and analysis through MLAs. Will it be possible to infer students' strategies from static gazeplots and heatmaps (Figures 4, 5, and 6)?

For students' solution strategies, it appeared, both from the qualitative study and the quantitative machine learning analysis, not to be relevant if the horizontal gaze pattern we found is a bit higher or lower on the graph area; if a student looks from left to right or vice versa or if there is a little slope in this gaze pattern or not. The only importance seemed to be whether this pattern was mainly horizontal or mainly vertical. The irrelevance of such specific order and position on the screen has a potential advantage for future webcam usage. However, webcams are still much less precise in their calibration. When a horizontal or vertical shift of a horizontal gaze pattern still results in a horizontal gaze pattern on the graph area, an MLA could possibly still recognize this.

A possible implication of our work is that the order of areas of interest and scanpath similarity or idiosyncrasy might be less important for uncovering some task-specific solving strategies. We have some indications for that. First, from our first attempts with the machine learning

algorithms (MLAs)—not further reported—we suspect that reading axes titles and graph titles are less important than we thought. Adding gaze data on those areas seemed to add noise and reduce the accuracy of the MLAs, or at least we did not find a fruitful way to use those areas for uncovering task-specific strategies. It is left for future research to decide on what spatial measures are relevant for task-specific strategies and when. Second, we saw in the videos of the gazes that students sometimes explicitly checked the titles of a graph and then still misinterpreted the graph. For education, the implication of our indications is that just telling students to carefully read the axes and titles probably might probably not be enough. More research is needed to figure out if and how looking at axes titles is related to considering the consequences of these titles as part of students' task-specific strategy.

## 4 | CONCLUDING REMARKS

This article discussed how can gaze data be used in statistics education. Rather than pretend to discuss all possibilities, ideas for those who are not yet familiar with gaze data were provided.

First, how students' solution strategies on statistical graph tasks can be inferred from their gaze patterns was discussed. To this end, initially, it was needed to interview students on their strategies in a cued recall to connect specific gaze patterns to strategies. Once this connection is established for specific tasks and solution strategies, these can be used in further education and in intelligent tutoring systems.

Second, gaze data can be used in a feedback tool for students, such as an intelligent tutoring system, or for teachers in a teacher dashboard, was discussed. To this end, teachers need to become familiar with gaze patterns stemming from eye-tracking. This article aims to be a first step in that. For intelligent tutoring systems, cheaper methods of gaze data collection are needed. In gaming, webcams are already used for showing gaze patterns but further research is needed to determine if this can be applied in statistics education. In addition to technical limitations, this also raises an ethical discussion.

Third, this paper discussed that gaze data can be used as a new form of data in tertiary education. Courses on eye-tracking are mostly given in the neurosciences or psychology department. Such courses tend to focus on general strategies and self-regulated learning. For statistics education, a focus on task-specific strategies might be more relevant both for teachers and future researchers. For this aim, this paper calls on statistics education teachers to include eye-tracking data as a new form of data in their courses. Previously collected data from

neuroscience or psychology research could be repurposed and re-used for this aim.

Fourth, for teachers and developers of educational materials, this paper discussed that gaze data can be used for revising an instructional design.

Fifth, for teachers who want to use gaze data—either themselves or as data to be analyzed by their students—some possible tools for using and analyzing gaze data were shown: heatmap, raw data, videos, and static gazeplots.

In our own country, several teachers indicated that our eye-tracking research created an awareness of the difficulties students have with distinguishing histograms from other types of graphs with bars and interpreting histograms. Moreover, it showed details of students' strategies that neither teachers nor students were aware of. With this discussion paper, we hope to contribute to further usage and implementation of gaze data in statistics education by teachers, researchers, educational and textbook designers, and students.

## ACKNOWLEDGEMENTS

The author thanks Arthur Bakker, Paul Drijvers, and Wim Van Dooren for their contribution to discussing the design, the data analysis, and the articles of the previous research with secondary school students and teachers. The author thanks Rutmer Ebbes for his contribution to all phases of the pilot study with university students. The author thanks Alex Lyford and Enrique Garcia-Moreno Esteva for their analyses with MLAs and an interpretable model and one Figure. The author thanks Ciera Lamb for proofreading the article. The studies summarized in this discussion paper were funded with a Doctoral Grant for Teachers from the Dutch Research Council (NWO), number 023.007.023 awarded to Lonneke Boels. Any opinions, findings, or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the Dutch Research Council.

## CONFLICT OF INTEREST STATEMENT

None.

## ORCID

Lonneke Boels  <https://orcid.org/0000-0002-1763-2691>

## REFERENCES

1. E. Garcia Moreno-Esteva, S. L. J. White, J. M. Wood, and A. A. Black, *Application of mathematical and machine learning techniques to analyse eye tracking data enabling better understanding of children's visual cognitive behaviours*, *Frontl. Learn. Res.* **6** (2018), no. 3, 72–84. [rg/10.14786/flr.v6i3.365](https://doi.org/10.14786/flr.v6i3.365).
2. K. A. I. Khalil. Expert-novice differences: visual and verbal responses in a two-group comparison task. Master thesis. 2005.
3. A. R. Strohmaier, K. J. MacKay, A. Obersteiner, and K. M. Reiss, *Eye-tracking methodology in mathematics education research: A systematic literature review*, *Educ. Stud. Math.* **104** (2020), 147–200. <https://doi.org/10.1007/s10649-020-09948-1>.
4. L. Boels, A. Bakker, and P. Drijvers, *Eye tracking secondary school students' strategies when interpreting statistical graphs*, *Proceedings of the Forty-Third Conference of the International Group for the Psychology of Mathematics Education 2* (2019), 113–120.
5. L. Boels, R. Ebbes, A. Bakker, W. Van Dooren, and P. Drijvers, “Revealing conceptual difficulties when interpreting histograms: An eye-tracking study. Invited paper, refereed,” *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics*, M. A. Sorto, A. White, and L. Guyot (eds.), International Statistical Institute, Kyoto, Japan. Voorburg, The Netherlands, 2018, pp. 1–4. [https://iase-web.org/icots/10/proceedings/pdfs/ICOTS10\\_8E2.pdf](https://iase-web.org/icots/10/proceedings/pdfs/ICOTS10_8E2.pdf).
6. A. L. Cohen and A. Staub, *Within-subject consistency and between-subject variability in Bayesian reasoning strategies*, *Cogn. Psychol.* **81** (2015), 26–47. <https://doi.org/10.1016/j.cogpsych.2015.08.001>.
7. H. Fleig, T. Meiser, F. Ettl, and J. Rummel, *Statistical numeracy as a moderator of (pseudo)contingency effects on decision behavior*, *Acta Psychol.* **174** (2017), 68–79. <https://doi.org/10.1016/j.actpsy.2017.01.002>.
8. L. Boels, A. Bakker, W. Van Dooren, and P. Drijvers, *Secondary school students' strategies when interpreting histograms and case-value plots: An eye-tracking study*, Freudenthal Institute, Utrecht University, Utrecht, The Netherlands, 2022, [Manuscript submitted for publication].
9. I. Gal, *Statistical tools and statistical literacy: The case of the average*, *Teach. Stat.* **17** (1995), no. 3, 97–99.
10. J. Mokros and S. J. Russell, *Children's concepts of average and representativeness*, *J. Res. Math. Educ.* **26** (1995), no. 1, 20–39. <https://doi.org/10.2307/749226>.
11. S. A. Peters, V. M. Bennett, M. Young, and J. D. Watkins, *A fair and balanced approach to the mean*, *Math. Teach. Middle School* **21** (2016), no. 6, 364–375.
12. L. Boels, E. Garcia Moreno-Esteva, A. Bakker, and P. Drijvers, *Automatic gaze-based identification of students' strategies in histogram tasks: Machine learning algorithm and interpretable model*, Freudenthal Institute, Utrecht University, Utrecht, The Netherlands, 2022, [Manuscript submitted for publication].
13. C. A. N. Knoop-Van Campen, E. Kok, R. V. Doornik, P. D. Vries, M. Immink, H. Jarodzka, and T. Van Gog, *How teachers interpret displays of students' gaze in reading comprehension assignments*, *Frontl. Learn. Res.* **9** (2021), no. 4, 116–140. <https://doi.org/10.14786/flr.v9i4.881>.
14. A. J. Lyford. Investigating Undergraduate Student Understanding of Graphical Displays of Quantitative Data Through Machine Learning Algorithms, Doctoral thesis. 2017. <https://iase-web.org/documents/dissertations/17.AlexanderLyford.Dissertation.pdf>.
15. A. Lyford and L. Boels, “Using machine learning to understand students' gaze patterns on graphing tasks. Invited paper—refereed,” *Bridging the gap: Empowering & educating today's learners in statistics. Proceedings of the 11th International Conference on Teaching Statistics (ICOTS11)*, S. A.

- Peters, L. Zapata-Cardona, F. Bonafini, and A. Fan (eds.), ISI/IASE, Rosario, Argentina, 2022. <https://doi.org/10.52041/iase.icots11.T8D2>.
16. L. Boels and W. Van Dooren, *Secondary school students interpreting and comparing dotplots: An eye-tracking study*. University of Applied Science Utrecht, Utrecht, The Netherlands, 2023, [Manuscript accepted for publication].
  17. J. K. Kaakinen, *What can eye movements tell us about visual perception processes in classroom contexts? Commentary on a special issue*, *Educ. Psychol. Rev.* **33** (2021), 169–179. <https://doi.org/10.1007/s10648-020-09573-7>.
  18. L. Cooke and E. Cuddihy, *Using eye tracking to address limitations in think-aloud protocol, IPCC 2005*. Proceedings International Professional Communication Conference, 2005, pp. 653–658. <https://doi.org/10.1109/ipcc.2005.1494236>
  19. E. M. Kok and H. Jarodzka, *Before your very eyes: The value and limitations of eye tracking in medical education*, *Med. Educ.* **51** (2017), no. 1, 114–122. <https://doi.org/10.1111/medu.13066>.
  20. M. Schindler and A. J. Lilienthal, *Domain-specific interpretation of eye-tracking data: Towards a refined use of the eye-mind hypothesis for the field of geometry*, *Educ. Stud. Math.* **101** (2019), 123–139. <https://doi.org/10.1007/s10649-019-9878-z>.
  21. M. Schindler, E. Schaffernicht, and A. J. Lilienthal, “Identifying student strategies through eye tracking and unsupervised learning: The case of quantity recognition,” *Proceedings of the forty-fourth conference of the International Group for the Psychology of Mathematics Education*, M. Inprasitha, N. Changsri, and N. Boonsena (eds.), PME, Khon Kaen, Thailand, 2021, pp. 4–10–4–19.
  22. L. Boels, A. Bakker, and P. Drijvers, “Unravelling teachers’ strategies when interpreting histograms: An eye-tracking study,” *Proceedings of the Eleventh Congress of the European Society for Research in Mathematics Education*, U. T. Jankvist, M. van den Heuvel-Panhuizen, and M. Veldhuis (eds.), Freudenthal Group and Freudenthal Institute, Utrecht University and ERME, Utrecht, The Netherlands, 2019, pp. 888–895. <https://hal.archives-ouvertes.fr/hal-02411575/document>.
  23. A. J. King, N. Bol, R. Glenn Cummins, and K. K. John, *Improving visual behavior research in communication science: An overview, review, and reporting recommendations for using eye-tracking methods*, *Commun. Methods Meas.* **13** (2019), no. 3, 149–177. <https://doi.org/10.1080/19312458.2018.1558194>.
  24. M. Lai, M. Tsai, F. Yang, C. Hsu, T. Liu, S. W. Lee, M.-H. Lee, G.-L. Chiou, J.-C. Liang, and C. A. Tsai, *Review of using eye-tracking technology in exploring learning from 2000 to 2012*, *Educ. Res. Rev.* **10** (2013), 90–115. <https://doi.org/10.1016/j.edurev.2013.10.001>.
  25. P. Beach and J. McConnel, *Eye tracking methodology for studying teacher learning: A review of the research*, *Int. J. Res. Meth. Educ.* **42** (2019), no. 5, 485–501. <https://doi.org/10.1080/1743727X.2018.1496415>.
  26. H. Ashraf, M. H. Sodergren, N. Merali, G. Mylonas, H. Singh, and A. Darzi, *Eye-tracking technology in medical education: A systematic review*, *Med. Teach.* **40** (2018), no. 1, 62–69. <https://doi.org/10.1080/0142159X.2017.1391373>.
  27. J. H. Goldberg and J. I. Helfman, “Comparing information graphics: A critical look at eye tracking,” *Proceedings of the 3rd BELIV’10 Workshop: Beyond time and errors: novel Evaluation methods for Information Visualization*, Atlanta, Georgia, Association for Computing Machinery, New York, NY, 2010, pp. 71–78. <https://doi.org/10.1145/2110192.2110203>.
  28. C. Godau, H. Haider, S. Hansen, T. Schubert, P. A. Frensch, and R. Gaschler, *Spontaneously spotting and applying shortcuts in arithmetic—A primary school perspective on expertise*, *Front. Psychol. Cogn.* **5** (2014), (art. e556). <https://doi.org/10.3389/fpsyg.2014.00556>.
  29. J. Hyönä, *The use of eye movements in the study of multimedia learning*, *Learn. Instr.* **20** (2010), no. 2, 172–176. <https://doi.org/10.1016/j.learninstruc.2009.02.013>.
  30. E. Alemdag and K. Cagiltay, *A systematic review of eye tracking research on multimedia learning*, *Comput. Educ.* **125** (2018), 413–428. <https://doi.org/10.1016/j.compedu.2018.06.023>.
  31. E. Garcia Moreno-Esteva, A. Kervinen, M. S. Hannula, and A. Uitto, *Scanning signatures: A graph theoretical model to represent visual scanning processes and a proof of concept study in biology education*, *Educ. Sci.* **10** (2020), (Article 141), no. 5, 1–14. <https://doi.org/10.3390/educsci10050141>.
  32. S. Schreiter and M. Vogel, *Eye-tracking measures as indicators for a local vs. global view of data*, *Front. Educ.* **7** (2023), 1–11. <https://doi.org/10.3389/educ.2022.1058150>.
  33. M. Abt, K. Loibl, T. Leuders, and F. Reinhold, *Considering conceptual change as a source for errors when comparing data sets with boxplots*, *Poster presented at EARLI SIG3*, Zwolle, The Netherlands, 2022.
  34. H. Fry, *Hello world, How to be human in the age of the machine*, Black Swan, 2019.
  35. M. Król and M. Król, *Learning from peers’ eye movements in the absence of expert guidance: A proof of concept using laboratory stock trading, eye tracking, and machine learning*, *Cogn. Sci.* **43** (2019), no. 2, e12716. <https://doi.org/10.1111/cogs.12716>.
  36. E. M. Kok and C. A. N. Knoop-van Campen, “Using webcam-based eye-tracking to uncover reading strategies,” *Presentation at the EARLI SIG 27 conference*, EARLI, Southampton, UK, 2022. <https://www.thomas-martens.de/sig-27-conference-2022/>.
  37. J. Kaplan, J. Gabrosek, P. Curtiss, and C. Malone, *Investigating student understanding of histograms*, *J. Stat. Educ.* **22** (2014), no. 2, 1–29.
  38. D. Whitaker and T. Jacobbe, *Students’ understanding of bar graphs and histograms: Results from the LOCUS assessments*, *J. Stat. Educ.* **25** (2017), no. 2, 90–102. <https://doi.org/10.1080/10691898.2017.1321974>.
  39. D. Abrahamson, *Grasp actually: An evolutionist argument for enactivist mathematics education*, *Hum. Dev.* **65** (2021), no. 2, 77–93.
  40. R. A. Alberto, A. Shvarts, P. Drijvers, and A. Bakker, *Action-based embodied design for mathematics learning: A decade of variations on a theme*, *Int. J. Child Comput. Interact* **32** (2022), (Article 100419), 1–23. <https://doi.org/10.1016/j.ijcci.2021.100419>.
  41. A. Shvarts, R. A. Alberto, A. Bakker, M. Doorman, and P. Drijvers, *Embodied instrumentation in learning mathematics as the genesis of a body-artifact functional system*, *Educ. Stud. Math.* **107** (2021), 447–469. <https://doi.org/10.1007/s10649-021-10053-0>.
  42. S. Shayan, D. Abrahamson, A. Bakker, C. A. Duijzer, and M. van der Schaaf, “Eye-tracking the emergence of attentional anchors in a mathematics learning tablet activity,” *Eye-tracking*

- technology applications in educational research, IGI Global, Hershey, Pennsylvania, 2017, pp. 166–194.
43. R. B. Church and S. Goldin-Meadow, *The mismatch between gesture and speech as an index of transitional knowledge*, *Cognition* **23** (1986), no. 1, 43–71. [https://doi.org/10.1016/0010-0277\(86\)90053-3](https://doi.org/10.1016/0010-0277(86)90053-3).
44. L. S. Vygotsky, *Educational psychology*, CRC Press, London, UK, 1997. <https://doi.org/10.4324/9780429273070-9/>.
45. A. Shvarts, M. Doorman, and R. Alberto, “Concrete-abstract-new-concrete: Freudenthal and Davydov in advancing embodied design framework,” *Paper presented at the 12th Congress of the European Society for Research in Mathematics Education*, Free University of Bozen-Bolzano and ERME, Bozen-Bolzano, Italy, 2022.
46. A. Shvarts, “Eye movements in emerging conceptual understanding of rectangle area,” *Proceedings of the 41st Conference of the International Group for the Psychology of Mathematics Education, Vol. 1*, B. Kaur, W. K. Ho, T. L. Toh, and B. H. Choy (eds.), PME, Singapore, 2017, pp. 268.
47. R. A. Alberto, A. Bakker, O. W. van Aalst, P. B. J. Boon, and P. H. M. Drijvers, “Networking theories in design research—An embodied instrumentation case study in trigonometry,” *Proceedings of the Eleventh Congress of the European Society for Research in Mathematics Education*, U. T. Jankvist, M. van den Heuvel-Panhuizen, and M. Veldhuis (eds.), Freudenthal Group and Freudenthal Institute, Utrecht University and ERME, Utrecht, the Netherlands, 2019, pp. 3088–3095 <https://hal.archives-ouvertes.fr/hal-02418076>.
48. T. Erickson, M. Wilkerson, W. Finzer, and F. Reichsman, *Data moves*, *Techn. Innov. Stat. Educ.* **12** (2019), (Article 1), no. 1, 25 p. <https://escholarship.org/uc/item/0mg8m7g6>.
49. Tobii. Tobii Studio. Users' Manual. Version 3.4.5. n.d. <https://www.tobii.com/siteassets/tobii-pro/user-manuals/tobii-pro-studio-user-manual.pdf?v=3.4.5>.
50. C. J. Wild and M. Pfannkuch, *Statistical thinking in empirical enquiry*, *Int. Stat. Rev.* **67** (1999), no. 3, 223–248. <https://doi.org/10.1111/j.1751-5823.1999.tb00442.x>.

**How to cite this article:** L. Boels, *Reflections on gaze data in statistics education*, *Teach. Stat.* **45** (2023), S40–S51. <https://doi.org/10.1111/test.12340>