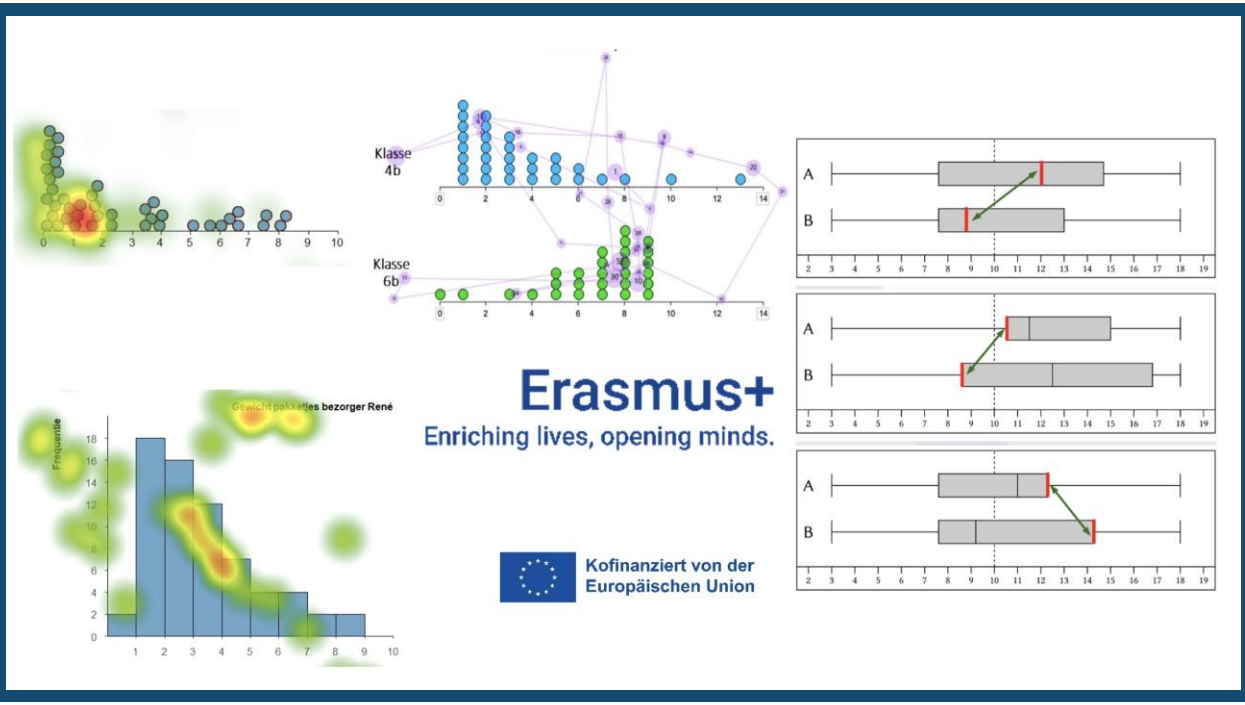


Eye-Teach-Stats

ProDaBi Colloquium

14.01.2026

Ayline Heursen[1] , Vuslat Seker[2] & Lonneke Boels[2]

Bridging research and practice: Integrating eye-tracking insights into teacher education on students' misinterpretation

[1] University of Education, Heidelberg; [2]HU University of Applied Sciences Utrecht

The eye-teach stats project

Markus Vogel, Ayline Heursen

Saskia Schreiter

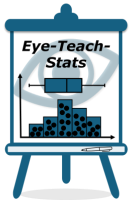
Wim Van Dooren

Frank Reinhold, Martin Abt

Lonneke Boels, Vuslat Seker, Anita Alexander

Gefördert aus Mitteln der Europäischen Union

The eye-teach stats project



Eye-teach-stats: Supporting teachers to diagnose and deal with students' statistical misinterpretations – innovative teacher training with eye-tracking vignettes"

- Support pre- and in-service mathematics teachers in (further) developing their statistical PCK and CK
- Enable teachers to prepare students to become data-literate citizens
- Offer innovative and effective learning opportunities to:
  - Diagnose students' statistical (mis)interpretations
  - Deal with these (mis)interpretations
  - Overcome their own statistical difficulties

- The development of a teacher training course using eye-tracking vignettes
  Designed based on eye-tracking research findings
  Focus: student (mis)interpretations of statistical graphs; dotplots, boxplots, case-value plots, histograms.

The eye-teach stats project

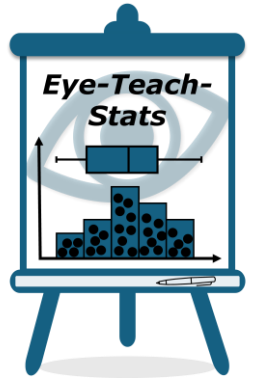
Eye-Tracking Studies on Learners' Strategies and Difficulties in Statistical Graphs

Systematic Literature Review

On systematic errors when interpreting statistical graphs

Development of eye-tracking vignettes

Evaluation of eye-tracking vignettes



Systematic errors

- Systematic errors are recurring, explainable reasoning—not mere random lapses

  Cox, 1975; Padberg, 1996; Shimizu & Kang, 2025)

- Based on strategies and conceptual difficulties
- Often traceable to specific heuristics or overgeneralizations, such as transferring bar graph reasoning to histograms or boxplots

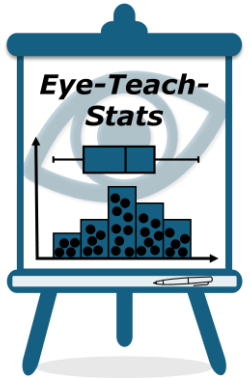
  (Boels et al., 2019b; Lem et al., 2013; Padberg, 1996)

- Persist without targeted instruction—practice alone is insufficient

  (Padberg, 1996; Radatz, 1979)

- Distinction between:
  - Misinterpretation: a systematic error evident in an incorrect answer
  - Conceptual difficulty: incomplete or flawed understanding of an underlying concept

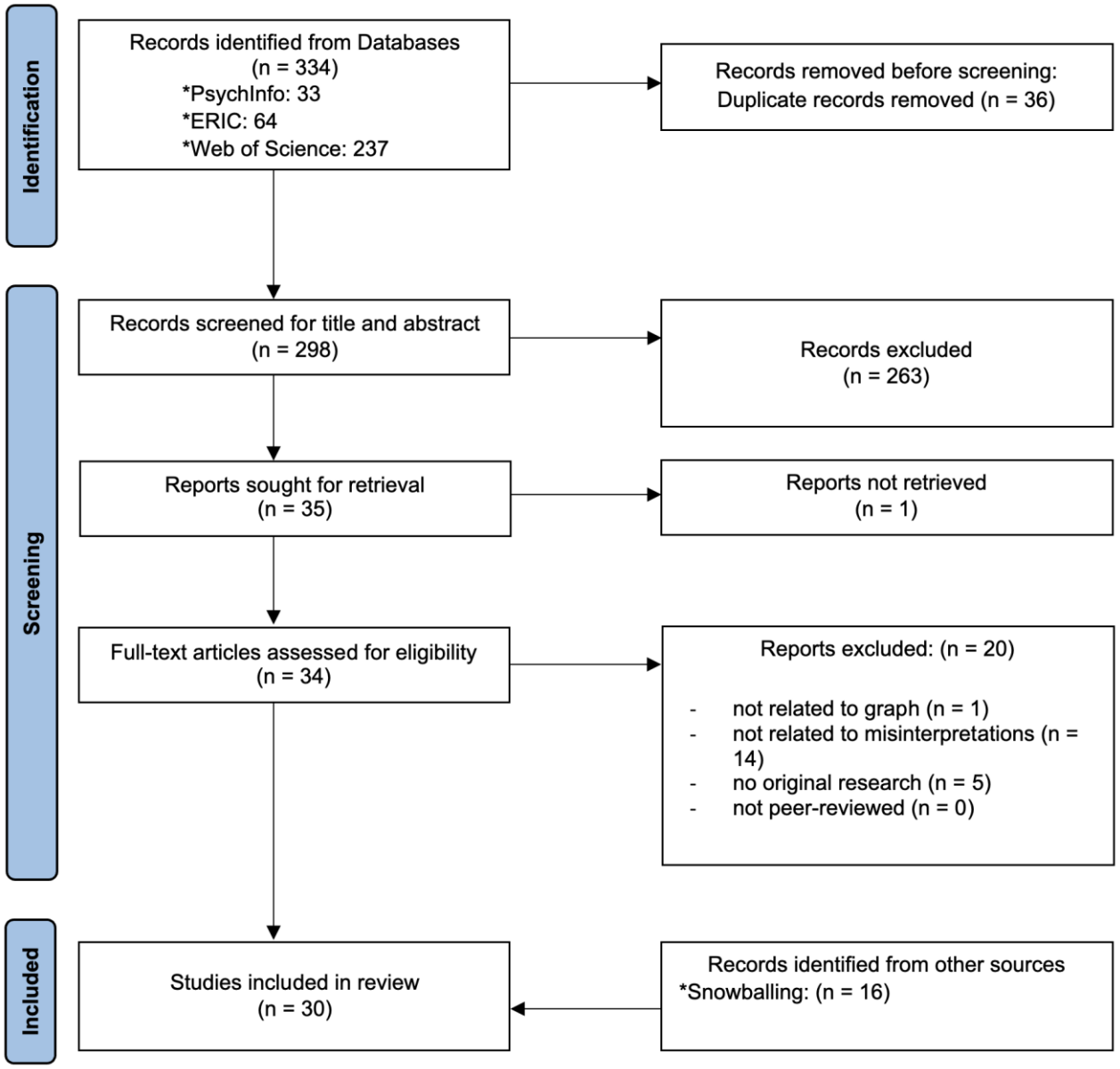
  (Boels et al., 2019b)



Systematic literature review

1. What **systematic errors** are reported in the interpretation of statistical graphs depicting univariate distributions in boxplots, case-value plots, dotplots and histograms?
2. What **meaningful categories** can be distinguished in systematic errors in the interpretation of statistical graphs depicting univariate distributions?

- Graph types: boxplots, case-value plots, dotplots, histogram
- Sample: school students, university students, teachers, expert



Systematic errors when interpreting statistical graphs

Variability shape confusion

Bar height differences
Flat shape – less variability
Flat shape – more variability
Symmetric shape less variability

Graph mechanics representation confusion

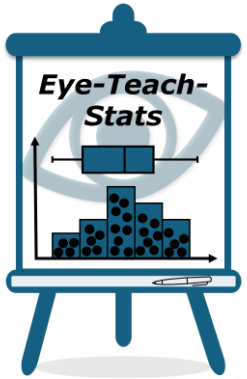
Frequency – value confusion
More area (length) – higher frequency
More area (height) – higher frequency
Whiskers irrelevant

Overemphasizing local feature: - Number of bars
- Personal experience
- Outlier
- Number of data points
- Range

Not reasoning about data as set

Sum of frequencies
Sum of measured values
Overemphasizing measures of central tendency
Height for mean
Variability for mean

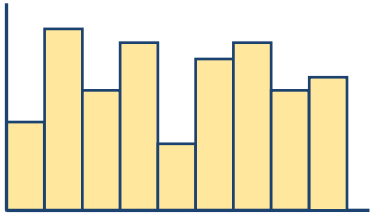
Looking at a wrong parameter for measures of central tendancy



Systematic errors when interpreting statistical graphs

Variability shape confusion

Bar height differences
Flat shape – less variability
Flat shape – more variability
Symmetric shape less variability

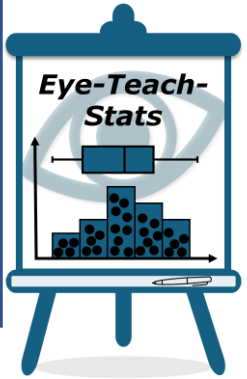


- Reasoning about variability is based on shape of the graph (e.g. flatness, symmetry)
- especially observed in the interpretation of histograms,
- but also for dotplots, boxplots and case-value plots

Graph mechanics representation confusion

Frequency – value confusion
More area (length) – higher frequency
More area (height) – higher frequency
Whiskers irrelevant

- Reasoning caused by misunderstandings of how statistical concepts are represented within specific graphical representations
- especially observed in the interpretation of boxplots,
- but also for dotplots and histograms

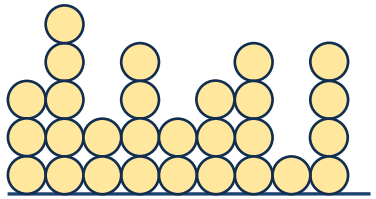


Systematic errors when interpreting statistical graphs

- Distributions are not considered as a whole, instead individual data points or specific features are considered

- especially observed in the interpretation of dotplots,
- but also for boxplots and case-value plots and histograms

Overemphasizing local feature:
- Number of bars
- Personal experience
- Outlier
- Number of data points
- Range



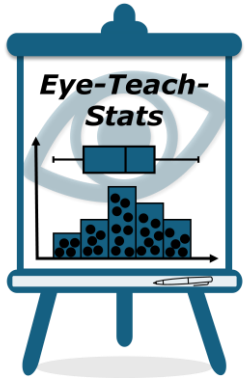
Not reasoning about data as set

- Reasoning caused by mixing up measures of central tendency or inferring from other features of data distribution

- especially observed in the interpretation of dotplots, histograms and boxplots

Sum of frequencies
Sum of measured values
Overemphasizing measures of central tendency
Height for mean
Variability for mean

Looking at a wrong parameter for measures of central tendancy



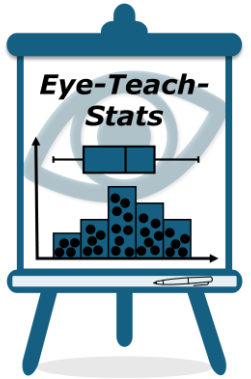
Eye-Tracking Studies on Learners' Strategies and Difficulties in Statistical Graphs

Systematic Literature Review

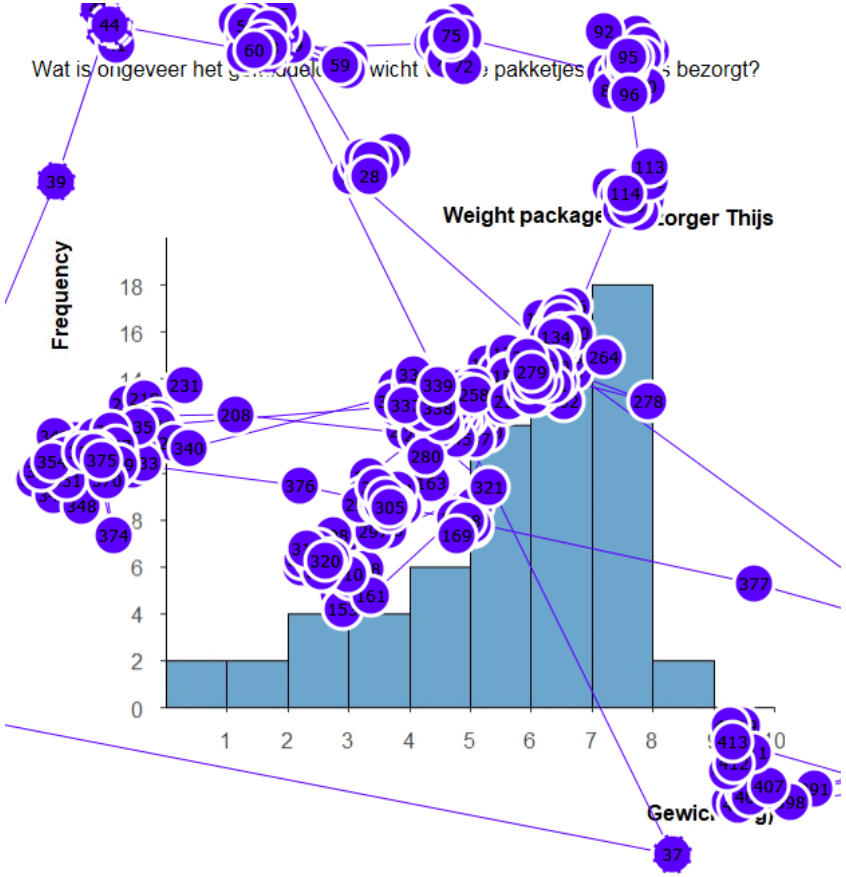
On systematic errors when interpreting statistical graphs

Development of eye-tracking vignettes

Evaluation of eye-tracking vignettes



Eye-tracking data



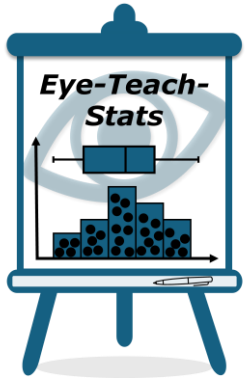
- Eye movements are frequent motor actions that support cognitive processes

  (Spivey & Dale, 2011).

- Eye-tracking data can reveal students' visual search patterns, which are normally hidden and challenging for students to explain.

- Gaze data offer insights into students' thinking, linking actions to specific concepts and revealing strategies used for answering.

  (e.g., Chumachemko et al., 2014)



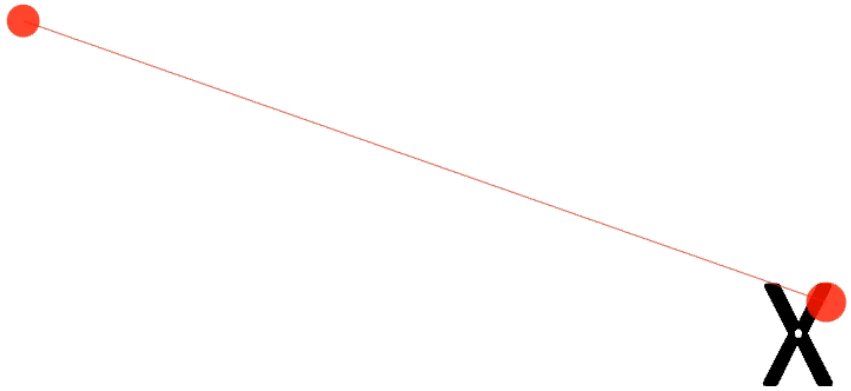
Eye-tracking data

Video of the eye movements of student L18 (approx. 19 seconds)

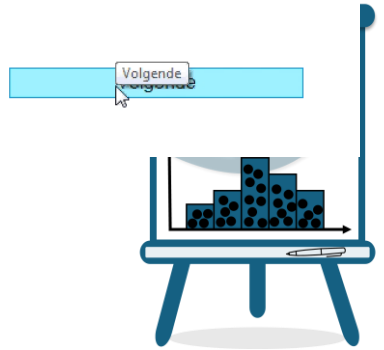
Interview: L18: I was mostly just kind of looking at the small numbers. […] Because the first one was quite long anyway[…] Not very far from the middle. […] The middle is at four-and-a-half.
Translated with deepl.com

Correct strategy - Item 02

Boels (2023)



Fixation cross between every task



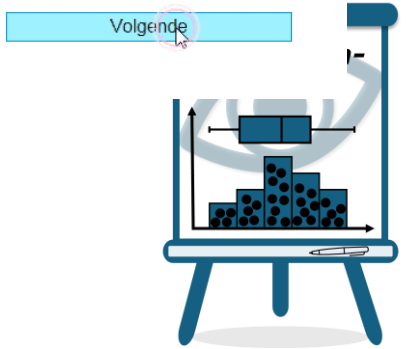
Eye-tracking data

Video of the eye movements of student L18 (approx. 19 seconds)

Interview: no interview data (time was up)
But for item19 (similar to item20) student noticed that they made a mistake

L18: Incorrect strategy - Item 20

Boels (2023)

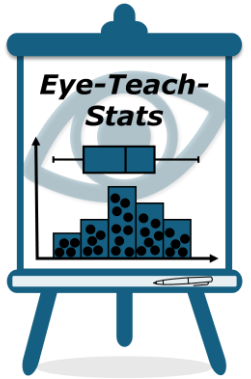


Eye-tracking vignettes

Vignettes present

- authentic scenes from teaching practice and
- provide rich opportunities for analysis and reflection without the pressure to act
- Their use has proven highly beneficial in mathematics teacher education

(Friesen et al., 2018; Rutsch et al., 2018)



Eye-tracking vignettes

Eye-tracking vignettes:

- innovative vignette format, extending traditional vignettes with students' visual expressions (e.g., gaze plots) alongside verbal expressions (e.g., utterances).
- Are accompanied by questions that prompt teachers to reflect on the student's solution process.
- Developed from recent research showing that students' graph-interpretation difficulties are reflected in their gaze behavior.

(Abt et al., 2023; Boels et al., 2019a,c, 2022; Schreiter & Vogel, 2023)

- The combination of visual and verbal expressions helps uncover students' underlying mental models when interpreting distributions

(Dvir & Ben-Zvi, 2021)

- Research suggests that teachers can interpret gaze displays and incorporate them into their teaching

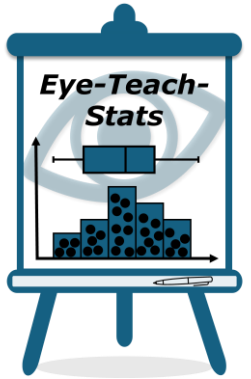
(Knoop-Van Campen et al., 2021)

Provide powerful stimuli for teachers to:

- Diagnose students' statistical (mis)interpretations.
- Address these (mis)interpretations effectively.

Contribute to learner-centred teaching, which relies on accurate diagnosis

(Loibl et al., 2020)



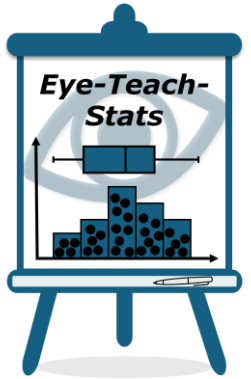
Development of eye-tracking vignettes

The Four-Component Instructional Design model (**the 4C/ID model)** which provides a framework for designing complex learning tasks like eye-tracking vignettes

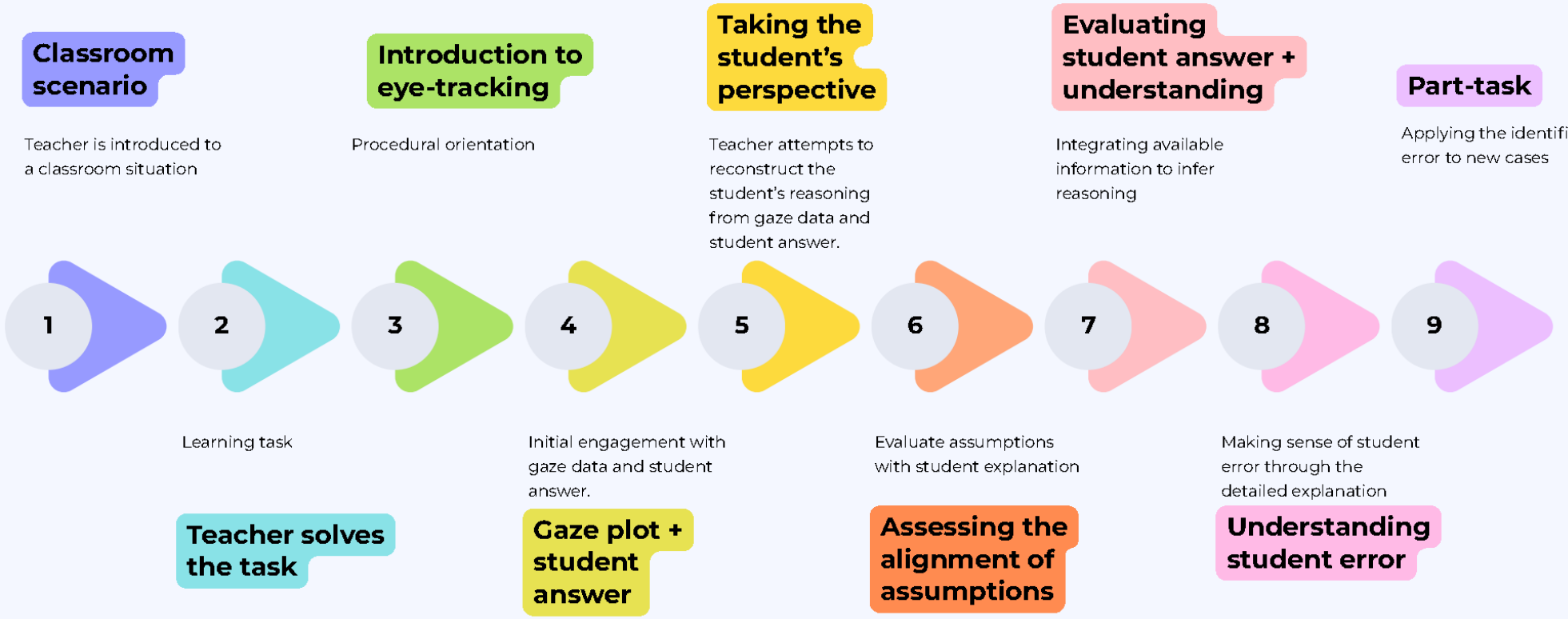
(van Merriënboer et al., 2024; see also Kreutz et al., 2020)

Each vignette includes:

- A learning task that focuses on:
  - A specific systematic error and the associated conceptual difficulty,
  - The graph type in which it occurs — for example, histograms, case-value plots boxplots, or dotplots.
- Supportive information on graph interpretation
- Procedural information on interpreting gaze data
  - Different ways for presenting gaze data
- Practice tasks for reinforcing diagnostic skills

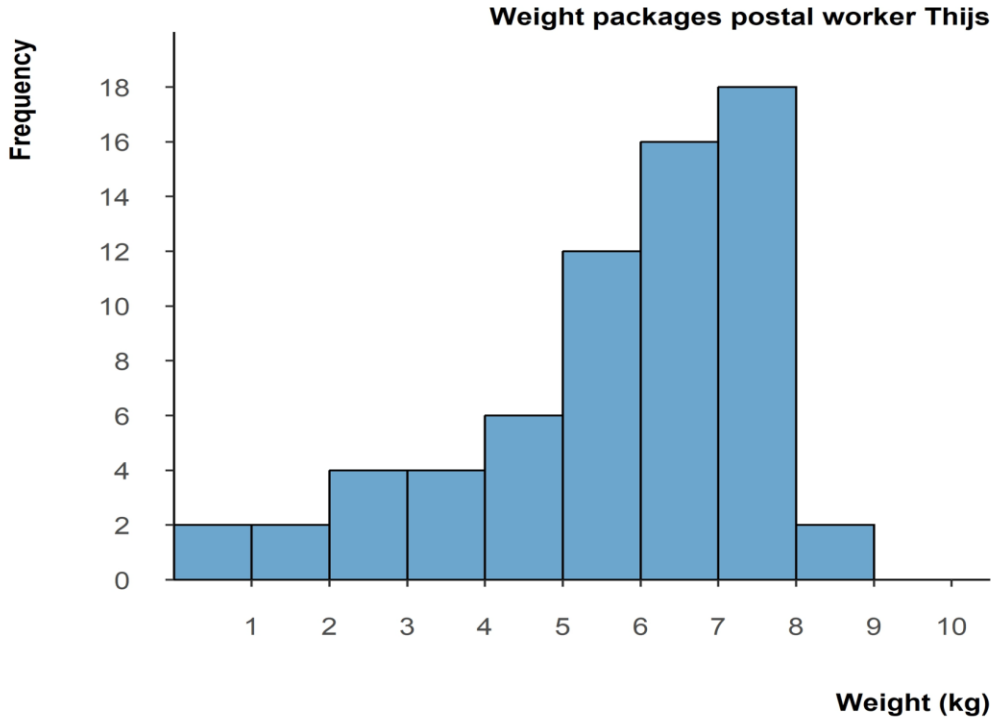


Vignette Structure



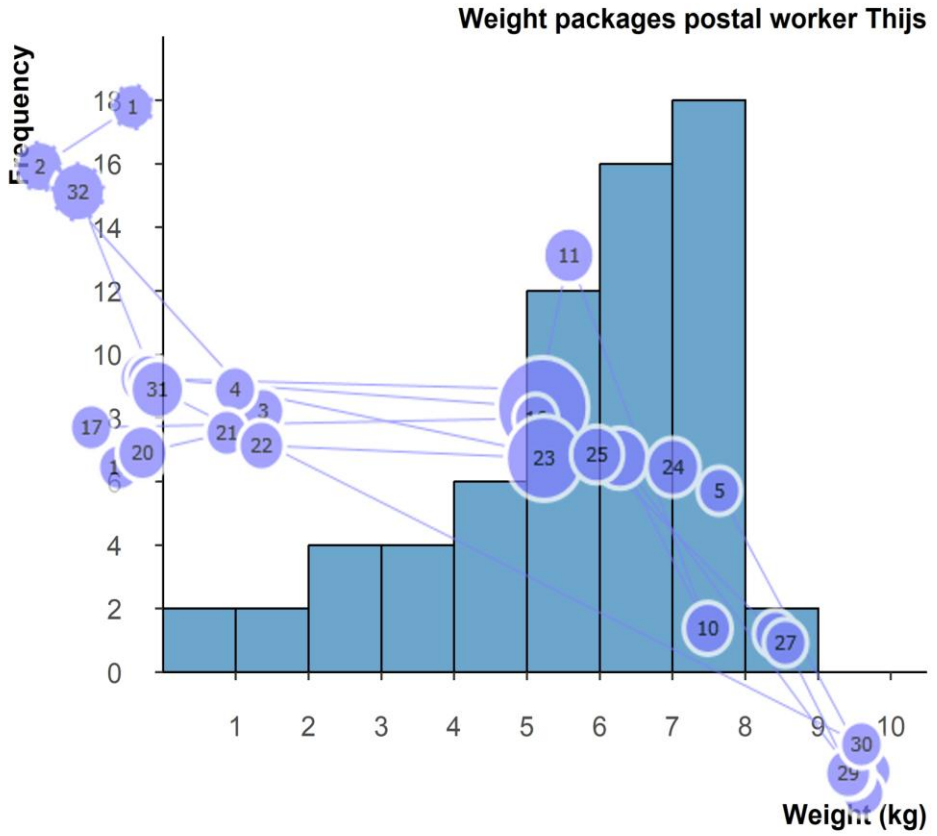
Example- histogram

What is approximately the mean weight of the packages that Thijs delivers?



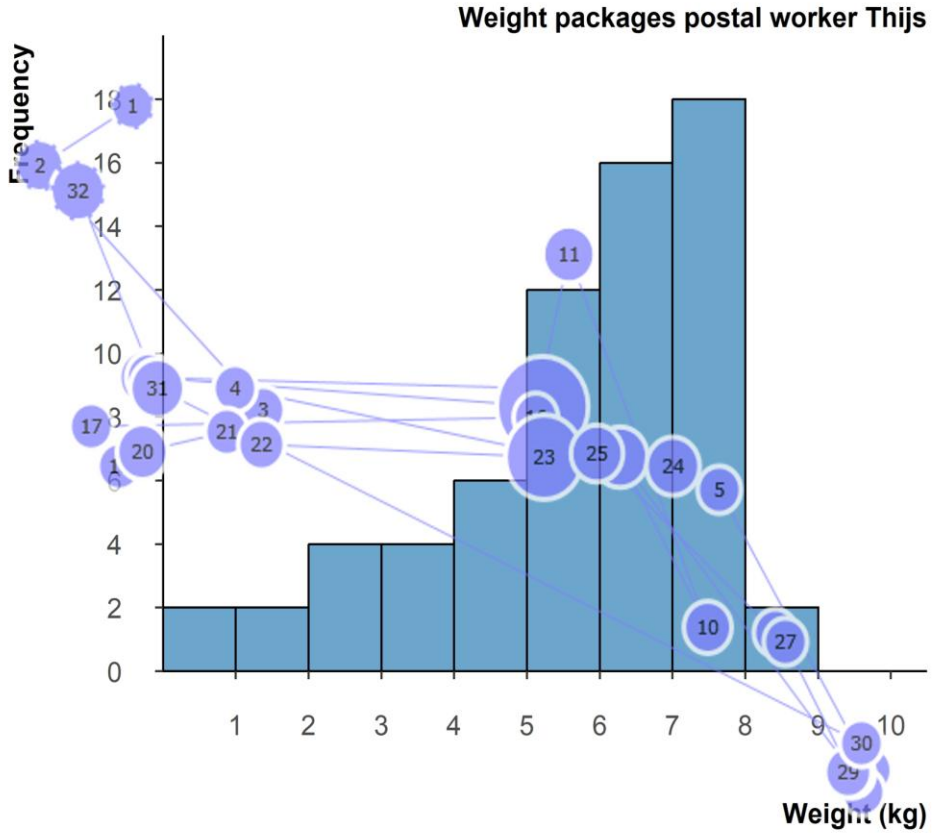
Example- histogram

Student's Answer: "the arithmetic mean is about eight."



Example- histogram

Student's Comment: "I looked at how tall the bars were on the graph to see how often things showed up. Then, I tried to guess where the mean would be based on that. I also tried to make the bars all the same height to make it easier to figure out."



Example-histogram

Understanding student errors:

Sophie has an underlying conceptual difficulty regarding the interpretation of histograms. This leads to an observable systematic error, called "**Frequency Value Confusion**."

Observable behavior: Sophie said she looked at how tall the bars were and that she tried to understand how often values appeared. She tried to make all bars of the same height to estimate the mean. The gaze plot substantiates her description of making all bars equally high as her gaze moved horizontally—more or less along a horizontal line—as she searched for the point along the vertical axis that in her opinion represents the mean value. This indicates an observable systematic error, which we term as "Frequency Value Confusion" in the interpretation of histograms.

Interpretation by the teacher/researcher: It appears that Sophie applied a strategy to the histogram that is typically used for case-value plots. From the explanation and gazes of Sophie it appears that she searched for a horizontal line that equalizes all bars. However, in a histogram, the heights of bars indicate frequencies, not measured values. Sophie's approach reflects a confusion about what the data are and what the axes represent. In a histogram, the horizontal axis shows the measured values (in this case, weight in kg) while the vertical axis represents the frequency of measurements within a specific interval (the bin width of the bar). The error occurs when students interpret the height of the bars (frequency) as the measured value rather than as a representation of how many measured values there are within a particular range. Although incorrect for histograms, this strategy is correct for finding the arithmetic mean in a case-value plot.

Underlying conceptual difficulty: Sophie's systematic error is likely due to an underlying conceptual difficulty with the concept of data (where, how and what data are presented) by not distinguishing between frequencies and measured values in a histogram. Such conceptual difficulties are common when students fail to grasp that the vertical axis in a histogram represents how many cases fall within each bin, not the measured value itself.

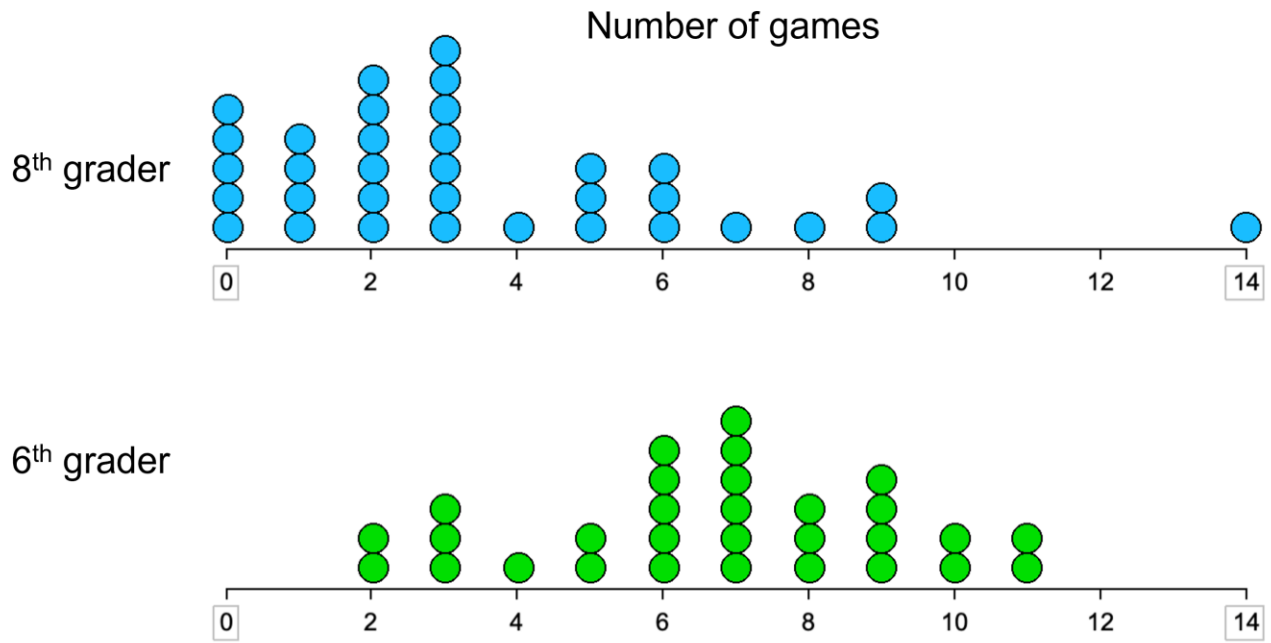
Correct strategies for histograms: In a histogram, the mean can be estimated by finding the point along the horizontal axis where the histogram (the distribution) is in balance. In such case, their gaze would move up and down between a point on the horizontal axis and the tops of the bars, creating a scanpath along an imagined vertical line. The mean can also be estimated through a calculation: by using the midpoint of each bin (representing the measured values within that bin) multiplied by its corresponding frequency (the bar height), summing these products, and then dividing by the sum of the frequencies (the total number of data points).

Sophie's solution

Zurück zur Bearbeitung | Nächste Folie

Example– dotplot

Do 8th graders and 6th graders differ in the number of games they have on their cell phones? A survey on this question was done at the Park School with 34 8th grader and 30 6th grader. Compare the two samples and then evaluate:

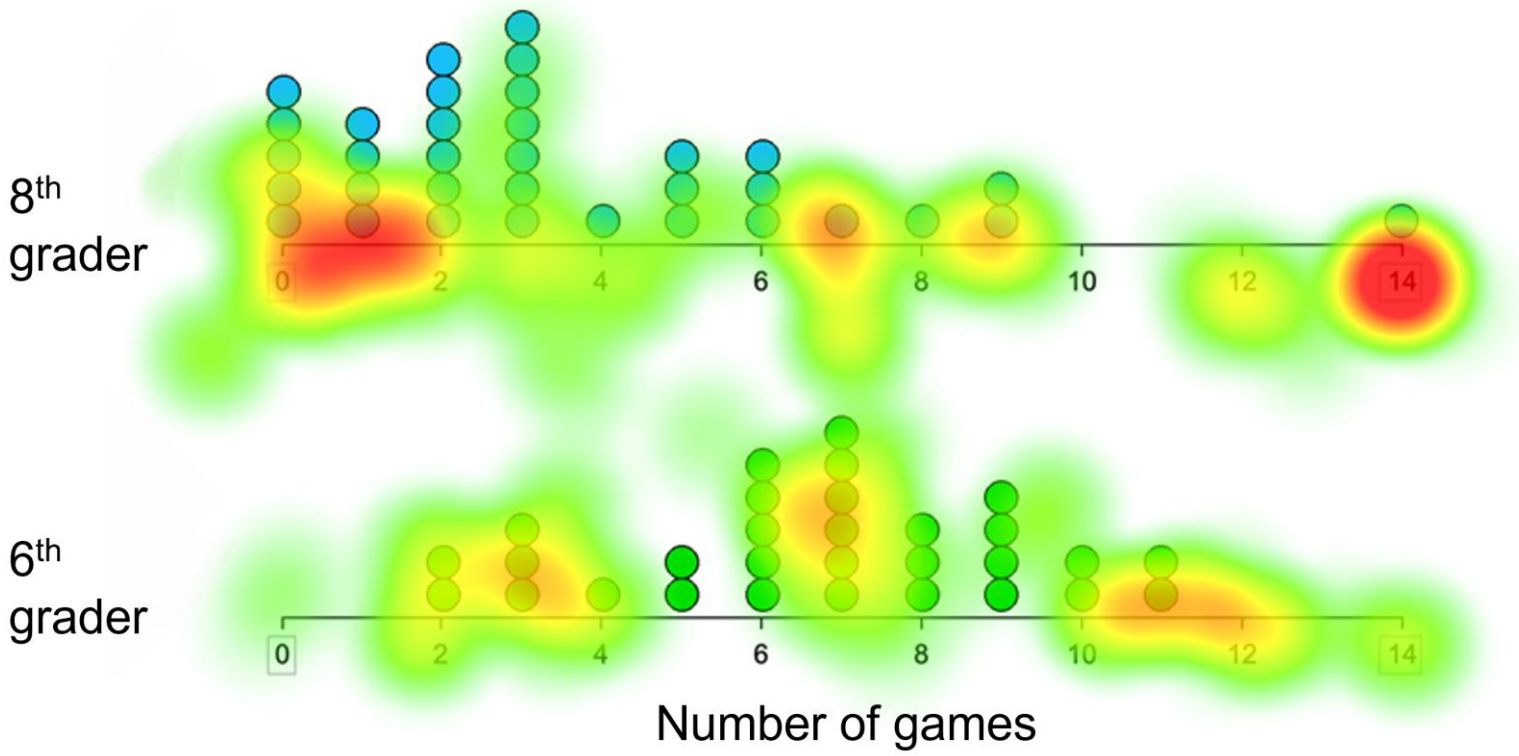


The data from the samples fits the assumption that

a) 8th grader have a lot more games on their phones.
b) 6th grader have a lot more games on their phones.
c) 8th and 6th grader have about the same number of games on their phone.

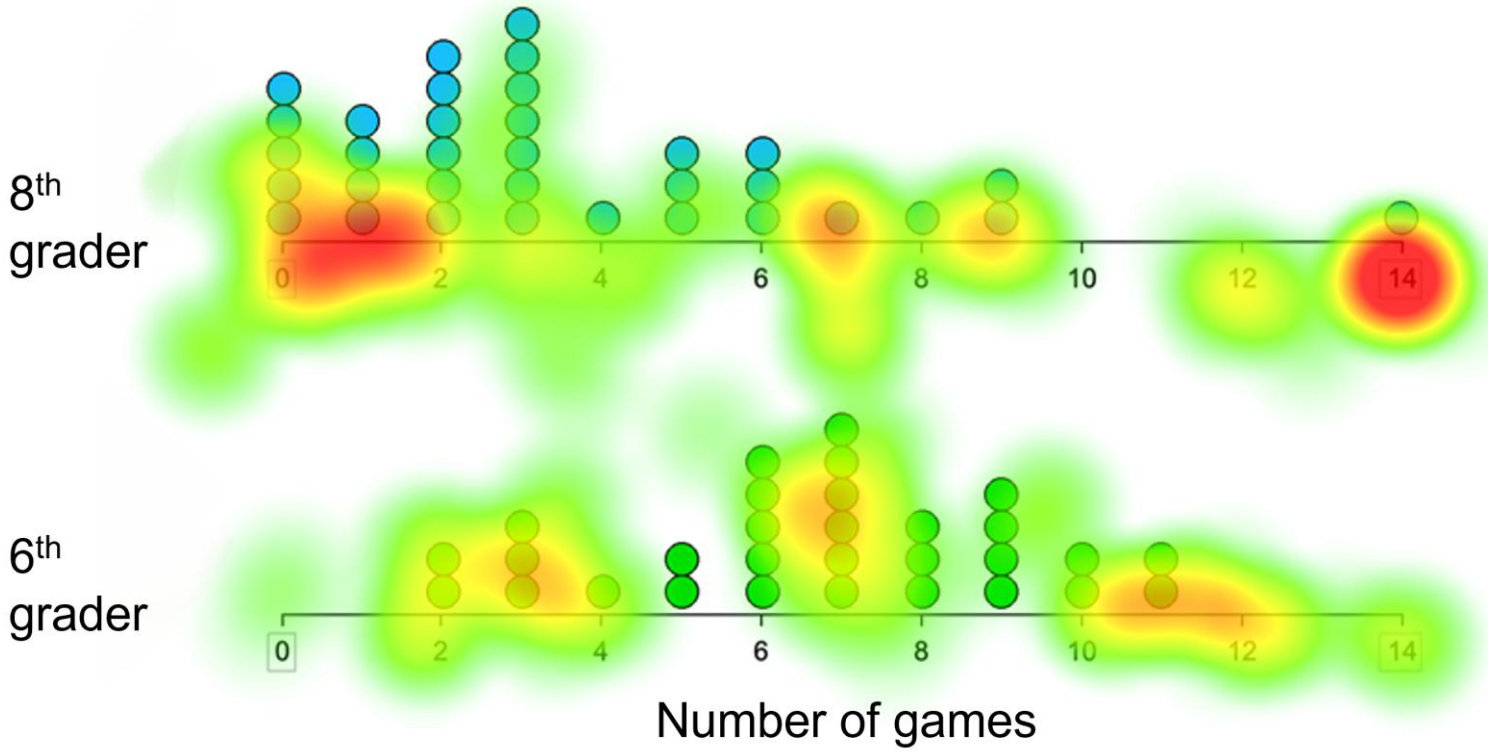
Example- dotplot

Student's answer: „I think a) is correct: the 8th graders have more games".



Example- dotplot

Student's comment: „Here I saw that one 8th grader has 14 games. This is really a lot, that is a very big number and that matters a lot. So I think that the 8th graders have more games."



Example-dotplot

Understanding student errors:

Amaya has an underlying conceptual difficulty regarding the interpretation of dotplots. This leads to an observable systematic error, called "**Overemphasizing of outliers**".

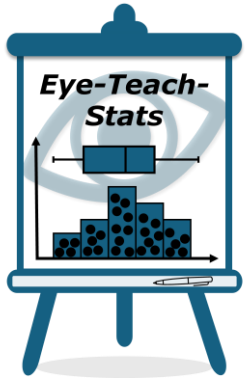
Interpreting behavior: Amaya's heatmap and her comment suggest that she focuses a lot on the outlier as an individual data point.

Understanding error and underlying conceptual difficulty: In a dotplot, each dot stands for one occurrence of a value (in this case the number of games) on the horizontal axis. The height of the stacks (how many dots are piled up) shows how many times each value occurs.

A central aim in teaching statistics is to help students develop a "global view" of distributions. In this perspective, students perceive distributions as conceptual entities with overall features such as center, spread, and shape. However, many students have a "local view" and perceive distributions as mere collections of individual data points. This leads them to focus too much on local details of the distribution. The systematic error occurs in this inappropriate use of a local view.

On the next slide, you will see an overview of local and global features that can be taken into account when comparing distributions.

Zurück zur Bearbeitung | Nächste Folie



Systematic errors when interpreting statistical graphs

Variability shape confusion

Bar height differences
Flat shape – less variability
Flat shape – more variability
Symmetric shape less variability

Graph mechanics representation confusion

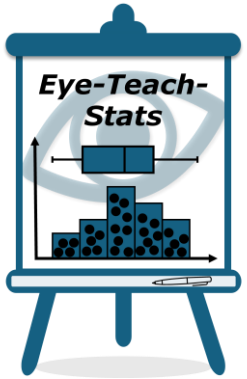
Frequency – value confusion
More area (length) – higher frequency
More area (height) – higher frequency
Whiskers irrelevant

Overemphasizing local feature: - Number of bars
- Personal experience
- Outlier
- Number of data points
- Range

Not reasoning about data as set

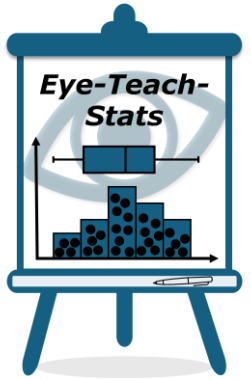
Sum of frequencies
Sum of measured values
Overemphasizing measures of central tendency
Height for mean
Variability for mean

Looking at a wrong parameter for measures of central tendancy



Piloting Eye-Tracking Vignettes -Case Studies-

- The research questions:
  Study 1: *How do pre-service mathematics teachers engage with eye-tracking vignettes when diagnosing students' systematic errors in histogram interpretation? (Seker et al., 2025 - ECER 2025)*
- *Study 2: How do pre-service teachers engage with students' eye-tracking data when diagnosing students' systematic errors in histogram interpretation*? (Seker *et al.*, 2025- IASE Satellite 2025)



Common Method for STUDY 1 & STUDY 2

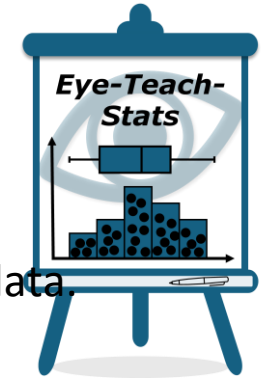
Research Design
Qualitative case study

Participant
a female pre-service mathematics teacher from Türkiye
qualifying to teach mathematics at the middle school level (grades 5–8, ages 10–14).
completed two courses related to statistics: one on inferential statistics and another on teaching statistics and probability
no prior experience working with or interpreting eye-tracking data

Data collection tools
An histogram eye tracking vignette

Data collection procedure
a think-aloud task and a follow-up semi-structured interview

Data analysis
The data were analyzed qualitatively with a focus on how the participant engaged with the eye-tracking data.



Study 1: Using Eye-Tracking Vignettes to Improve Pre-Service Teachers' Diagnostic Skills in Histogram Interpretation

Interpreting the student's answer

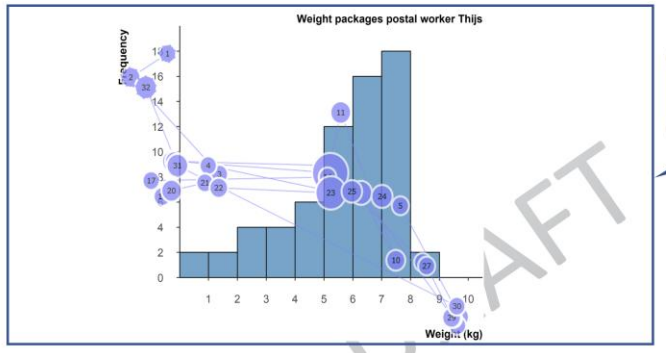
- Student's answer ("≈ 8") was unexpected
- Linked to attention on the tallest bar
- Assumptions vs. histogram with eye-tracking data

Exploring supportive & procedural information

- Participant consulted supportive (histogram info) and procedural (gaze info) texts
- These helped clarify what gaze plots represent

Making sense of the explanation

- Explanation ("making bars the same height") → confusing, insufficient
- Explanation alone did **not** allow diagnosis



Sophie's Answer:

"the arithmetic mean is about eight".

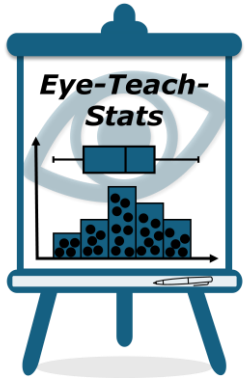
Have a look at the gaze plot that shows your student Sophie's gaze and her answer.

Information on gaze plot (Procedural Information):

In a gaze plot, gaze data collected from eye-tracking devices is superimposed onto the original stimulus, creating a combined view of both the stimulus and the participants' eye movements.

Eye Tracking Gaze plot with three participants

The visualization includes dots and lines representing fixations and saccades, respectively. Fixations are depicted as dots, with each dot indicating a phase when the participant's gaze remains relatively stable on a specific area of the stimulus. The size of the dot here corresponds to the duration of the fixation, with larger dots representing longer durations of fixation. Saccades, on the other hand, are represented by lines connecting two consecutive fixations. These lines indicate rapid eye movements between different areas of the stimulus.



Study 1: Using Eye-Tracking Vignettes to Improve Pre-Service Teachers' Diagnostic Skills in Histogram Interpretation

Combining multiple sources

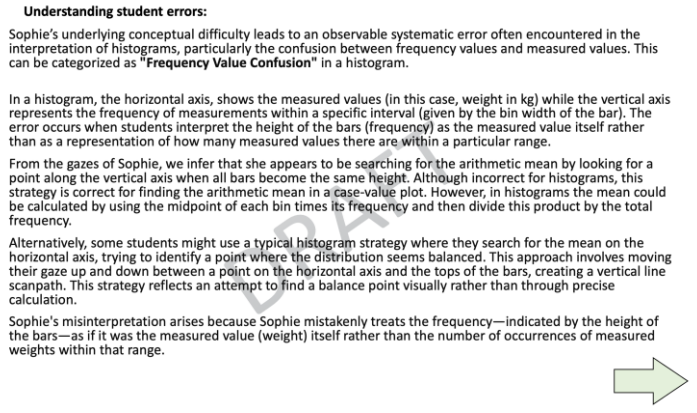
- Diagnosis developed only when explanation, gaze data, and supportive & procedural information were integrated
- Triangulation supported a more confident judgment

Diagnosing systematic error

- Participant had noticed **axes confusion**
- Diagnosis consolidated labeled the error as **frequency–value confusion and** explained *why* the student made it
- Even confusing explanations helped decide whether the error was systematic

The follow-up task

- Participant applied what they had learned
- Correctly distinguished which histograms contained the same error
- Gaze data is mostly sufficient; in one case, an explanation also needed
- Evidence of **dev**eloping diagnostic skills through the vignette



Understanding student errors:

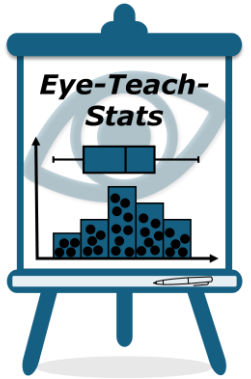
Sophie's underlying conceptual difficulty leads to an observable systematic error often encountered in the interpretation of histograms, particularly the confusion between frequency values and measured values. This can be categorized as **"Frequency Value Confusion"** in a histogram.

In a histogram, the horizontal axis, shows the measured values (in this case, weight in kg) while the vertical axis represents the frequency of measurements within a specific interval (given by the bin width of the bar). The error occurs when students interpret the height of the bars (frequency) as the measured value itself rather than as a representation of how many measured values there are within a particular range.

From the gazes of Sophie, we infer that she appears to be searching for the arithmetic mean by looking for a point along the vertical axis when all bars become the same height. Although incorrect for histograms, this strategy is correct for finding the arithmetic mean in a case-value plot. However, in histograms the mean could be calculated by using the midpoint of each bin times its frequency and then divide this product by the total frequency.

Alternatively, some students might use a typical histogram strategy where they search for the mean on the horizontal axis, trying to identify a point where the distribution seems balanced. This approach involves moving their gaze up and down between a point on the horizontal axis and the tops of the bars, creating a vertical line scanpath. This strategy reflects an attempt to find a balance point visually rather than through precise calculation.

Sophie's misinterpretation arises because Sophie mistakenly treats the frequency—indicated by the height of the bars—as if it was the measured value (weight) itself rather than the number of occurrences of measured weights within that range.



STUDY 2:Pre-Service Teachers' Use of Eye-Tracking Data to Diagnose Students' Misinterpretations in Statistical Graphs

Making assumptions before engaging with gaze data

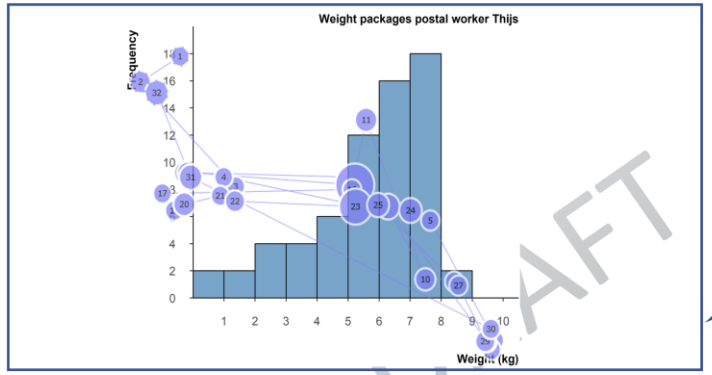
- Initial reasoning based on the student's answer, not gaze data
- Assumed the student focused on the tallest bar ("8")
  - *"Because the frequency of eight is higher, I think it could be eight."*

Misread fixation number as attention count

  - *"I said I had looked at these sixteen parts (referring to the frequency 16 on the y-axis) 32 times…"*
- Gaze plot used to **confirm assumptions**, not to question them

Revisiting gaze data with new information

- Explanation made vague gaze data more meaningful
- Student focused on **Y-axis** (bar height), teacher used **X-axis** (balance point)
- Shift from assumption to **interpretive use** of gaze data



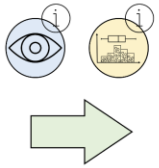
Sophie's Comment:

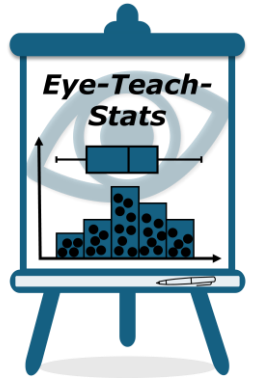
"I looked at how tall the bars were on the graph to see how often things showed up. Then, I tried to guess where the mean would be based on that. I also tried to make the bars all the same height to make it easier to figure out."

Now you can read the original comment of Sophie.

Indicate how well your assumptions align to those presented here.

To greater extend
Somewhat
Very little
Not at all

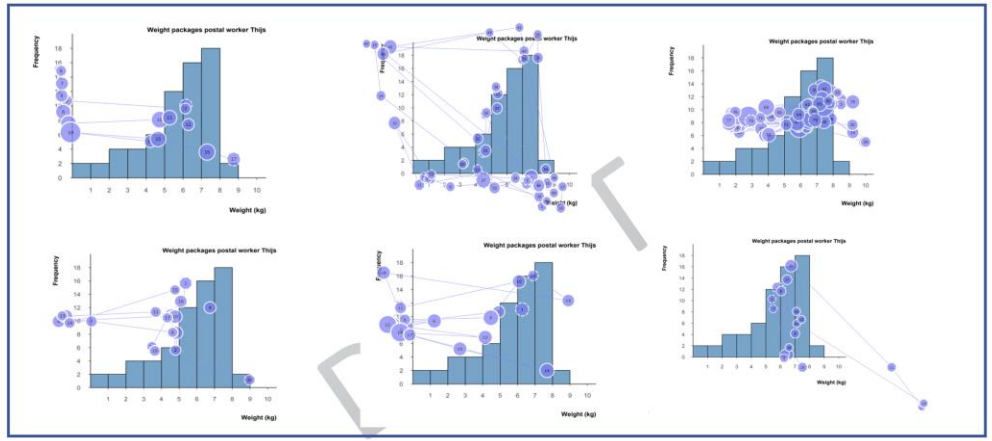




STUDY 2:Pre-Service Teachers' Use of Eye-Tracking Data to Diagnose Students' Misinterpretations in Statistical Graphs

- **Recognising the limits of gaze-only evidence**

- Reviewed 6 anonymous gaze plots without explanations
- Detected familiar patterns (e.g. frequency–value confusion)
- One plot showed scattered fixations → caused uncertainty
- Realised gaze plots can be useful, but not always conclusive
- Diagnosis became possible only through triangulation
  (gaze data + student explanation + supportive/procedural information)



Given the gaze plots of six students for the same item, identify which students are likely to make a frequency-value confusion systematic error.

Open answer field:

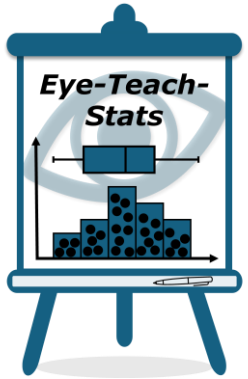
Eye-Tracking Studies on Learners' Strategies and Difficulties in Statistical Graphs

Systematic Literature Review

On systematic errors when interpreting statistical graphs

Development of eye-tracking vignettes

Evaluation of eye-tracking vignettes



STUDY 3: How Teachers Make Sense of Students' Misinterpreting Statistical Graphs: Teachers' Responses to Eye-tracking Vignettes

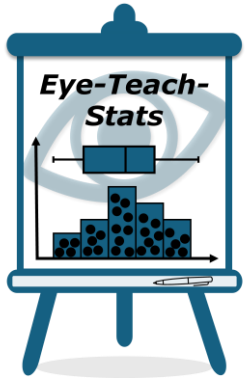
•Guiding research question:
How do teachers make sense of students' misinterpretations of histograms and case-value plots when provided with information about students' gaze patterns?

METHOD

Context: E-learning course with eye-tracking vignettes (NL & Flanders; Dutch-speaking teachers).

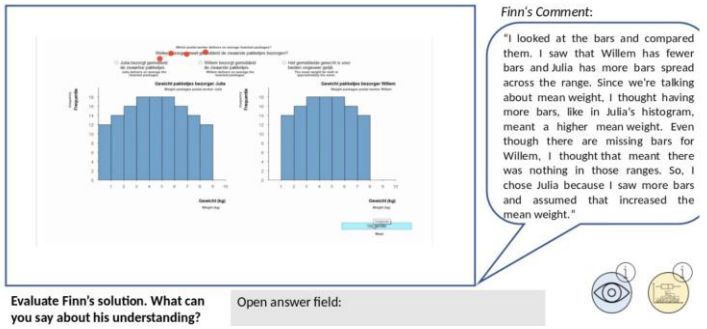
- **Participants:** 18 pre-/in-service teachers.
- **Materials:** 3 vignettes analysed here (2 histogram, 1 CVP) + 4C/ID-based structure.
- **Data:** teachers' open responses (steps 5,7,9) + closed answers (step 2).
- *Preliminary analysis — paper under review.*

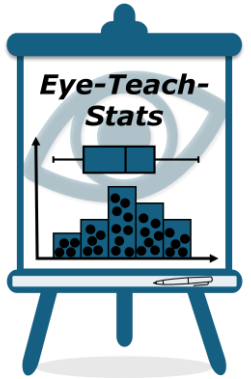
Boels et al., – ICOTS 2026



STUDY 3: How Teachers Make Sense of Students' Misinterpreting Statistical Graphs: Teachers' Responses to Eye-tracking Vignettes

- Teachers often explained student thinking via **local features** (e.g., number/height/positions of bars) and sometimes via **global features** (shape/distribution).
- Some teachers showed **similar misinterpretations** themselves while solving the tasks.
- After the vignette's **systematic error explanation**, teachers' diagnoses became **more accurate/precise** in follow-up tasks.
- Overall pattern: productive diagnosis seems to require attending to **both local and global features**.
- *Ongoing data collection & coding; findings preliminary.*

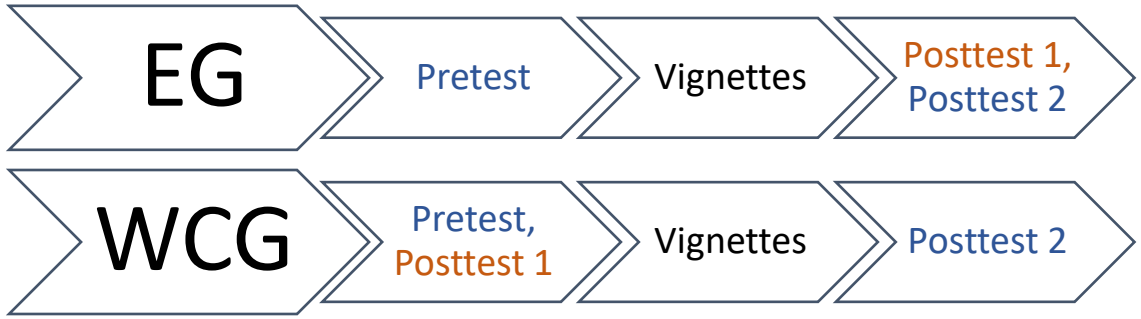




STUDY 4: Preliminary look at pre- and post-tests.

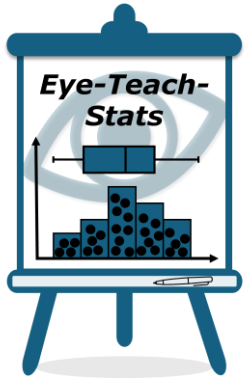
N= 52 (DE= 41; BE= 10, NL= 1);
Mean age: 24.4 years (67% female);
Pre-service teacher & in-training teacher

- Two groups
- Three test points
- Two different tests



What was measured?

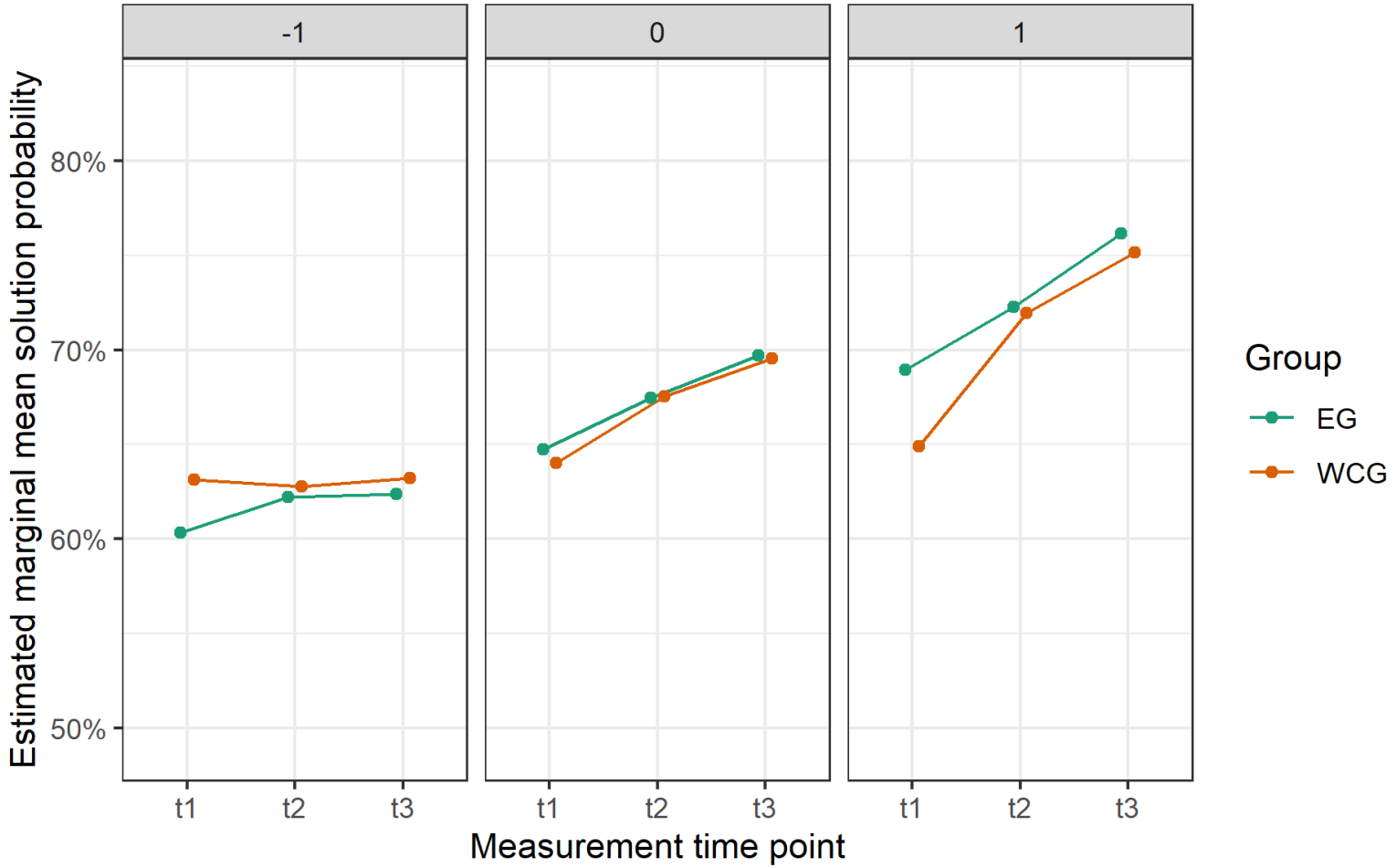
- CK (Content Knowledge): Subject matter knowledge
- PCK (Pedagogical Content Knowledge): Knowledge required to teach statistics



STUDY 4: Preliminary look at pre- and post-tests.

Grouping based on the students' initial Content Knowledge (CK) at the start of the study:

- Panel -1: Students with below-average CK.
- Panel 0: Students with average CK.
- Panel 1: Students with above-average CK.



The eye-teach stats project

Eye-teach-stats: Supporting teachers to diagnose and deal with students' statistical misinterpretations - innovative teacher training with eye-tracking vignettes"

Modules for teacher training with eye-tracking vignettes
developed, tested in practice and evaluated for effectiveness

Objective: Establishment and further development of the statistical PCK and CK of (prospective) teachers

Gefördert aus Mitteln der Europäischen Union

You would like to support us? Share this link with pre- and in-service teacher:

https://www.soscisurvey.de/pretest_ckpck/?q=base

Thank you!

Contact: heursen@ph-heidelberg.de & vuslat.seker@hu.nl
For more information: https://eye-teach-stats.eu/

References

Abt, M., Reinhold, F., & Van Dooren, W. (2023). Revealing cognitive processes when comparing box plots using eye-tracking data—A pilot study. In M. Ayalon, B. Koichu, R. Leikin, L. Rubel, & M. Tabach (Eds.), Proceedings of the 46th Conference of the International Group for the Psychology of Mathematics Education (pp. xx–xx). PME.

Boels, L. (2023). 1_Raw_Data_Students (Version V1) [Data set]. DataverseNL. https://doi.org/10.34894/WEKAYE

Boels, L., & Van Dooren, W. (2023). Secondary school students interpreting and comparing dotplots: An eye-tracking study. Utrecht University of Applied Sciences.

Boels, L., Bakker, A., & Drijvers, P. (2019a). Eye tracking secondary school students' strategies when interpreting statistical graphs. In M. Graven, H. Venkat, A. A. Essien, & P. Vale (Eds*.), Proceedings of the Forty-third Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 113–120). PME. http://www.igpme.org/publications/

Boels, L., Bakker, A., Van Dooren, W., & Drijvers, P. (2019b). Conceptual difficulties when interpreting histograms: A review. *Educational Research Review, 28*, 100291. https://doi.org/10.1016/j.edurev.2019.100291

Boels, L., Bakker, A., & Drijvers, P. (2019c). Unravelling teachers' strategies when interpreting histograms: An eye-tracking study. In U. T. Jankvist, M. Van den Heuvel-Panhuizen, & M. Veldhuis (Eds.), *Proceedings of the Eleventh Congress of the European Society for Research in Mathematics Education* (pp. 888–895). Freudenthal Group & Freudenthal Institute, Utrecht University, and ERME. https://hal.archives-ouvertes.fr/hal-02411575/document

Boels, L., Bakker, A., Van Dooren, W., & Drijvers, P. (2022). Secondary school students' strategies when interpreting histograms and case-value plots: An eye-tracking study. Freudenthal Institute, Utrecht University.

Chumachenko, D., Shvarts, A., & Budanov, A. (2014). The development of the visual perception of the Cartesian coordinate system: An eye tracking study. In C. Nicol, P. Liljedahl, S. Oesterle, & D. Allan (Eds.), *Proceedings of the Joint Meeting of PME 38 and PME-NA 36* (Vol. 2, pp. 313–320). PME. https://files.eric.ed.gov/fulltext/ED599779.pdf

Cox, L. S. (1975). Systematic Errors in the Four Vertical Algorithms in Normal and Handicapped Populations. *Journal for Research in Mathematics Education*, *6*(4), 202–220. https://doi.org/10.2307/748696

Dvir, M., & Ben-Zvi, D. (2018). The role of model comparison in young learners' reasoning with statistical models and modeling. *ZDM Mathematics Education, 50*, 1183–1196.

References

Friesen, M., Kuntze, S., & Vogel, M. (2018). Videos, Texte oder Comics? Die Rolle des Vignettenformats bei der Erhebung fachdidaktischer Analysekompetenz zum Umgang mit Darstellungen im Mathematikunterricht. In J. Rutsch, M. Rehm, M. Vogel, M. Seidenfuß, & T. Dörfler (Eds.*), Effektive Kompetenzdiagnose in der Lehrerbildung: Professionalisierungsprozesse angehender Lehrkräfte untersuchen* (pp. 153–177). Springer.

Knoop-Van Campen, C. A. N., Kok, E., Van Doornik, R., De Vries, P., Immink, M., Jarodzka, H., & Van Gog, T. (2021). How teachers interpret displays of students' gaze in reading comprehension assignments. *Frontline Learning Research, 9(4)*, 116–140. https://doi.org/10.14786/flr.v9i4.881

Kreutz, J., Leuders, T., & Hellmann, K. (Hrsg.). (2020). *Professionsorientierung in der Lehrerbildung*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-25046-1

Lem, S., Onghena, P., Verschaffel, L. & van Dooren, W. (2013). External Representations For Data Distributions: In Search Of Cognitive Fit. *STATISTICS EDUCATION RESEARCH JOURNAL, 12(1),* 4-19. https://doi.org/10.52041/serj.v12i1.319

Loibl, K., Leuders, T., & Dörfler, T. (2020). A framework for explaining teachers' diagnostic judgements by cognitive modeling (DiaCoM). *Teaching and Teacher Education, 91,* 103059. https://doi.org/10.1016/j.tate.2020.103059

Padberg, F. (1996). Learning from mistakes: Improving mathematics teaching through error analysis*. Friedrich-Jahresheft XIV: Testing and Assessment*, 56-59.

Radatz, H. (1979). Error analysis in mathematics education. *Journal for Research in Mathematics Education, 10(3),* 163–172. https://doi.org/10.2307/748804

Rutsch, J., Rehm, M., Vogel, M., Seidenfuß, M., & Dörfler, T. (Eds.). (2018). *Effektive Kompetenzdiagnose in der Lehrerbildung: Professionalisierungsprozesse angehender Lehrkräfte untersuchen*. Springer.

Schreiter, S., & Vogel, M. (2023). Eye-tracking measures as indicators for a local vs. global view of data*. Frontiers in Education, 7*, 1–11. https://doi.org/10.3389/feduc.2022.1058150

Shimizu, Y., & Kang, H. (2025). Research on classroom practice and students' errors in mathematics education: A scoping review of recent developments for 2018–2023. *ZDM – Mathematics Education.* https://doi.org/10.1007/s11858-025-01704-0

Spivey, M. J., & Dale, R. (2011). Eye movements both reveal and influence problem solving. In S. P. Liversedge, I. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 551–562). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199539789.013.0030

van Merriënboer, J. J. G., Kirschner, P. A., & Frèrejean, J. (2024). *Ten Steps to Complex Learning: A Systematic Approach to Four-Component Instructional Design.* Routledge. https://doi.org/10.4324/9781003322481