



Eye-Teach-Stats Eye-Tracking Vignettes © 2026

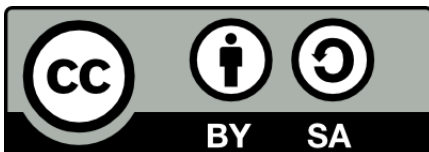


Co-funded by
the European Union

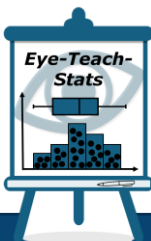
Erasmus+
Enriching lives, opening minds.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor the granting authority can be held responsible for them.

Eye-Tracking Vignettes



Eye-Tracking Vignettes © 2026 by Eye-Teach-Stats Project is licensed under CC BY-SA 4.0. To view a copy of this license, visit creativecommons.org/licenses/by-sa/4.0/



Overview



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

01 Dotplot

02 Boxplot

03 Boxplot

04 Case-Value Plot

05 Case-Value Plot

06 Dotplot

07 Histogram

08 Case-Value Plot

09 Histogram

10 Dotplot

11 Dotplot

12 Histogram

13 Histogram

14 Boxplot

15 Boxplot

16 Boxplot

17 Histogram

18 Case-Value Plot

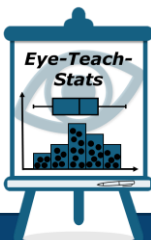
19 Histogram

20 Histogram

Supportive Information

Procedural Information

Students' solutions



Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor the granting authority can be held responsible for them.

01 Dotplot

Vignette on an error caused by outlier overgeneralization

Introducing situation



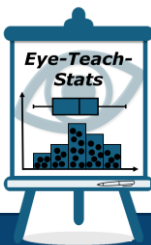
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Imagine you are teaching a 6th-grade class on data interpretation, and you introduce your students to comparing data distributions using dotplots.

You notice your student Amaya is having difficulties with a task you've assigned. She gives an incorrect answer.

You will now see the task Amaya is working on.



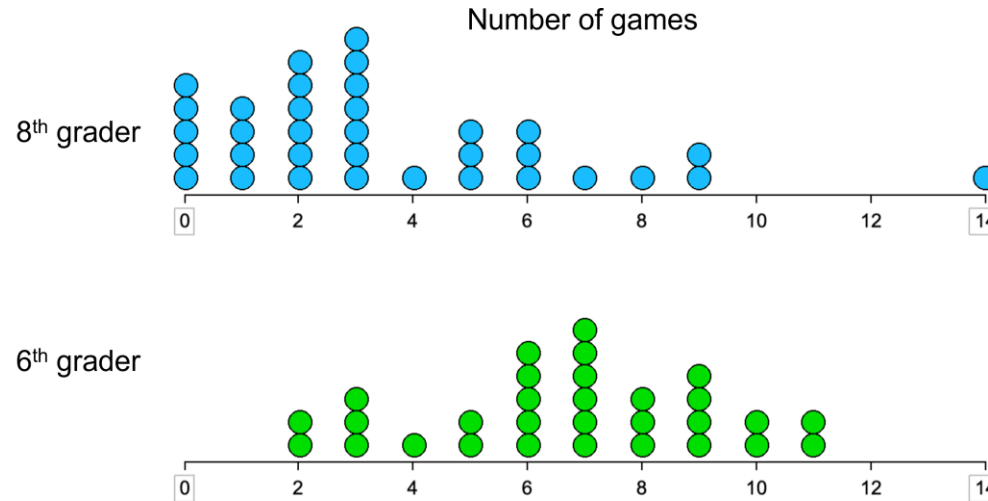
Student's task



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

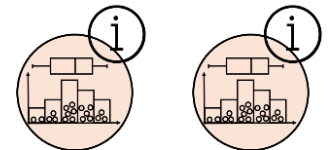
Do 8th graders and 6th graders differ in the number of games they have on their cell phones? A survey on this question was done at the Park School with 34 8th grader and 30 6th grader. Compare the two samples and then evaluate:



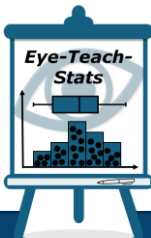
The data from the samples fits the assumption that

- 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.

This is the task Amaya is working on. **Please complete the task on your own first and enter your solution.**



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



Introducing eye-tracking

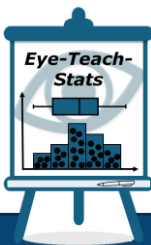


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

A group of researchers in mathematics education is pursuing the idea of supporting teachers with additional diagnostic material: To this end, eye movement data from students is recorded using eye-tracking technology and made available as supplementary material.

You will receive a heatmap that visualizes gaze data of your student Amaya that was collected while she was working on the task. This heatmap indicates where her visual focus lies. Often it is possible to infer what Amaya's thinking process might have been.



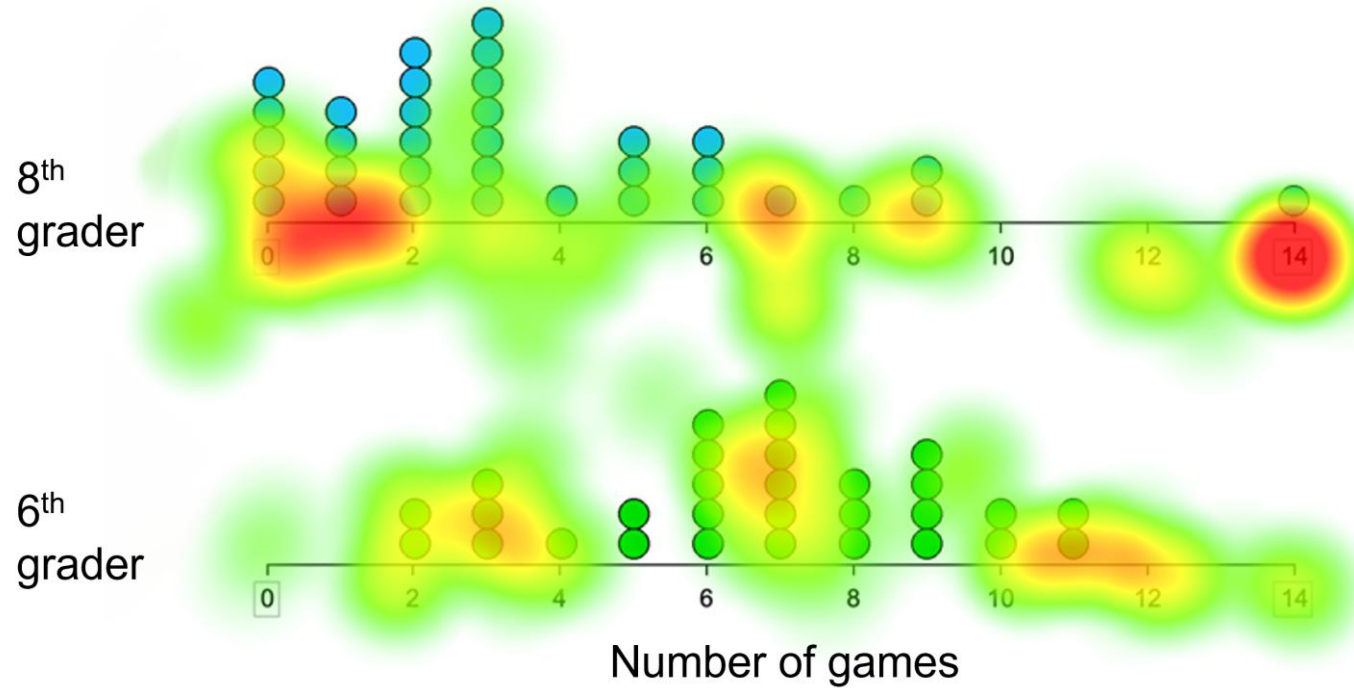
First look



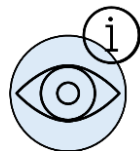
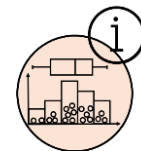
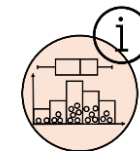
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Amaya's Answer: „I think a) is correct: the 8th graders have more games.”



Have a look at the heatmap that shows your student Amaya's gaze data and her answer.



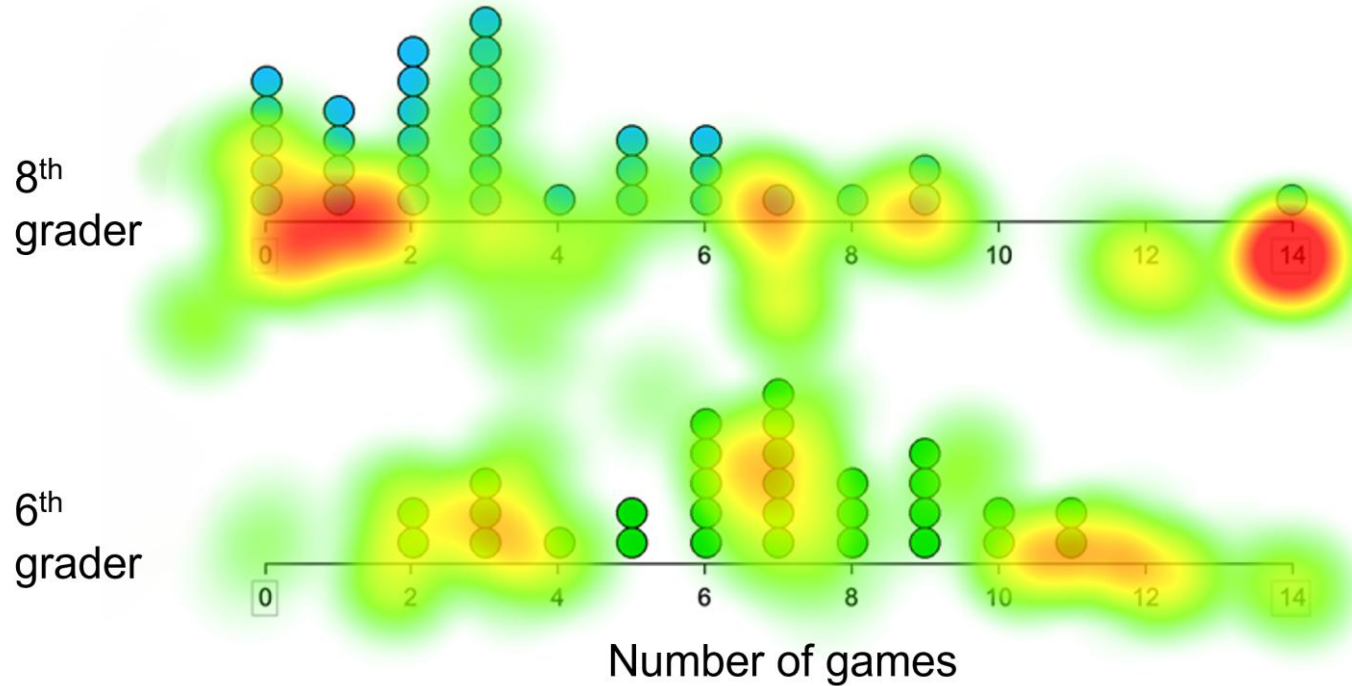
Take perspective



Co-funded by
the European Union

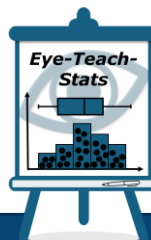
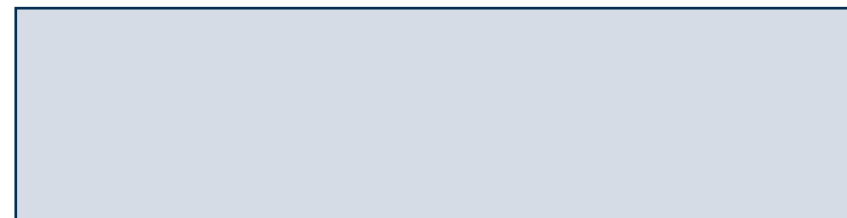
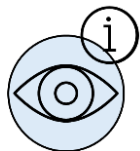
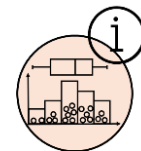
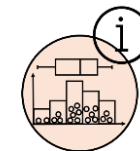
Erasmus+
Enriching lives, opening minds.

Amaya's Answer: „I think a) is correct: the 8th graders have more games.”



After the task Amaya was shown her own heatmap. She explains retrospectively what she was thinking while working on the task.

Try to take Amaya's perspective, and formulate what she may comment.



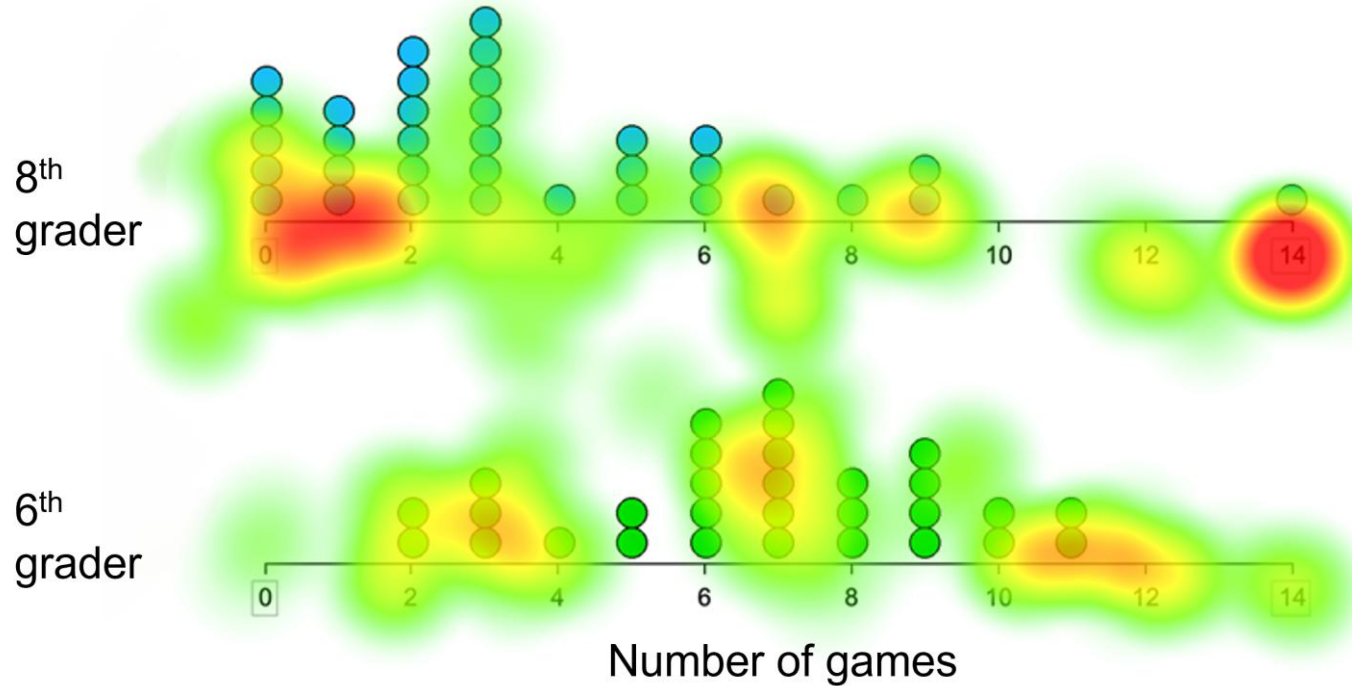
Indicate assumption



Co-funded by
the European Union

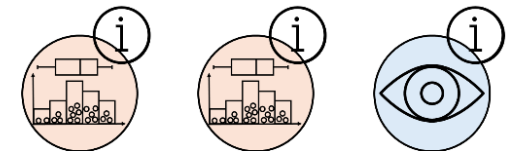
Erasmus+
Enriching lives, opening minds.

Amaya'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.“



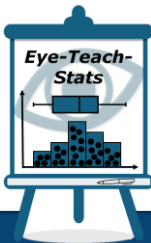
Now you can read the original comment of Amaya.

Indicate to what extent Amaya's comment agrees with what you assumed.



- ☐ No alignment at all
- ☐ Little alignment

- ☐ Quite some alignment
- ☐ Complete alignment



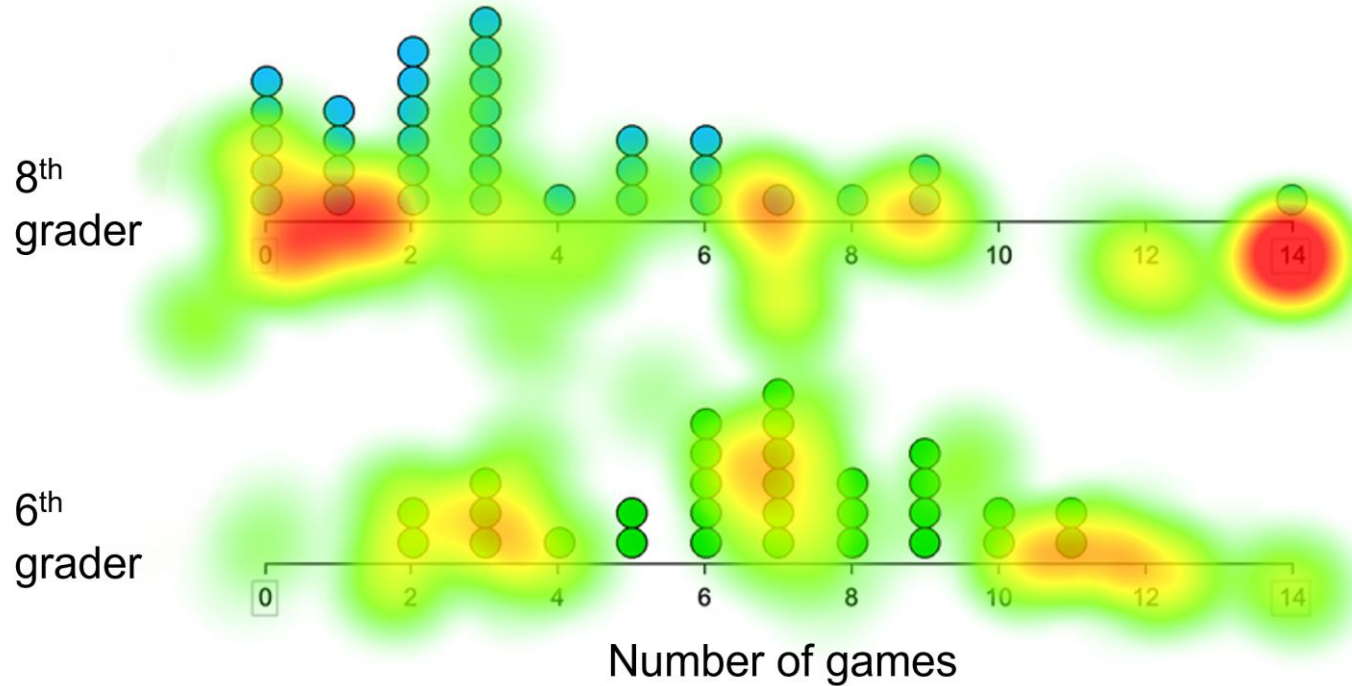
Evaluate solution



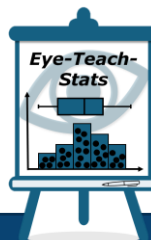
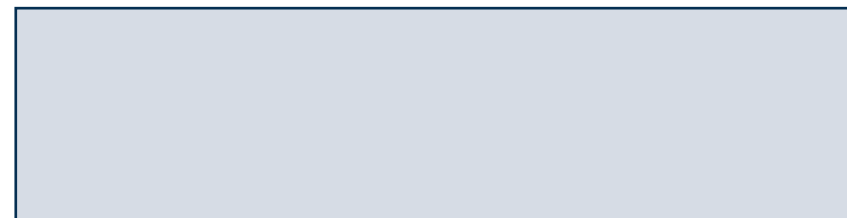
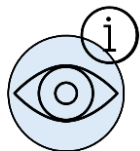
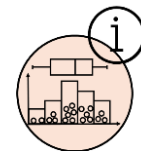
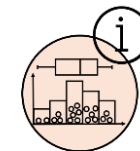
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Amaya'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.“



Evaluate Amaya's solution. What can you say about her understanding?



Introducing systematic error



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Understanding student errors:

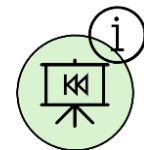
Amaya has an underlying conceptual difficulty regarding the interpretation of dotplots. This leads to an observable systematic error, called “**Overemphasizing of outliers**”.

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

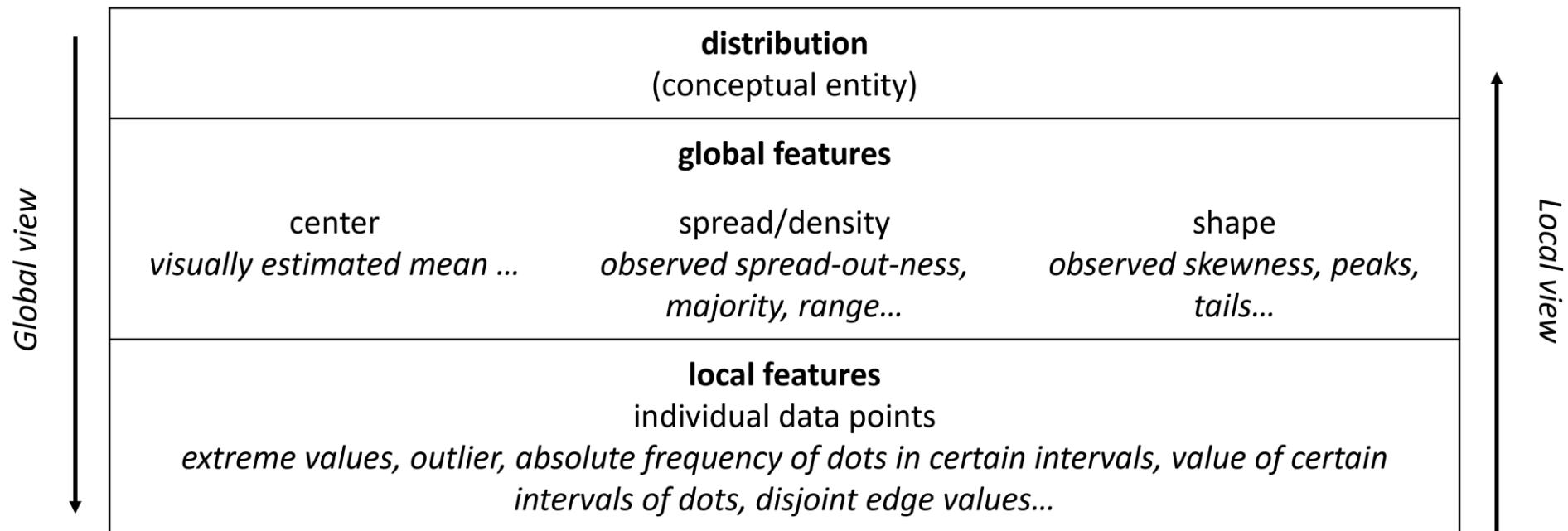
Understanding error and underlying conceptual difficulty: In a dotplot, each dot stands for one occurrence of a value (in this case the number of games) on the horizontal axis. The height of the stacks (how many dots are piled up) shows how many times each value occurs.

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.

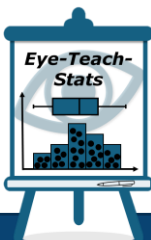
On the next slide, you will see an overview of local and global features that can be taken into account when comparing distributions.



Perspectives on data distribution: local vs. global view



Adapted according to Schreiter & Vogel, 2023





Eye-Teach-Stats Eye-Tracking Vignettes © 2026



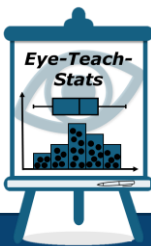
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor the granting authority can be held responsible for them.

02 Boxplot

Vignette on an error caused by area bias



Introducing situation



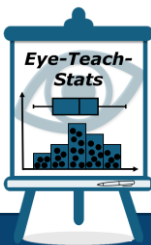
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Imagine you are teaching an 8th-grade class on data interpretation, and you introduce your students to comparing data distributions using boxplots.

You notice your student Quinn is having difficulties with a task you've assigned. He gives an incorrect answer.

You will now see the task Quinn is working on.



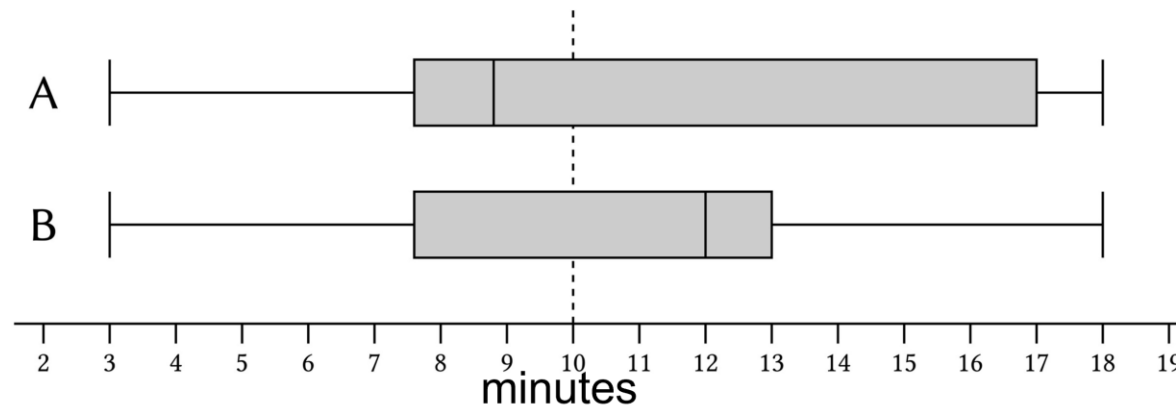
Student's task



Co-funded by
the European Union

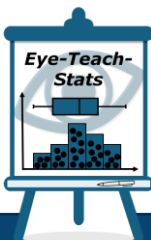
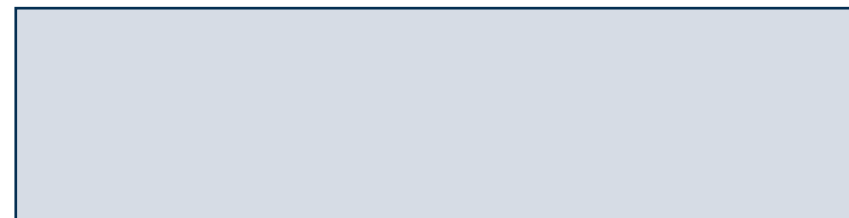
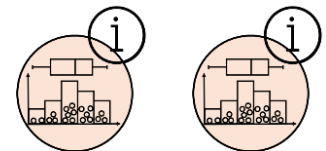
Erasmus+
Enriching lives, opening minds.

Two schools A and B are attended by the same number of students. All students have indicated in survey how many minutes they need to get to school in the morning. The results are presented by two box plots. One school (A or B) is attended by more students with a walk to school of more than 10 minutes. Decide which one.



This is the task Quinn is working on.

Please complete the task on your own first and enter your solution.



Introducing eye-tracking

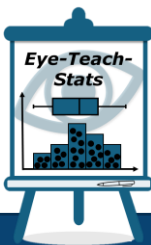


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

A group of researchers in mathematics education is pursuing the idea of supporting teachers with additional diagnostic material: To this end, eye movement data from students is recorded using eye-tracking technology and made available as supplementary material.

You will receive a gaze video sequence of your student Quinn that was collected while he was working on the task. This gaze video indicates where his visual focus lies. Often it is possible to infer what your student's thinking process might have been.



First look



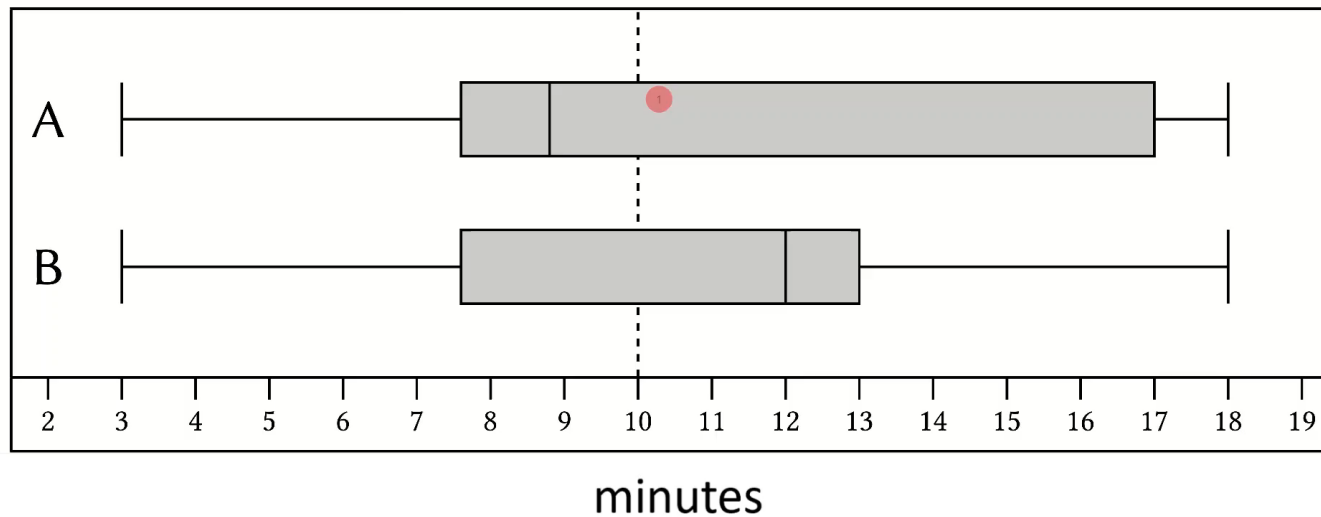
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

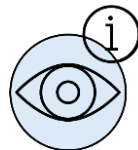
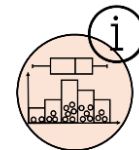
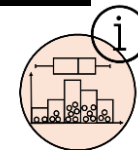
Quinn's Answer: „For me, School A is clearly the right answer. You can easily see from the box plot A that many more pupils stated that they walked more than 10 minutes to school.“

One school has a larger number of children with a walk to school longer than 10 minutes.

Decide which one.



Have a look at the video that shows your student Quinn's gaze movement and his answer.



Take perspective



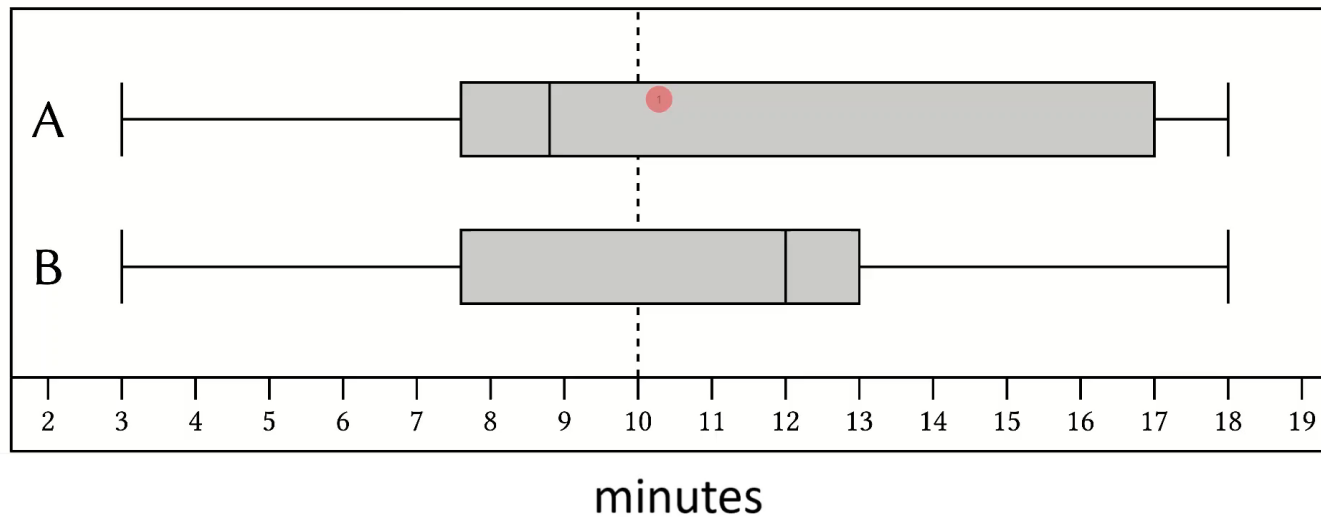
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Quinn's Answer: „For me, School A is clearly the right answer. You can easily see from the box plot A that many more pupils stated that they walked more than 10 minutes to school.“

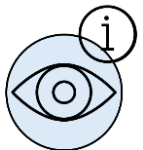
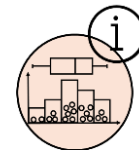
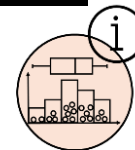
One school has a larger number of children with a walk to school longer than 10 minutes.

Decide which one.



After the task Quinn was shown his own video. He explains retrospectively what he was thinking while working on the task.

Try to take Quinn's perspective, and formulate what he may comment.



Indicate assumption



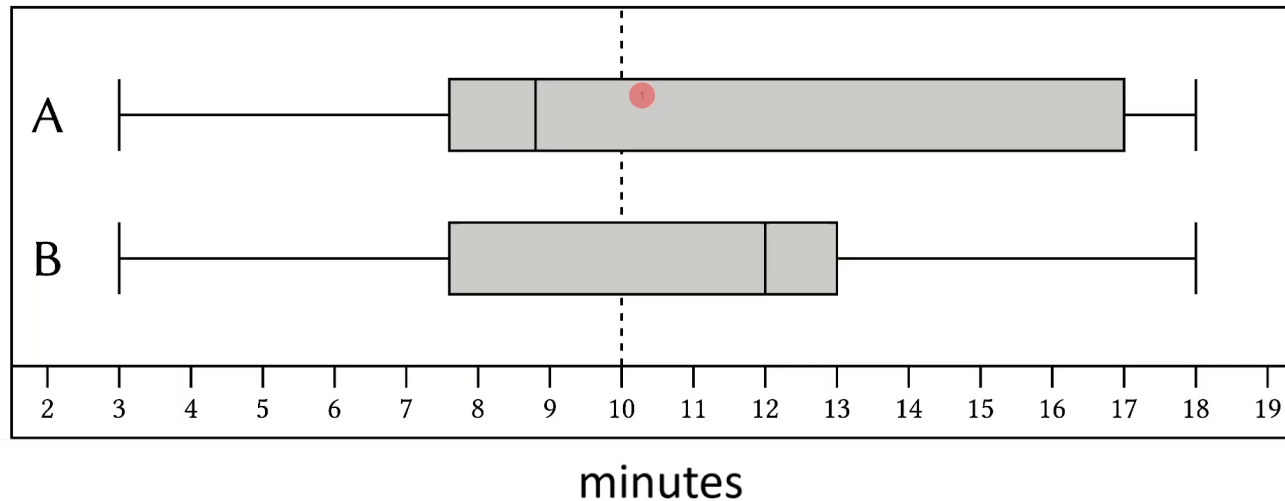
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Quinn's Comment: "The box plot represents all the responses, so to speak. So I checked where more of the box is above the 10-minute limit. This is quite clearly the case with box plot A. So I also opted for school A."

One school has a larger number of children with a walk to school longer than 10 minutes.

Decide which one.

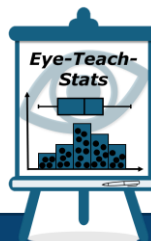
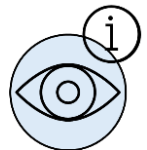
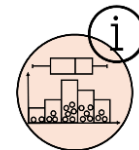
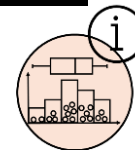


Now you can read the original comment of Quinn.

Indicate to what extent Quinn's comment agrees with what you assumed.

- ☐ No alignment at all
- ☐ Little alignment

- ☐ Quite some alignment
- ☐ Complete alignment



Evaluate solution



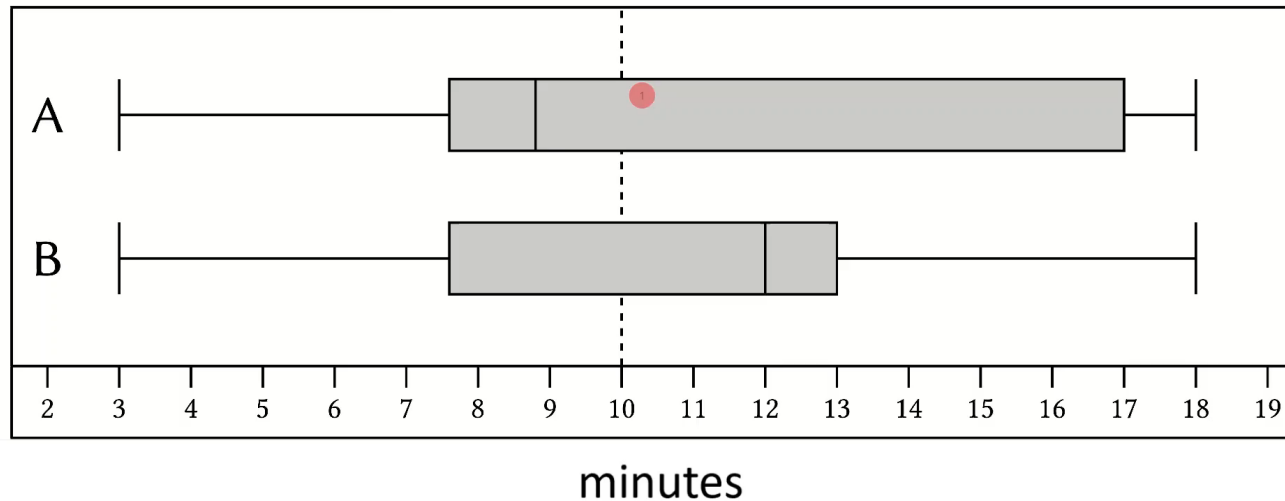
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

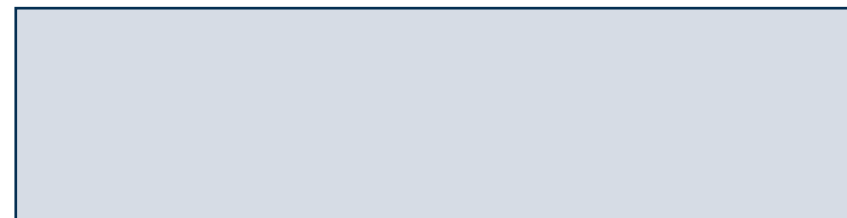
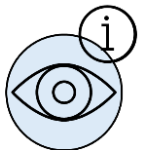
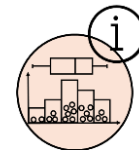
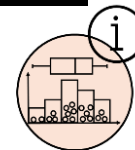
Quinn's Comment: "The box plot represents all the responses, so to speak. So I checked where more of the box is above the 10-minute limit. This is quite clearly the case with box plot A. So I also opted for school A."

One school has a larger number of children with a walk to school longer than 10 minutes.

Decide which one.



Evaluate Quinn's solution. What can you say about his understanding?



Introducing systematic error



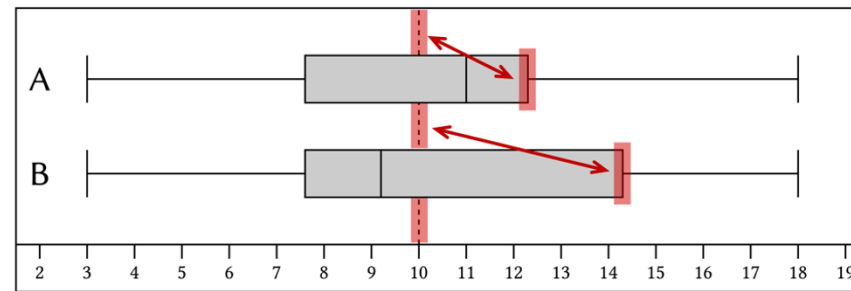
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Understanding student errors:

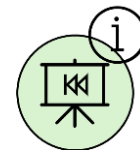
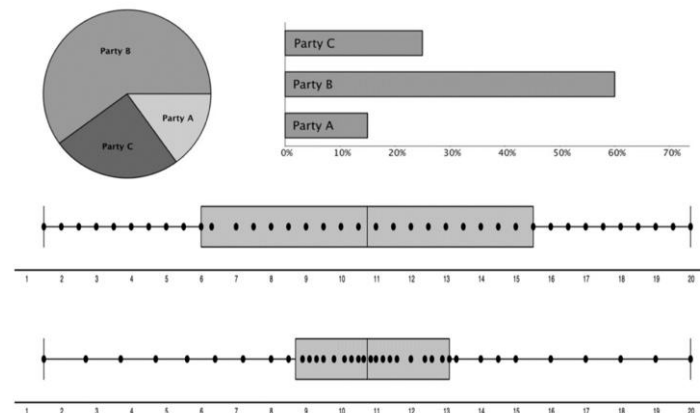
Quinn has an underlying conceptual difficulty regarding the interpretation of boxplots. This leads to an observable systematic error, called **area bias**.

Observable behavior: In his comment, Quinn describes that he decides the question of which of two boxplots represents more data points above a certain value based on which of the two boxplots shows more box area above this critical value. His eye movements also show a noticeable focus on the third quartiles, which represent the right-hand boundary of the box.



Interpretation by the teacher/researcher: A plausible explanation for Quinn's approach and his eye movements is that he incorrectly assumes that a larger box area in the box plot also corresponds to a larger proportion of the represented sample. Not only Quinn's comment fits this hypothesis, but also his eye movements, as the focus here is on the third quartiles, which indicate the amount of box area above the critical value.

Underlying conceptual difficulty: When learners encounter box plots for the first time, they are usually already familiar with other types of statistical representations such as circle diagrams or bar charts. In these types of representation, Quinn's assumption that there is a proportional relationship between the area and the represented sample fraction is correct. This is not the case with box plots, however, as the box always represents the middle half of the data and the box area or box width is therefore a measure of the variability of the data set and indicates how much the middle half of the data scatters (interquartile range). If learners inappropriately transfer their prior knowledge of other statistical representations when interpreting box plots, then the systematic error occurs that you have just observed with your student Quinn.

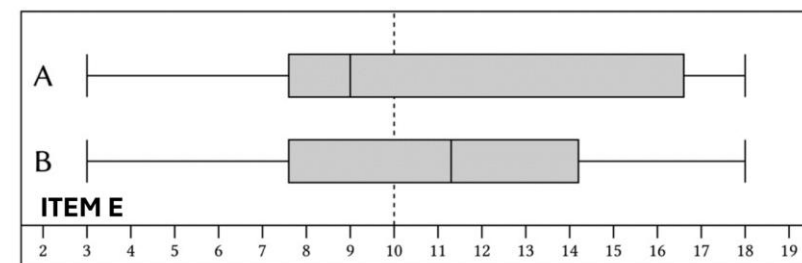
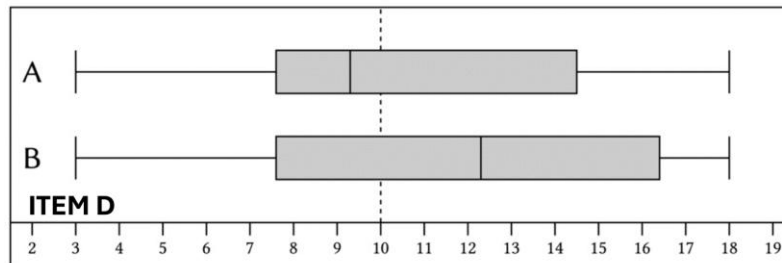
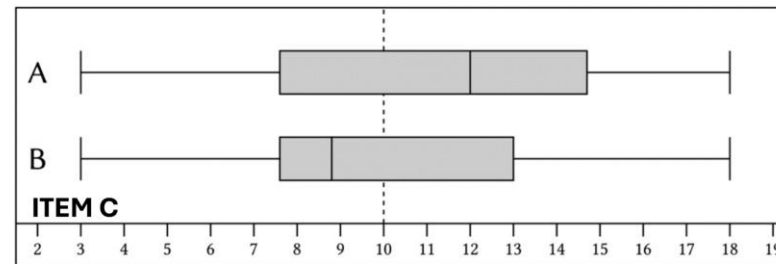
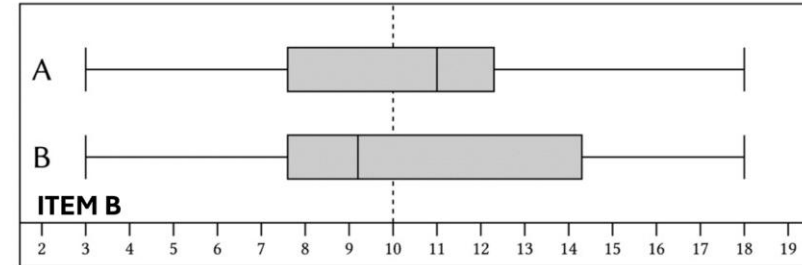
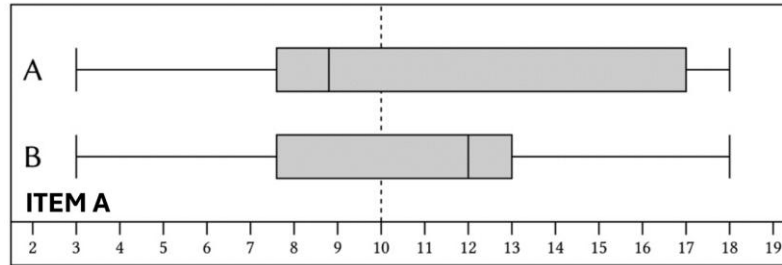


Practice



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.



Which of the following items (task and question are the same) would a student with conceptual difficulties answer incorrectly?

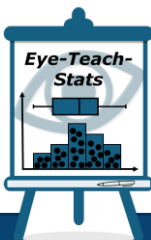
☐ A

☐ B

☐ C

☐ D

☐ E

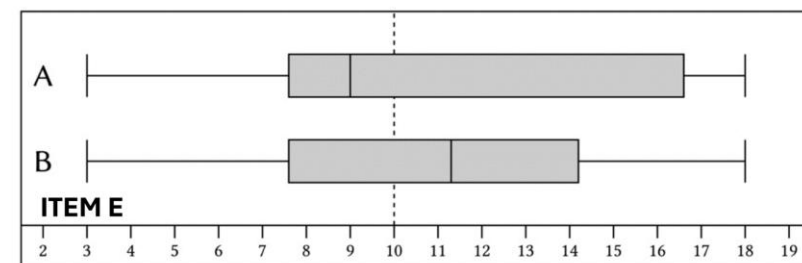
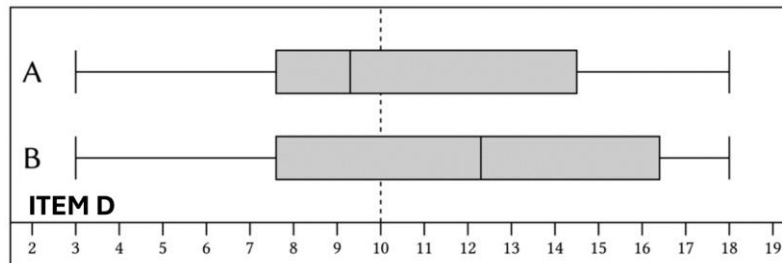
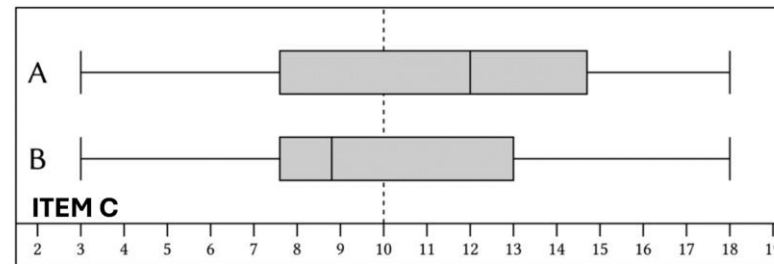
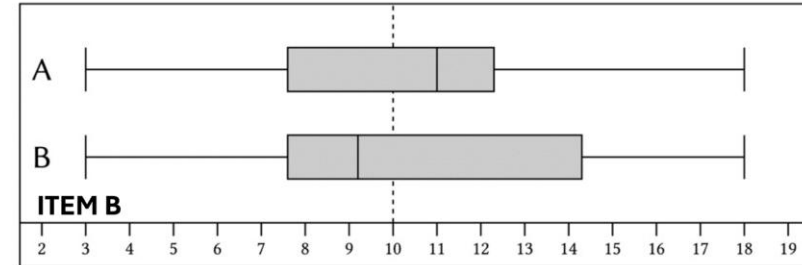
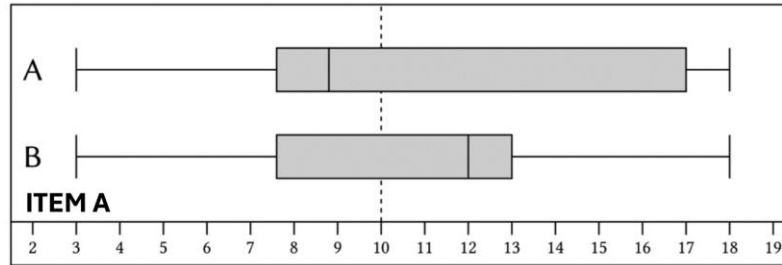


Practice II



Co-funded by
the European Union

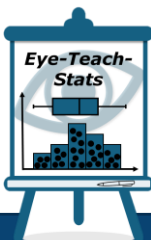
Erasmus+
Enriching lives, opening minds.



For the same items you see the gaze-pattern of five different students.

Which of these five different students – do you think – provided most likely an incorrect answer?

- ☐ Student A
- ☐ Student B
- ☐ Student C
- ☐ Student D
- ☐ Student E





Eye-Teach-Stats Eye-Tracking Vignettes © 2026



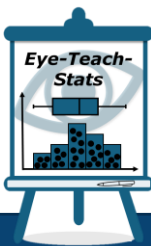
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor the granting authority can be held responsible for them.

03 Boxplot

Vignette on an error caused by median bias



Introducing situation



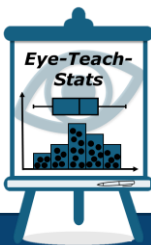
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Imagine you are teaching an 8th-grade class on data interpretation, and you introduce your students to comparing data distributions using box plots.

You notice your student Elodie is having difficulties with a task you've assigned. She gives an incorrect answer.

You will now see the task Elodie is working on.



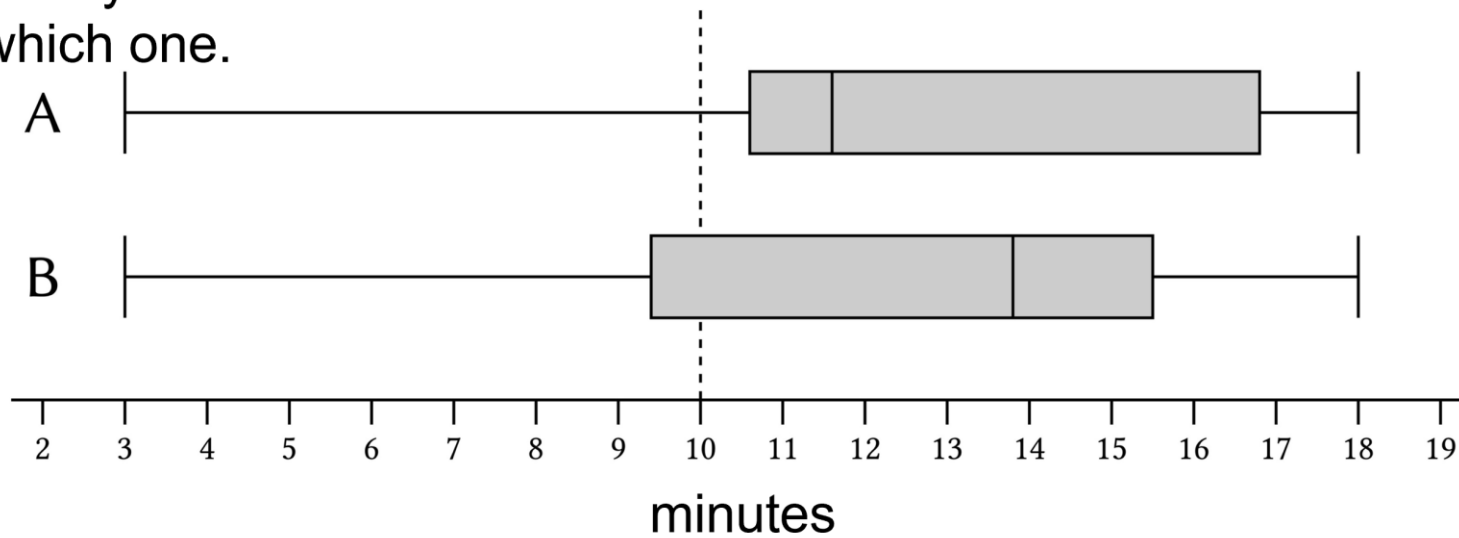
Student's task



Co-funded by
the European Union

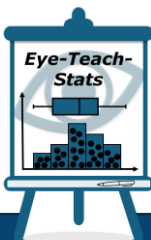
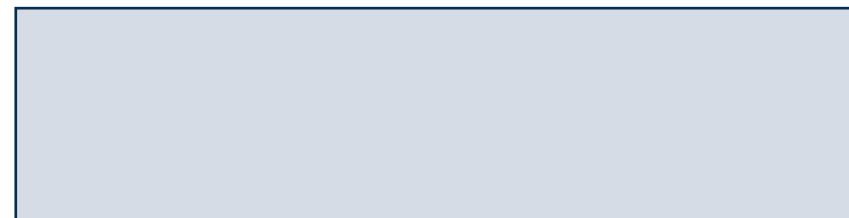
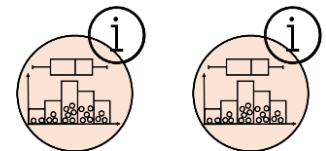
Erasmus+
Enriching lives, opening minds.

Two schools A and B are attended by the same number of students. All students have indicated in survey how many minutes they need to get to school in the morning. The results are presented by two box plots. One school (A or B) is attended by more students with a walk to school of more than 10 minutes. Decide which one.



This is the task Elodie is working on.

Please complete the task on your own first and enter your solution.



Introducing eye-tracking

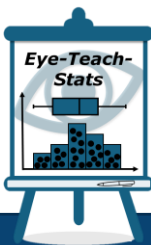


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

A group of researchers in mathematics education is pursuing the idea of supporting teachers with additional diagnostic material: To this end, eye movement data from students is recorded using eye-tracking technology and made available as supplementary material.

You will receive a heatmap that visualizes fixation points of your student Elodie that was collected while she was working on the task. This heatmap indicates where her visual focus lies. Often it is possible to infer what your student's thinking process might have been.



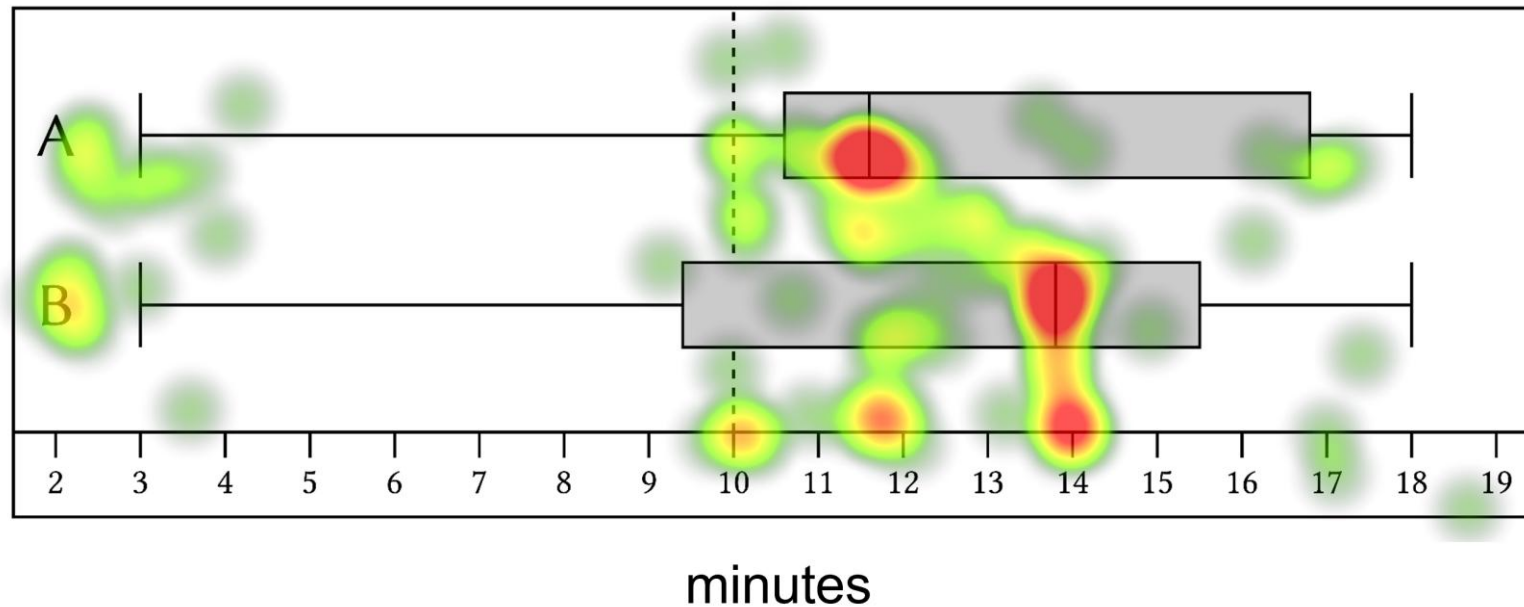
First look



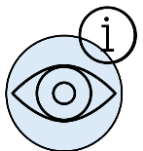
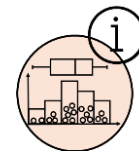
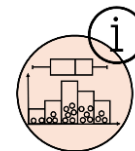
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Elodie's Answer: „For me, School B is clearly the right answer. You can easily see from the boxplot B that many more pupils stated that they walked more than 10 minutes to school.”



Have a look at the heatmap that shows your student Elodie's gaze movement and her answer.



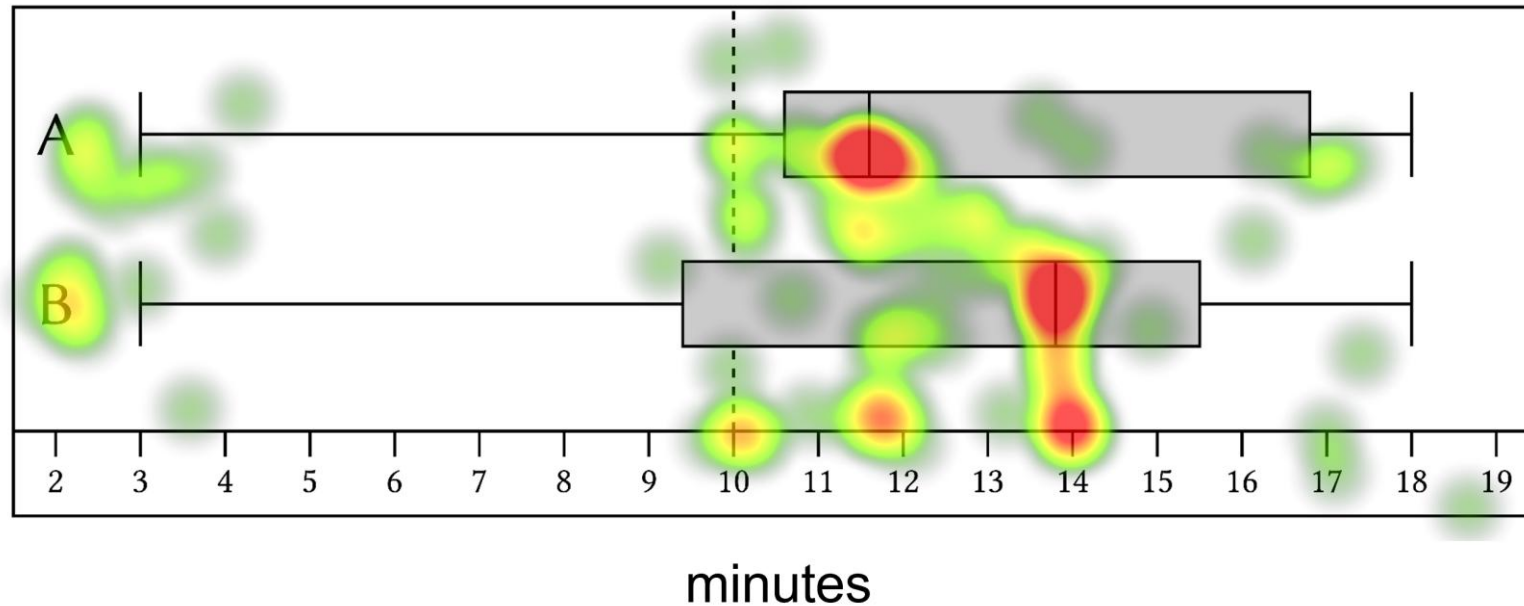
Take perspective



Co-funded by
the European Union

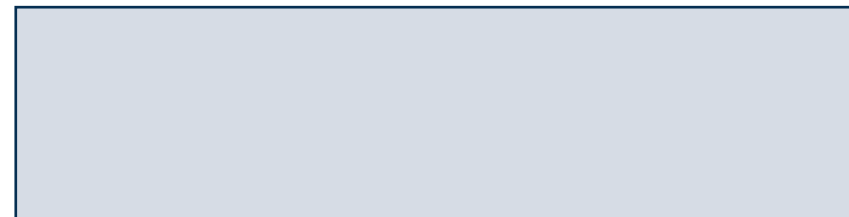
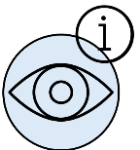
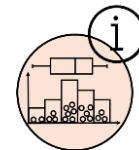
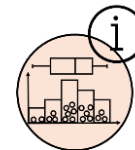
Erasmus+
Enriching lives, opening minds.

Elodie's Answer: „For me, School B is clearly the right answer. You can easily see from the boxplot B that many more pupils stated that they walked more than 10 minutes to school.“



After the task Elodie was shown her own heatmap. She explains retrospectively what she was thinking while working on the task.

Try to take Elodie's perspective, and formulate what she may comment.



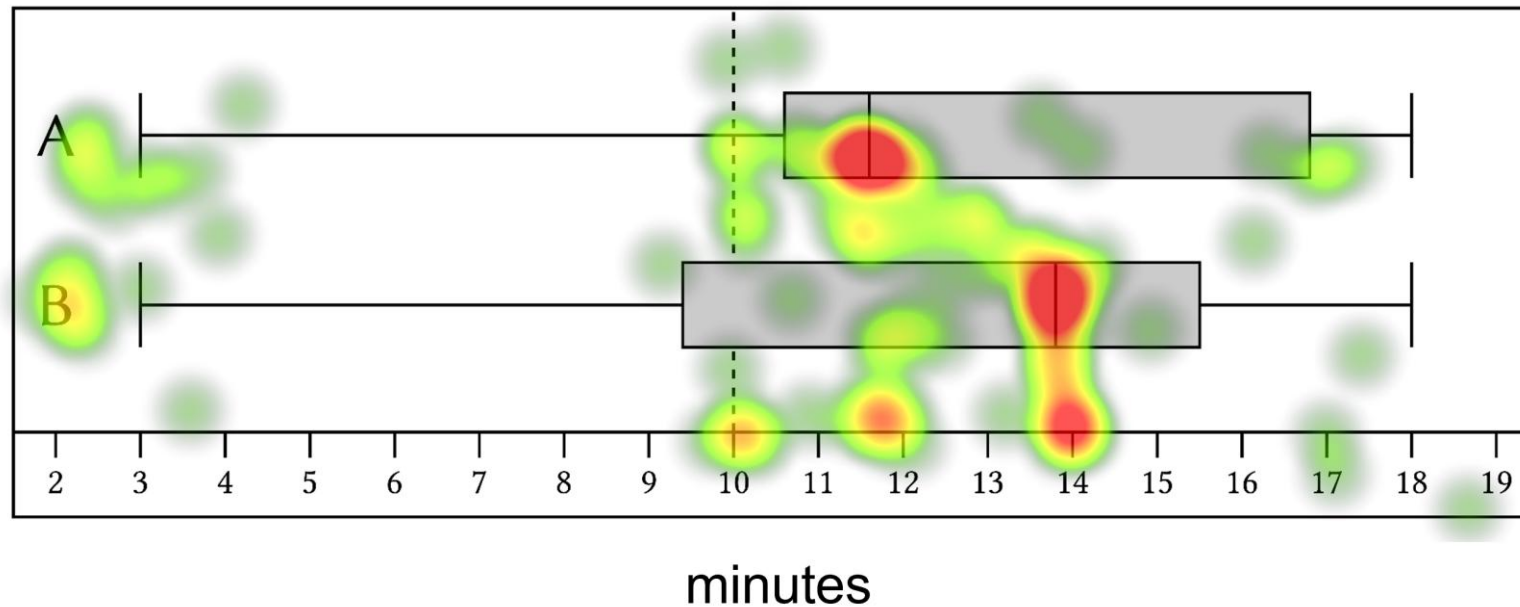
Indicate assumption



Co-funded by
the European Union

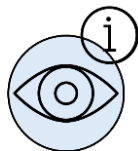
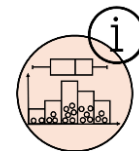
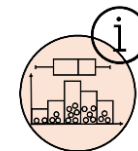
Erasmus+
Enriching lives, opening minds.

Elodie's Comment: "The line in the middle of the box is the median, which is an average value. So I mainly compared the two medians. You can clearly see that the children from school B take much longer on average to get to school than those from school A. So I selected B."



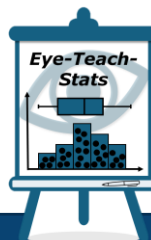
Now you can read the original comment of Elodie.

Indicate to what extent Elodie's comment agrees with what you assumed.



- ☐ No alignment at all
- ☐ Little alignment

- ☐ Quite some alignment
- ☐ Complete alignment



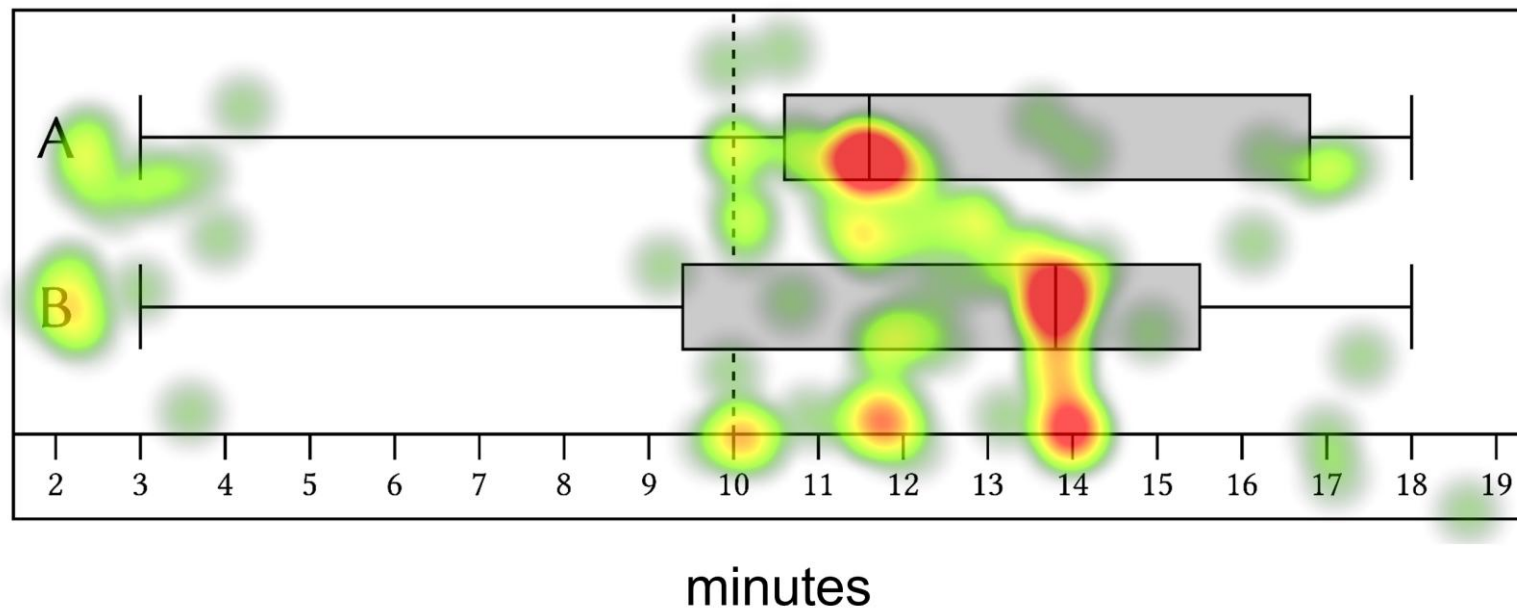
Evaluate solution



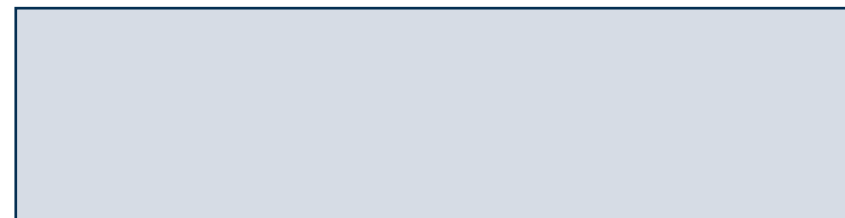
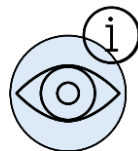
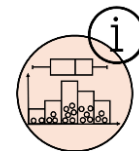
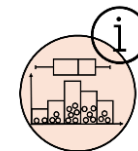
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Elodie's Comment: "The line in the middle of the box is the median, which is an average value. So I mainly compared the two medians. You can clearly see that the children from school B take much longer on average to get to school than those from school A. So I selected B."



Evaluate Elodie's solution. What can you say about her understanding?



Introducing systematic error



Co-funded by
the European Union

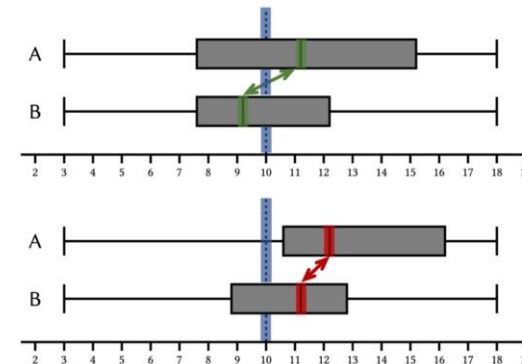
Erasmus+
Enriching lives, opening minds.

Understanding student errors:

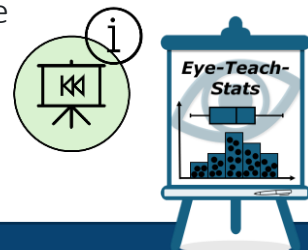
Elodie has an underlying conceptual difficulty regarding the interpretation of boxplots. This leads to an observable systematic error, called median bias.

Observable behavior: In her comment, Elodie describes that she decides the question of which of two boxplots represents more data points above a certain value based on which of the two boxplots has the higher median. Her eye movements also show a clear focus on the two medians.

Interpretation by the teacher/researcher: A plausible explanation for Elodie's approach and her eye movements is that she does not interpret the median primarily as a cut-off point that indicates where the boundary between the upper and lower half of the data is, but uses the median to decide which of the two statistical distributions tends to have the higher values and uses this central tendency for her decision. Her eye movements, which indicate that she is comparing the two medians directly with each other and not their position relative to th



Underlying conceptual difficulty: When learners encounter box plots for the first time, they are usually already familiar with various mean values such as the median. In particular, averages are not only familiar from school contexts, but also occur in different contexts outside of school. If learners identify the (well-known) median in the boxplot representation, it is plausible to assume that they use this familiar parameter to make decisions especially when they have not (yet) acquired the boxplot-specific knowledge that would be necessary for a correct solution. This overgeneralization of the median then leads to the systematic error that you have just observed with your student Elodie.

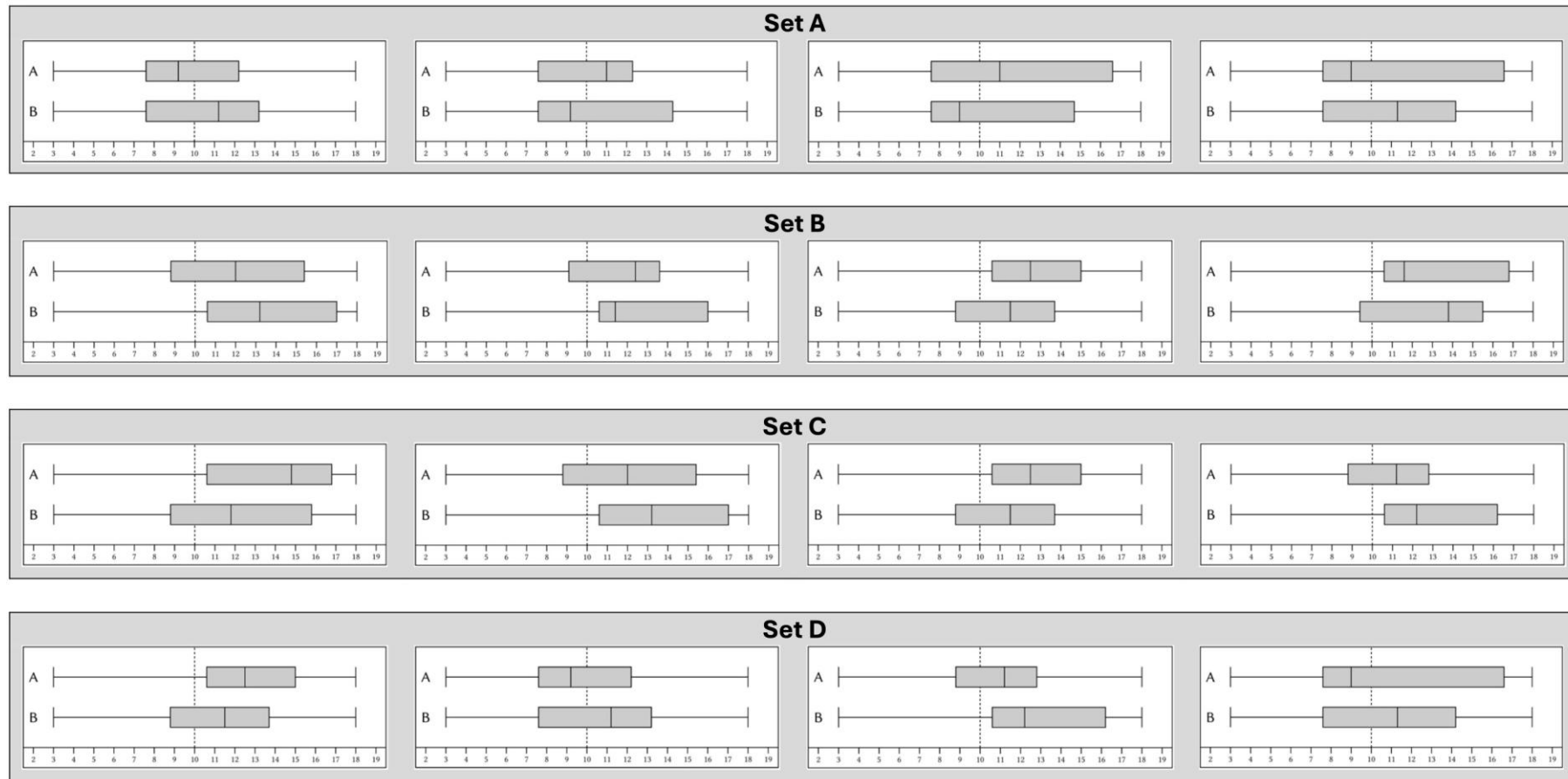


Practice



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.



Suppose you suspect that one of your students has the conceptual difficulty previously explained.

Which of the following item sets would you let your student solve so that you are sure?

☐ Set A

☐ Set B

☐ Set C

☐ Set D

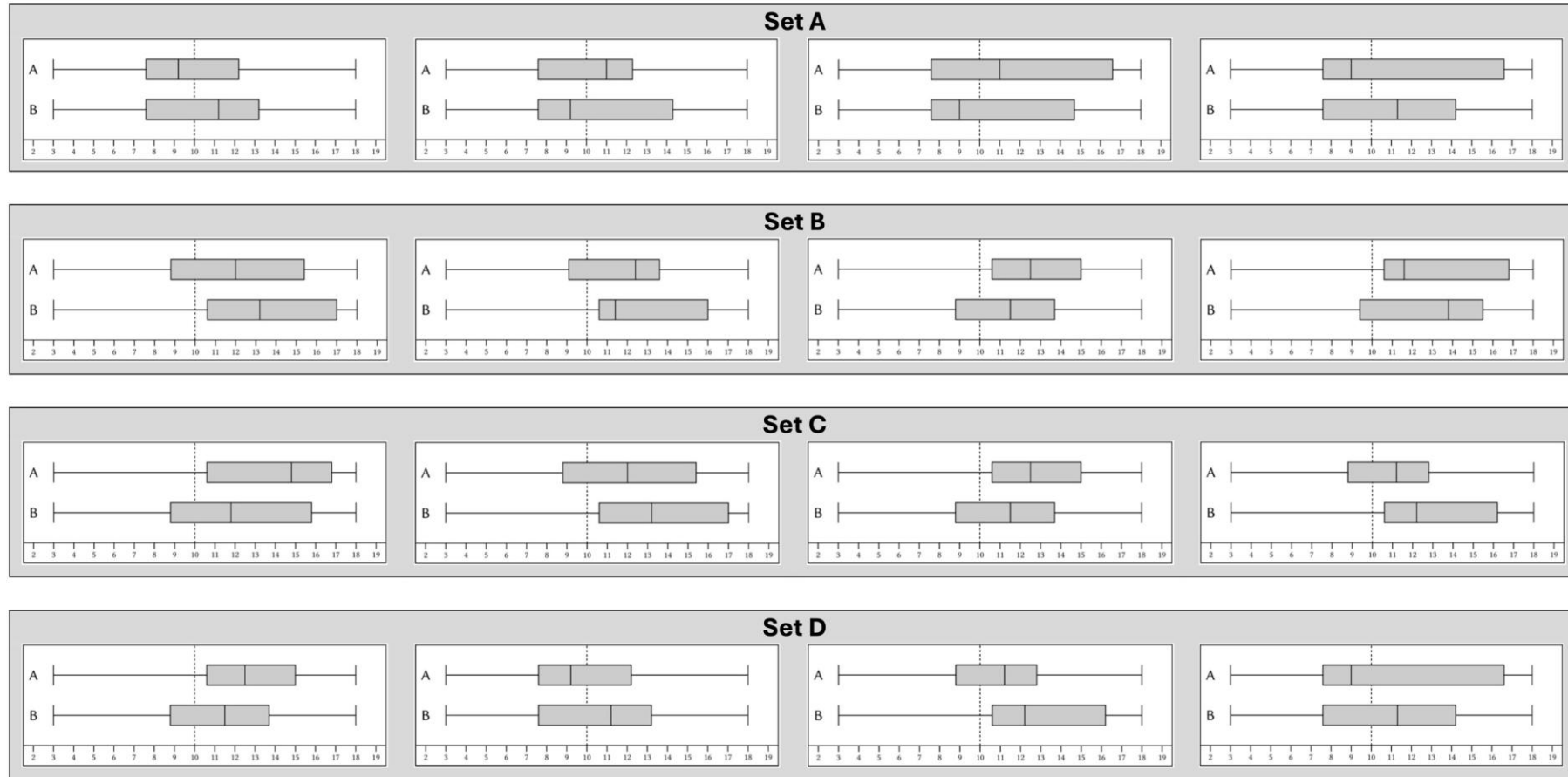


Practice II



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.



Suppose you suspect that one of your students has the conceptual difficulty previously explained. Which of the following item sets would you let your student solve so that you are sure?

Please provide detailed reasons why you have a) opted for item set [xy], and b) decided against the other item sets.



Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor the granting authority can be held responsible for them.

04 Case-Value Plot

Vignette on an error caused by an overgeneralization of a histogram approach

Introducing situation



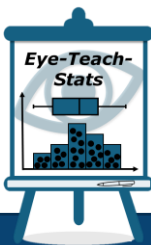
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Imagine you are teaching a 10th grade classroom on data interpretation, and you introduce your students to comparing data distributions using case-value plots.

You notice your student Roos is having difficulties with a task you've assigned. She gives an incorrect answer.

You will now see the task Roos is working on.

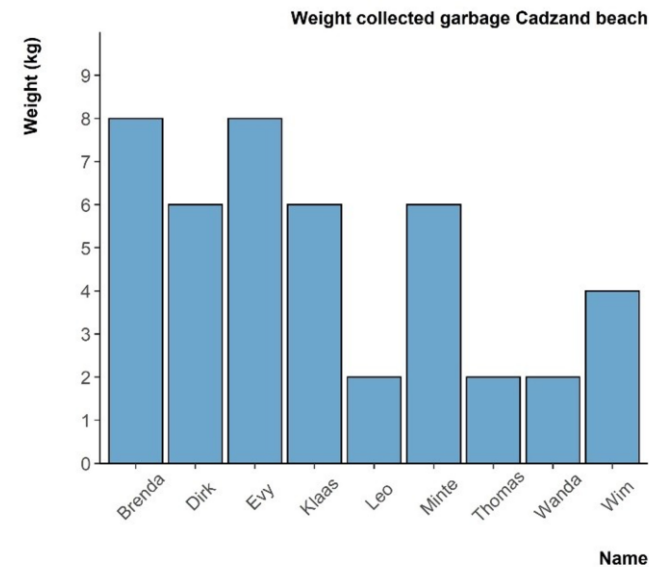
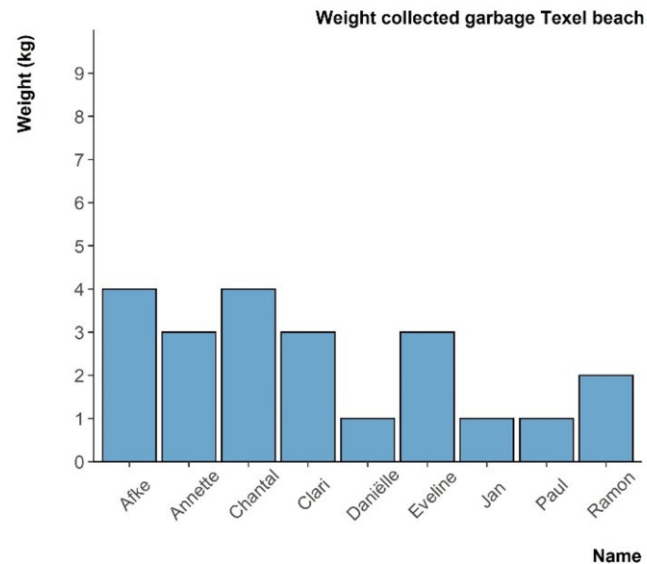


Student's task



Co-funded by
the European Union

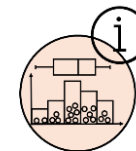
Erasmus+
Enriching lives, opening minds.



On which beach is the mean weight of the collected beach waste higher?

- a) On Texel the mean weight is the highest.
- b) On Cadzand the mean weight is the highest.
- c) The mean weight for both is approximately the same.

This is the task Roos is working on. **Please complete the task on your own first and enter your solution.**



- a) On Texel the mean weight is the highest.
- b) On Cadzand the mean weight is the highest.
- c) The mean weight for both is approximately the same.



Introducing eye-tracking

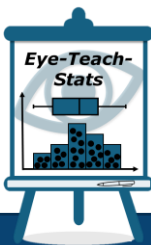


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

A group of researchers in mathematics education is pursuing the idea of supporting teachers with additional diagnostic material: To this end, eye movement data from students is recorded using eye-tracking technology and made available as supplementary material.

You will receive a short gaze video that shows the eye movements of your student Roos that were collected while she was working on the task. The gaze data indicates where her visual focus lies. Often it is possible to infer what Roos' thinking process might have been.



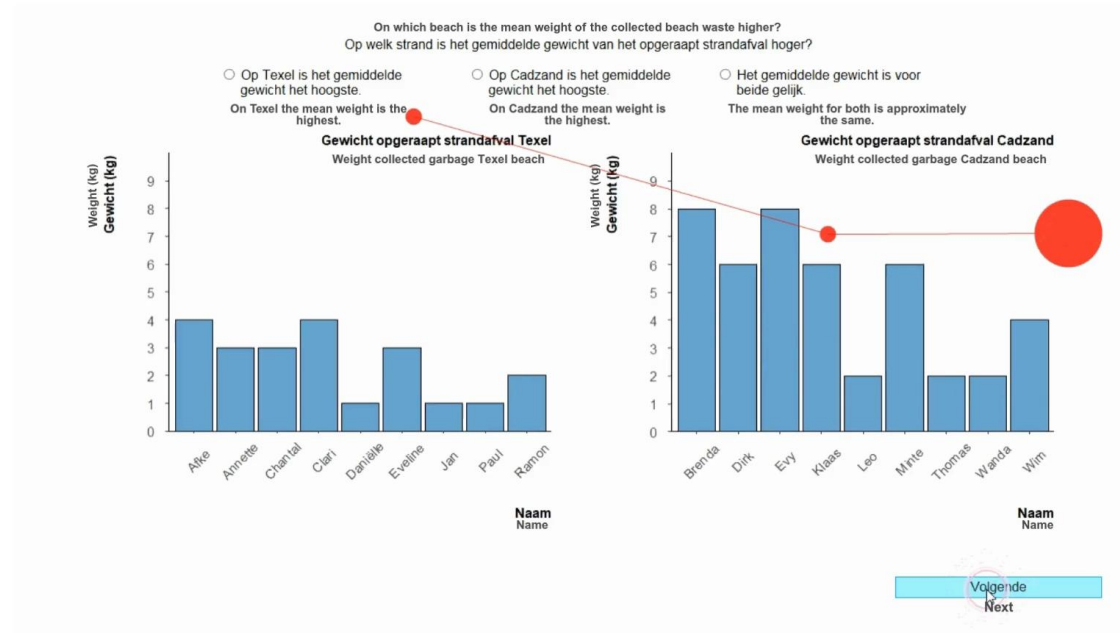
First look



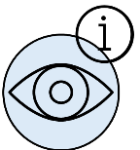
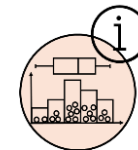
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Roos' answer: "The mean weight for both is approximately the same."



Have a look at the video sequence that shows your student Roos' gaze movements and her answer.



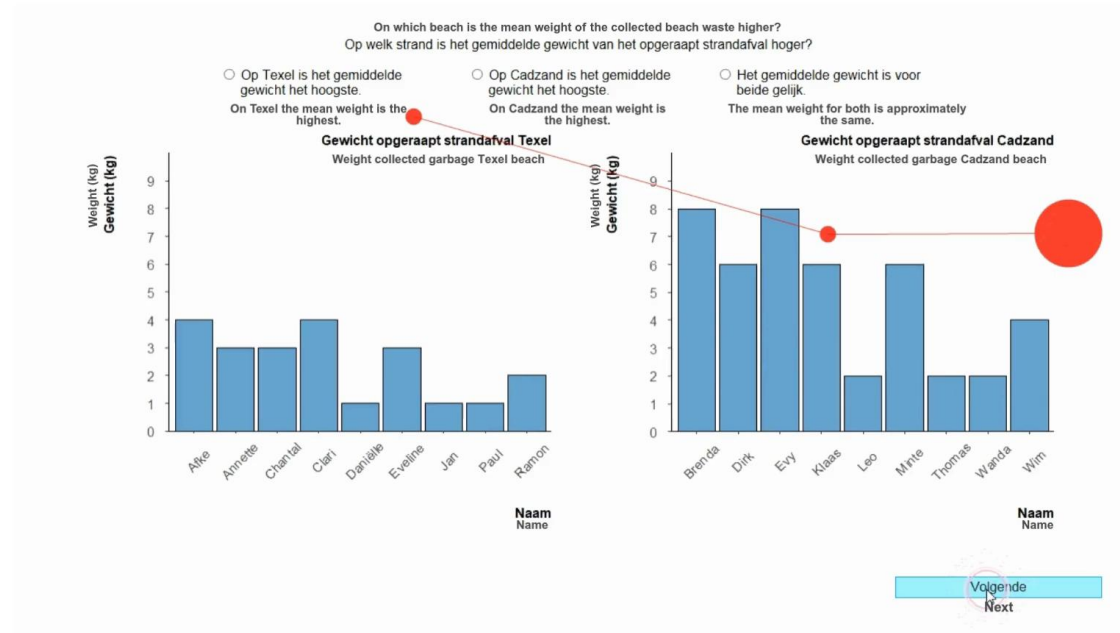
Take perspective



Co-funded by
the European Union

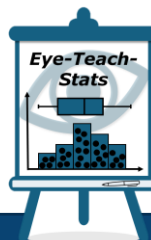
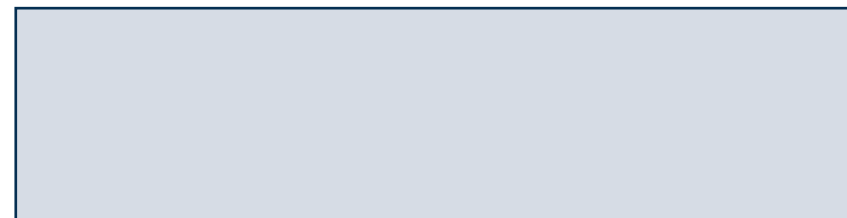
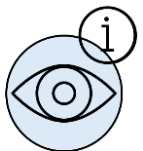
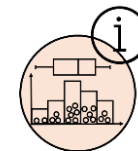
Erasmus+
Enriching lives, opening minds.

Roos' answer: "The mean weight for both is approximately the same."



After the task Roos was shown her own gaze video. She explains retrospectively what she was thinking while working on the task.

Try to take Roos' perspective, and formulate what she may comment.



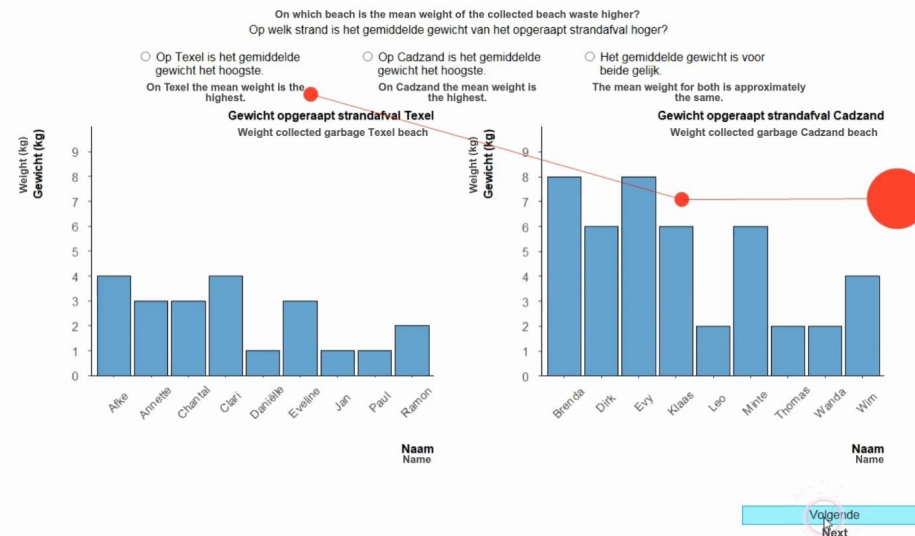
Indicate assumption



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Roos' comment: "I thought the means were the same because I looked at the weights in both graphs. When I saw that the weights in one graph were twice as high, I figured it was the same pattern but with higher values. So, I thought the means would still be the same."

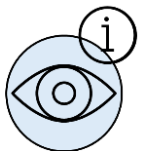
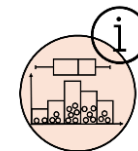


Now you can read the original comment of Roos.

Indicate to what extent Roos' comment agrees with what you assumed.

- ☐ No alignment at all
- ☐ Little alignment

- ☐ Quite some alignment
- ☐ Complete alignment



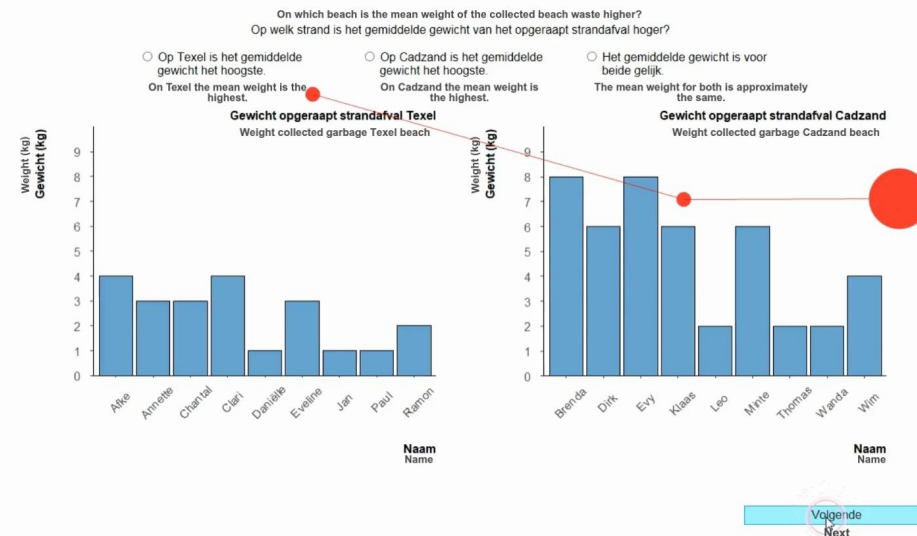
Evaluate solution



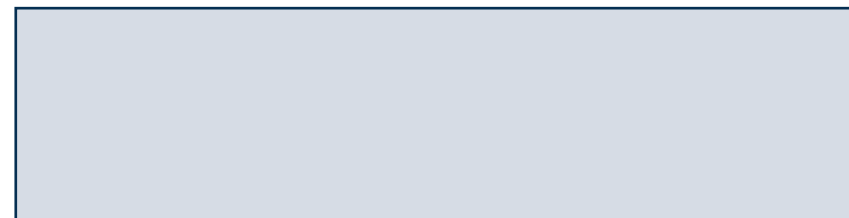
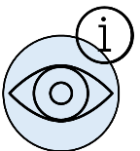
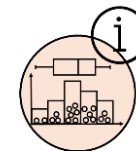
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Roos' comment: "I thought the means were the same because I looked at the weights in both graphs. When I saw that the weights in one graph were twice as high, I figured it was the same pattern but with higher values. So, I thought the means would still be the same."



Evaluate Roos' solution. What can you say about her understanding?



Introducing systematic error



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Understanding student errors:

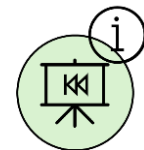
Roos has an underlying conceptual difficulty regarding the interpretation of case-value plots. This leads to an observable systematic error, called „**overgeneralization of a histogram approach**”.

Observable behavior: Roos said that the values in one case-value plot were doubled compared to the other but the pattern was the same. She interpreted this doubling as an indication of similar distributions and, therefore, similar means. Her gaze data show that Roos focused on visual similarity and equivalent bars in the case-value plots. Together this observable behavior points at a systematic error, which we term overgeneralization of a histogram approach when she interprets case-value plots.

Interpretation by the teacher/researcher: It appears that Roos applied a reasoning typically used for histograms—visually assessing bar heights and distribution shape—and interpreted the apparent doubling as a doubling of frequencies. Such strategy is appropriate for histograms, which display aggregated data in bins, but was here incorrectly transferred to case-value plots, where each bar represents only one individual data point.

Underlying conceptual difficulty: The systematic error is likely due to an underlying conceptual difficulty with the relation between the concept of distribution and the concept of center; therefore, in understanding how the mean is determined and compared in case-value plots. In a case-value plot, changing a bar's height is changing a measured value which in turn directly influences the mean. In histograms, doubling the heights of all bars (frequencies) does not change the mean as the overall distribution remains the same, hence the balance, thus the mean, remains the same.

Alternative Strategies and Correct Approaches for Case-value Plots: There are different ways students might approach case-value plots. For instance, some students use "area" as an approach for estimating the mean (more area is a higher mean). Although this can be correct in many case-value plots, it is an incomplete strategy (e.g., when there are more data points in one graph compared to another). A correct method for estimating the mean in case-value plots involves understanding that each bar's height is the measured value. Estimation of the mean could, for example, be done by visually drawing a horizontal line that would make all bars equally high. In histograms, bar heights represent frequencies, and the balance point of the graph along the horizontal axis approximates the mean.

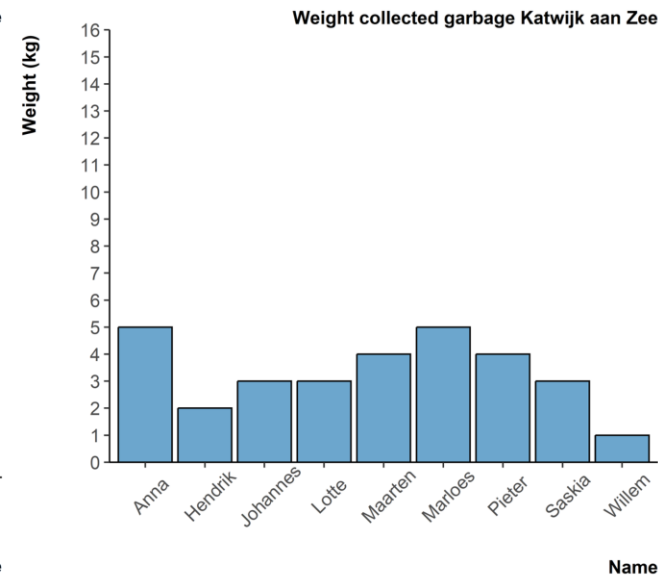
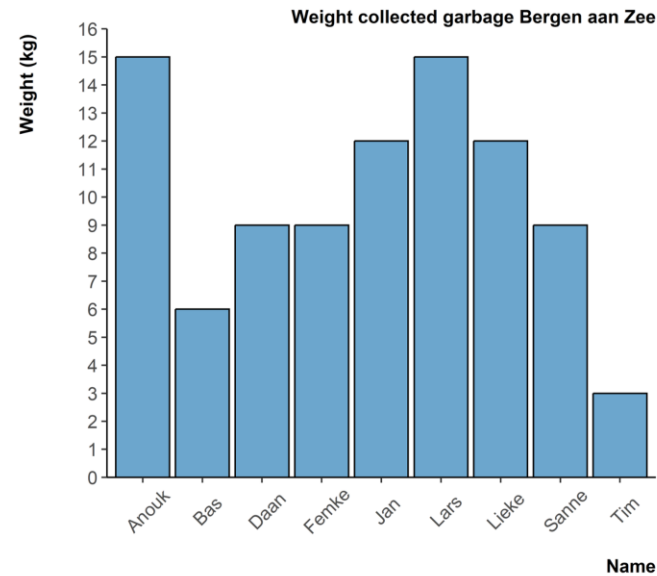


Practice



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

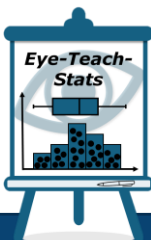
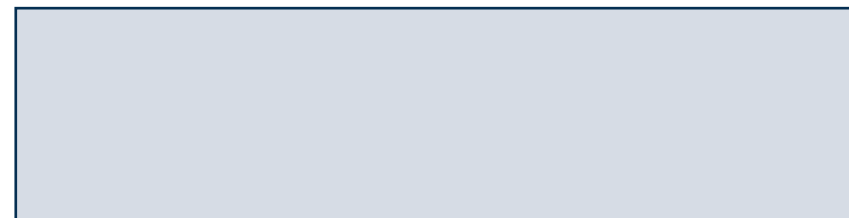


On which beach is the mean weight of the collected beach waste higher?

- a) On Bergen aan Zee the mean weight is the highest.
- b) On Katwijk aan Zee the mean weight is the highest.
- c) The mean weight for both is approximately the same.

Imagine a student makes the same systematic error of applying a histogram strategy to case-value plots.

Describe the possible gaze pattern the student may show. Do you expect this error to be used here? If so, why? If not, why not?



Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor the granting authority can be held responsible for them.

05 Case-Value Plot

Vignette on an error caused by an overgeneralization of a histogram approach

Introducing situation



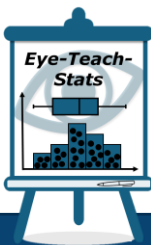
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Imagine you are teaching a 10th grade classroom on data interpretation, and you introduce your students to comparing data distributions using case-value plots.

You notice your student Aline is having difficulties with a task you've assigned. She gives an incorrect answer.

You will now see the task Aline is working on.

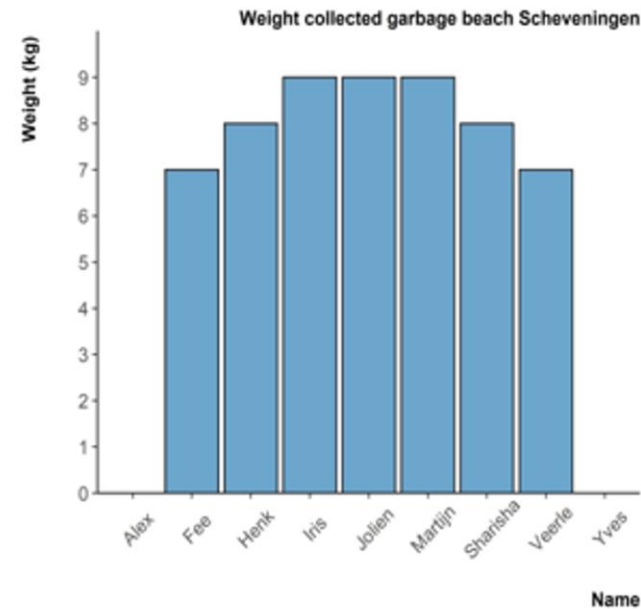
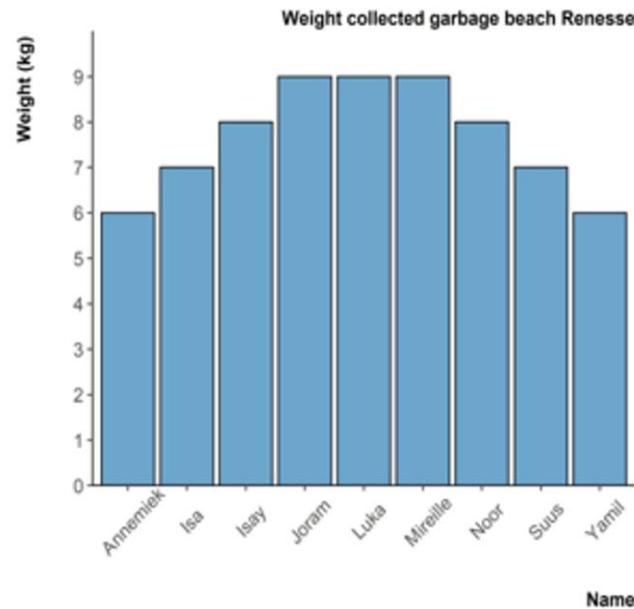


Student's task



Co-funded by
the European Union

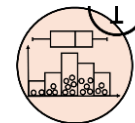
Erasmus+
Enriching lives, opening minds.



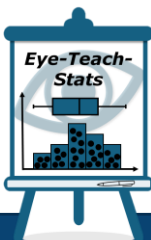
On which beach is the mean weight of the collected beach waste higher?

- a) On Renesse the mean weight is the highest.
- b) On Scheveningen the mean weight is the highest.
- c) The mean weight for both is approximately the same.

This is the task Aline is working on. **Please complete the task on your own first and enter your solution.**



- a) On Renesse the mean weight is the highest.
- b) On Scheveningen the mean weight is the highest.
- c) The mean weight for both is approximately the same.



Introducing eye-tracking

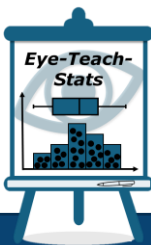


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

A group of researchers in mathematics education is pursuing the idea of supporting teachers with additional diagnostic material: To this end, eye movement data from students is recorded using eye-tracking technology and made available as supplementary material.

You will receive a short gaze video that shows the eye movements of your student Aline that were collected while she was working on the task. The gaze data indicates where her visual focus lies. Often it is possible to infer what Aline's thinking process might have been.



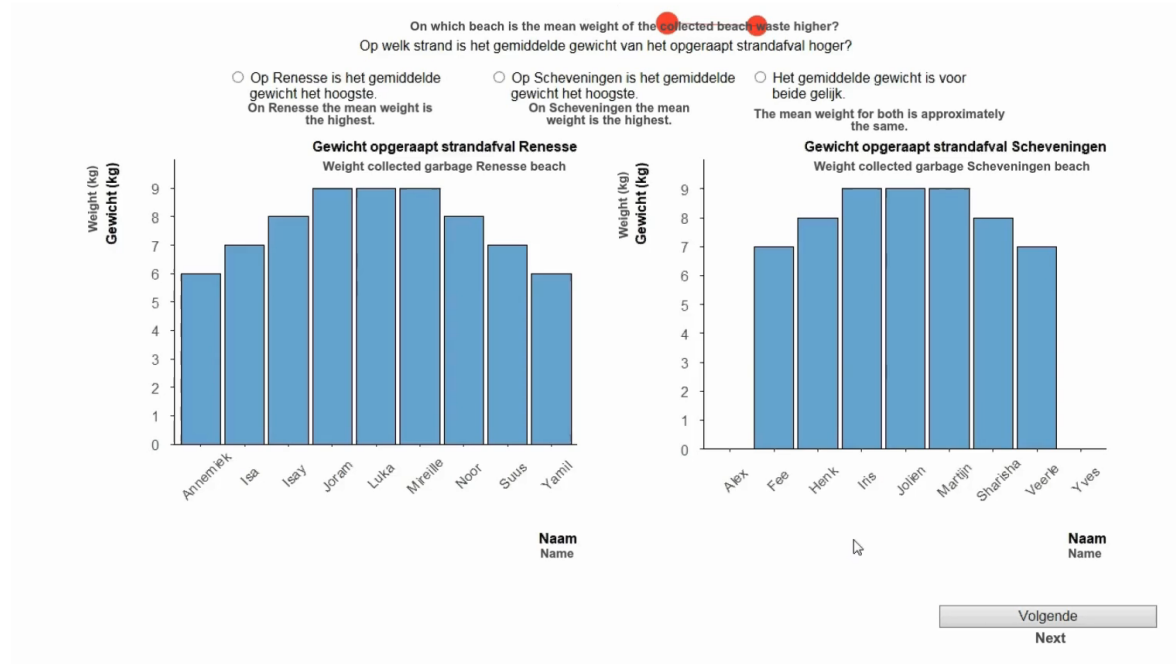
First look



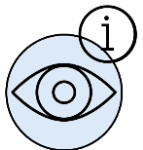
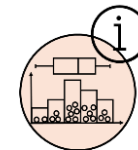
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Aline's Answer: "The mean weight for both is approximately the same."



Have a look at the video sequence that shows your student Aline's gaze movements and her answer.



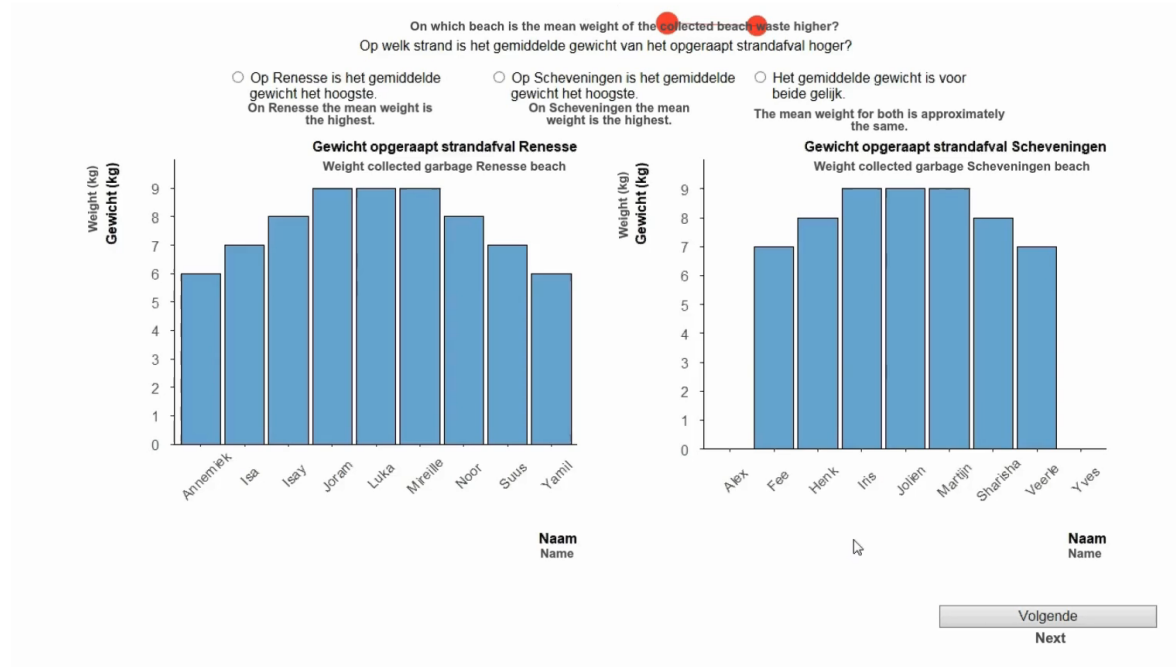
Take perspective



Co-funded by
the European Union

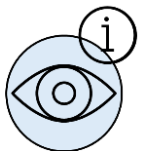
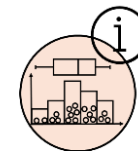
Erasmus+
Enriching lives, opening minds.

Aline's Answer: "The mean weight for both is approximately the same."



After the task Aline was shown her own gaze video. She explains retrospectively what she was thinking while working on the task.

Try to take Aline's perspective, and formulate what she may comment.



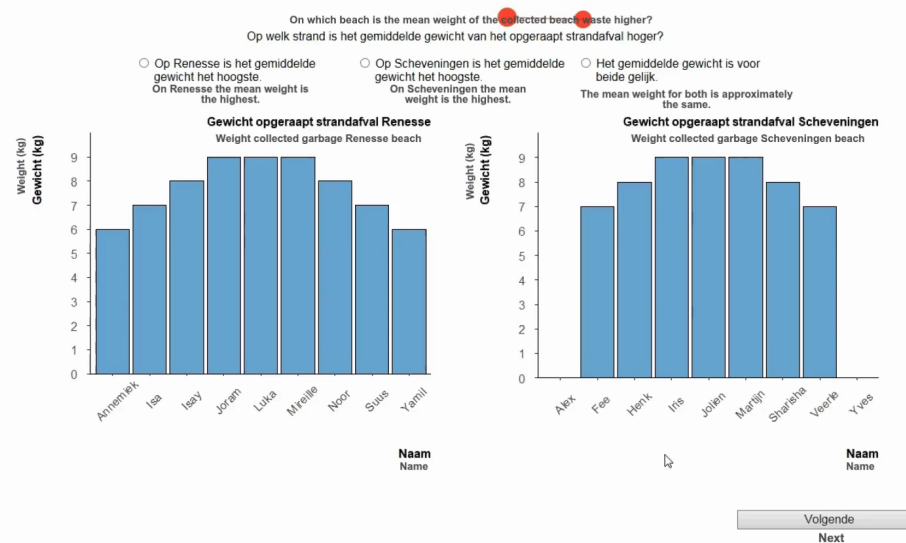
Indicate assumption



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

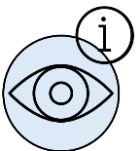
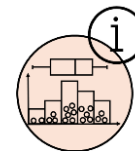
Aline's comment: "When I compared the graphs of the weights collected from the two beaches, I thought the mean weights were pretty much the same. I looked at the graphs and tried to see if the weights balanced out on both sides. I figured that if the really heavy and really light weights looked even on each end, they would cancel each other out, so the means would be about the same. I saw that the bars on both ends of the graphs were the same distance from the middle, so I thought they balanced each other out."



Now you can read the original comment of Aline.

Indicate to what extent Aline's comment agrees with what you assumed.

- ☐ No alignment at all
- ☐ Little alignment
- ☐ Quite some alignment
- ☐ Complete alignment



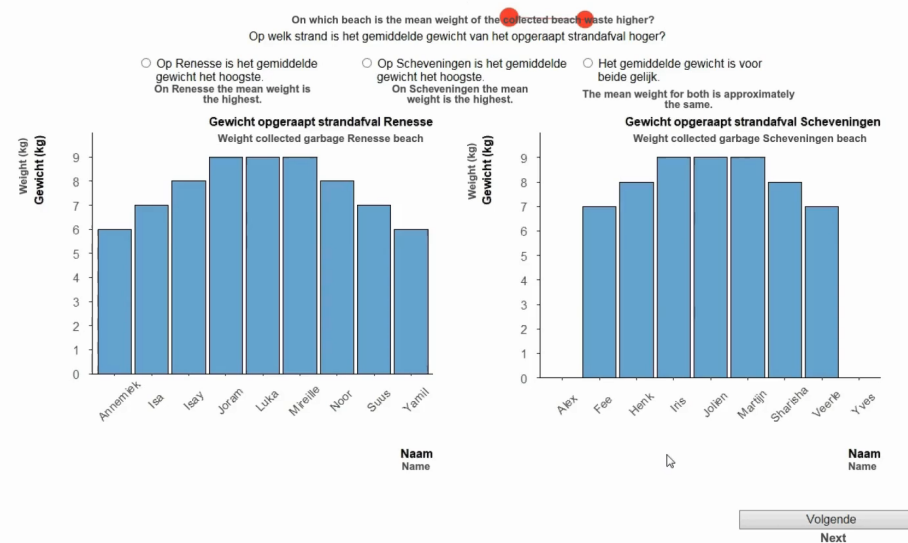
Evaluate solution



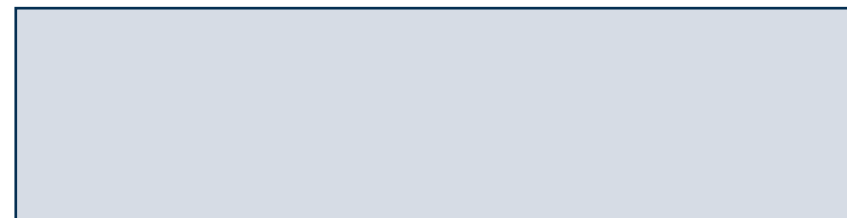
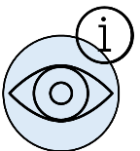
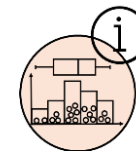
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Aline's comment: "When I compared the graphs of the weights collected from the two beaches, I thought the mean weights were pretty much the same. I looked at the graphs and tried to see if the weights balanced out on both sides. I figured that if the really heavy and really light weights looked even on each end, they would cancel each other out, so the means would be about the same. I saw that the bars on both ends of the graphs were the same distance from the middle, so I thought they balanced each other out."



Evaluate Aline's solution. What can you say about her understanding?



Introducing systematic error



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Understanding student errors:

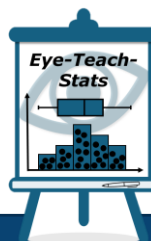
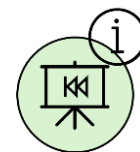
Aline has an underlying conceptual difficulty regarding the interpretation of case-value plots. This leads to an observable systematic error, called „**overgeneralization of a histogram approach**”.

Observable behavior: Aline said that the weights [bars] balanced out and cancel each other out in each graph. Her gaze data show that she compared similar bars and paid attention to the zero bars as well. Together this observable behavior points at a systematic error, which we term overgeneralization of a histogram approach when she interprets case-value plots.

Interpretation by the teacher/researcher: It appears that Aline applied a reasoning typically used for histograms—searching for a balance point and visually assessing the distribution shape and symmetry—and interpreted this as similar thus the mean is approximately the same, despite the two zero bars in one graph. Such a strategy is appropriate for histograms, which display aggregated data in bins and where the point on the horizontal axis where the graph is in balance is indeed the mean, but was here incorrectly transferred to case-value plots, where each bar represents only one individual data point.

Underlying conceptual difficulty: The systematic error is likely due to an underlying conceptual difficulty with the relation between the concept of distribution and the concept of center; therefore, in understanding how the mean is determined and compared in case-value plots. In a case-value plot, changing a bar's height is changing a measured value which in turn directly influences the mean. In histograms, doubling the heights of all bars (frequencies) does not change the mean as the overall distribution remains the same, hence the balance, thus the mean, remains the same.

Alternative Strategies and Correct Approaches for Case-value Plots: There are different ways students might approach case-value plots. For instance, some students use "area" as an approach for estimating the mean (more area is a higher mean). Although this can be correct in many case-value plots, it is an incomplete strategy (e.g., when there are more data points in one graph compared to another). A correct method for estimating the mean in case-value plots involves understanding that each bar's height is the measured value. Estimation of the mean could, for example, be done by visually drawing a horizontal line that would make all bars equally high. In histograms, bar heights represent frequencies, and the balance point of the graph along the horizontal axis approximates the mean.

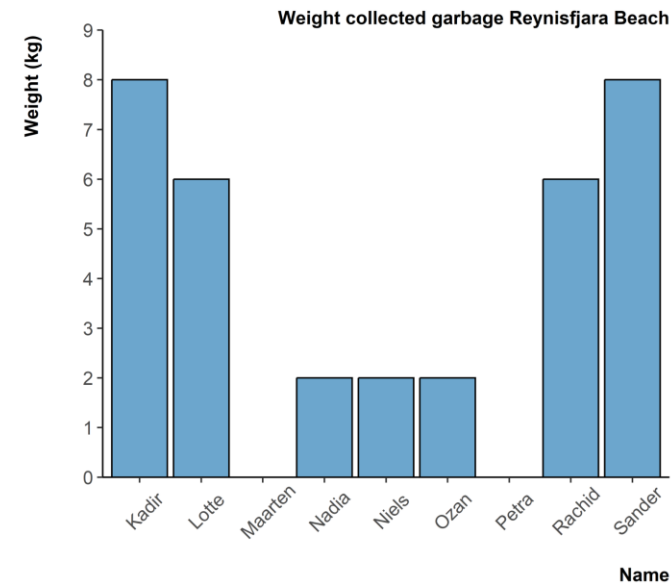
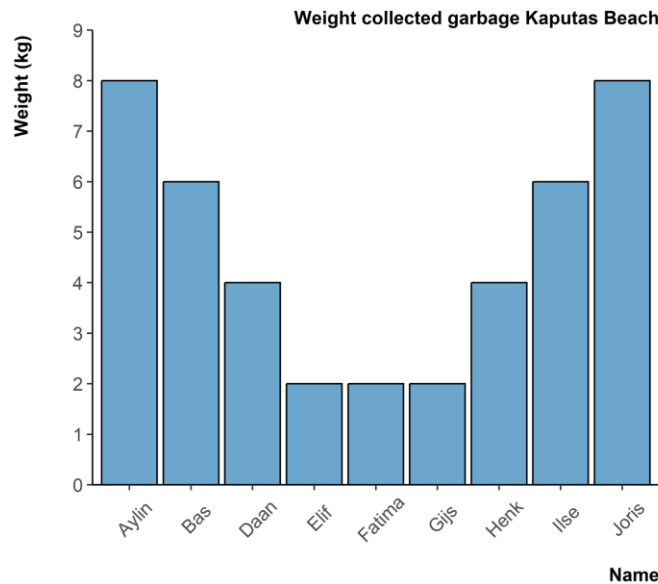


Practice



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

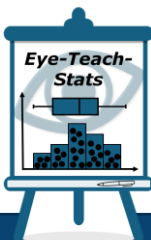
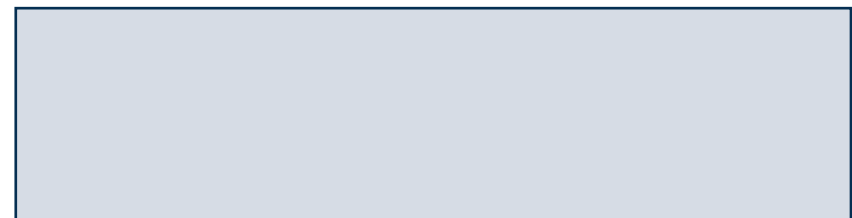


On which beach is the mean weight of the collected beach waste higher?

- a) On Kaputas the mean weight is the highest.
- b) On Reynisfjara the mean weight is the highest.
- c) The mean weight for both is approximately the same.

Imagine a student makes the same systematic error of applying a histogram strategy to case-value plots.

Describe the possible gaze pattern the student may show. Do you expect this error to be used here? If so, why? If not, why not?





Eye-Teach-Stats Eye-Tracking Vignettes © 2026



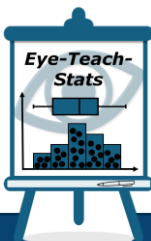
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor the granting authority can be held responsible for them.

06 Dotplot

Vignette on an error caused by pairwise stack comparison



Introducing situation



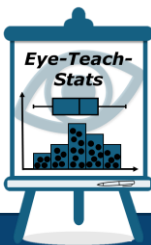
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Imagine you are teaching a 6th-grade classroom on data interpretation, and you introduce your students to comparing data distributions using dotplots.

You notice your student Zeyd is having difficulties with a task you've assigned. He gives an incorrect answer.

You will now see the task Zeyd is working on.



Student's task



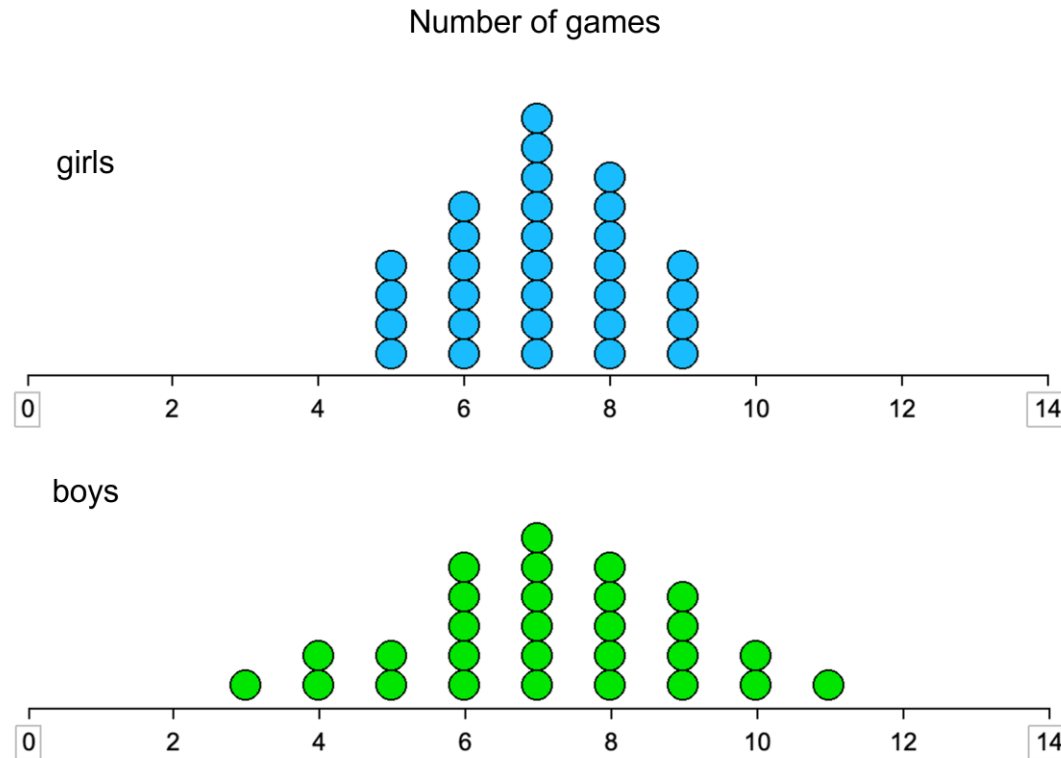
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

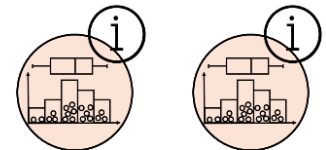
Do girls and boys differ in the number of games they have on their cell phones? A survey on this question was done at the Park School with 30 girls and 28 boys. Compare the two samples and then evaluate:

The data from the samples fits the assumption that

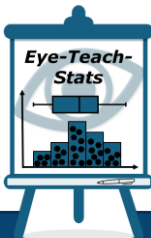
- a) Girls have a lot more games on their phones.
- b) Boys have a lot more games on their phones.
- c) Girls and boys have about the same number of games on their phone.



This is the task Zeyd is working on. Please complete the task on your own first and enter your solution.



- a) Girls have a lot more games on their phones.
- b) Boys have a lot more games on their phones.
- c) Girls and boys have about the same number of games on their phones.



Introducing eye-tracking

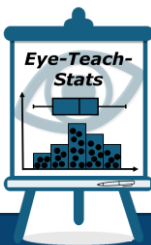


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

A group of researchers in mathematics education is pursuing the idea of supporting teachers with additional diagnostic material: To this end, eye movement data from students is recorded using eye-tracking technology and made available as supplementary material.

You will receive a short gaze video that shows the eye movements of your student Zeyd that were collected while he was working on the task. This gaze video indicates where his visual focus lies. Often it is possible to infer what Zeyd's thinking process might have been.



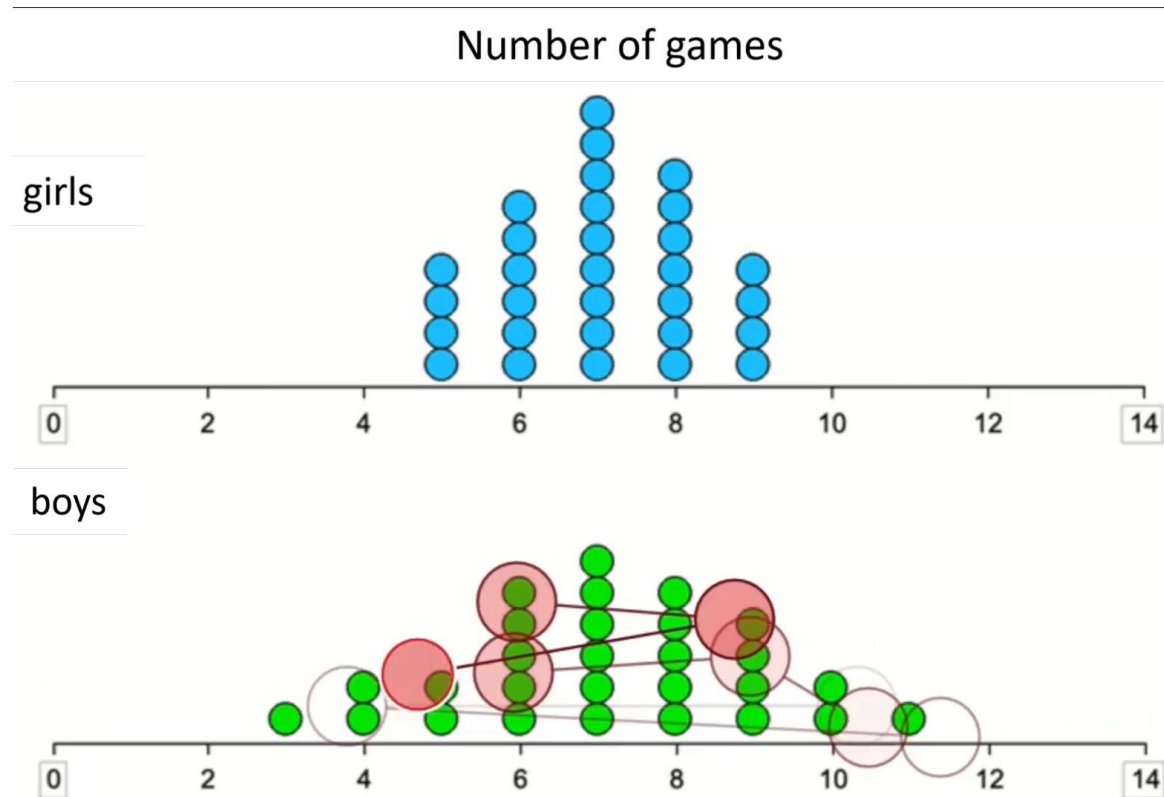
First look



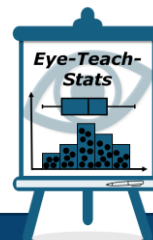
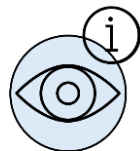
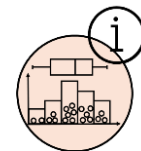
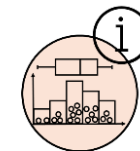
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Zeyd's answer: „I would say a). The girls have a lot more games on their phones.”



Have a look at the video sequence that shows your student Zeyd's gaze movements and his answer.



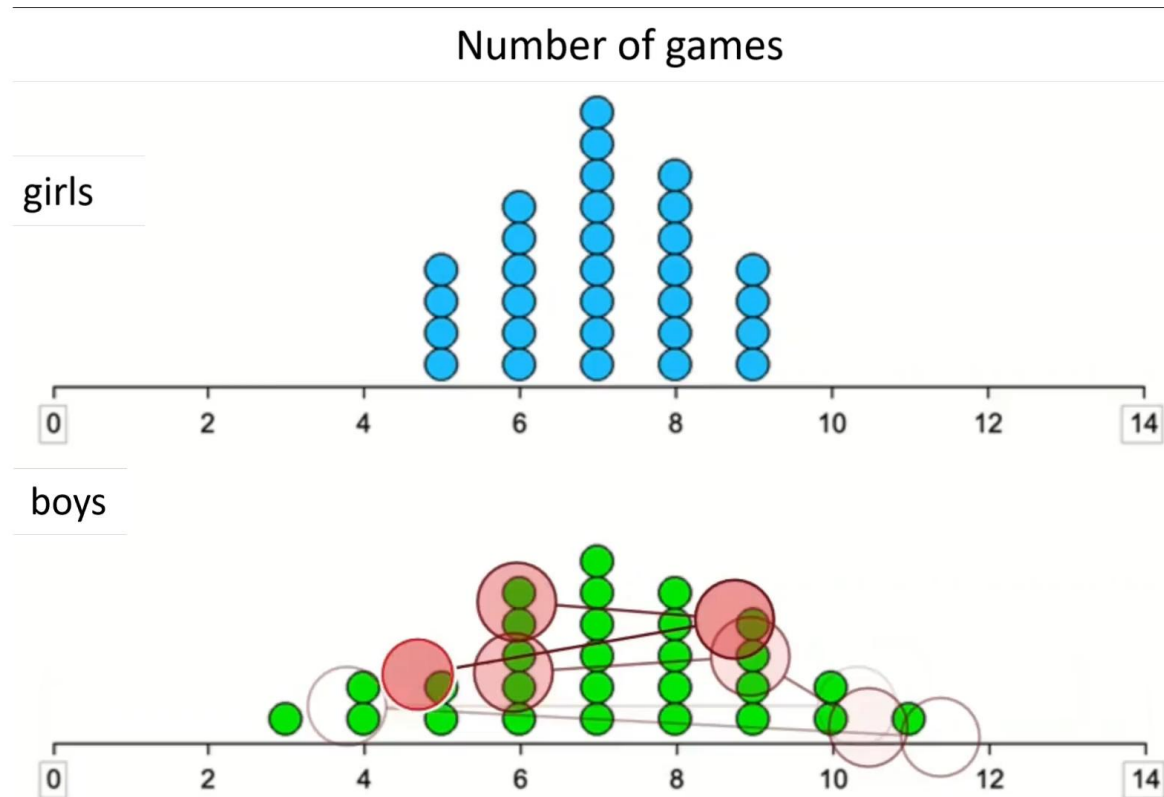
Take perspective



Co-funded by
the European Union

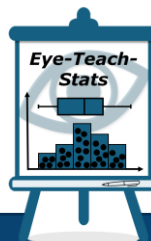
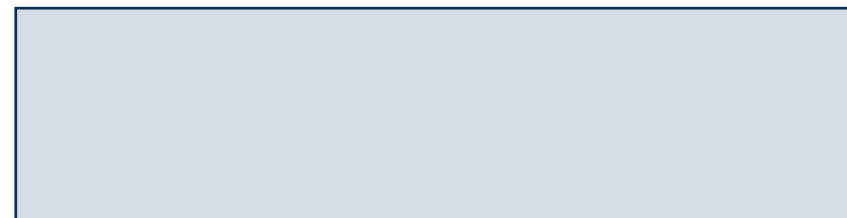
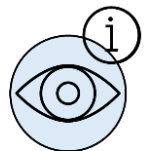
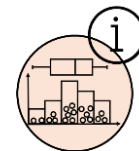
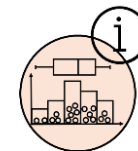
Erasmus+
Enriching lives, opening minds.

Zeyd's answer: „I would say a). The girls have a lot more games on their phones.”



After the task Zeyd was shown his own gaze video. He explains retrospectively what he was thinking while working on the task.

Try to take Zeyd's perspective, and formulate what he may comment.



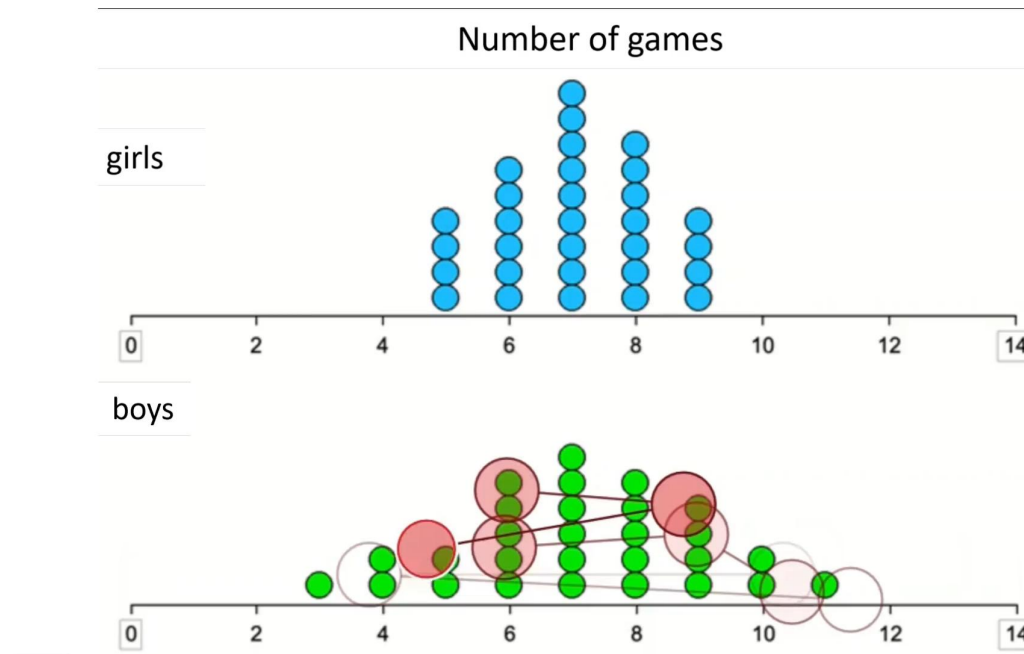
Indicate assumption



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

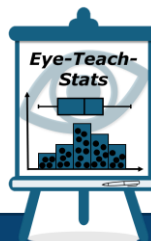
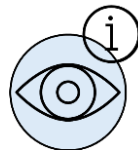
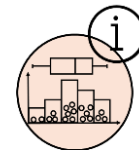
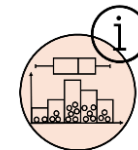
Zeyd's comment: „So I looked at how many dots there were and then I compared them. For example, boys and girls have the same number of dots at nine, but the girls have more dots at seven and eight and so on. The little stacks at the outside from the boys is not so important. That's why I think the girls have more games.“



Now you can read the original comment of Zeyd.

Indicate to what extent Zeyd's comment agrees with what you assumed.

- ☐ No alignment at all
- ☐ Little alignment
- ☐ Quite some alignment
- ☐ Complete alignment



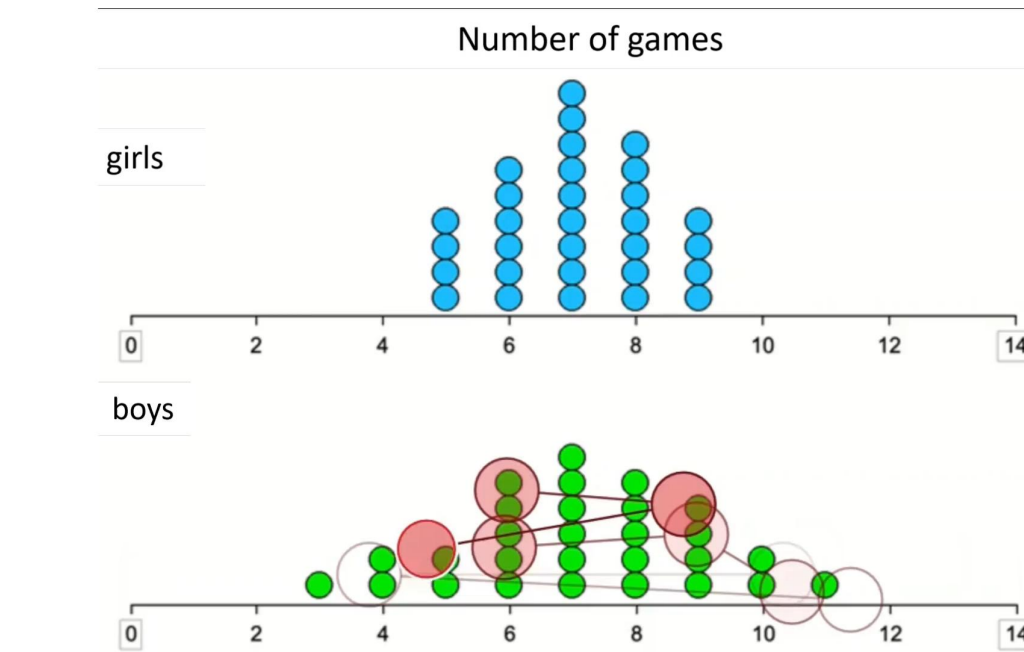
Evaluate solution



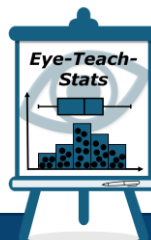
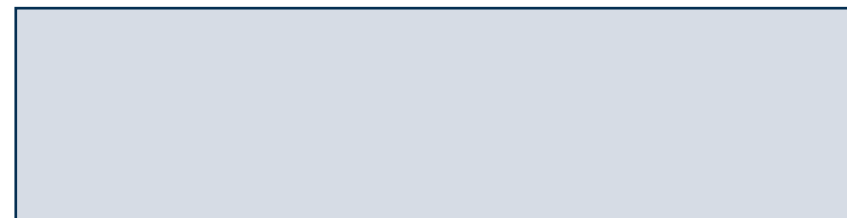
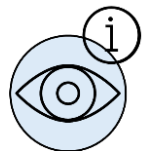
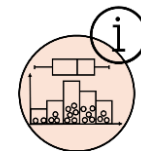
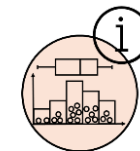
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Zeyd's comment: „So I looked at how many dots there were and then I compared them. For example, boys and girls have the same number of dots at nine, but the girls have more dots at seven and eight and so on. The little stacks at the outside from the boys is not so important. That's why I think the girls have more games.“



Evaluate Zeyd's solution. What can you say about his understanding?



Introducing systematic error



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

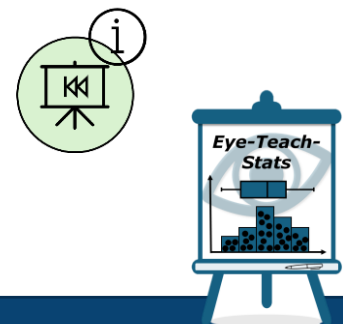
Understanding student errors:

Zeyd has an underlying conceptual difficulty regarding the interpretation of dotplots. This leads to an observable systematic error, called “**pairwise stack comparison**”.

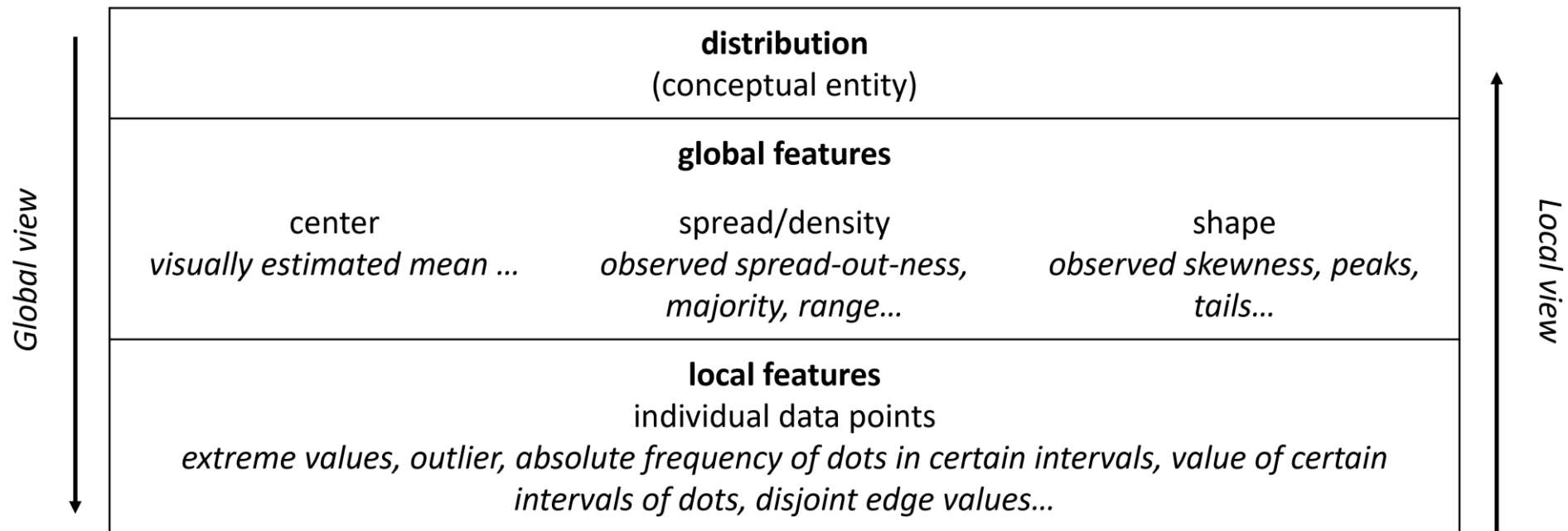
Interpreting behavior: From the gaze videos and his comment, we can see that Zeyd’s gaze moves from one distribution to the other focusing on comparable stacks mostly. He counts the number of dots as individual data points and compares the stacks with each other. He focuses on the distribution as a collection of individual data points.

Understanding error and underlying conceptual difficulty: However, when comparing samples of different sizes, the frequencies must be set in relation to each other. The distribution should not be seen merely as a collection of individual data points; global characteristics should also be taken into account. The systematic error occurs in this inappropriate use of a local view. 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.

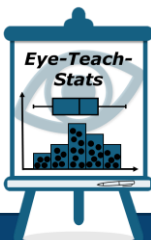
On the next slide, you will see an overview of local and global features that can be taken into account when comparing distributions.



Perspectives on data distribution: local vs. global view



Adapted according to Schreiter & Vogel, 2023



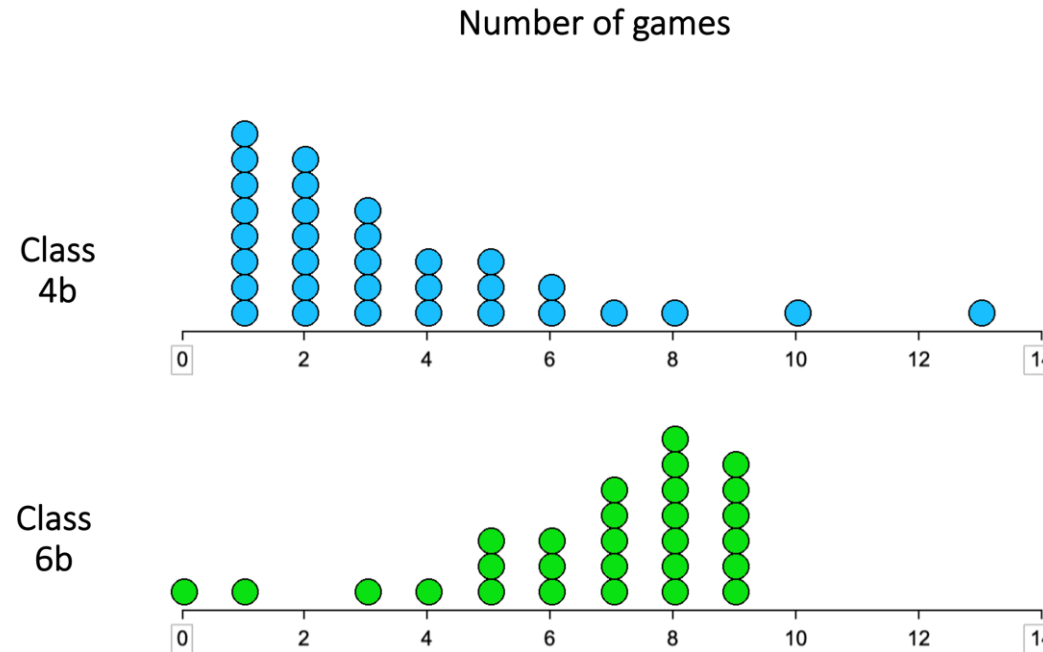
Practice



Co-funded by
the European Union

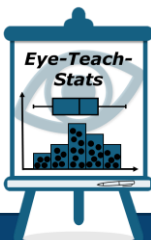
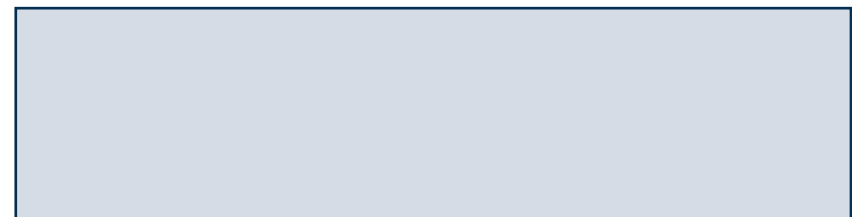
Erasmus+
Enriching lives, opening minds.

Do 4th grader and 6th grader differ in the number of games they have on their smartphones? A survey on this question was done at the Park School with 32 4th grader and 28 6th grader. Compare the two samples and then evaluate:



Imagine a student makes the same systematic error of pairwise stack comparison.

Describe the possible gaze pattern the student may show. Would the strategy lead to a correct answer? Why, or why not?





Eye-Teach-Stats Eye-Tracking Vignettes © 2026



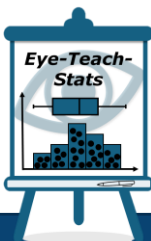
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor the granting authority can be held responsible for them.

07 Histogram

Vignette on an error caused by pairwise stack comparison



Introducing situation



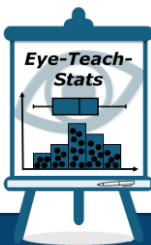
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Imagine you are teaching a 10th grade class on data interpretation, and you introduce your students to comparing data distributions using histograms.

You notice your student Sem is having difficulties with a task you've assigned. He gives an incorrect answer.

You will now see the task Sem is working on.

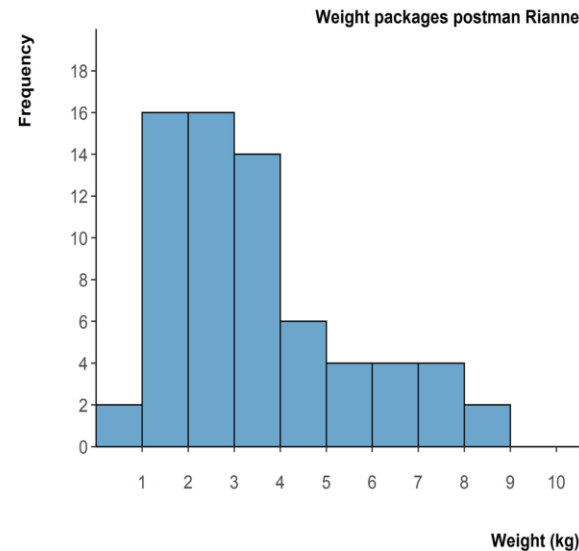
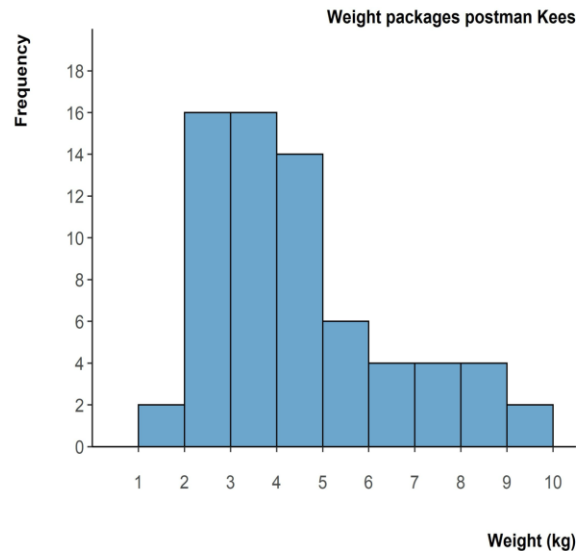


Student's task



Co-funded by
the European Union

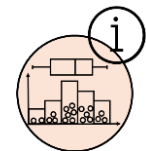
Erasmus+
Enriching lives, opening minds.



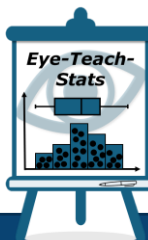
Which postal worker delivers on average heaviest packages?

- a) Kees delivers on average the heaviest packages.
- b) Rianne delivers on average the heaviest packages.
- c) The mean weight for both is approximately the same.

This is the task Sem is working on. Please complete the task on your own first and enter your solution.



- a) Kees delivers on average the heaviest packages.
- b) Rianne delivers on average the heaviest packages.
- c) The mean weight for both is approximately the same.



Introducing eye-tracking

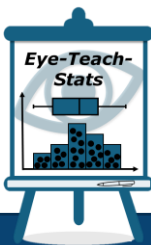


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

A group of researchers in mathematics education is pursuing the idea of supporting teachers with additional diagnostic material: To this end, eye movement data from students is recorded using eye-tracking technology and made available as supplementary material.

You will receive a short gaze video that shows the eye movements of your student Sem that were collected while he was working on the task. The gaze data indicates where his visual focus lies. Often it is possible to infer what Sem's thinking process might have been.



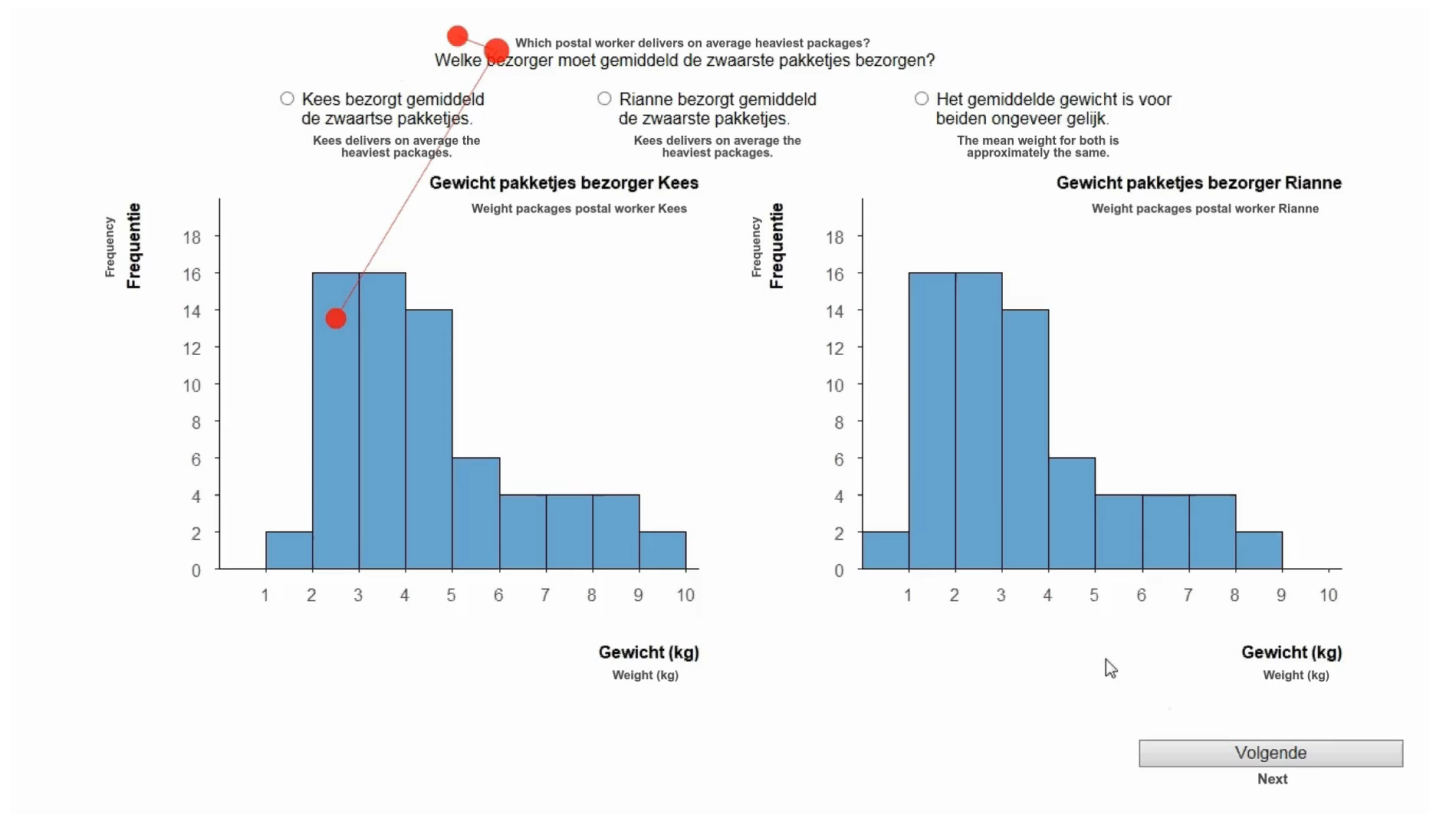
First look



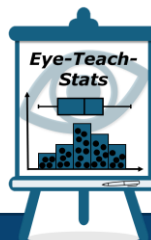
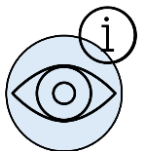
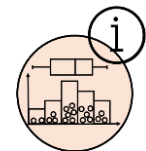
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Sem's answer: „The mean weight for both is approximately the same.”



Have a look at the video sequence that shows your student Sem's gaze and his answer.



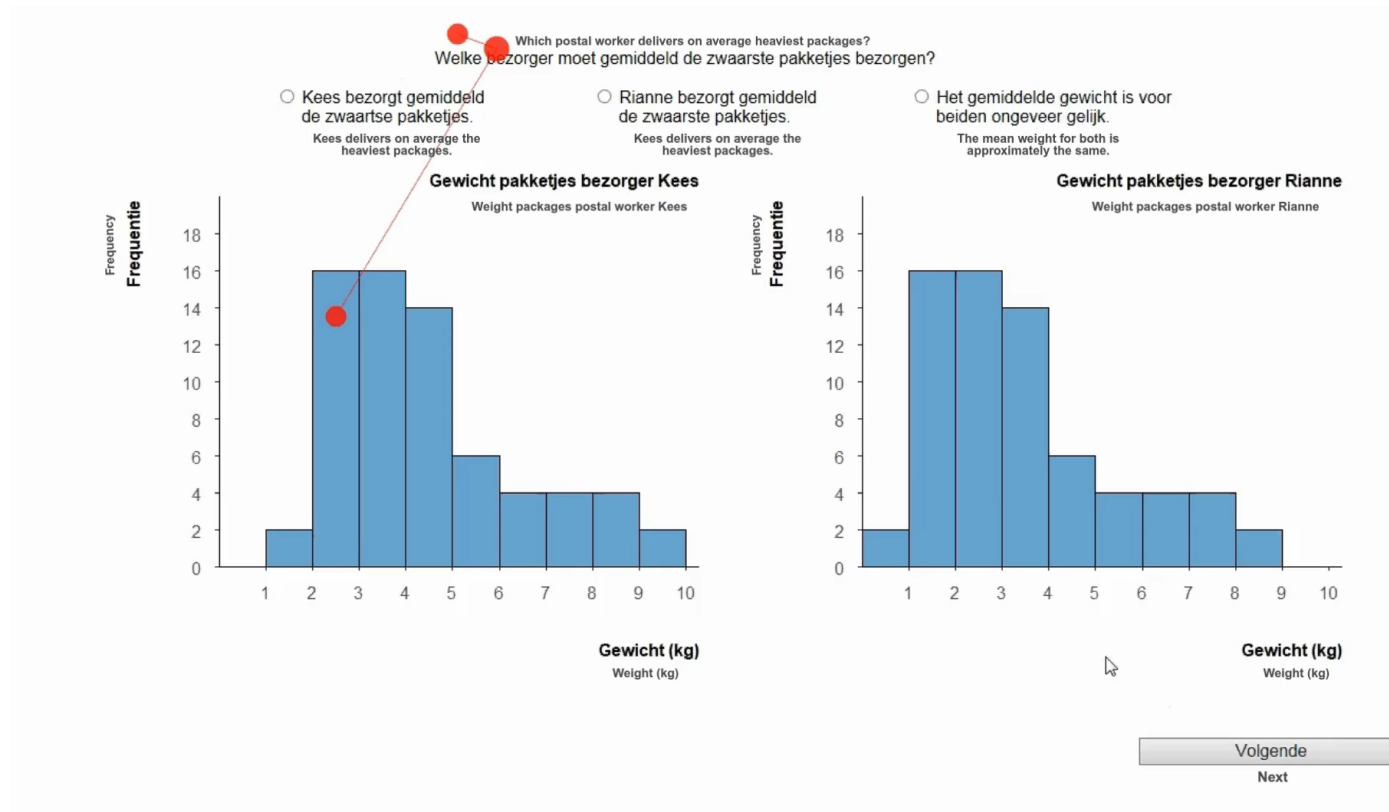
Take perspective



Co-funded by
the European Union

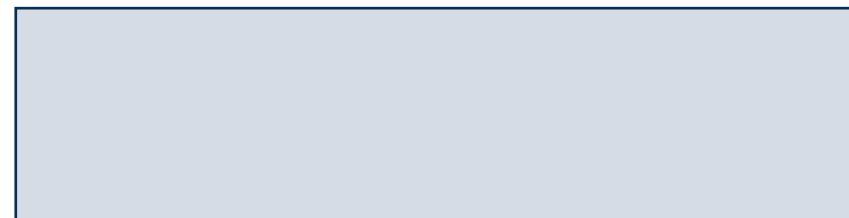
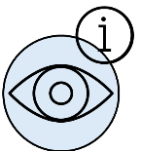
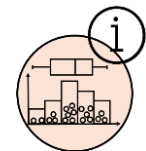
Erasmus+
Enriching lives, opening minds.

Sem's answer: „The mean weight for both is approximately the same.”



After the task Sem was shown his own gaze video. He explains retrospectively what he was thinking while working on the task.

Try to take Sem's perspective and formulate what he may comment.



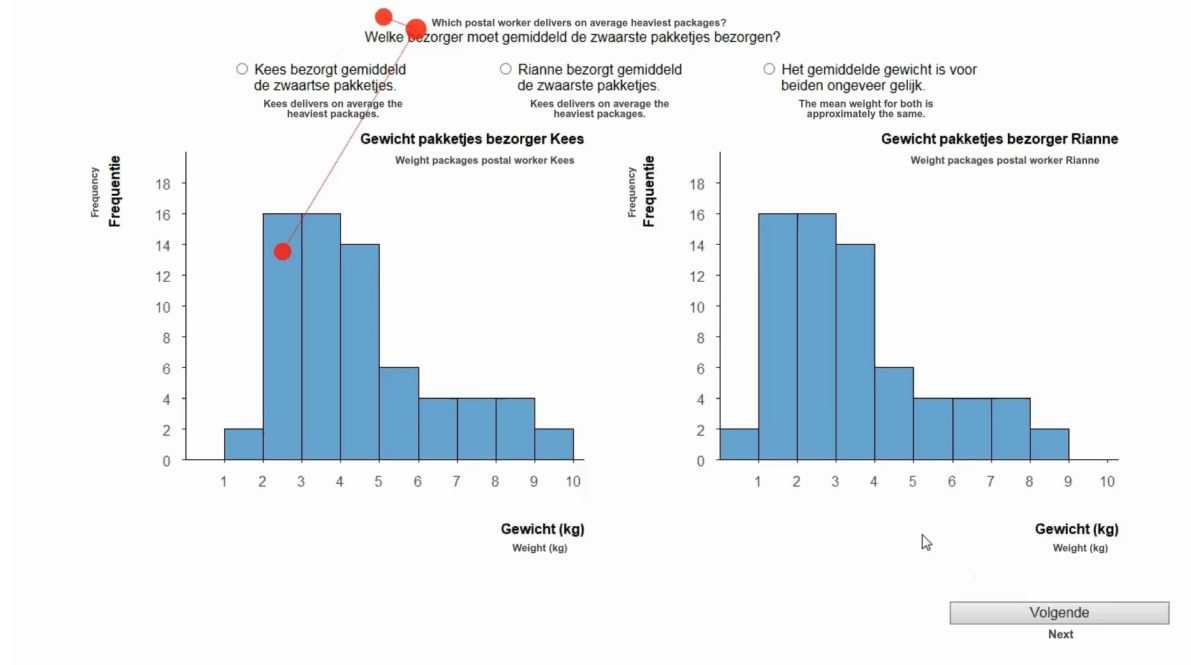
Indicate assumption



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Sem's Comment: "I saw that Kees had heavier packages with fewer light ones, and Rianne had mostly lighter packages with fewer heavy ones. But the distributions were similar. I compared each pair of stacks across the histograms, and it looked like the differences in stacks could cancel each other out, so I think the approximate means were the same."

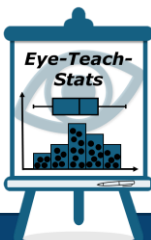
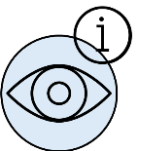
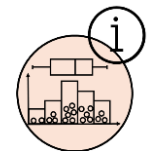


Now you can read the original comment of Sem.

Indicate to what extent Sem's comment agrees with what you assumed.

- ☐ No alignment at all
- ☐ Little alignment

- ☐ Quite some alignment
- ☐ Complete alignment



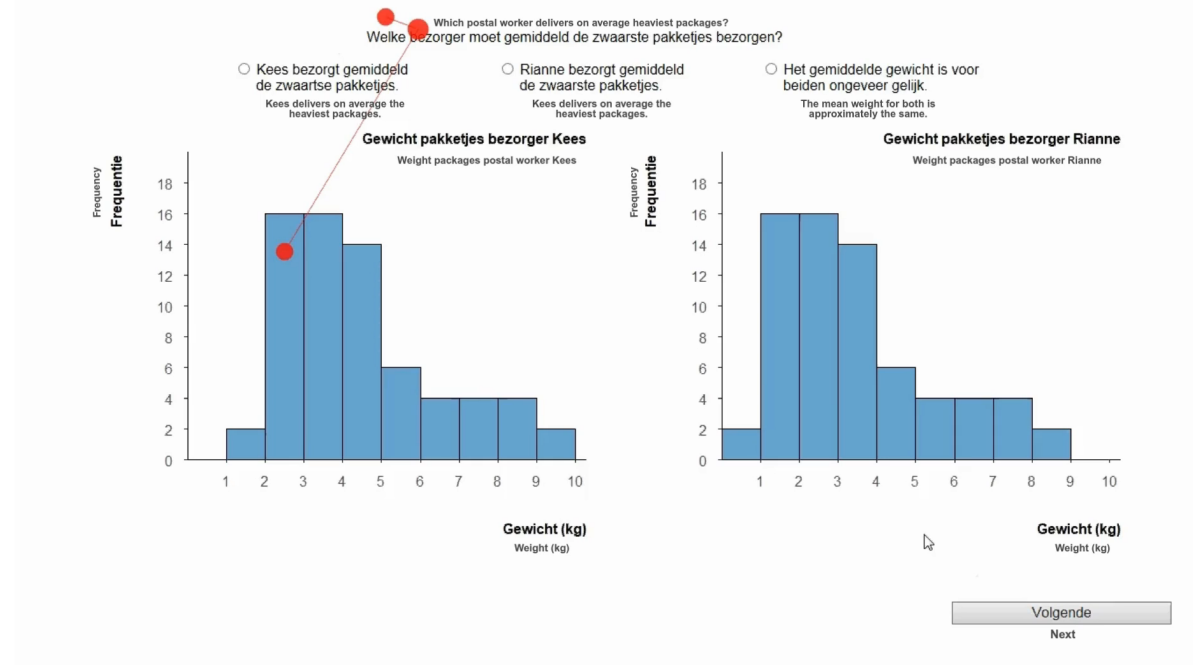
Evaluate solution



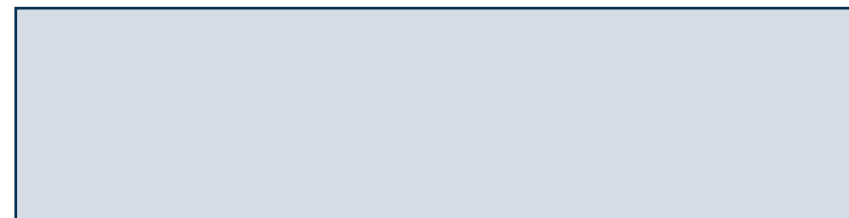
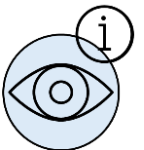
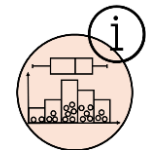
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Sem's Comment: "I saw that Kees had heavier packages with fewer light ones, and Rianne had mostly lighter packages with fewer heavy ones. But the distributions were similar. I compared each pair of stacks across the histograms, and it looked like the differences in stacks could cancel each other out, so I think the approximate means were the same."



Evaluate Sem's solution. What can you say about his understanding?



Introducing systematic error



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

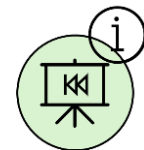
Understanding student errors

Sem has an underlying conceptual difficulty regarding the interpretation of histograms. This leads to an observable systematic error, called **pairwise stack comparison**.

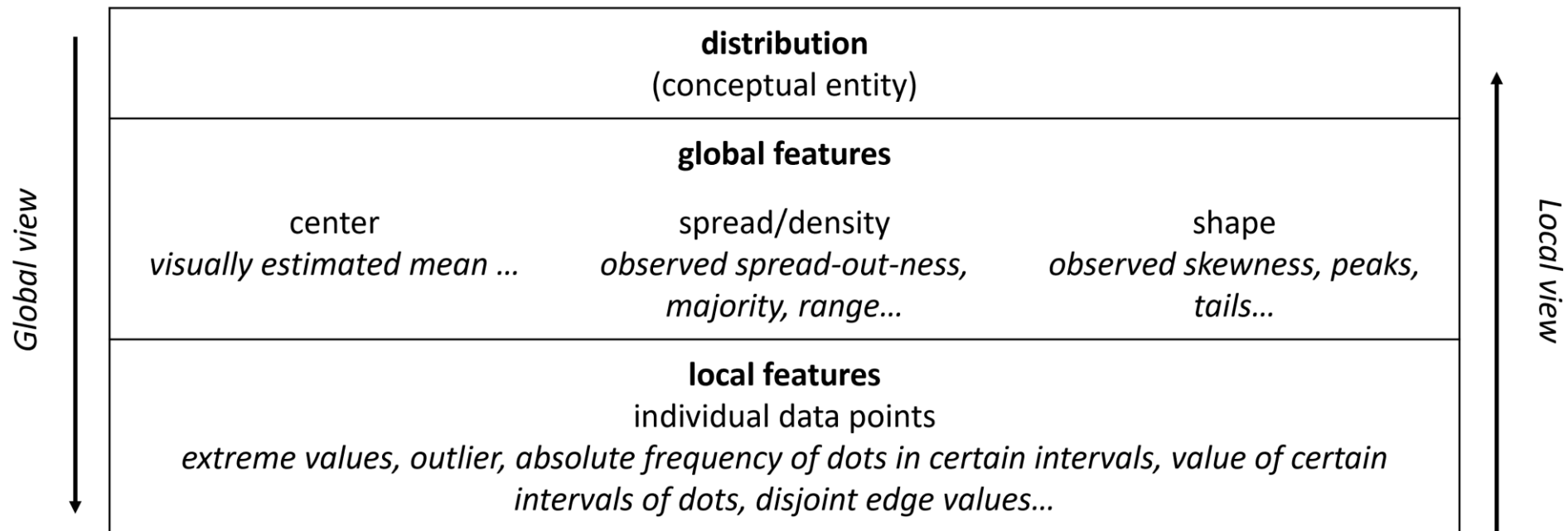
Observable behavior: Sem explained that he compared the stacks between both histograms and assumed the differences would cancel out, leading him to conclude the means were similar. We can observe that Sem's gaze moved from one stack in a histogram to a comparable stack in the other histogram. This together indicates an observable systematic error, which we term "overemphasizing local features: a pairwise stack comparison" in the interpretation of histograms.

Interpretation by the teacher/researcher: It appears that Sem engaged in a pairwise comparison of individual stacks across the two histograms. His gaze pattern reflects a focused comparison of individual bars rather than considering the histograms as wholes. In a histogram, data is aggregated into bins and comparing individual bars without considering the overall distribution leads to flawed conclusions about the approximate means. Sem ignored that the shift to the right (of the left-hand histogram) leads to a higher arithmetic mean, hence, he did not consider the global shift in the distributions. However, interpreting the mean accurately requires evaluating the overall distribution and shifts of the data rather than isolated bar comparisons.

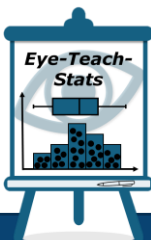
Underlying conceptual difficulty: Sem's systematic error is likely due to an underlying conceptual difficulty with the concepts of data and distribution. He used local features (pairwise stack comparison) and ignored global features (the shift of the distribution) when interpreting histograms. A correct interpretation requires attention to the overall distribution and shifts in it rather than isolated bar comparisons.



Perspectives on data distribution: local vs. global view



Adapted according to Schreiter & Vogel, 2023

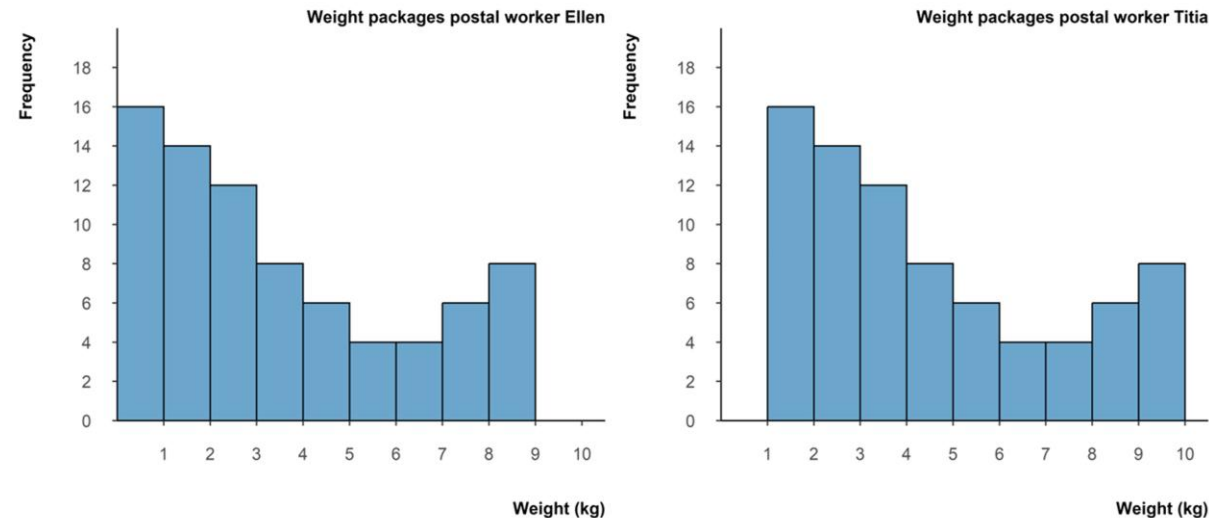


Practice



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

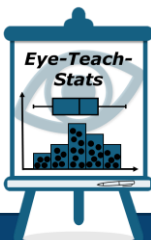
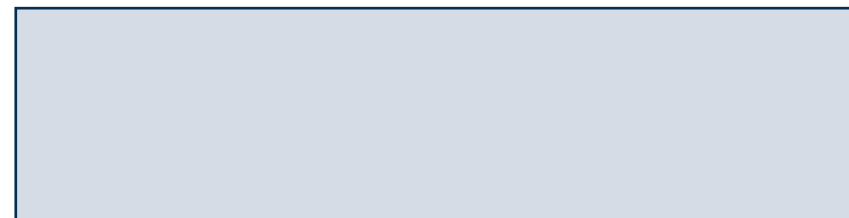


Which postal worker delivers on average the heaviest packages?

- a) Ellen delivers on average the heaviest packages.
- b) Titia delivers on average the heaviest packages.
- c) The mean weight is approximately the same for both.

Imagine a student makes the same systematic error by overemphasizing local features: pairwise stack comparison.

Describe the possible gaze pattern the student may show. Do you expect this error to be used here? If so, why? If not, why not?



Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor the granting authority can be held responsible for them.

08 Case-Value Plot

Vignette on an error caused by pairwise stack comparison

Introducing situation



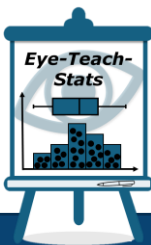
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Imagine you are teaching a 10th grade classroom on data interpretation, and you introduce your students to comparing data distributions using case-value plots.

You notice your student Eva is having difficulties with a task you've assigned. She gives an incorrect answer.

You will now see the task Eva is working on.

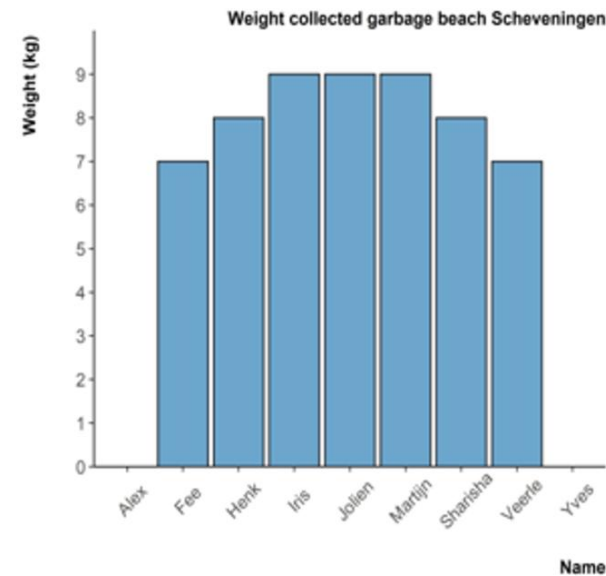
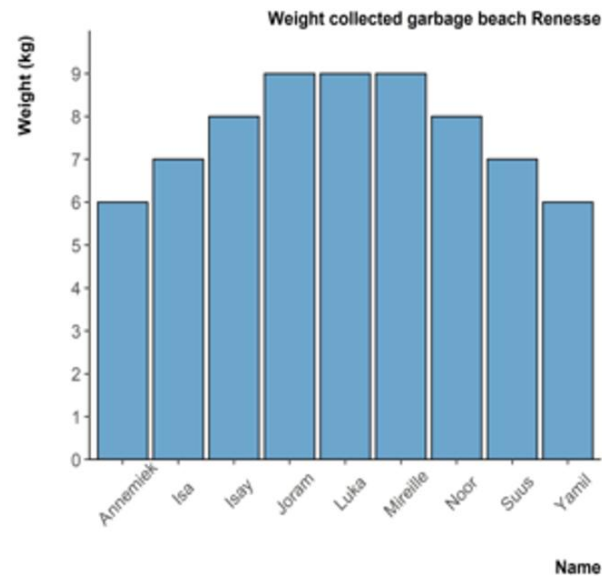


Student's task



Co-funded by
the European Union

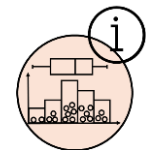
Erasmus+
Enriching lives, opening minds.



On which beach is the mean weight of the collected beach waste higher?

- a) On Renesse the mean weight is the highest.
- b) On Scheveningen the mean weight is the highest.
- c) The mean weight for both is approximately the same.

This is the task Eva is working on. **Please complete the task on your own first and enter your solution.**



- a) On Renesse the mean weight is the highest.
- b) On Scheveningen the mean weight is the highest.
- c) The mean weight for both is approximately the same.



Introducing eye-tracking

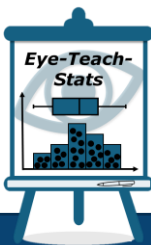


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

A group of researchers in mathematics education is pursuing the idea of supporting teachers with additional diagnostic material: To this end, eye movement data from students is recorded using eye-tracking technology and made available as supplementary material.

You will receive a short gaze video that shows the eye movements of your student Eva that were collected while she was working on the task. The gaze data indicates where her visual focus lies. Often it is possible to infer what Eva's thinking process might have been.



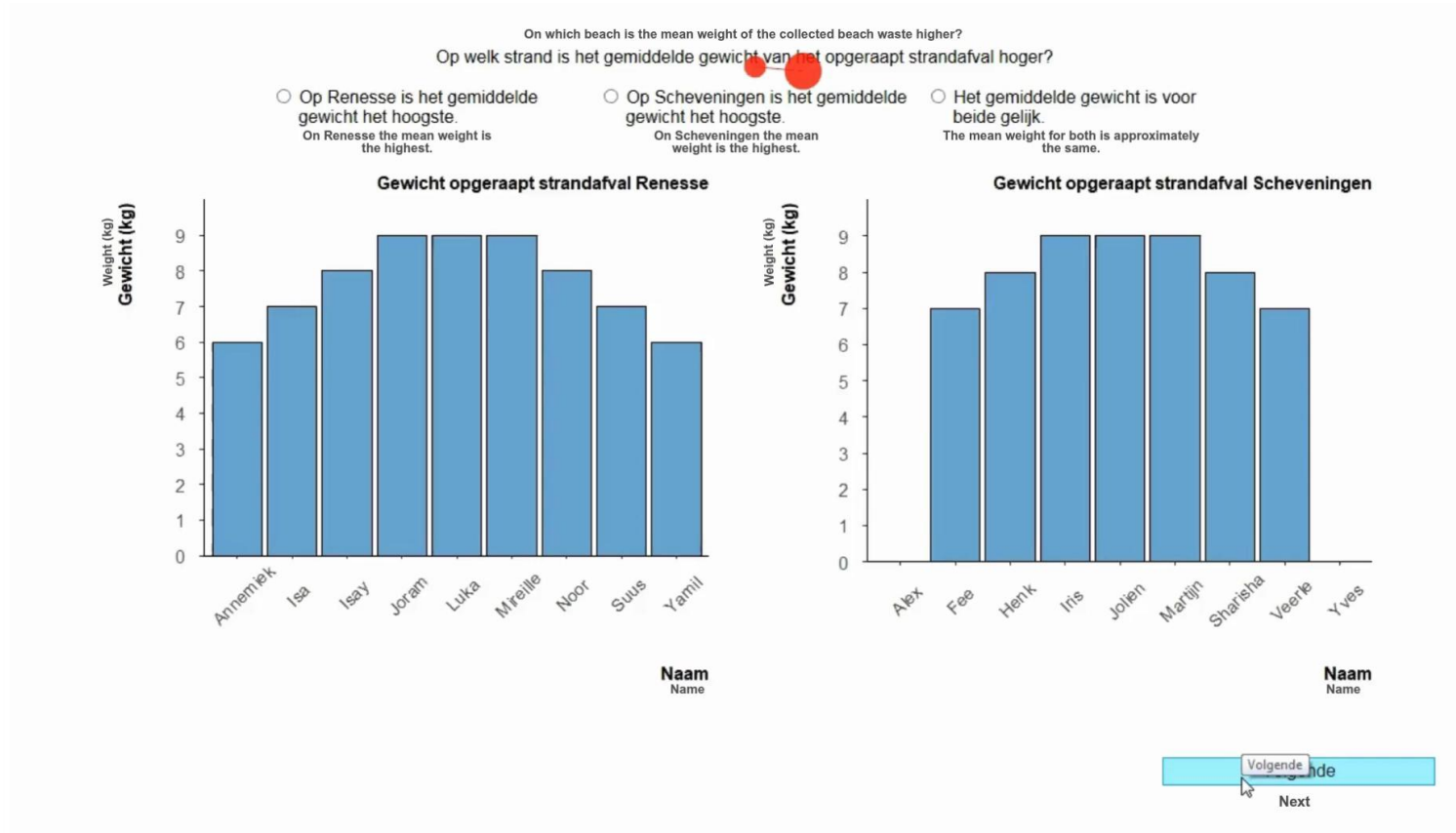
First look



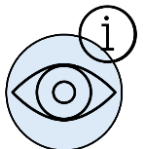
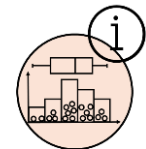
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Eva's answer: "On Scheveningen the mean weight is the highest."



Have a look at the video sequence that shows your student Eva's gaze movements and her answer.



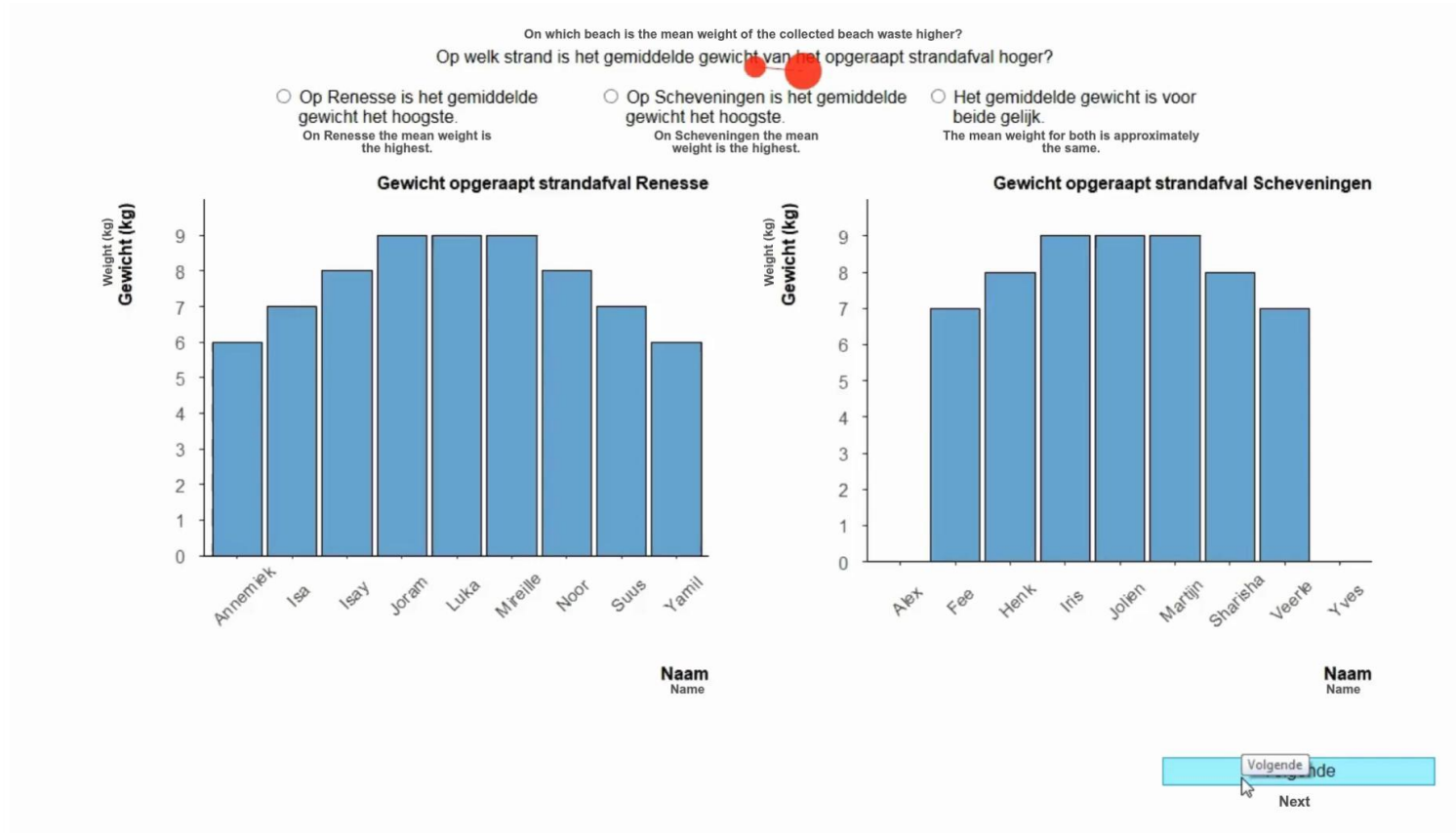
Take perspective



Co-funded by
the European Union

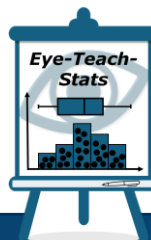
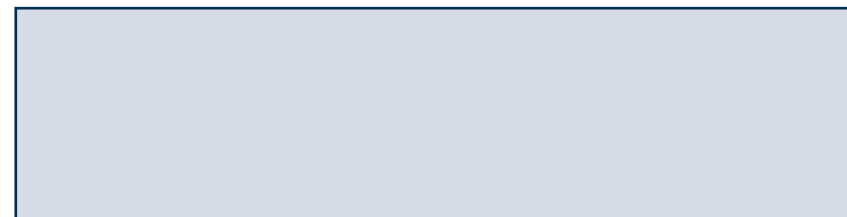
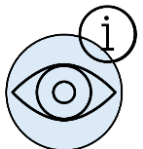
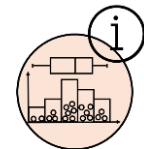
Erasmus+
Enriching lives, opening minds.

Eva's answer: "On Scheveningen the mean weight is the highest."



After the task Eva was shown her own gaze video. She explains retrospectively what she was thinking while working on the task.

Try to take Eva's perspective, and formulate what she may comment.



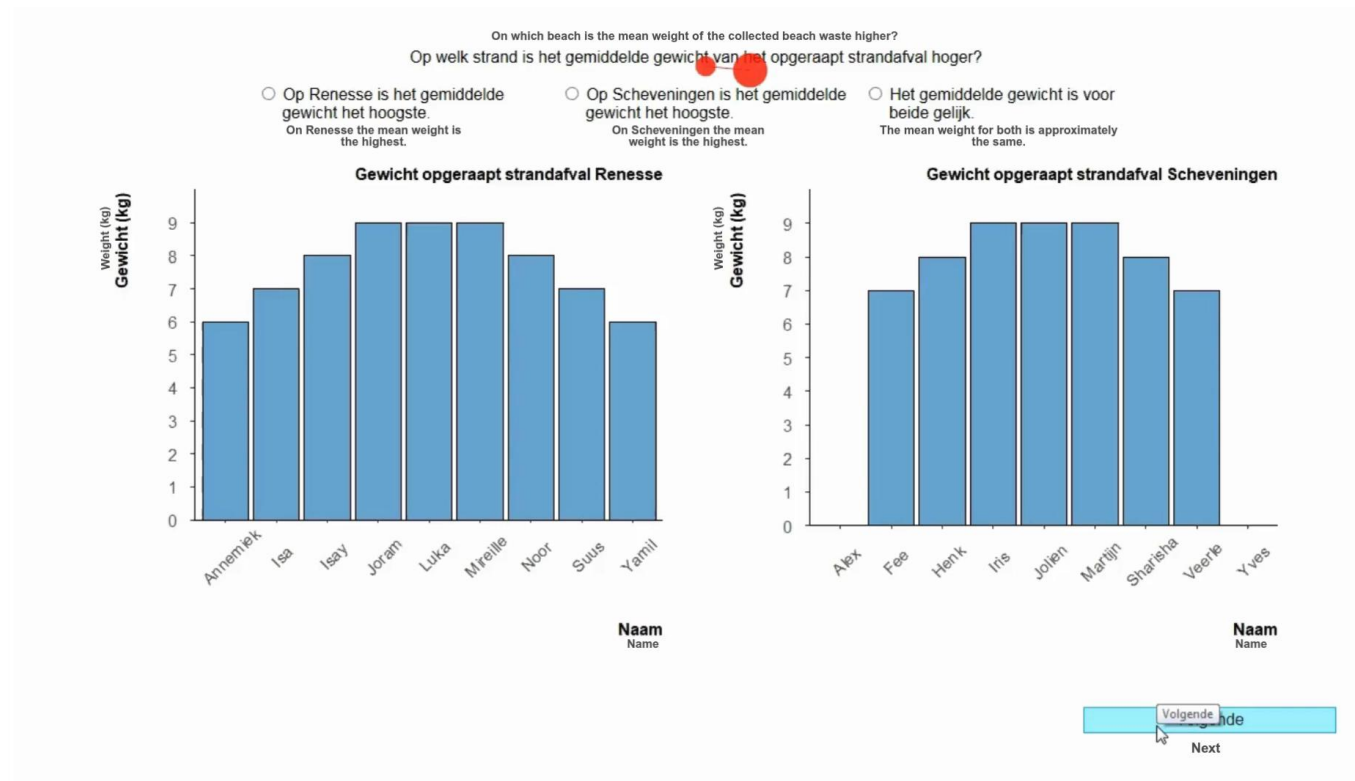
Indicate assumption



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

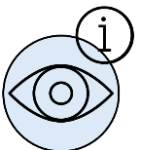
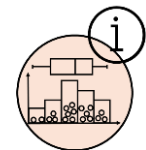
Eva's comment: "I compared some bars from each graph one by one and saw that Renesse had more bars with lower values. So, I thought Scheveningen must have a higher beach waste weight."



Now you can read the original comment of Eva.

Indicate to what extent Eva's comment agrees with what you assumed.

- ☐ No alignment at all
- ☐ Little alignment
- ☐ Quite some alignment
- ☐ Complete alignment



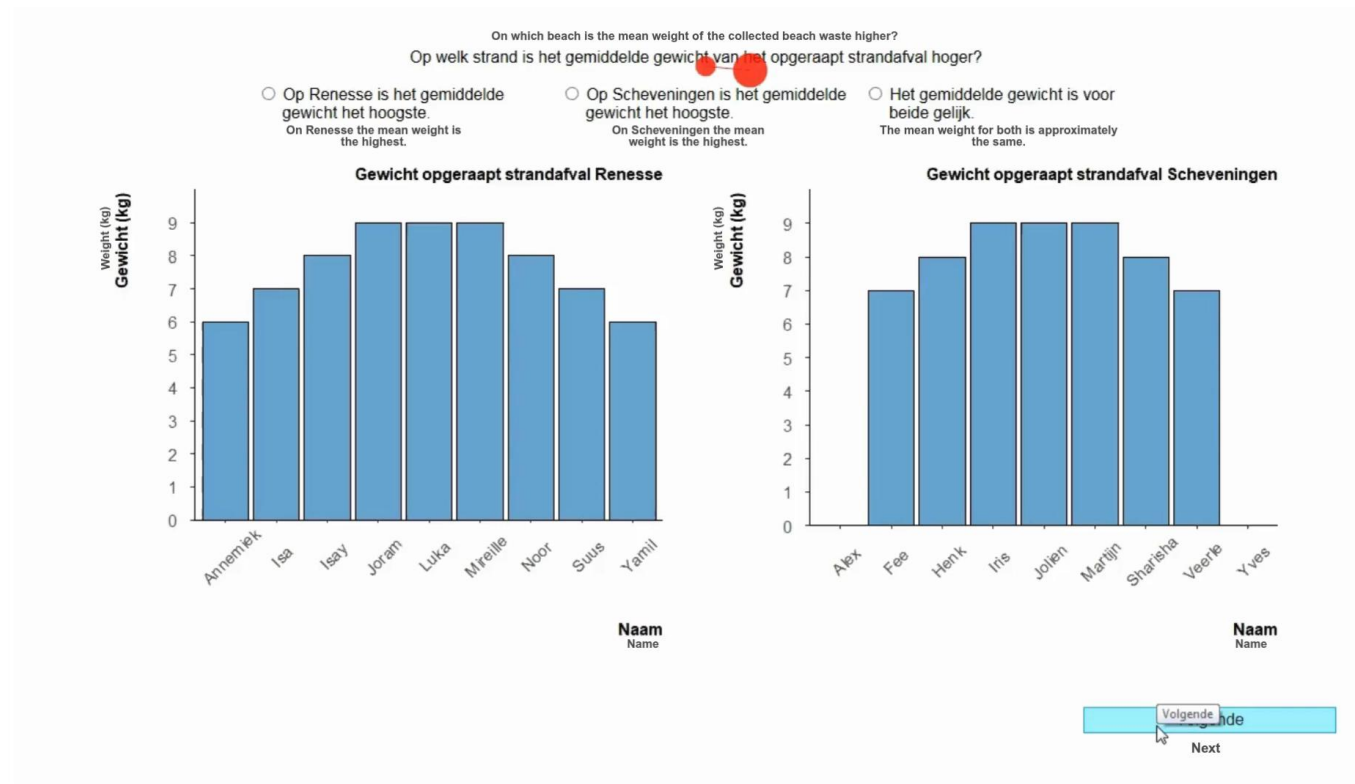
Evaluate solution



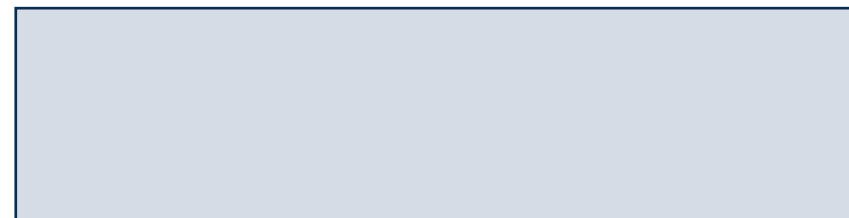
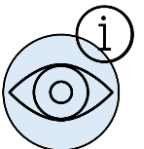
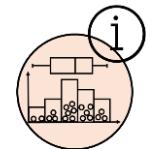
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Eva's comment: "I compared some bars from each graph one by one and saw that Renesse had more bars with lower values. So, I thought Scheveningen must have a higher beach waste weight."



Evaluate Eva's solution. What can you say about her understanding?



Introducing systematic error



Co-funded by
the European Union

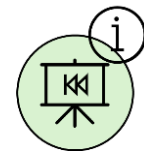
Erasmus+
Enriching lives, opening minds.

Understanding student error

Observable behavior: Eva said she compared some bars in and between graphs one by one and mentioned that Renesse had more bars with lower values. Her gaze indeed moved within and between case-value plots, focusing mainly on comparable bars; she also briefly focused on the name below one “zero” bar in the right-hand graph. This together indicates an observable systematic error, which we term **overemphasizing local features: a pairwise stack comparison** approach to estimating the mean. Ignoring the influence of “zero” bars is a common error.

Interpretation by the teacher/researcher: Although Eva seems to correctly interpret the height of bars as representing measured values, she neglected the influence of the zero bars—the measured values of “zero”. She overlooked that in both graphs the same number of measured values are represented.

Underlying conceptual difficulty: Eva shows an underlying conceptual difficulty with data, hence, in understanding how measured values of zero are represented in case-value plots. Eva needs to consider that, in this scenario, two students did not collect any garbage at all; therefore, their measured value is zero. The presence of zero values lowers the mean in the right-hand graph (Scheveningen). By focusing on counting the number of visible bars and comparing them pairwise, Eva missed out on the significance of these zero values and the actual distribution of all data points. This leads to her incorrect conclusion that the graph with more bars has a lower mean due to the additional bars having lower values.

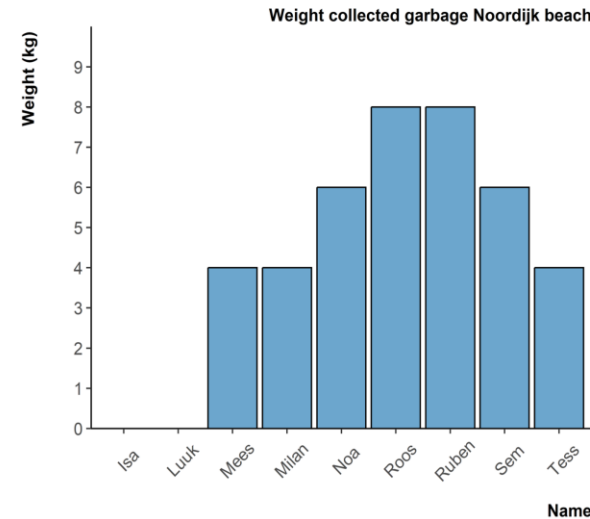
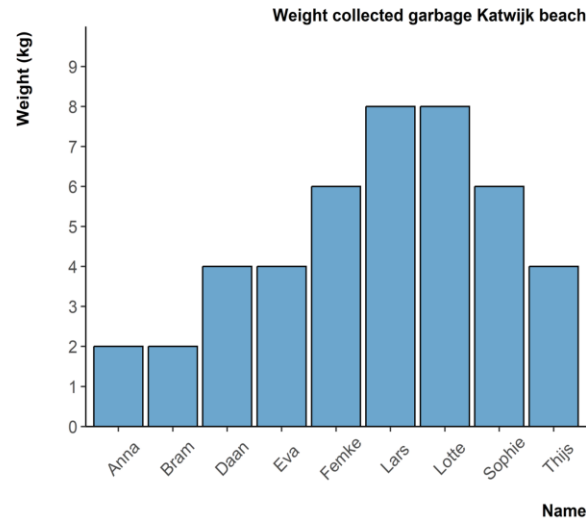


Practice



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

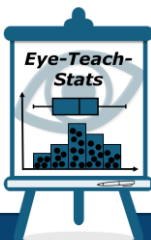
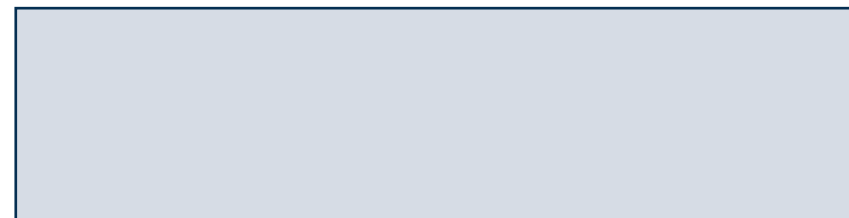


On which beach is the mean weight of the collected beach waste higher?

- a) On Katwijk the mean weight is the highest.
- b) On Noordijk the mean weight is the highest.
- c) The mean weight for both is approximately the same.

Imagine a student makes the same systematic error of overemphasizing local features: a pairwise stack comparison.

Describe the possible gaze pattern the student may show. Do you expect this error to be used here? If so, why? If not, why not?



Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor the granting authority can be held responsible for them.

09 Histogram

Vignette on an error caused by frequency value confusion

Introducing situation



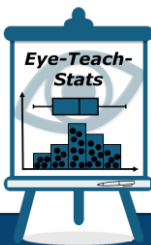
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Imagine you are teaching a 10th-grade classroom on data interpretation, and you introduce your students to find the approximate mean of weights represented in a histogram.

You notice your student Sophie is having difficulties with a task you've assigned. She gives an incorrect answer.

You will now see the task Sophie is working on.



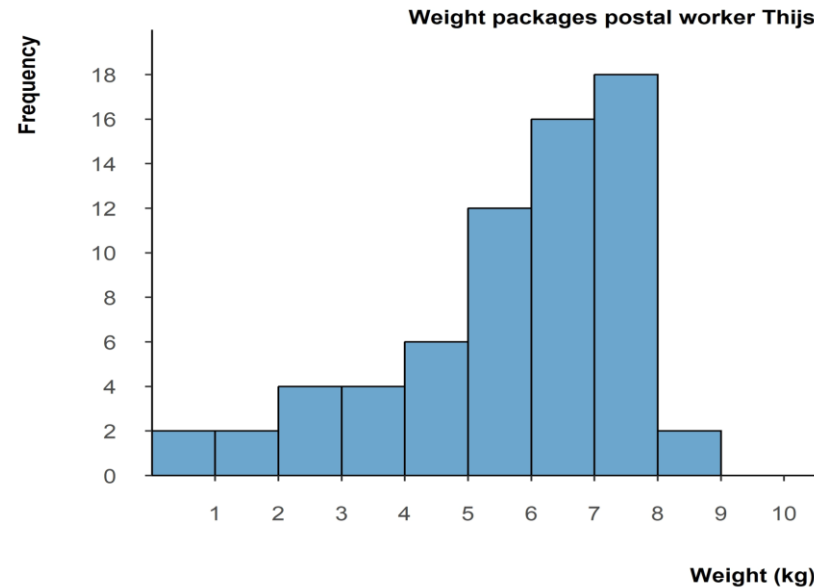
Student's task



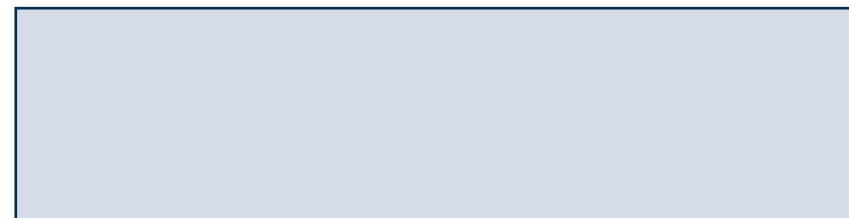
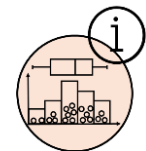
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

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



This is the task Sophie is working on. Please complete the task on your own first and enter your solution.



Introducing eye-tracking

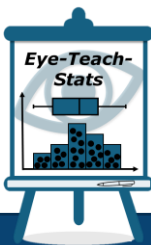


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

A group of researchers in mathematics education is pursuing the idea of supporting teachers with additional diagnostic material: To this end, eye movement data from students is recorded using eye-tracking technology and made available as supplementary material.

You will receive a gaze plot that visualizes gaze data of your student Sophie that was collected while she was working on the task. The gaze data indicates where her visual focus lies. Often it is possible to infer what Sophie's thinking process might have been.



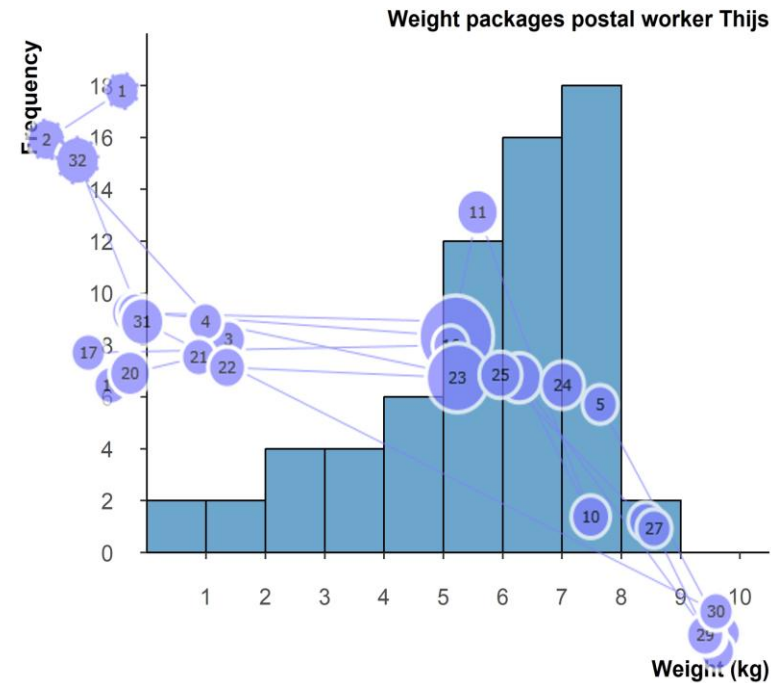
First look



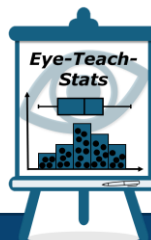
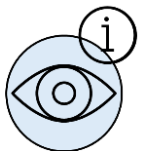
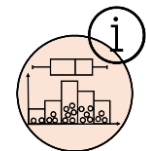
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

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.



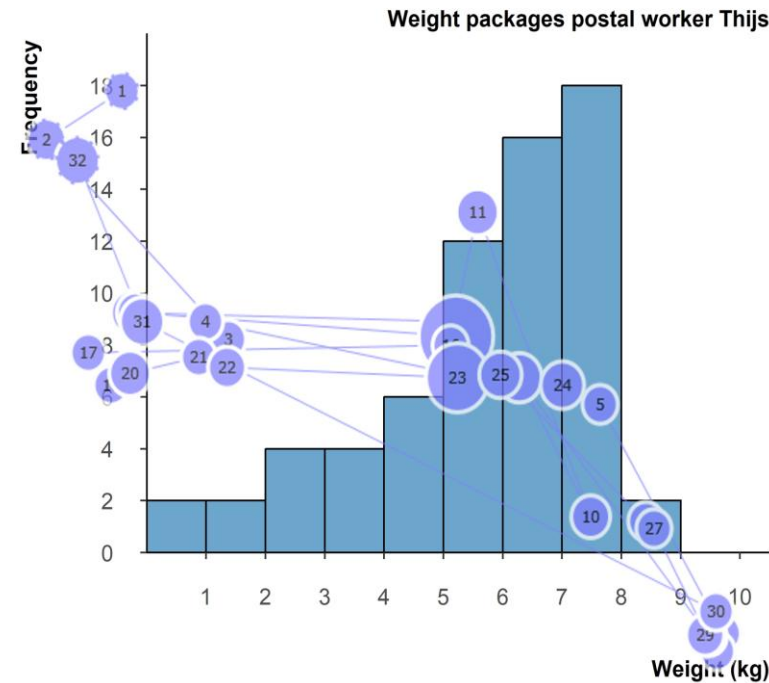
Take perspective



Co-funded by
the European Union

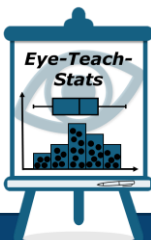
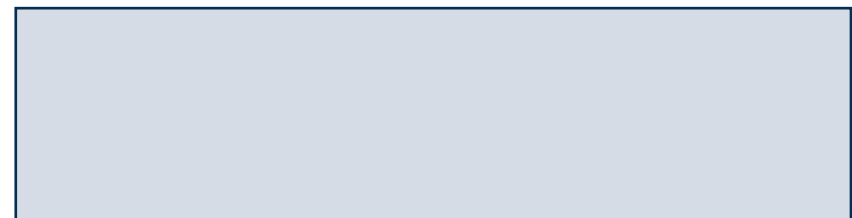
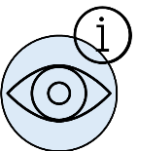
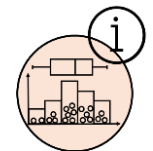
Erasmus+
Enriching lives, opening minds.

Sophie's Answer: "The arithmetic mean is about eight."



After the task Sophie was shown her own gaze plot. She explains retrospectively what she was thinking while working on the task.

Try to take Sophie's perspective, and formulate what she may comment.



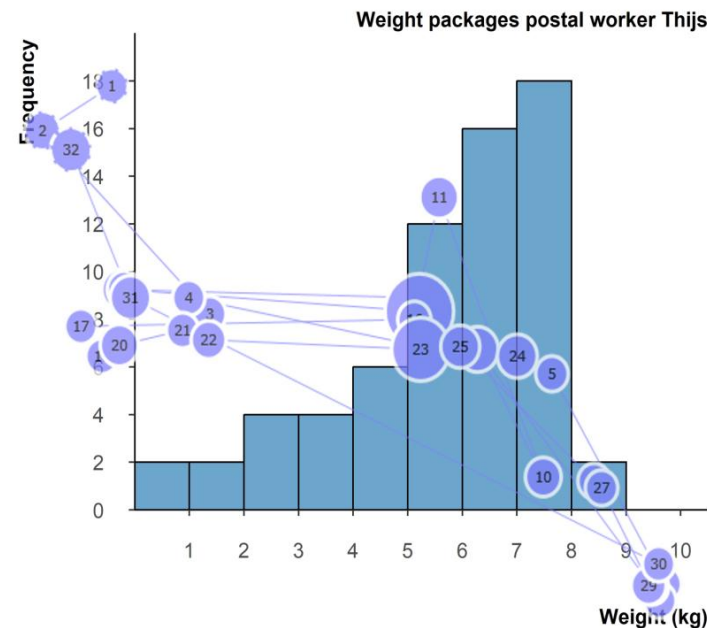
Indicate assumption



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

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."

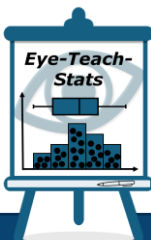
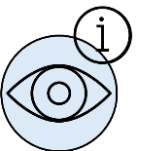
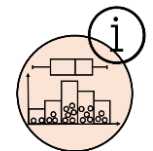


Now you can read the original comment of Sophie.

Indicate to what extent Sophie's comment agrees with what you assumed.

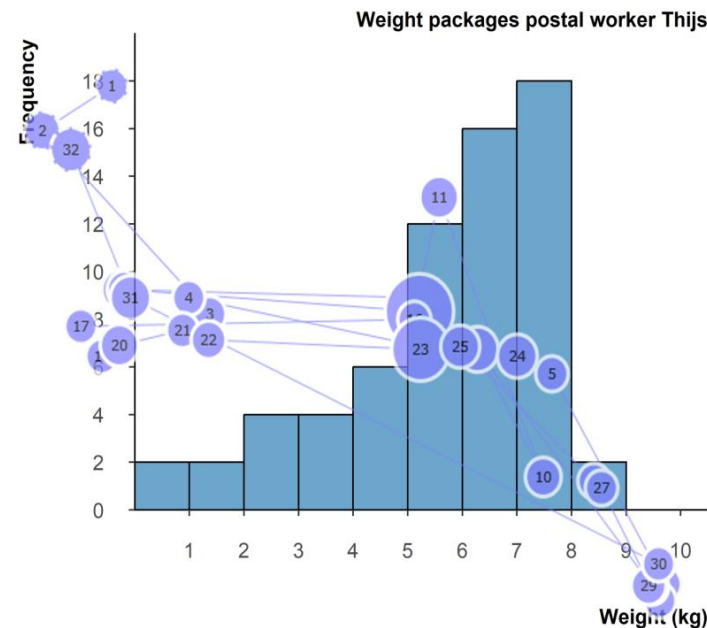
- ☐ No alignment at all
- ☐ Little alignment

- ☐ Quite some alignment
- ☐ Complete alignment

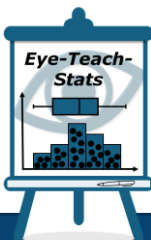
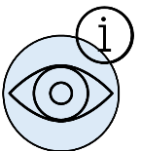
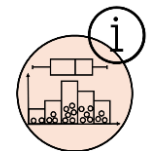


Evaluate solution

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."



Evaluate Sophie's solution. What can you say about her understanding?



Introducing systematic error



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Understanding student errors:

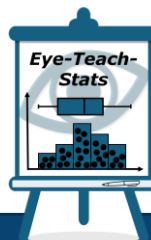
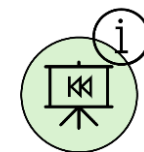
Sophie has an underlying conceptual difficulty regarding the interpretation of histograms. This leads to an observable systematic error, called “**frequency value confusion**”.

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 “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 a case, their gaze would move up and down between a point on the horizontal axis and the tops of the bars, creating a scan path 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).

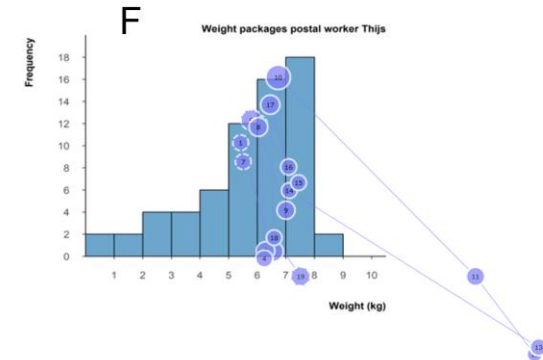
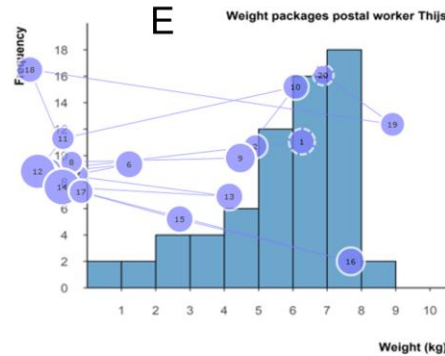
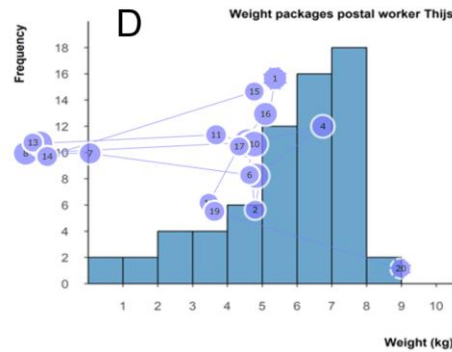
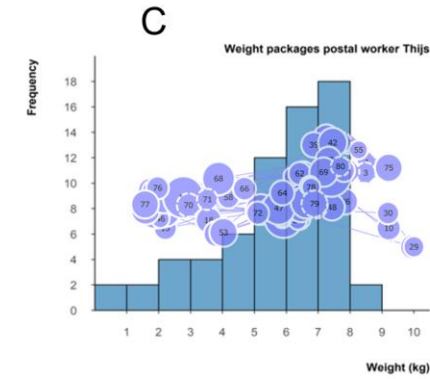
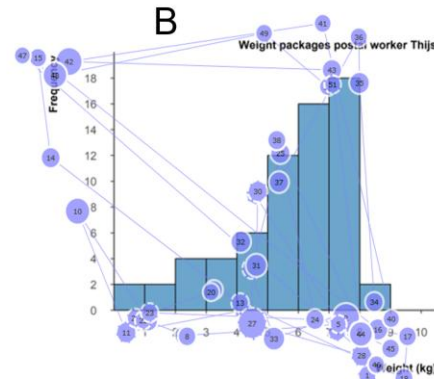
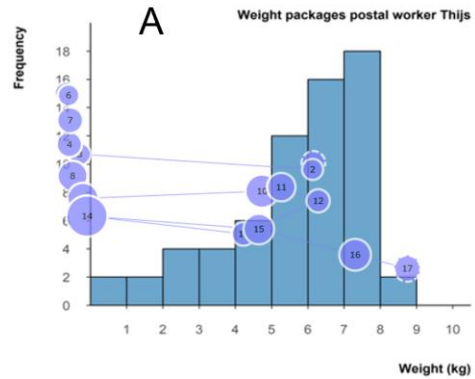


Practice



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.



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

☐ A

☐ B

☐ C

☐ D

☐ E

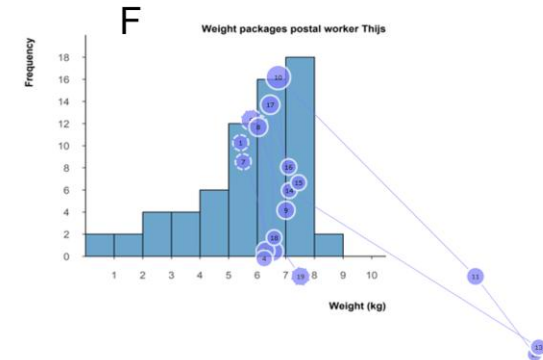
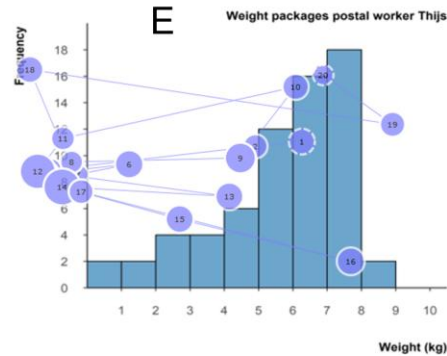
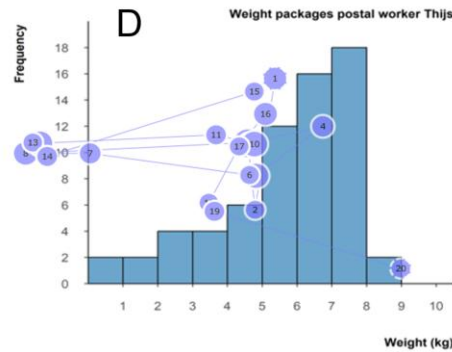
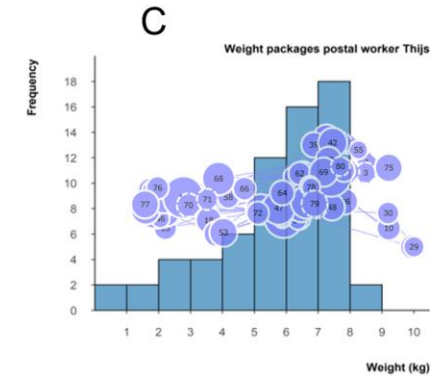
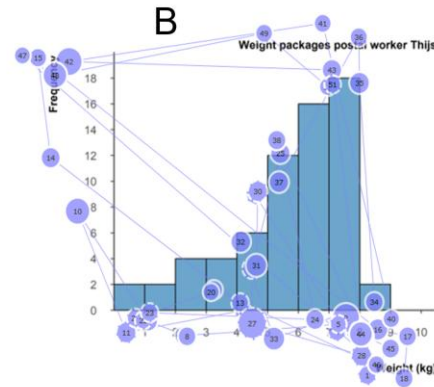
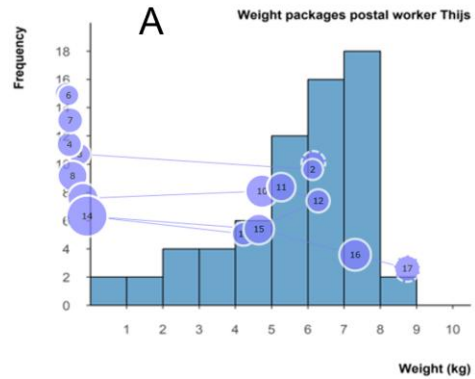
☐ F


Practice II



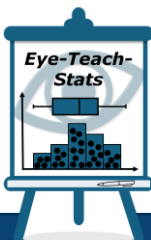
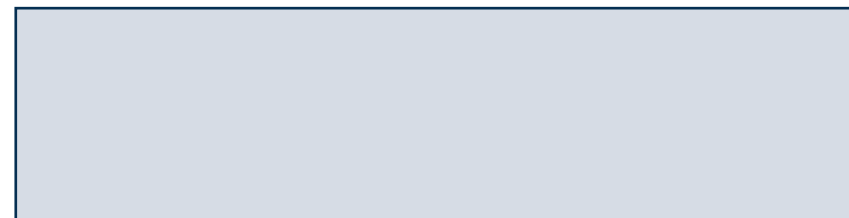
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.



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

Please provide detailed reasons why you have a) opted for student [xy], and b) decided against the other students.





Eye-Teach-Stats Eye-Tracking Vignettes © 2026



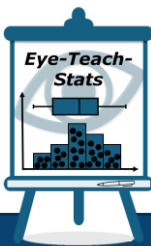
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor the granting authority can be held responsible for them.

10 Dotplot

Vignette on an error caused by frequency value confusion



Introducing situation



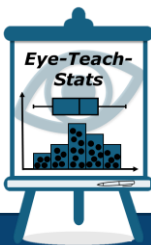
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Imagine you are teaching a 6th-grade class on data interpretation, and you introduce your students to comparing data distributions using dotplots.

You notice your student Noel is having difficulties with a task you've assigned. He gives an incorrect answer.

You will now see the task Noel is working on.



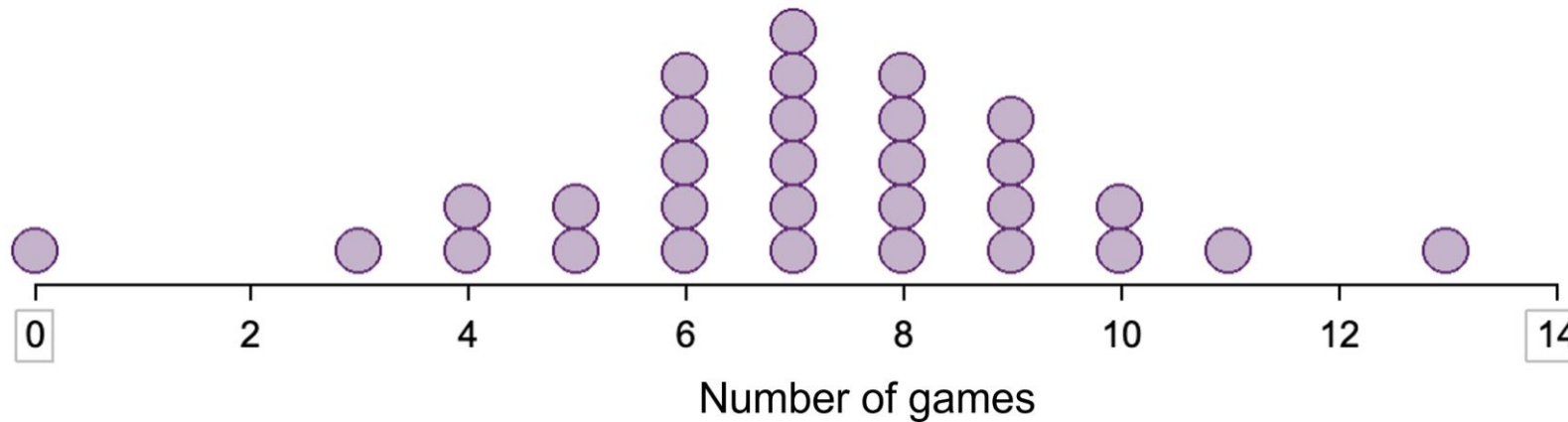
Student's task



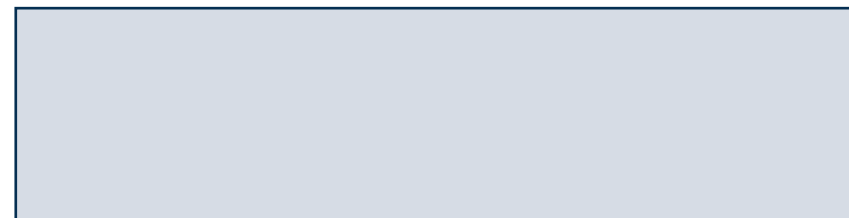
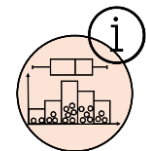
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

All data points together represent a distribution.
Estimate the arithmetic mean.



This is the task Noel is working on. Please complete the task on your own first and enter your solution.



Introducing eye-tracking

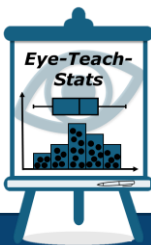


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

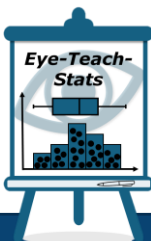
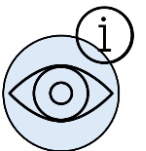
A group of researchers in mathematics education is pursuing the idea of supporting teachers with additional diagnostic material: To this end, eye movement data from students is recorded using eye-tracking technology and made available as supplementary material.

You will receive a gaze plot that visualizes gaze data of your student Noel that was collected while he was working on the task. This gaze plot indicates where his visual focus lies. Often it is possible to infer what Noel's thinking process might have been.



Erasmus+
Enriching lives, opening minds.

The scatter plot displays the relationship between the number of games played (x-axis) and the number of goals scored (y-axis). The x-axis is labeled 'Number of games' and ranges from 0 to 14. The y-axis represents the number of goals, ranging from 0 to 10. Data points are represented by circles of varying sizes and colors (green and purple). Lines connect some points, indicating a progression or comparison between different game counts for the same player.



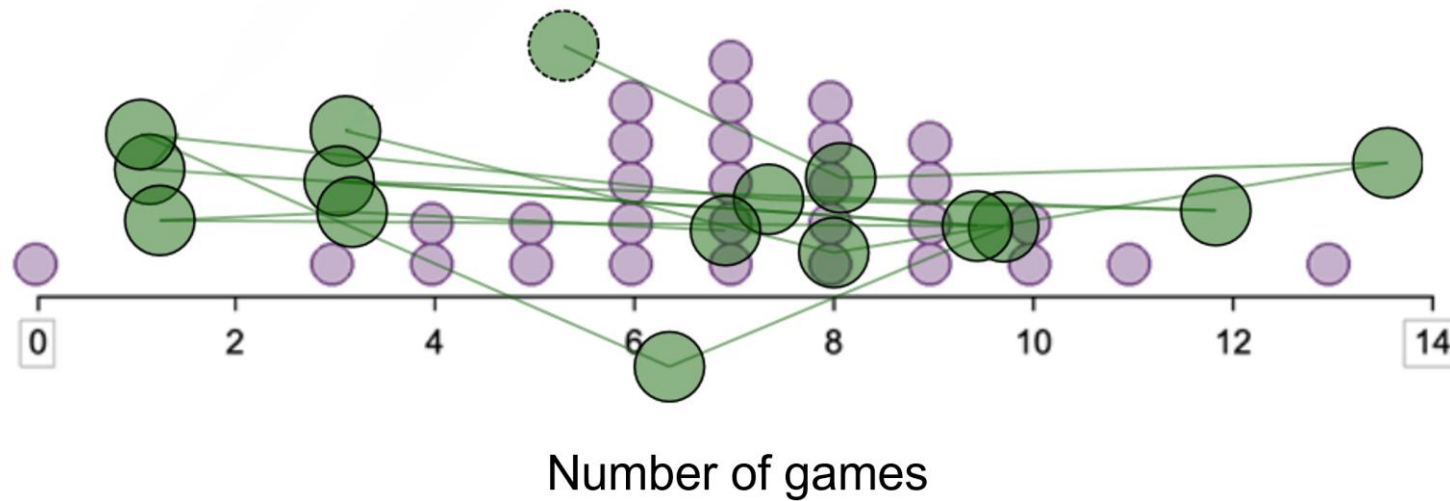
Take perspective



Co-funded by
the European Union

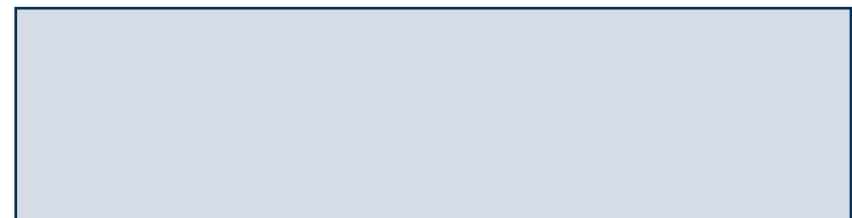
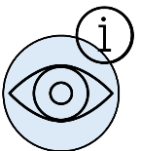
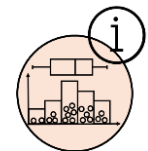
Erasmus+
Enriching lives, opening minds.

Noel's Answer: „The arithmetic mean is about two.“



After the task Noel was shown his own gaze plot. He explains retrospectively what he was thinking while working on the task.

Try to take Noel's perspective, and formulate what he may comment.



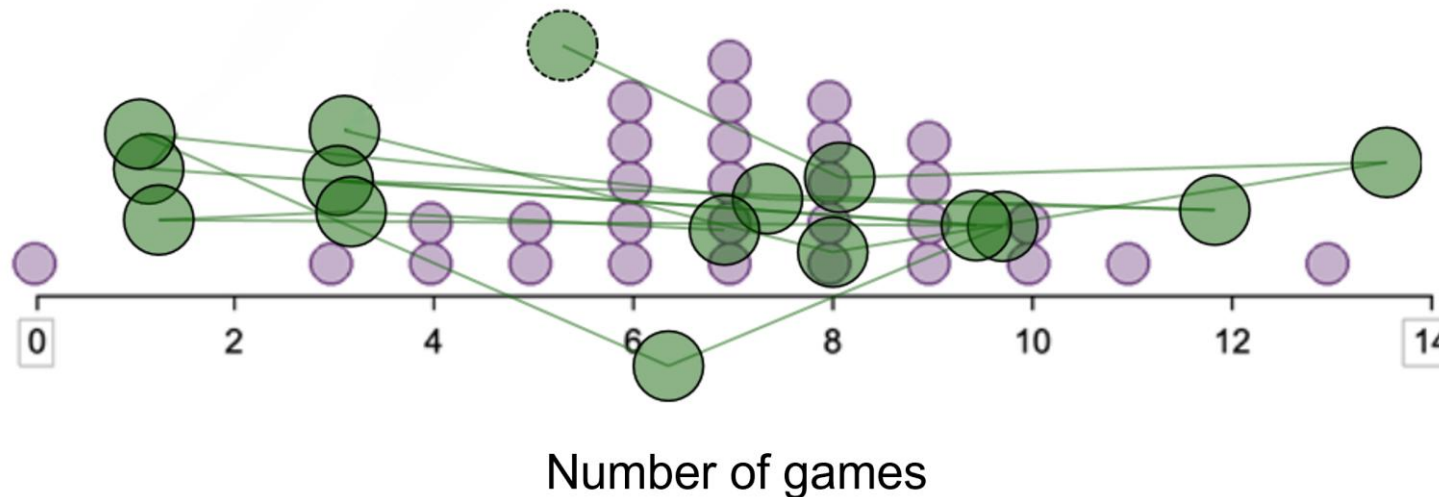
Indicate assumption



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Noel's Comment: "I think it's two because when I look at it, there are some with many games and some with no games and I tried to make the stacks equally high to estimate the arithmetic mean. Then there are about two dots everywhere."

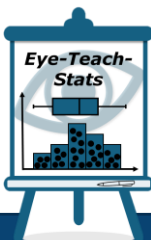
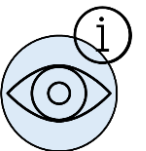
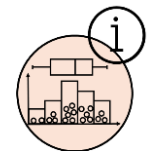


Now you can read the original comment of Noel.

Indicate to what extent Noel's comment agrees with what you assumed.

- ☐ No alignment at all
- ☐ Little alignment

- ☐ Quite some alignment
- ☐ Complete alignment



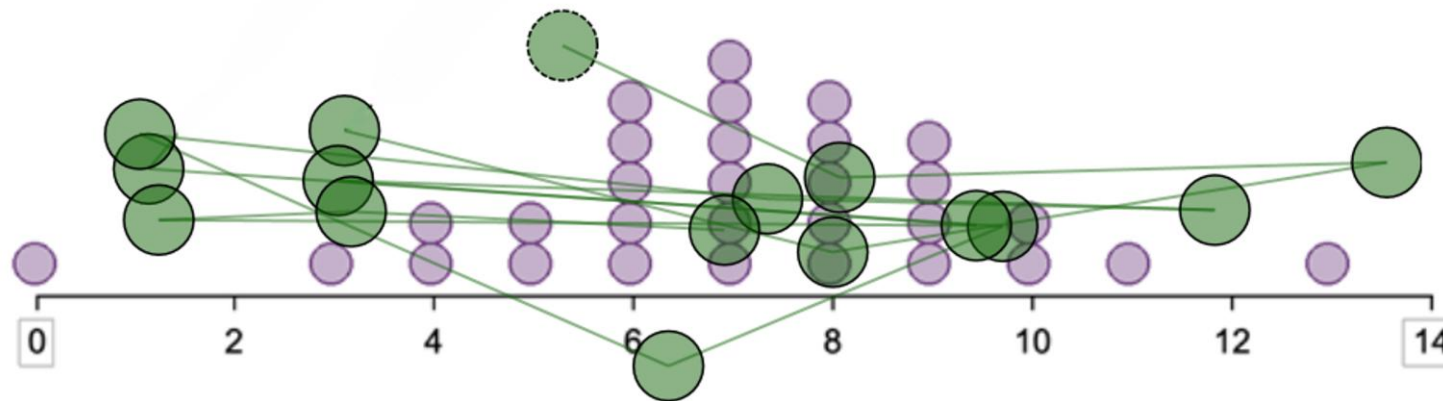
Evaluate solution



Co-funded by
the European Union

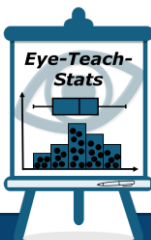
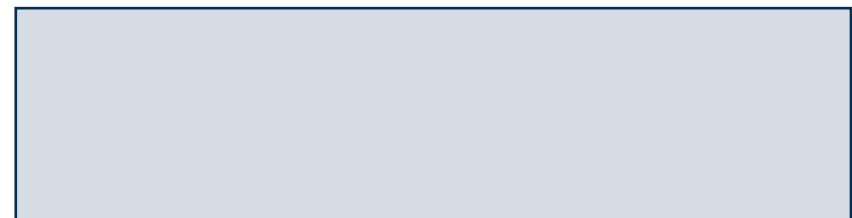
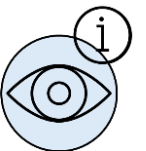
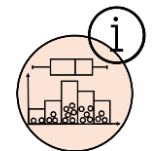
Erasmus+
Enriching lives, opening minds.

Noel's Comment: "I think it's two because when I look at it, there are some with many games and some with no games and I tried to make the stacks equally high to estimate the arithmetic mean. Then there are about two dots everywhere."



Number of games

Evaluate Noel's solution. What can you say about his understanding?



Introducing systematic error



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

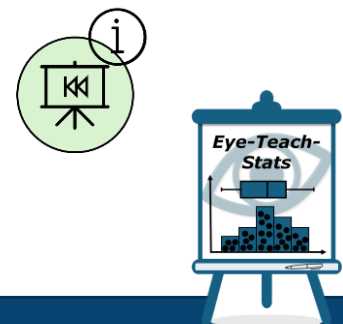
Understanding student errors:

Noel has an underlying conceptual difficulty regarding the interpretation of dotplots. This leads to an observable systematic error, called “**frequency value confusion**”.

Interpreting behavior: Noel’s eye movements form a horizontal line within the graph. This suggests that Noel is looking for the average by finding a point along an imaginary vertical axis where the stacks are equally high. In doing so, he considers the heights as values. This leads to his incorrect answer.

Understanding error and underlying conceptual difficulty: In a dotplot, each dot stands for one occurrence of a value on the horizontal axis (in this case the number of games). The height of the stacks (how many dots are stacked) shows how many times each value occurs. The error occurs when students consider frequency (the height of the stacks) as a measured value rather than simply as the number of times that value occurs.

This strategy of finding a horizontal equalization line is often used and correct when interpreting case-value plots. But to find the mean in a dotplot, you need to consider each value and how often it occurs. This can be done by determining the balance point of the dots according to their measured values. A possible strategy can be found on the next page.



Estimating mean

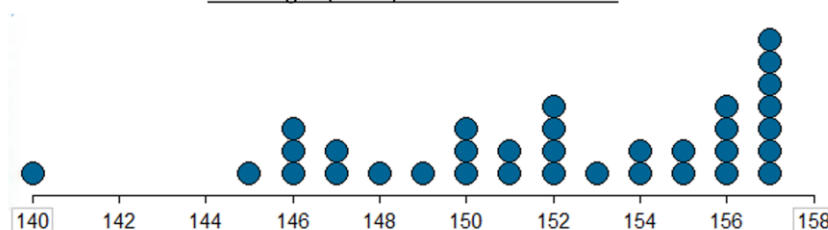


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

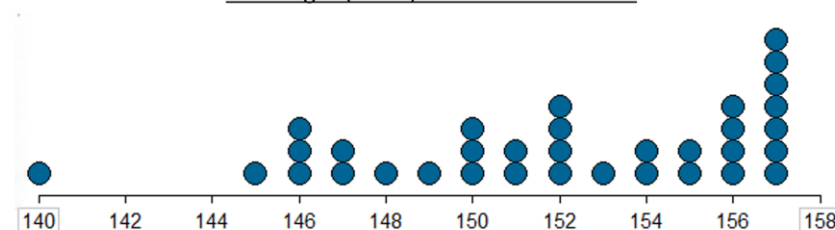
Strategy for estimating the mean in dotplots:

The height (in cm) of children in class 5



Step 1: Look at the distribution shown in the dotplot.

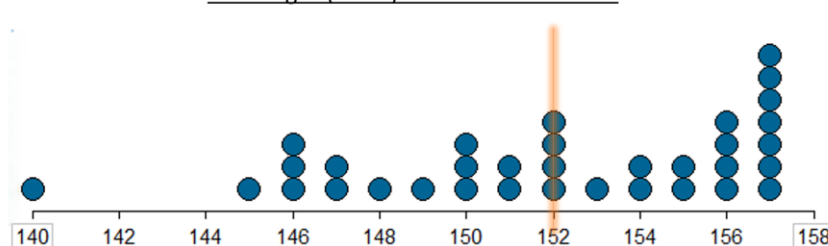
The height (in cm) of children in class 5



balance point

Step 2: Try to find the point at which the weighted dots balance each other out.

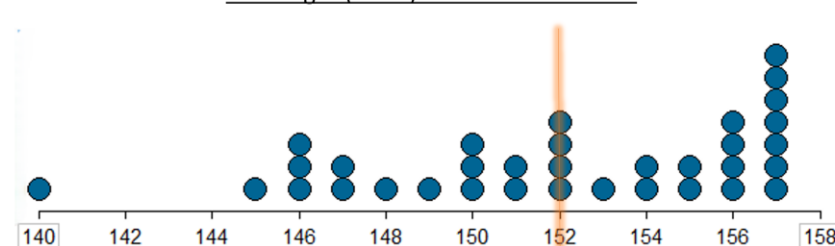
The height (in cm) of children in class 5



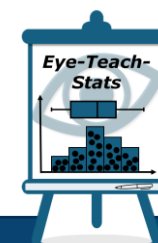
balance point

Step 3: This balance point is the mean of the distribution

The height (in cm) of children in class 5



estimated mean in dotplot



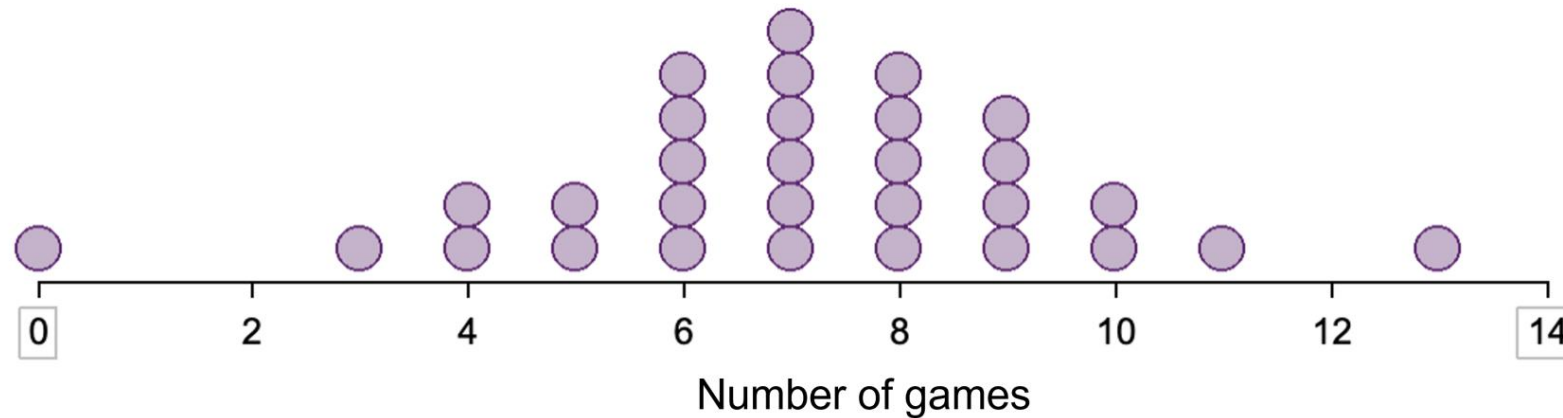
Practice



Co-funded by
the European Union

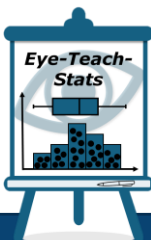
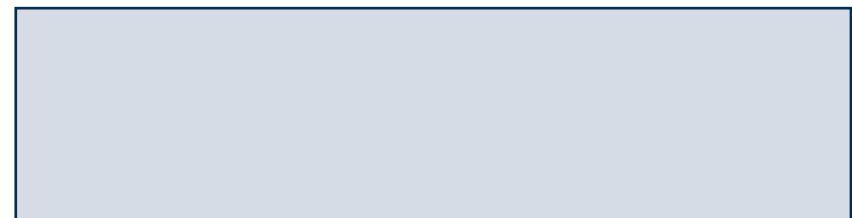
Erasmus+
Enriching lives, opening minds.

All data points together represent a distribution.
Estimate the arithmetic mean.



Imagine a student makes the same systematic error of frequency value confusion, what could be the answer to the following task?

Describe the possible gaze pattern the student may show. What may be the answer of the student and why?



Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor the granting authority can be held responsible for them.

11 Dotplot

Vignette on an error caused by overgeneralization of the sum of measured values for mean

Introducing situation



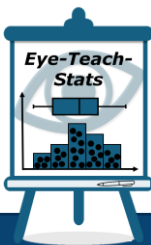
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Imagine you are teaching a 6th-grade class on data interpretation, and you introduce your students to comparing data distributions using dotplots.

You notice your student Talia is having difficulties with a task you've assigned. She gives an incorrect answer.

You will now see the task Talia is working on.



Student's task



Co-funded by
the European Union

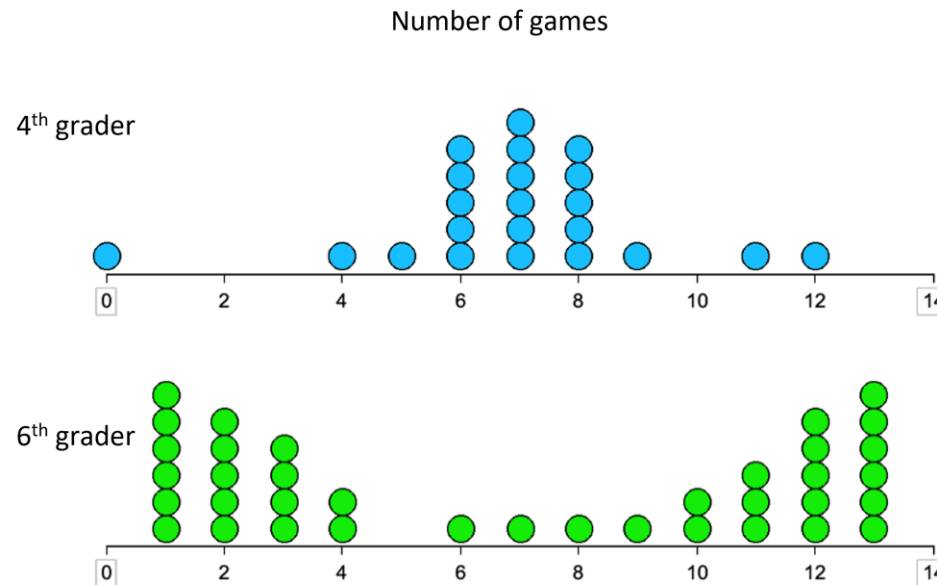
Erasmus+
Enriching lives, opening minds.

Do 4th and 6th graders differ in the number of games they have on their phones?

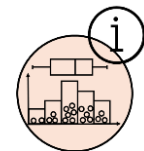
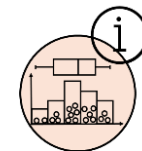
A survey was conducted on this question with children from classes 4 (22 students) and 6 (37 students). Judge from the graph.

The data from the samples fits the assumption that

- a) 4th graders have a lot more games on their phones.
- b) 6th graders have a lot more games on their phones.
- c) 4th and 6th graders have about the same number of games on their phone.



This is the task Talia is working on. **Please complete the task on your own first and enter your solution.**



- a) 4th graders have a lot more games on their phones.
- b) 6th graders have a lot more games on their phones.
- c) 4th and 6th graders have about the same number of games on their phones.



Introducing eye-tracking

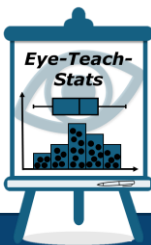


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

A group of researchers in mathematics education is pursuing the idea of supporting teachers with additional diagnostic material: To this end, eye movement data from students is recorded using eye-tracking technology and made available as supplementary material.

You will receive short gaze video sequences that show the eye movements of your student Talia as she was working on the task. This gaze video indicates where her visual focus lies. Often it is possible to infer what Talia's thinking process might have been.



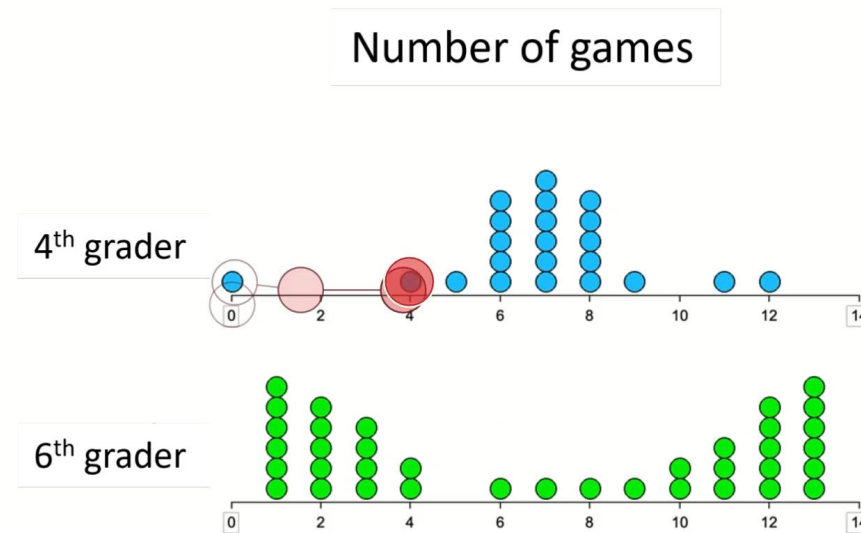
First look



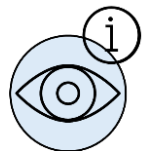
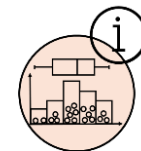
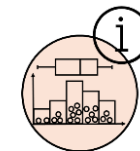
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Talia's Answer: „Therefore, I would say b) is the correct answer: 6th graders have a lot more games on their phones.”



Look at the video sequences that show the viewing behavior of Talia and her answer.



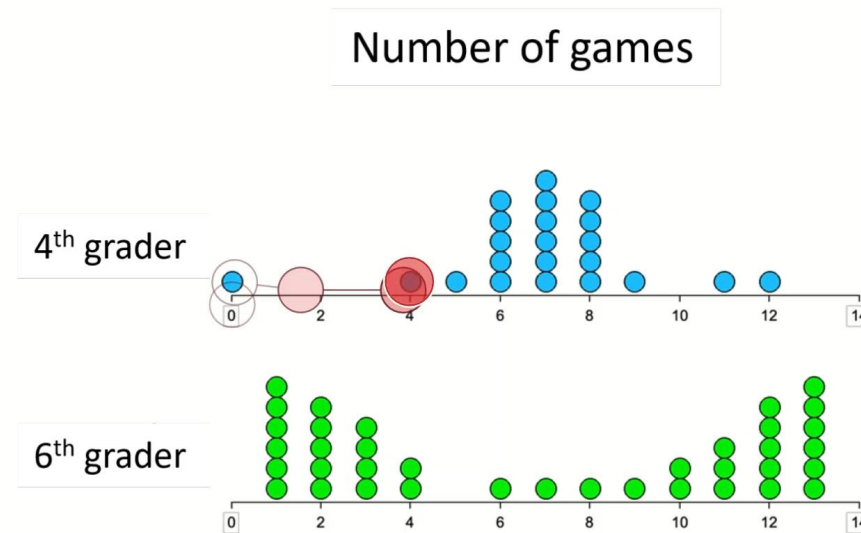
Take perspective



Co-funded by
the European Union

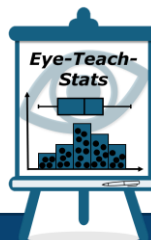
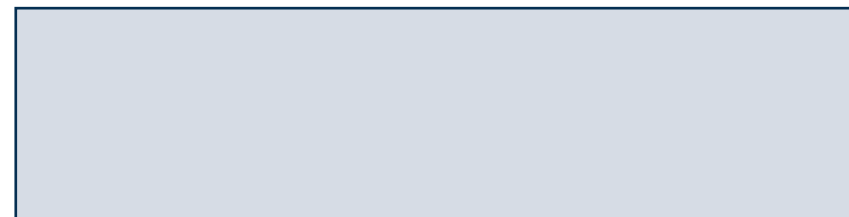
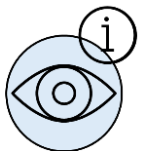
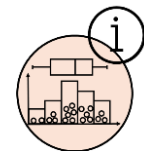
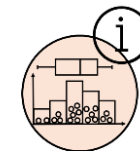
Erasmus+
Enriching lives, opening minds.

Talia's Answer: „Therefore, I would say b) is the correct answer: 6th graders have a lot more games on their phones.”



After the task Talia was shown her own videos. She explains retrospectively what she was thinking while working on the task.

Try to take Talia's perspective, and formulate what she may comment.



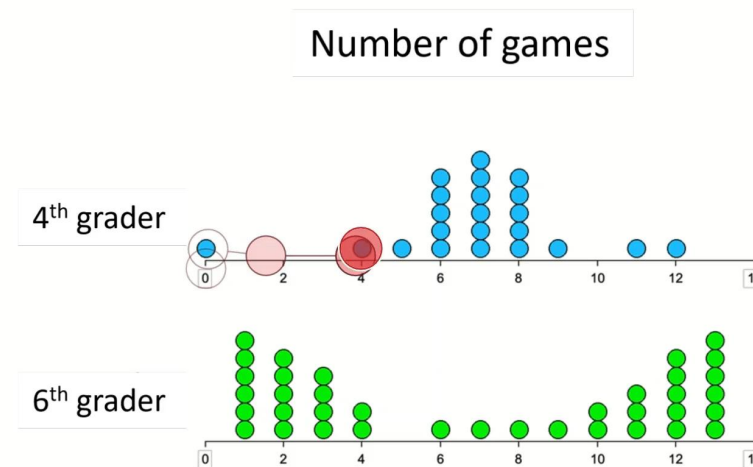
Indicate assumption



Co-funded by
the European Union

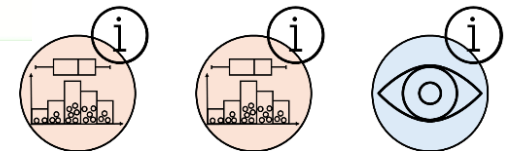
Erasmus+
Enriching lives, opening minds.

Talia's Comment: „First, I counted four plus five, which makes nine. Here, I counted five people who have six games, and calculated five times six, which makes thirty. Then I added the nine from before, making a total of 39. Then I counted... ..and here I added 9, 11, and 12, which makes a total of 134 games for the fourth graders. For the sixth graders, I just calculated 6 times 13 plus 5 times 12, and that was already more than the fourth graders have..”



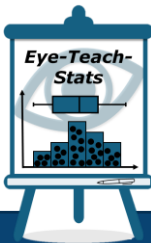
Now you can read the original comment of Talia.

Indicate to what extent Talia's comment agrees with what you assumed.



- ☐ No alignment at all
- ☐ Little alignment

- ☐ Quite some alignment
- ☐ Complete alignment



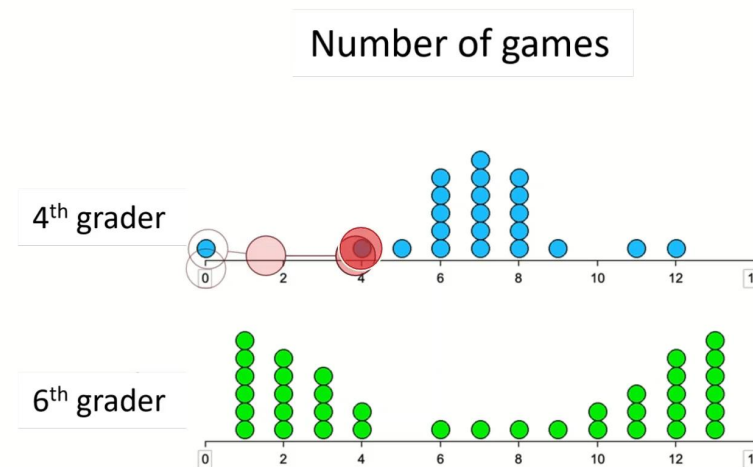
Evaluate solution



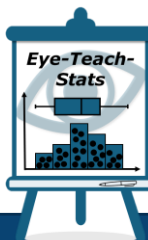
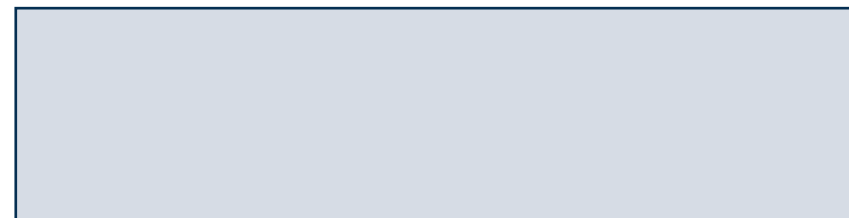
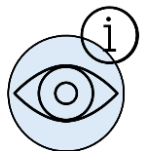
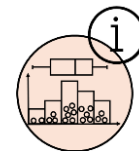
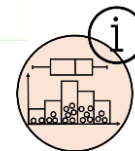
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Talia's Comment: „First, I counted four plus five, which makes nine. Here, I counted five people who have six games, and calculated five times six, which makes thirty. Then I added the nine from before, making a total of 39. Then I counted... ..and here I added 9, 11, and 12, which makes a total of 134 games for the fourth graders. For the sixth graders, I just calculated 6 times 13 plus 5 times 12, and that was already more than the fourth graders have..”



Evaluate Talia's solution. What can you say about her understanding?



Introducing systematic error



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Understanding student errors:

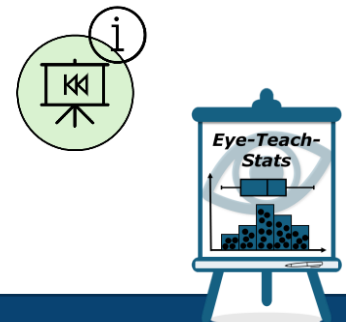
Talia has an underlying conceptual difficulty regarding the interpretation of dotplots. This leads to an observable systematic error, called “**sum of measured values for mean**”.

Interpreting behavior: She calculated and compared the total number of games (measured values) between two distributions with unequal sample sizes. Talia’s gaze data and her responses indicate a strong fixation on individual data points.

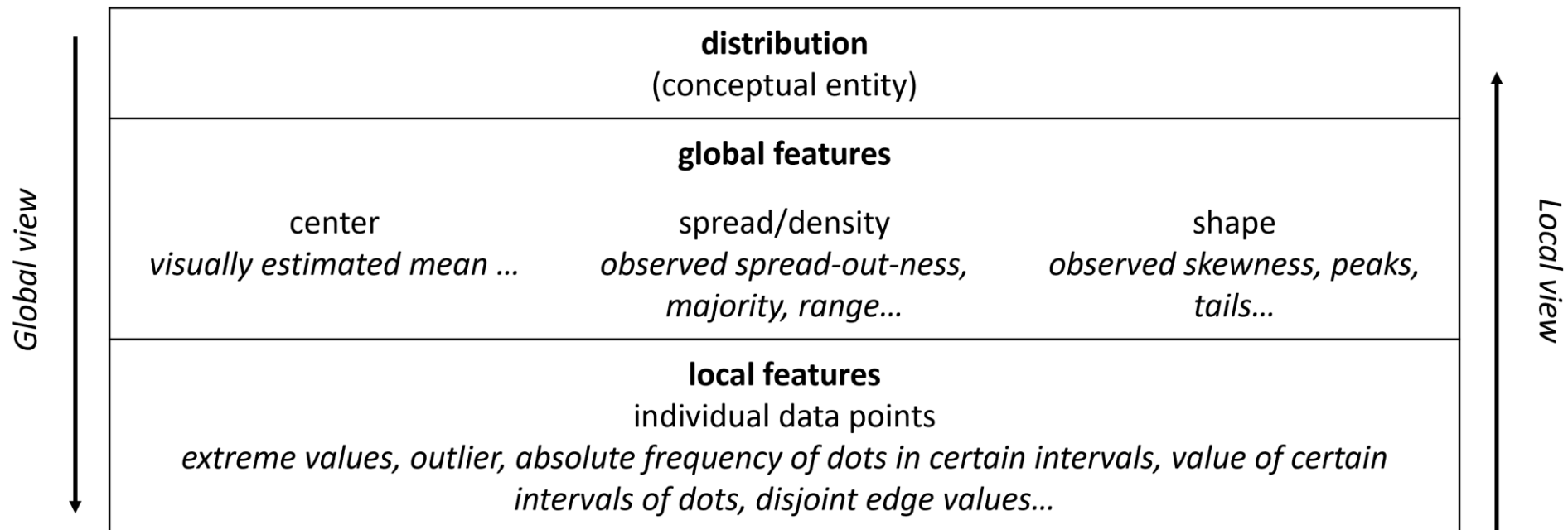
Understanding error and underlying conceptual difficulty: However, when comparing distributions of different sizes, the frequencies must be adjusted relative to each other, such as by using the mean. This type of error often occurs due to difficulties with the concept of distribution and selecting appropriate strategies for their comparison.

A central aim in teaching statistics is to help students develop also 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.

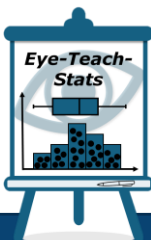
On the next slide, you will see an overview of local and global features that can be taken into account when comparing distributions.



Perspectives on data distribution: local vs. global view



Adapted according to Schreiter & Vogel, 2023



Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor the granting authority can be held responsible for them.

12 Histogram

Vignette on an error caused by overgeneralization of the sum of frequencies for mean

Introducing situation



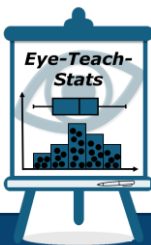
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Imagine you are teaching a 10th-grade classroom on data interpretation and you introduce your students to find the approximate mean of weights represented in a histogram.

You notice your student Mustafa is having difficulties with a task you've assigned. He gives an incorrect answer.

You will now see the task Mustafa is working on.



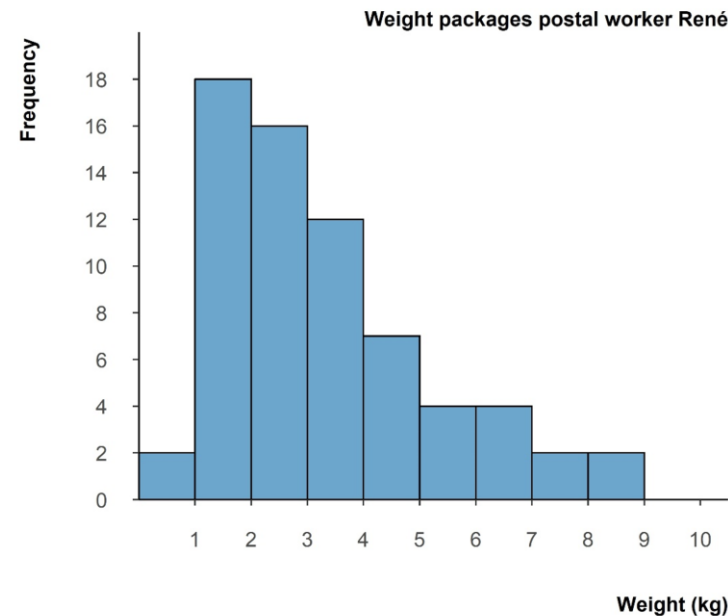
Student's task



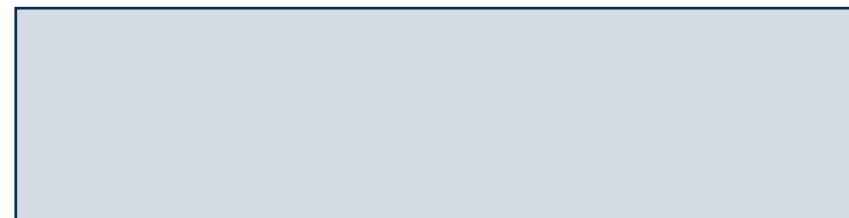
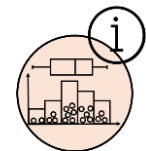
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

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



This is the task Mustafa is working on. Please complete the task on your own first and enter your solution.



Introducing eye-tracking

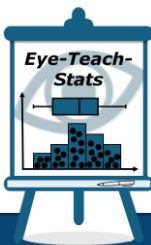


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

A group of researchers in mathematics education is pursuing the idea of supporting teachers with additional diagnostic material: To this end, eye movement data from students is recorded using eye-tracking technology and made available as supplementary material.

You will receive a short gaze video that shows the eye movements of your student Mustafa as he was working on the task. The gaze data indicates where his visual focus lies. Often it is possible to infer what Mustafa's thinking process might have been.



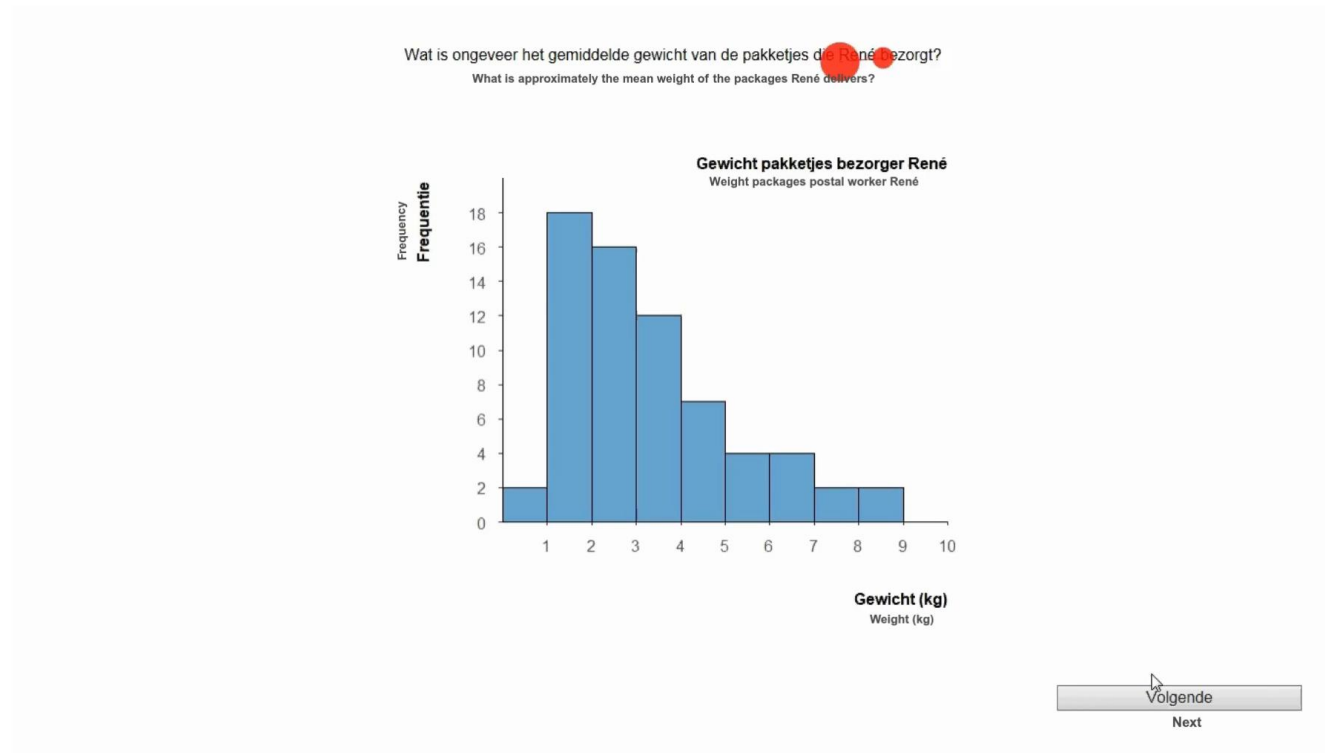
First look



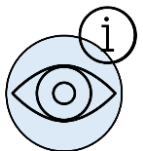
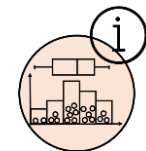
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Mustafa's Answer: "The arithmetic mean is about seven."



Have a look at the gaze video that shows your student Mustafa's gaze and his answer.



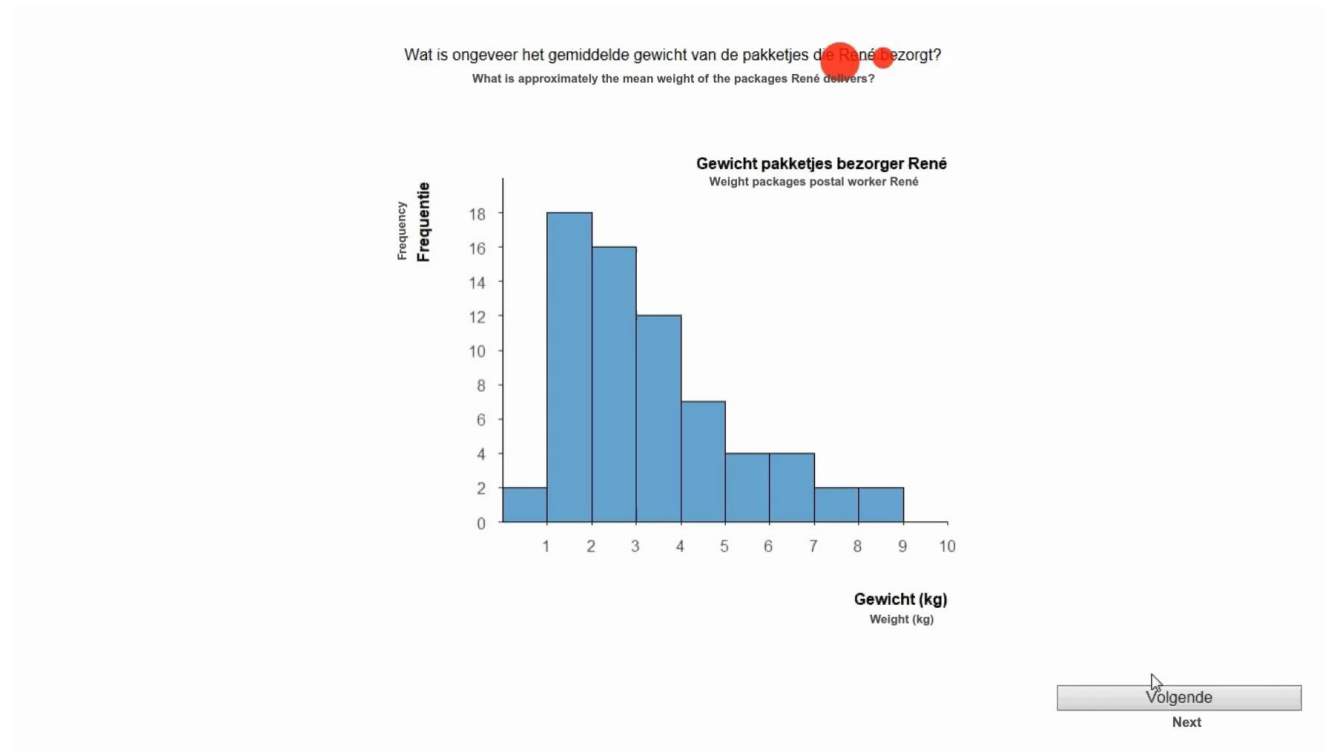
Take perspective



Co-funded by
the European Union

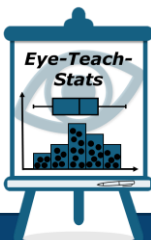
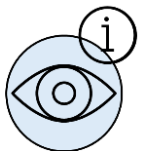
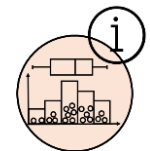
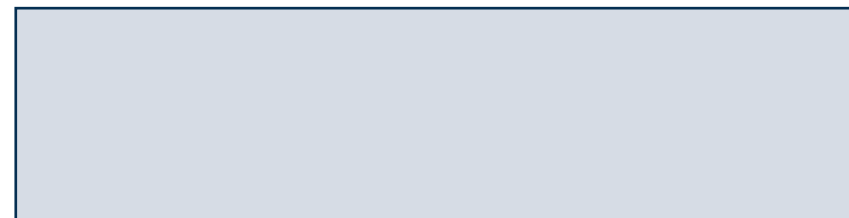
Erasmus+
Enriching lives, opening minds.

Mustafa's Answer: "The arithmetic mean is about seven."



After the task Mustafa was shown his own gaze video. He explains retrospectively what he was thinking while working on the task.

Try to take Mustafa's perspective and formulate what he may comment.



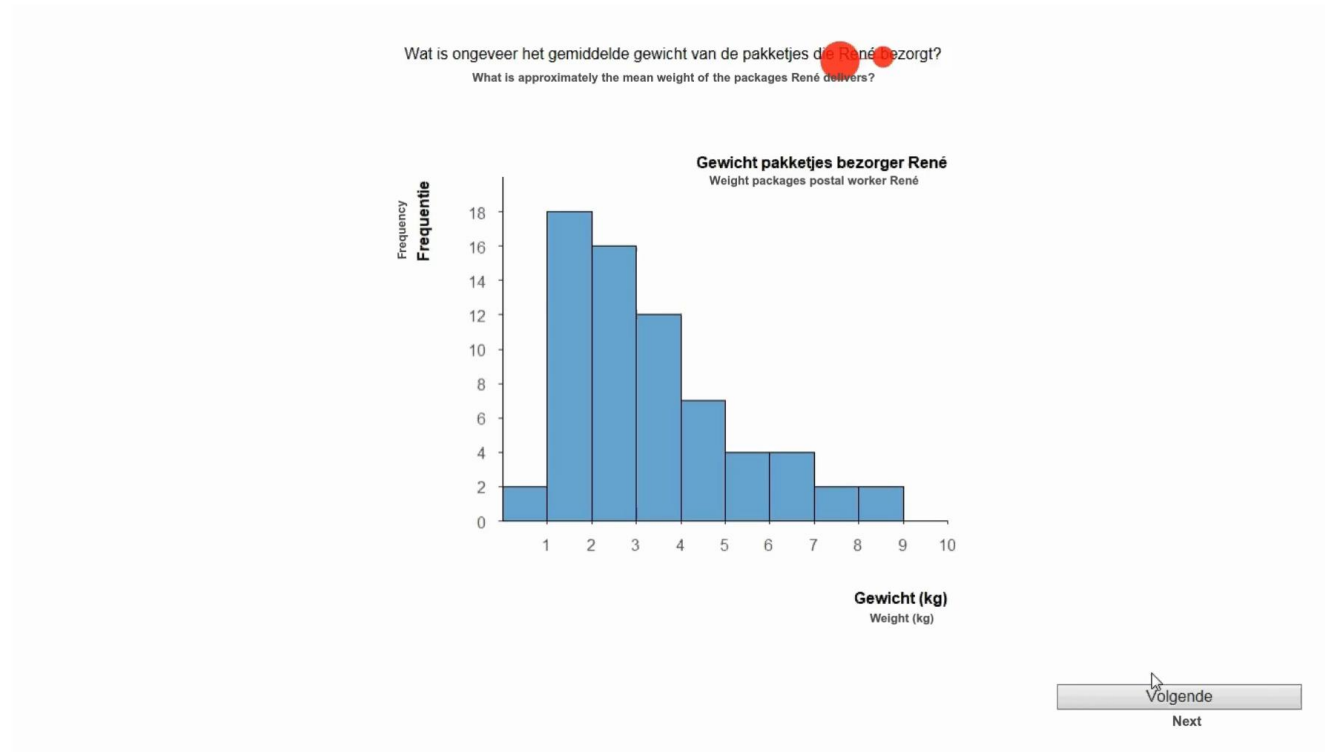
Indicate assumption



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Mustafa's Comment: "I added all the frequencies together and then divided by nine. And then I got a number that should be about the mean."

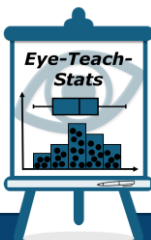
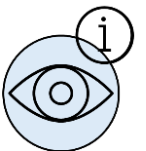
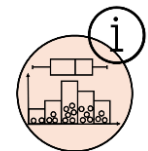


Now you can read the original comment of Mustafa.

Indicate to what extent Mustafa's comment agrees with what you assumed.

- ☐ No alignment at all
- ☐ Little alignment

- ☐ Quite some alignment
- ☐ Complete alignment



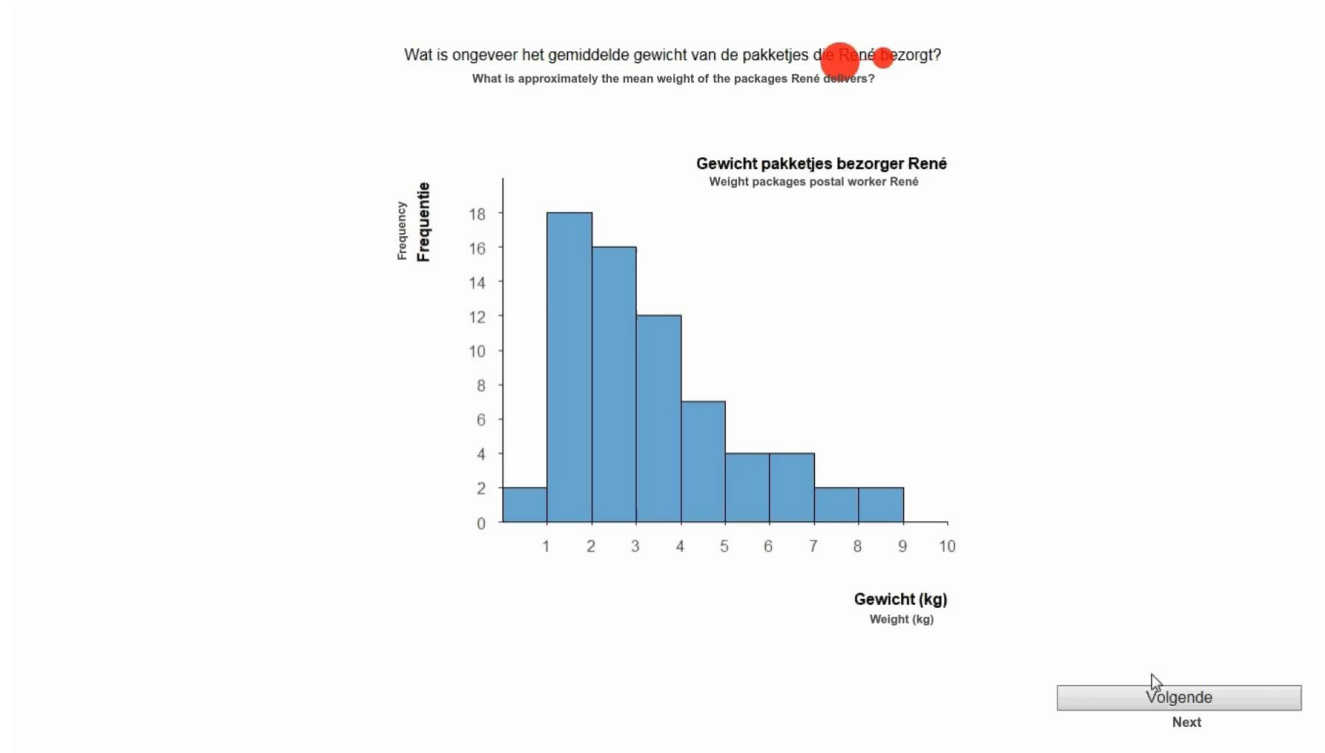
Evaluate solution



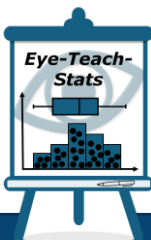
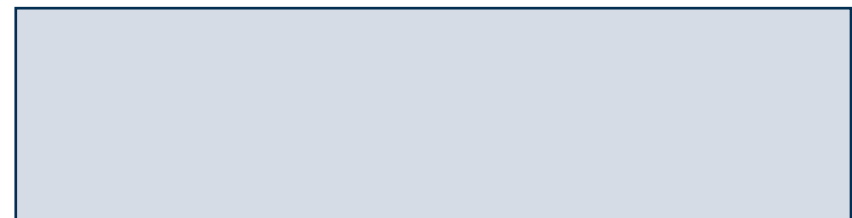
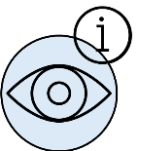
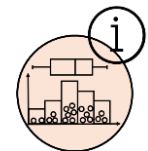
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Mustafa's Comment: "I added all the frequencies together and then divided by nine. And then I got a number that should be about the mean."



Evaluate Mustafa's solution. What can you say about his understanding?



Introducing systematic error



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

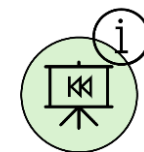
Understanding student errors:

Observable behavior: Mustafa stated that he added the frequencies together and divided by the number of bars, assuming this would result in the mean. We can also observe that Mustafa's gaze after some time moved back and forth between the heights of the bars and the frequency values along the y-axis. At some frequencies, the growing dot indicates that he took longer. The gaze then shifts to the x-axis, where he pays particular attention to the number 9, which corresponds to the total number of bars in the histogram. This behavior together indicates an observable systematic error, which we term **"sum for mean: sum of frequencies"**.

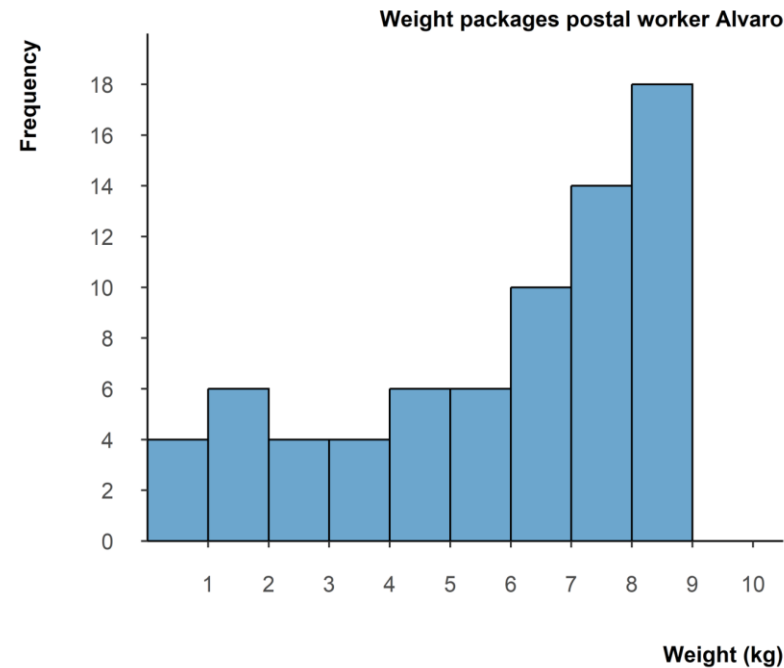
Interpretation by the teacher/researcher: His explanation suggests that he treated the bar heights—representing frequencies—as if they were the measured data values themselves. His gaze pattern suggests that he added up all frequencies (given by the bars' heights) and then divided by 9 to find what he believed to be the arithmetic mean. The growing dot most likely indicates Mustafa is thinking. Here, it is because Mustafa is adding the frequencies in his head. 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. Mustafa's approach reflects confusion between frequency and measured value.

Underlying conceptual difficulty: Mustafa'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 a case, their gaze would move up and down between a point on the horizontal axis and the tops of the bars, creating a scan path 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).

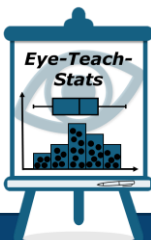
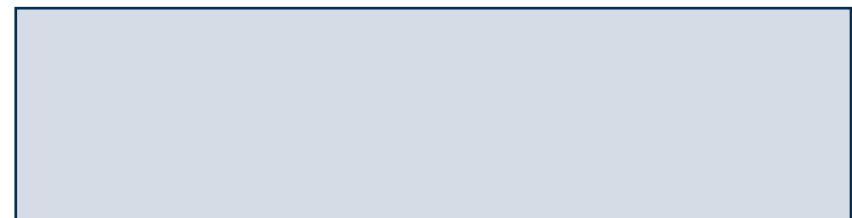


What is approximately the mean weight of the packages that Alvaro delivers?



Imagine a student makes the same systematic error of “sum for mean: sum of frequencies”, what could be the answer to the following task?

Describe the possible gaze pattern the student may show. Do you expect this error to be used here? If so, why? If not, why not?



Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor the granting authority can be held responsible for them.

13 Histogram

Vignette on an error caused by overemphasizing the number of bars

Introducing situation



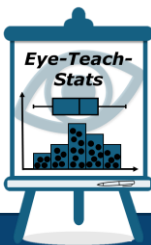
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Imagine you are teaching a 10th grade classroom on data interpretation, and you introduce your students to comparing data distributions using histograms.

You notice your student Finn is having difficulties with a task you've assigned. He gives an incorrect answer.

You will now see the task Finn is working on.

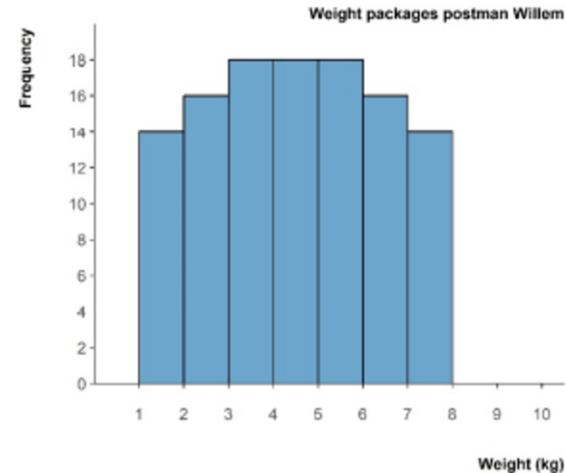
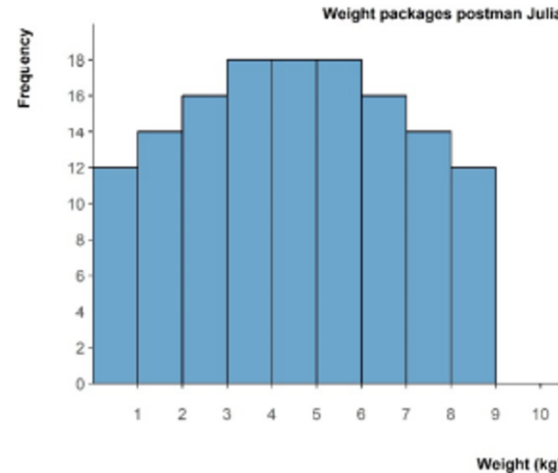


Student's task



Co-funded by
the European Union

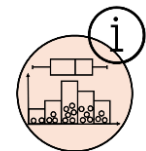
Erasmus+
Enriching lives, opening minds.



Which postal worker delivers on average heaviest packages?

- a) Julia delivers on average the heaviest packages.
- b) Willem delivers on average the heaviest packages.
- c) The mean weight for both is approximately the same.

This is the task Finn is working on. Please complete the task on your own first and enter your solution.



- a) Julia delivers on average the heaviest packages.
- b) Willem delivers on average the heaviest packages.
- c) The mean weight for both is approximately the same.



Introducing eye-tracking

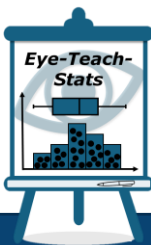


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

A group of researchers in mathematics education is pursuing the idea of supporting teachers with additional diagnostic material: To this end, eye movement data from students is recorded using eye-tracking technology and made available as supplementary material.

You will receive a short gaze video that shows the eye movements of your student Finn that were collected while he was working on the task. The gaze data indicates where his visual focus lies. Often it is possible to infer what Finn's thinking process might have been.



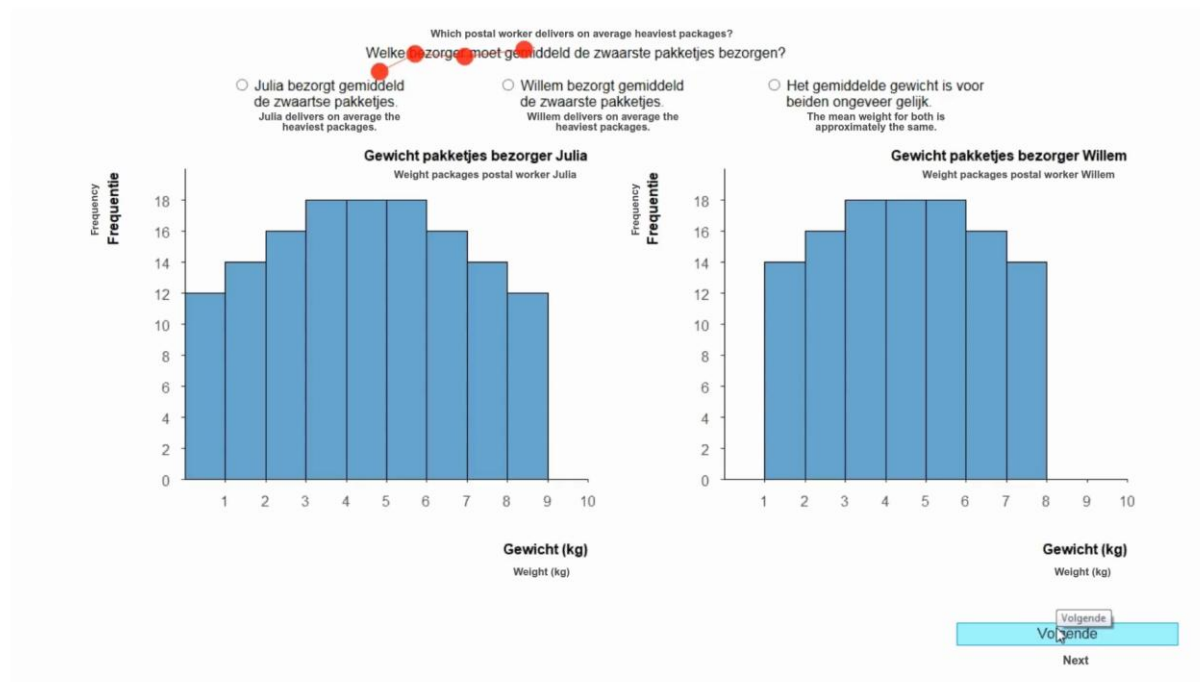
First look



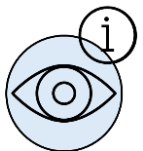
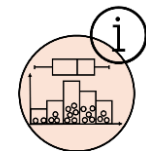
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Finn's Answer: "Julia delivers on average the heaviest packages."



Have a look at the video sequence that shows your student Finn's gaze and his answer.



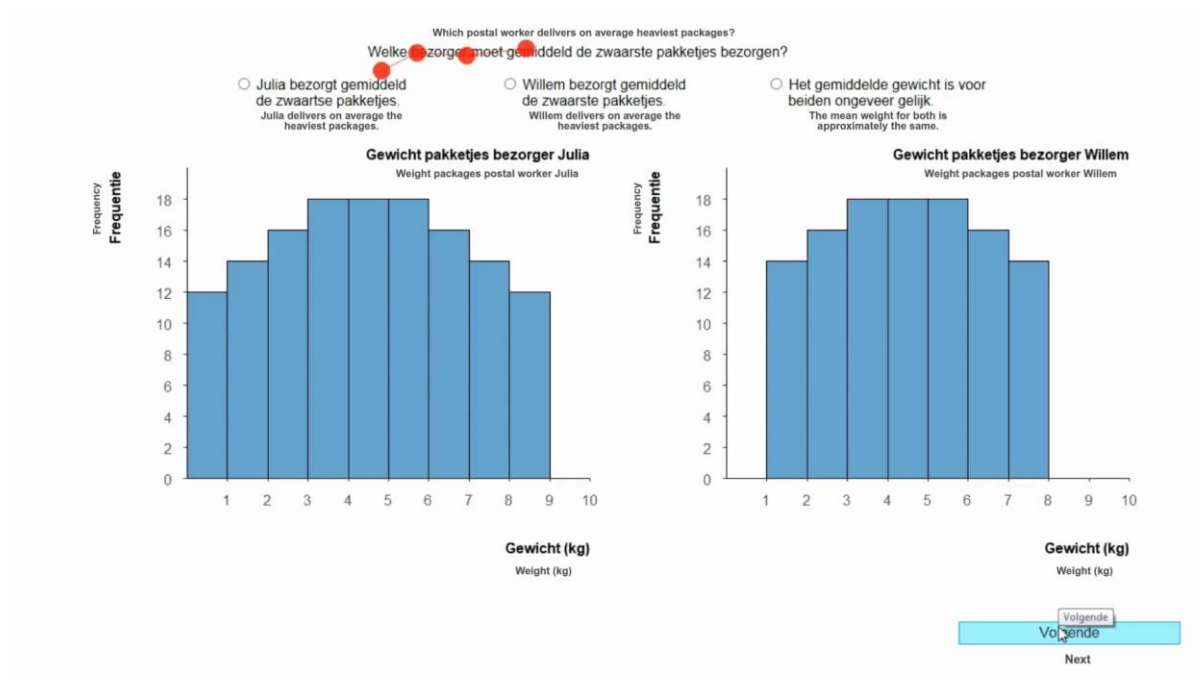
Take perspective



Co-funded by
the European Union

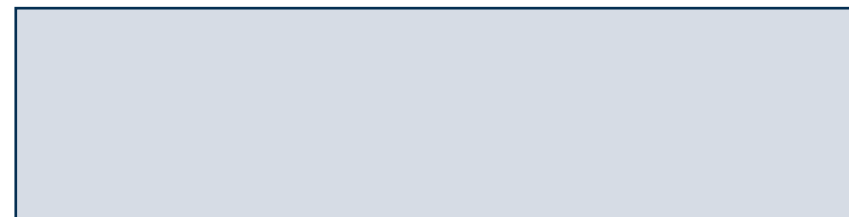
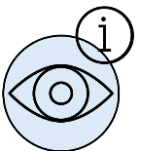
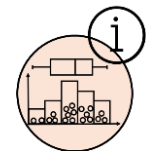
Erasmus+
Enriching lives, opening minds.

Finn's Answer: "Julia delivers on average the heaviest packages."



After the task Finn was shown his own gaze video. He explains retrospectively what he was thinking while working on the task.

Try to take Finn's perspective and formulate what he may comment.



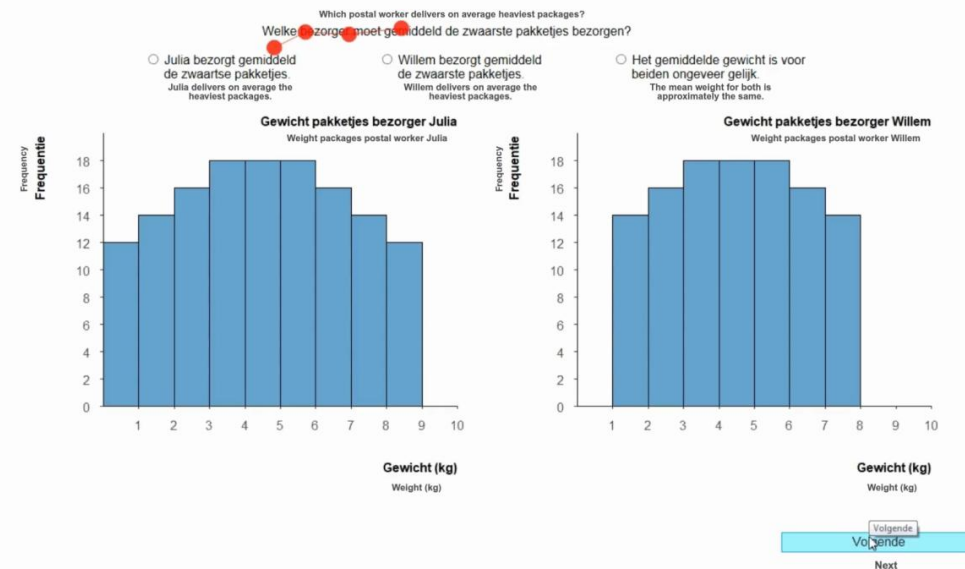
Indicate assumption



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

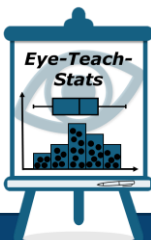
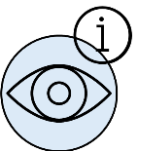
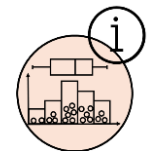
Finn's Comment: "I looked at the bars and compared them. I saw that Willem has fewer bars and Julia has more bars spread across the range. Since we're talking about mean weight, I thought having more bars, like in Julia's histogram, meant a higher mean weight. Even though there are missing bars for Willem, I thought that meant there was nothing in those ranges. So, I chose Julia because I saw more bars and assumed that increased the mean weight."



Now you can read the original comment of Finn.

Indicate to what extent Finn's comment agrees with what you assumed.

- ☐ No alignment at all
- ☐ Little alignment
- ☐ Quite some alignment
- ☐ Complete alignment



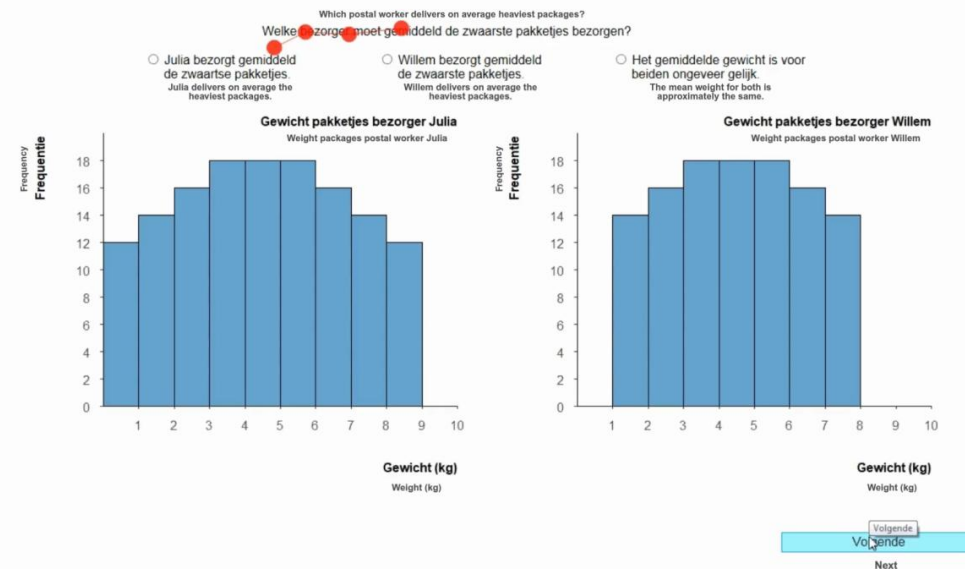
Evaluate solution



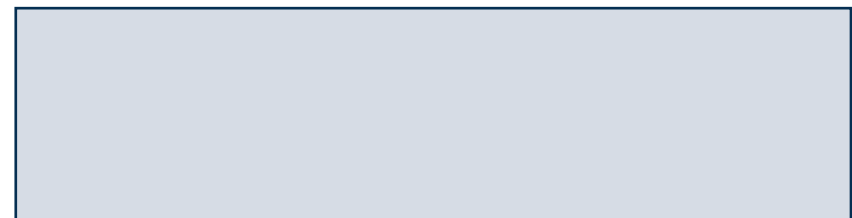
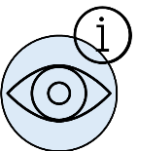
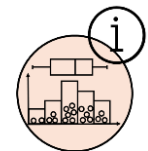
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Finn's Comment: "I looked at the bars and compared them. I saw that Willem has fewer bars and Julia has more bars spread across the range. Since we're talking about mean weight, I thought having more bars, like in Julia's histogram, meant a higher mean weight. Even though there are missing bars for Willem, I thought that meant there was nothing in those ranges. So, I chose Julia because I saw more bars and assumed that increased the mean weight."



Evaluate Finn's solution. What can you say about his understanding?



Introducing systematic error



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

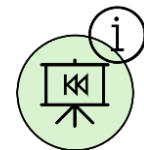
Understanding student errors

Finn has an underlying conceptual difficulty regarding the interpretation of histograms. This leads to an observable systematic error, called **"overemphasizing local features: number of bars"**.

Observable behavior: Finn stated that he compared the number of bars between the graphs and noticed that Julia's histogram had more bars across the range. He said that more bars increases the sum of the weights, hence, the mean. The gaze data show that Finn repeatedly moved his gaze from one bar in one histogram to a corresponding bar in the other histogram. He did not explicitly look at the "zero" bars in the Willem histogram. Together, this indicates an observable systematic error, which we term overemphasizing local features: number of bars, in the interpretation of histograms.

Interpretation by the teacher/researcher: Based on his observed behavior, it appears that Finn focused primarily on the number and visual spread of the bars in each histogram. He ignored that—due to its symmetry (the overall distribution)—the balance points of the data (along the horizontal axis) are the same in both histograms. It seems that he interpreted the graph as a case-value plot where each bar stands for one measured value. However, in histograms, the data is aggregated into bins where the position of this bar is the range of measured values in that bar and the height of each bar represents the number of measurements (hence, frequency). Comparing individual bars between histograms without considering the distributional characteristics (including their positions) leads to incorrect conclusions about the approximate means.

Underlying conceptual difficulty: The systematic error is likely due to an underlying conceptual difficulty with the concept of data and distribution; therefore, in determining and comparing means in histograms. Finn overemphasized the number of bars and failed to correctly account for zero-frequency intervals—that did not influence the mean in this pair of histograms. In histograms, adding bars in such a way that the balance point along the horizontal axis stays the same, does not influence the mean.

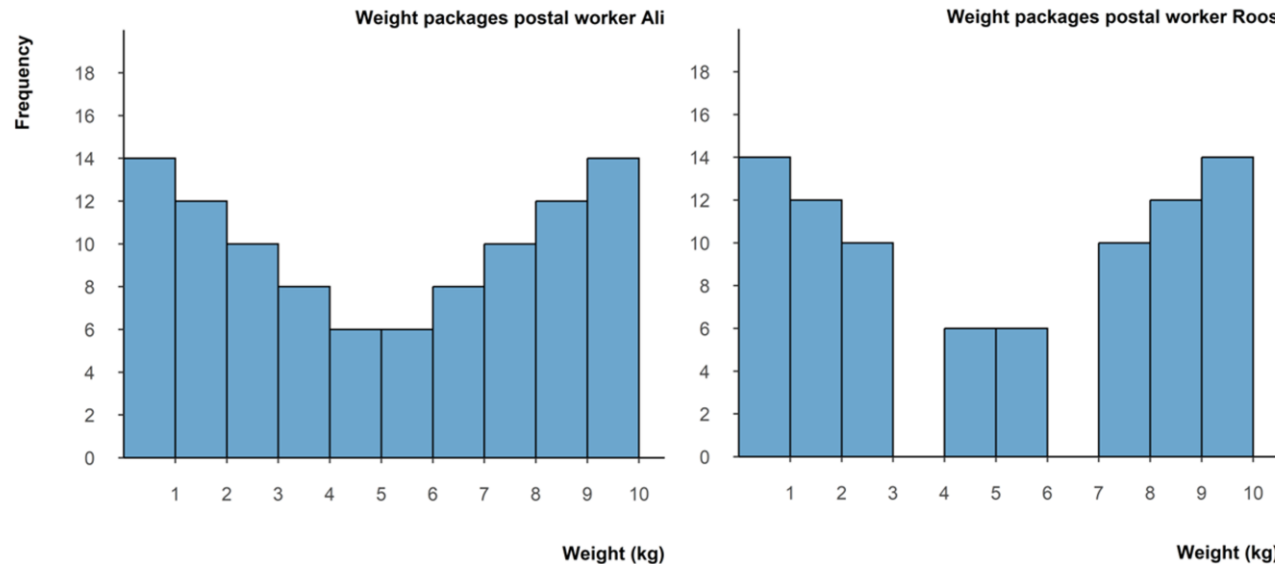


Practice



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

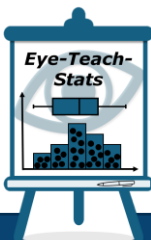
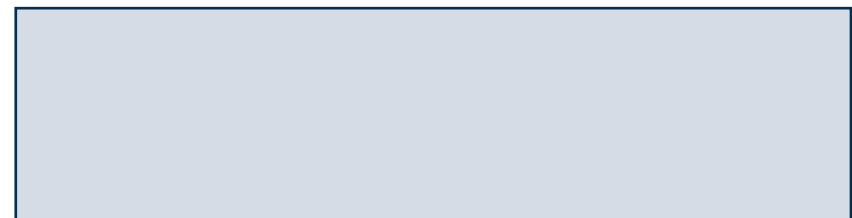


Which postal worker delivers on average the heaviest packages?

- a) Ali delivers on average the heaviest packages.
- b) Roos delivers on average the heaviest packages.
- c) The mean weight is approximately the same for both.

Imagine a student makes the same systematic error of overemphasizing local features: number of bars.

Describe the possible gaze pattern the student may show. Do you think this error occurs here? If so, why? If not, why not?





Eye-Teach-Stats Eye-Tracking Vignettes © 2026



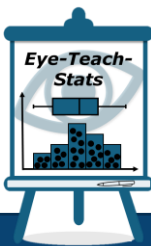
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor the granting authority can be held responsible for them.

14 Boxplot

Vignette on an error caused by area bias



Introducing situation



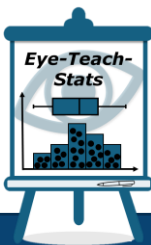
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Imagine you are teaching an 8th-grade class on data interpretation, and you introduce your students to comparing data distributions using boxplots.

You notice your student Linda is having difficulties with a task you've assigned. Linda gives an incorrect answer.

You will now see the task Linda is working on.



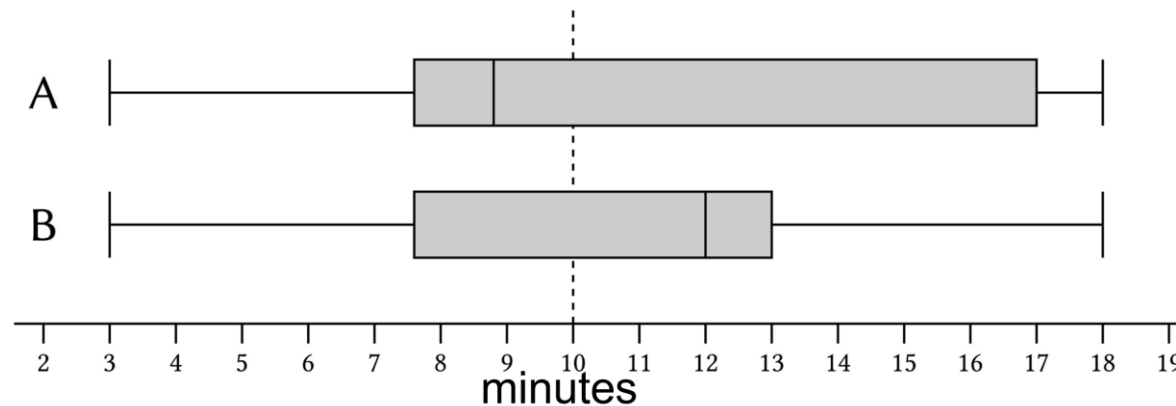
Student's task



Co-funded by
the European Union

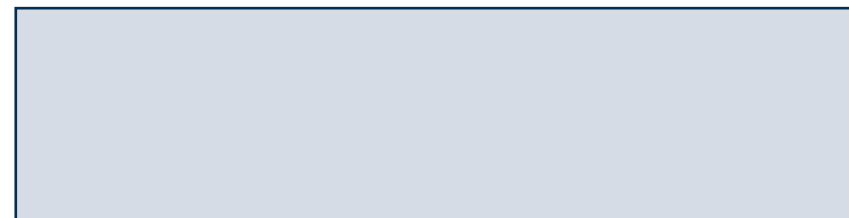
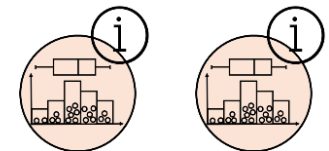
Erasmus+
Enriching lives, opening minds.

Two schools A and B are attended by the same number of students. All students have indicated in survey how many minutes they need to get to school in the morning. The results are presented by two box plots. One school (A or B) is attended by more students with a walk to school of more than 10 minutes. Decide which one.



This is the task Linda is working on.

Please complete the task on your own first and enter your solution.



Introducing eye-tracking

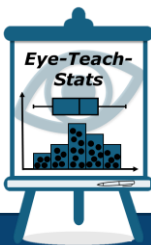


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

A group of researchers in mathematics education is pursuing the idea of supporting teachers with additional diagnostic material: To this end, eye movement data from students is recorded using eye-tracking technology and made available as supplementary material.

You will receive a heatmap that visualizes gaze data of your student Linda that was collected while Linda was working on the task. This heatmap indicates where her visual focus lies. Often it is possible to infer what Linda's thinking process might have been.



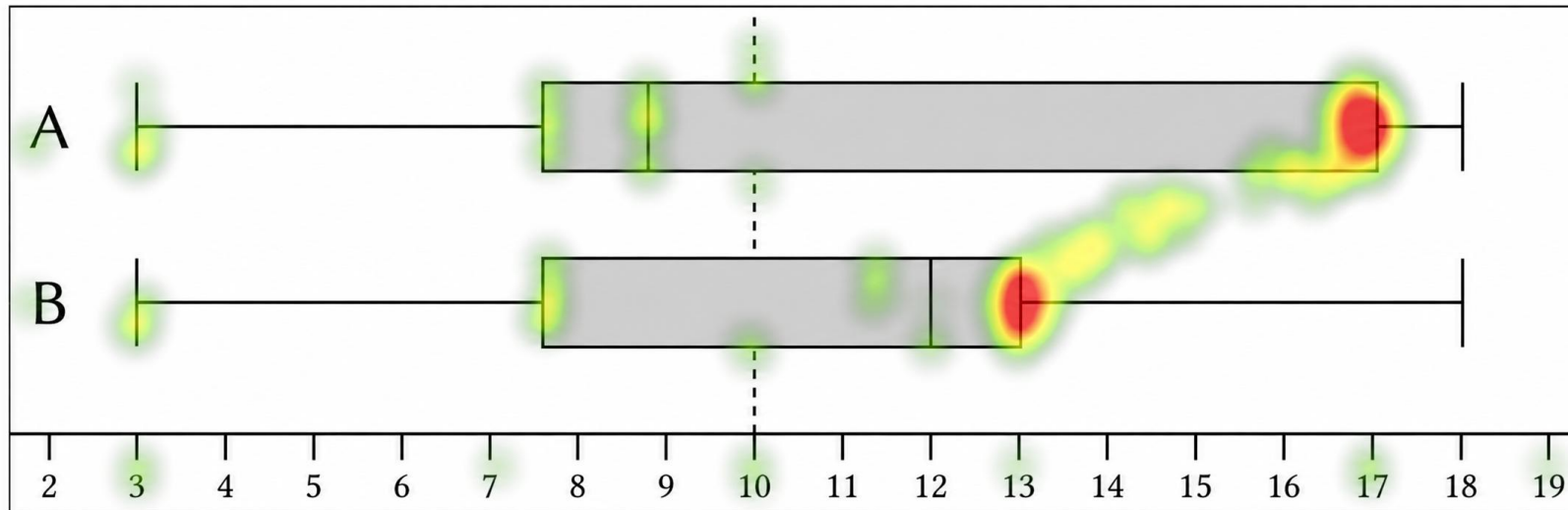
First look



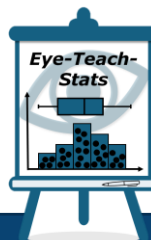
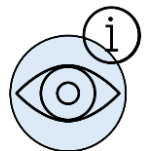
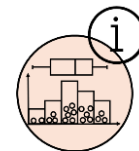
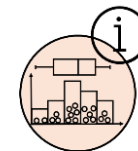
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Linda's Answer: „For me, School A is clearly the right answer. You can easily see from the box plot A that many more pupils stated that they walked more than 10 minutes to school.“



Have a look at the heatmap that shows your student Linda's gaze movement and her answer.



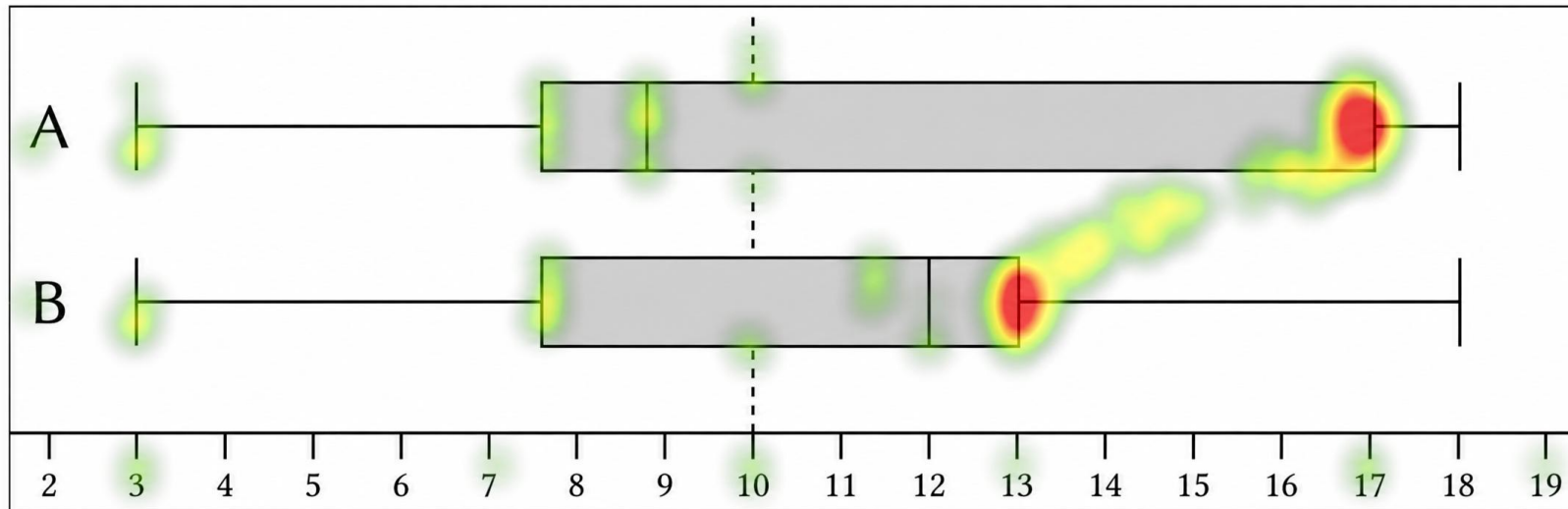
Take perspective



Co-funded by
the European Union

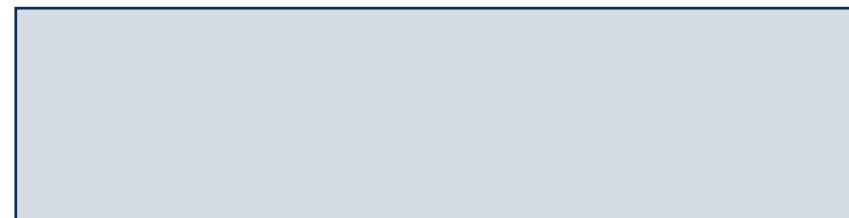
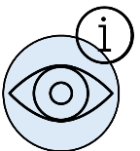
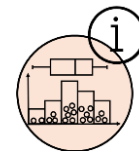
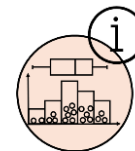
Erasmus+
Enriching lives, opening minds.

Linda's Answer: „For me, School A is clearly the right answer. You can easily see from the box plot A that many more pupils stated that they walked more than 10 minutes to school.“



After the task Linda was shown her own heatmap. She explains retrospectively what she was thinking while working on the task.

Try to take Linda's perspective, and formulate what she may comment.



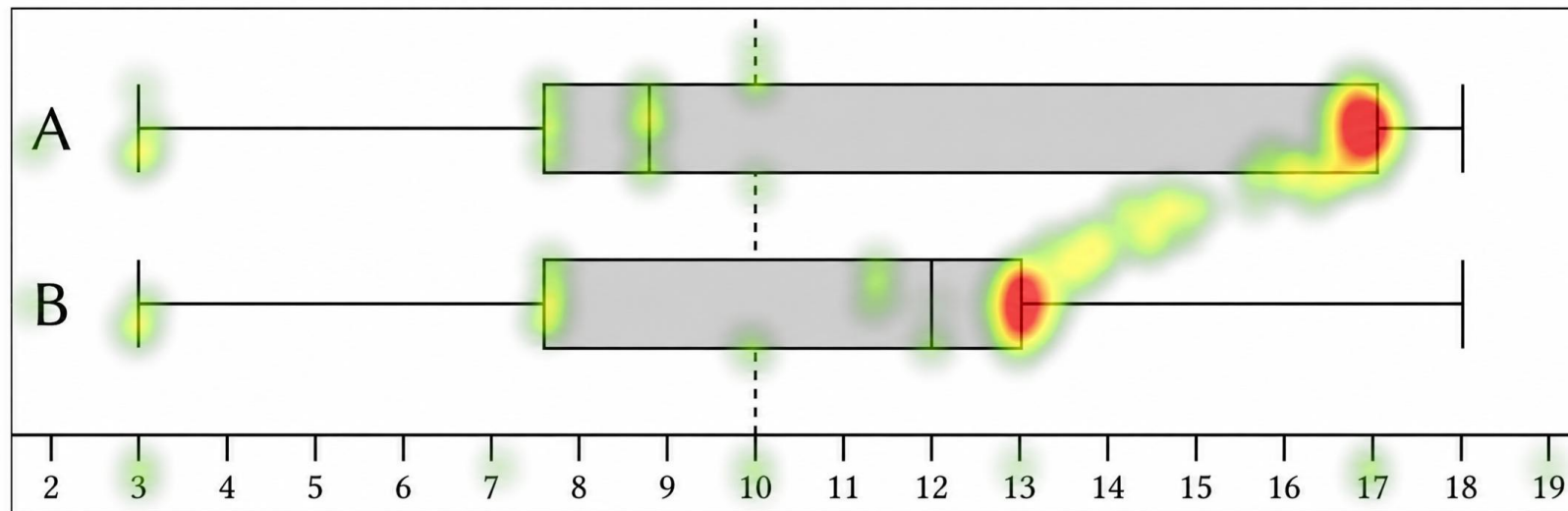
Indicate assumption



Co-funded by
the European Union

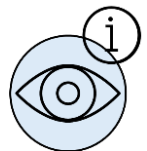
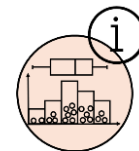
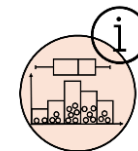
Erasmus+
Enriching lives, opening minds.

Linda's Comment: "The box plot represents all the responses, so to speak. So I checked where more of the box is above the 10-minute limit. This is quite clearly the case with box plot A. So I also opted for school A."



Now you can read the original comment of Linda.

Indicate to what extent Linda's comment agrees with what you assumed.

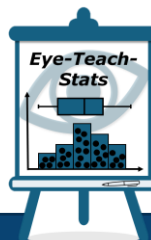


☐ No alignment at all

☐ Little alignment

☐ Quite some alignment

☐ Complete alignment



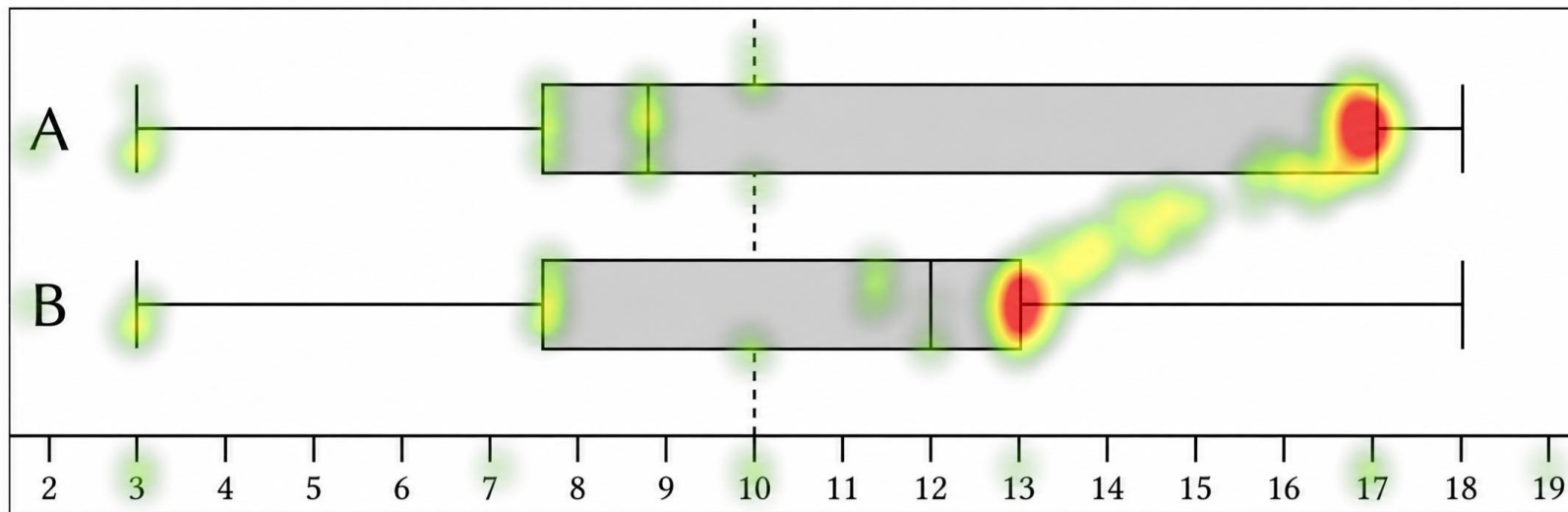
Evaluate solution



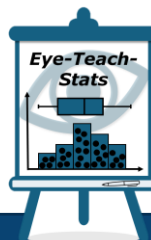
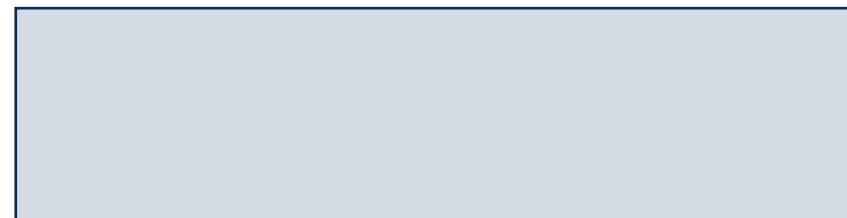
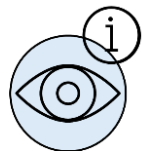
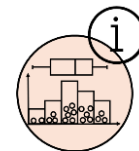
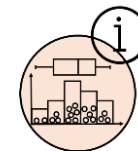
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Linda's Comment: "The box plot represents all the responses, so to speak. So I checked where more of the box is above the 10-minute limit. This is quite clearly the case with box plot A. So I also opted for school A."



Evaluate Linda's solution. What can you say about her understanding?



Introducing systematic error



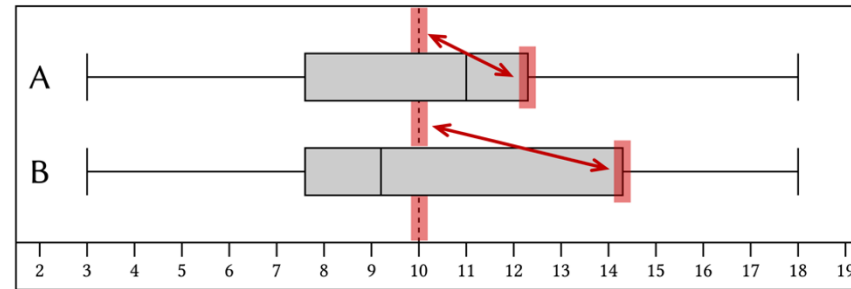
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Understanding student errors:

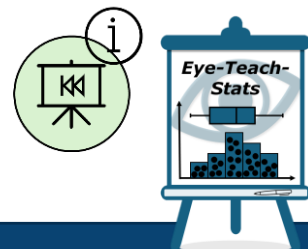
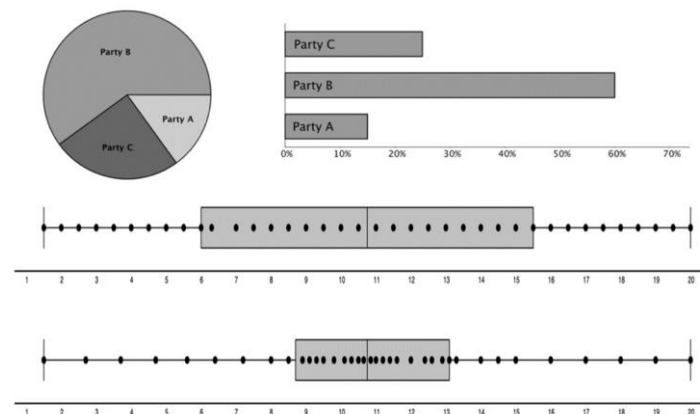
Linda has an underlying conceptual difficulty regarding the interpretation of boxplots. This leads to an observable systematic error, called **area bias**.

Observable behavior: In her comment, Linda describes that she decides the question of which of two boxplots represents more data points above a certain value based on which of the two boxplots shows more box area above this critical value. Her eye movements also show a noticeable focus on the third quartiles, which represent the right-hand boundary of the box.



Interpretation by the teacher/researcher: A plausible explanation for Linda's approach and her eye movements is that she incorrectly assumes that a larger box area in the boxplot also corresponds to a larger proportion of the represented sample. Not only Linda's comment fits this hypothesis, but also her eye movements, as the focus here is on the third quartiles, which indicate the amount of box area above the critical value.

Underlying conceptual difficulty: When learners encounter boxplots for the first time, they are usually already familiar with other types of statistical representations such as circle diagrams or bar charts. In these types of representation, Linda's assumption that there is a proportional relationship between the area and the represented sample fraction is correct. This is not the case with boxplots, however, as the box always represents the middle half of the data and the box area or box width is therefore a measure of the variability of the data set and indicates how much the middle half of the data scatters (interquartile range). If learners inappropriately transfer their prior knowledge of other statistical representations when interpreting boxplots, then the systematic error occurs that you have just observed with your student Linda.



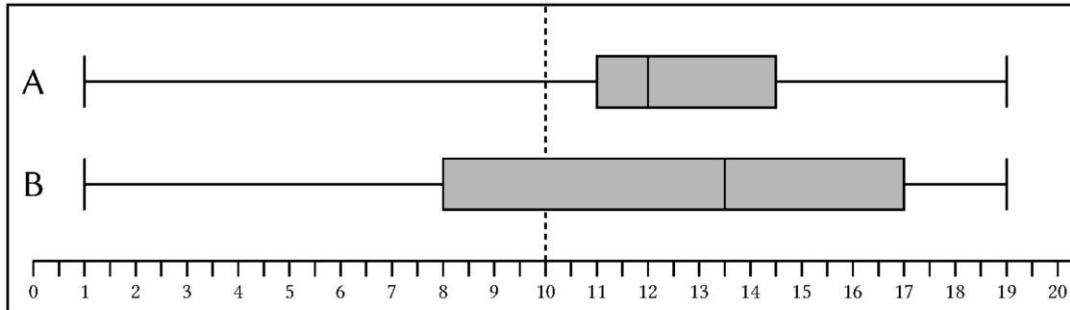
Practice



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Decide in which boxplot a higher percentage of the sample is above 10.



Three students described a different reasoning when solving the task:

Kira: “The quartiles split the data into four parts with about the same number of values. Thus, the position of the quartiles you can see in which areas the different quarters of the sample are located.”

Lionel: “The range shows how spread out all the data points are, but the box only includes the middle half of the data. By looking at where the box is placed, you can see where most of the data points are located.”

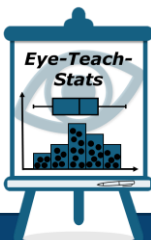
Tom: “The box represents the central data points, and we can use the box to make statements about the distribution of the sample, because depending on how the data points are arranged, the size of the box changes.”

Which of these arguments leads most likely to an incorrect answer?

☐ Kira's

☐ Lionel's

☐ Tom's



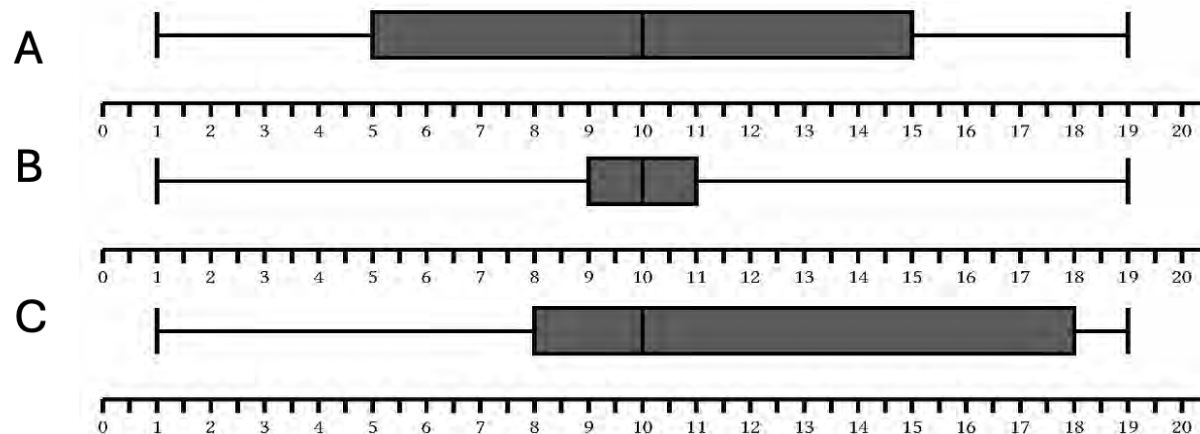
Practice II



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

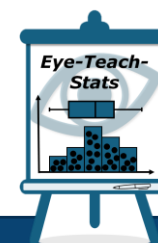
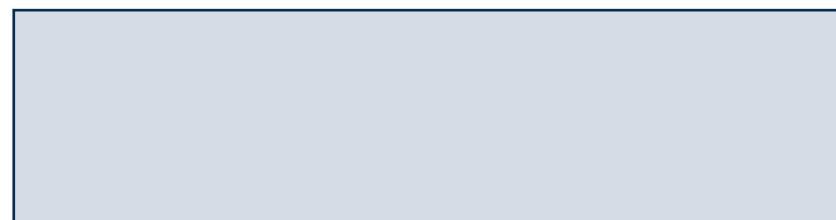
Amy has conceptual difficulties with the meaning of the box in a boxplot, as she appears to associate the size of the box with the frequency of data points. For each of the following three statements, she decided whether it was true or false. Predict Amy's response to each statement and briefly justify your prediction.



Statement 1: In all three Boxplots, the portion of data points above and below 10 is the same.

Statement 2: In boxplot A, the data points are clustered around the middle, while fewer data points are located toward the ends.

Statement 3: Boxplot A represents more data points than boxplot B.





Eye-Teach-Stats Eye-Tracking Vignettes © 2026



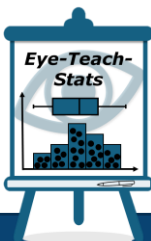
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor the granting authority can be held responsible for them.

15 Boxplot

Vignette on an error caused by area bias



Introducing situation



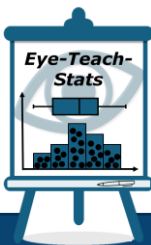
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Imagine you are teaching an 8th-grade class on data interpretation, and you introduce your students to comparing data distributions using boxplots.

You notice your student Paul is having difficulties with a task you've assigned. Paul gives an incorrect answer.

You will now see the task Paul is working on.



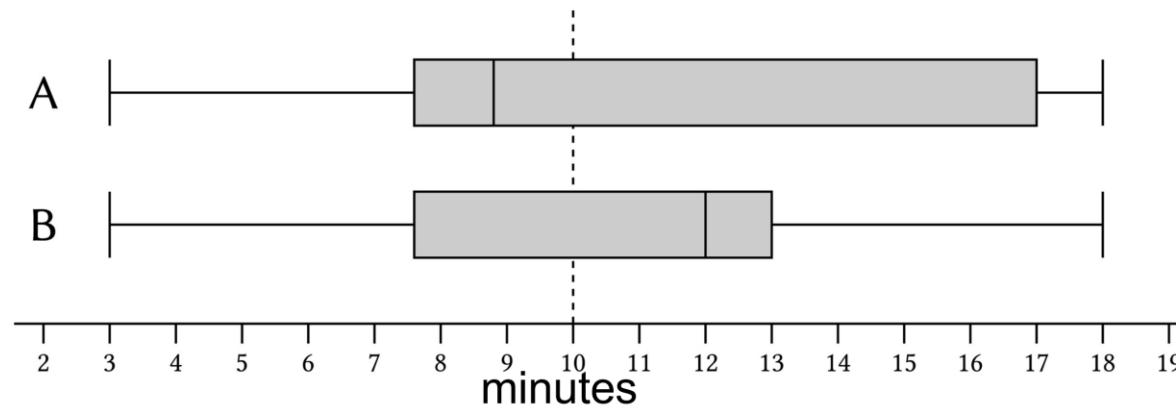
Student's task



Co-funded by
the European Union

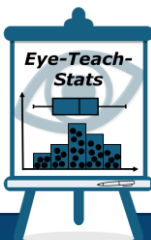
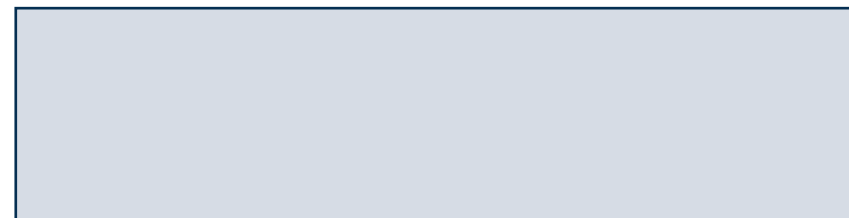
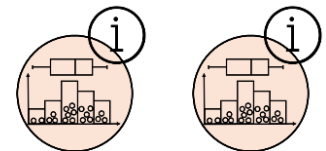
Erasmus+
Enriching lives, opening minds.

Two schools A and B are attended by the same number of students. All students have indicated in survey how many minutes they need to get to school in the morning. The results are presented by two box plots. One school (A or B) is attended by more students with a walk to school of more than 10 minutes. Decide which one.



This is the task Paul is working on.

Please complete the task on your own first and enter your solution.



Introducing eye-tracking

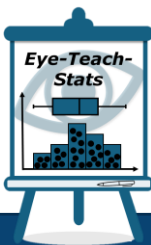


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

A group of researchers in mathematics education is pursuing the idea of supporting teachers with additional diagnostic material: To this end, eye movement data from students is recorded using eye-tracking technology and made available as supplementary material.

You will receive a gaze plot of your student Paul that was collected while Paul was working on the task. This gaze plot indicates where Paul's visual focus lies. Often it is possible to infer what your student's thinking process might have been.



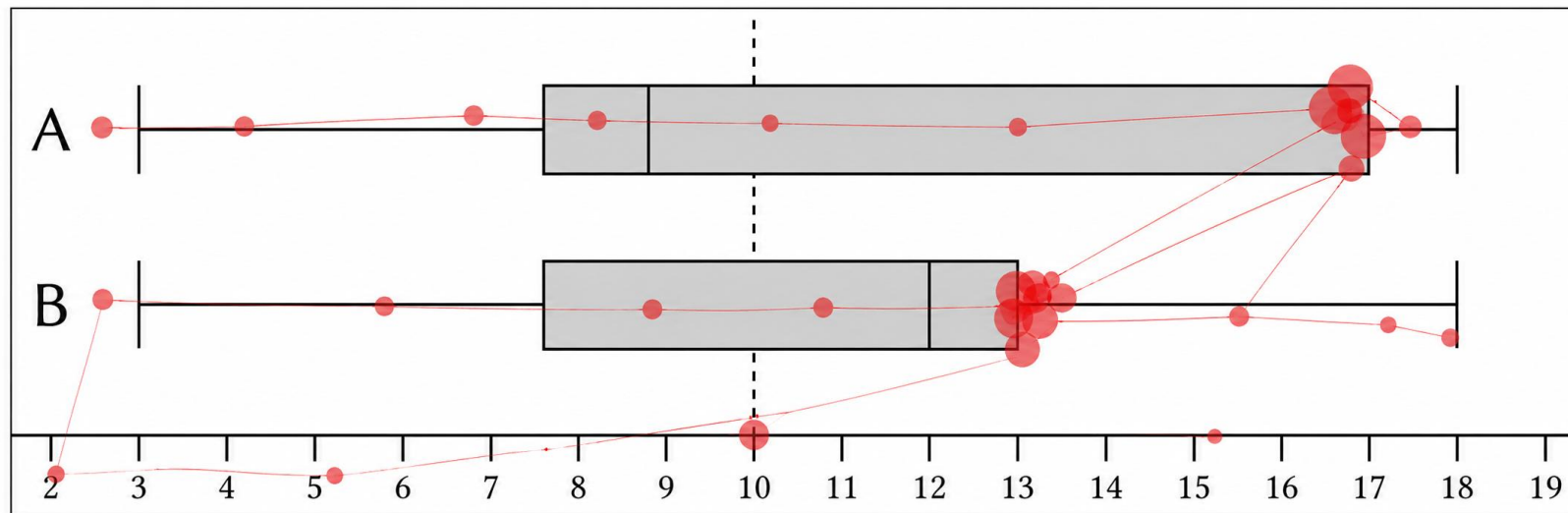
First look



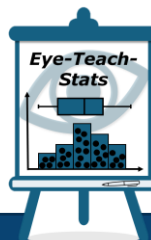
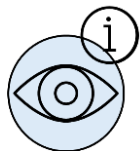
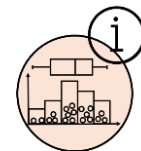
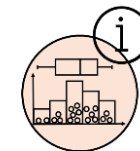
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Paul's Answer: „For me, School A is clearly the right answer. You can easily see from the box plot A that many more pupils stated that they walked more than 10 minutes to school.“



Have a look at the gaze plot that shows your student Paul's gaze movement and his answer.



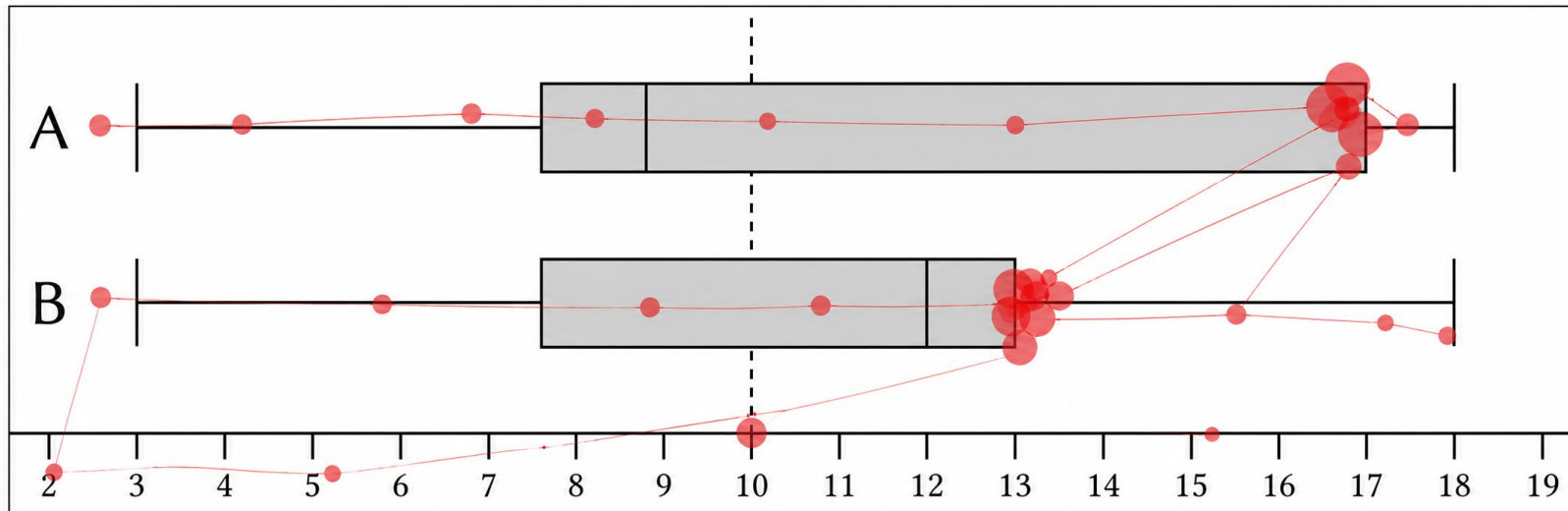
Take perspective



Co-funded by
the European Union

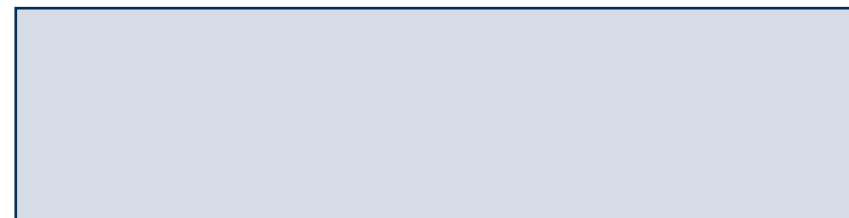
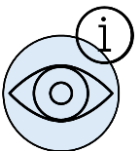
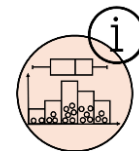
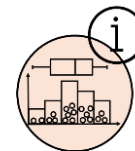
Erasmus+
Enriching lives, opening minds.

Paul's Answer: „For me, School A is clearly the right answer. You can easily see from the box plot A that many more pupils stated that they walked more than 10 minutes to school.“



After the task Paul was shown his own gaze plot. He explains retrospectively what he was thinking while working on the task.

Try to take Paul's perspective, and formulate what he may comment.



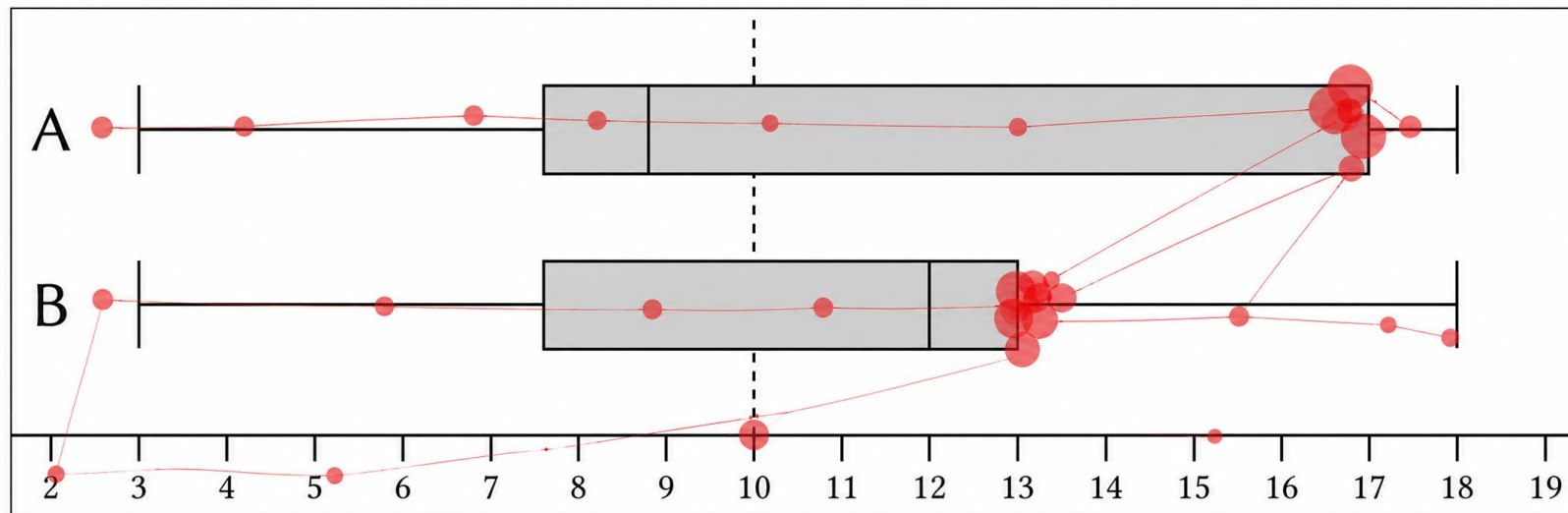
Indicate assumption



Co-funded by
the European Union

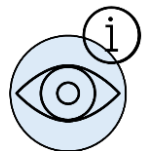
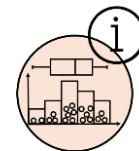
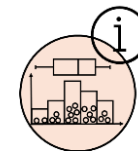
Erasmus+
Enriching lives, opening minds.

Paul's Comment: "The box plot represents all the responses, so to speak. So I checked where more of the box is above the 10-minute limit. This is quite clearly the case with box plot A. So I also opted for school A."



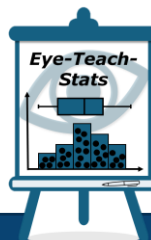
Now you can read the original comment of Paul.

Indicate to what extent Paul's comment agrees with what you assumed.



- ☐ No alignment at all
- ☐ Little alignment

- ☐ Quite some alignment
- ☐ Complete alignment



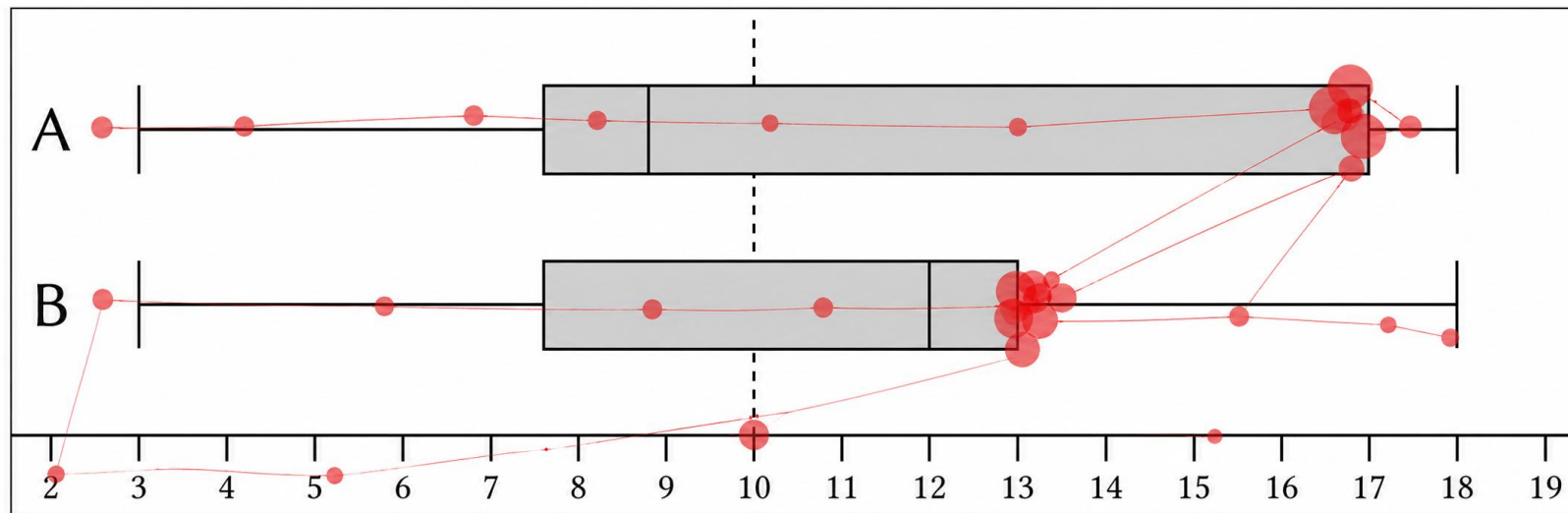
Evaluate solution



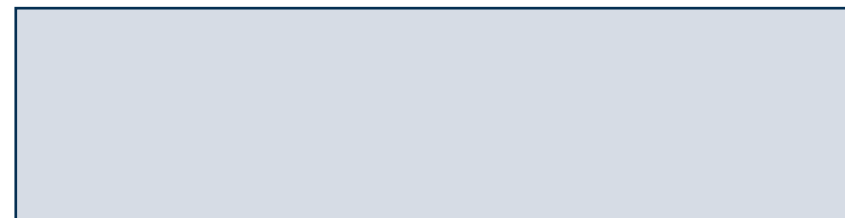
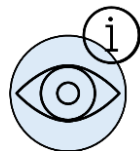
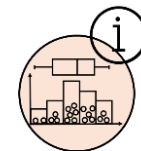
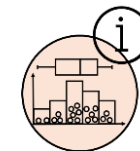
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Paul's Comment: "The box plot represents all the responses, so to speak. So I checked where more of the box is above the 10-minute limit. This is quite clearly the case with box plot A. So I also opted for school A."



Evaluate Paul's solution. What can you say about his understanding?



Introducing systematic error



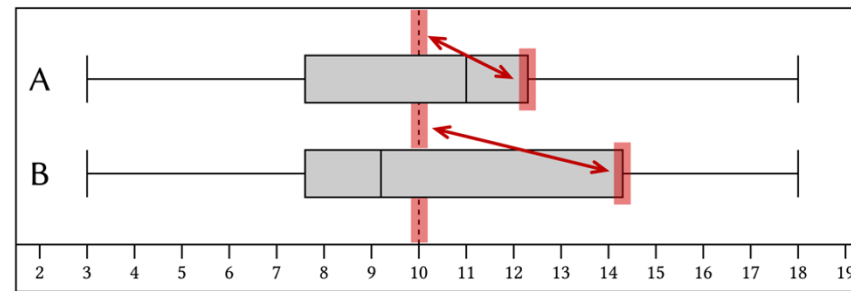
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Understanding student errors:

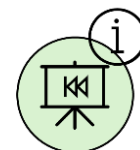
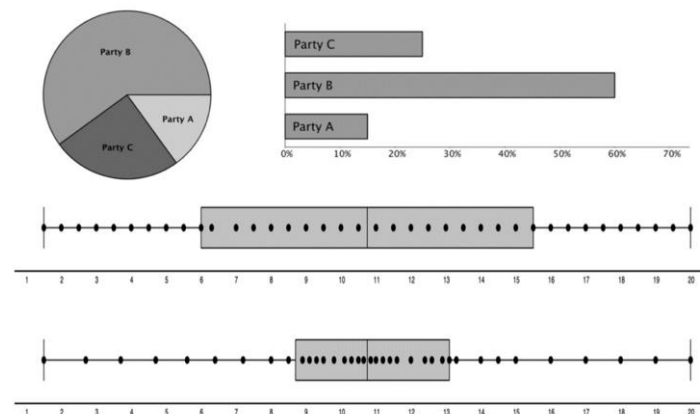
Paul has an underlying conceptual difficulty regarding the interpretation of boxplots. This leads to an observable systematic error, called **area bias**.

Observable behavior: In his comment, Paul describes that he decides the question of which of two boxplots represents more data points above a certain value based on which of the two boxplots shows more box area above this critical value. His eye movements also show a noticeable focus on the third quartiles, which represent the right-hand boundary of the box.



Interpretation by the teacher/researcher: A plausible explanation for Paul's approach and his eye movements is that he incorrectly assumes that a larger box area in the box plot also corresponds to a larger proportion of the represented sample. Not only Paul's comment fits this hypothesis, but also his eye movements, as the focus here is on the third quartiles, which indicate the amount of box area above the critical value.

Underlying conceptual difficulty: When learners encounter box plots for the first time, they are usually already familiar with other types of statistical representations such as circle diagrams or bar charts. In these types of representation, Paul's assumption that there is a proportional relationship between the area and the represented sample fraction is correct. This is not the case with box plots, however, as the box always represents the middle half of the data and the box area or box width is therefore a measure of the variability of the data set and indicates how much the middle half of the data scatters (interquartile range). If learners inappropriately transfer their prior knowledge of other statistical representations when interpreting box plots, then the systematic error occurs that you have just observed with your student Paul.



Practice

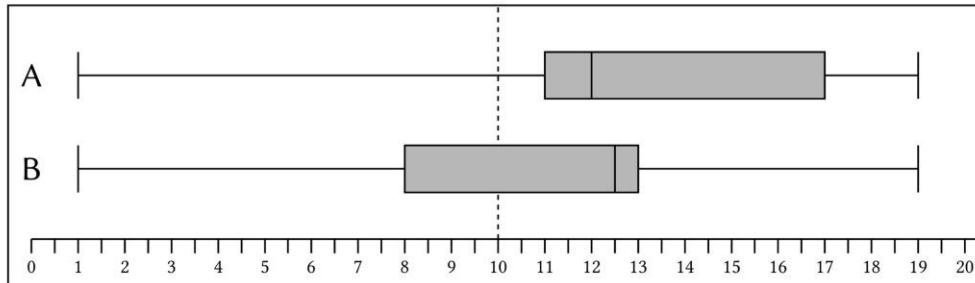


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

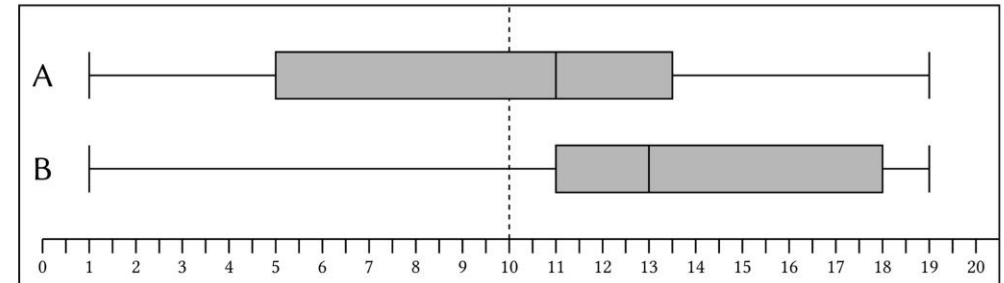
Suppose you suspect that one of your students has conceptual difficulties when interpreting the box area. Choose one of the following items, that you would let your student solve to diagnose this possible conceptual difficulty?

A



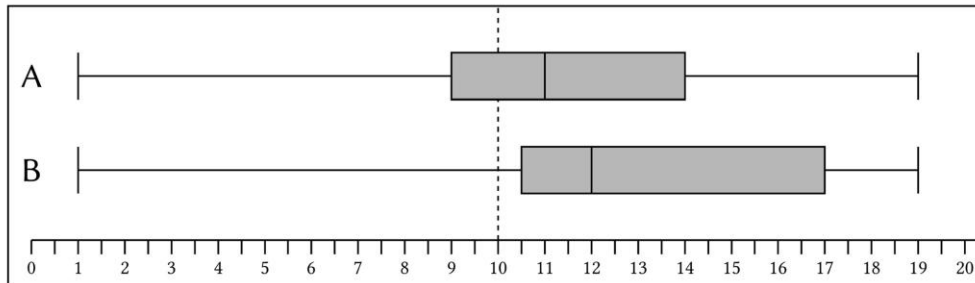
Decide in which boxplot a higher percentage of the sample is above 10.

B



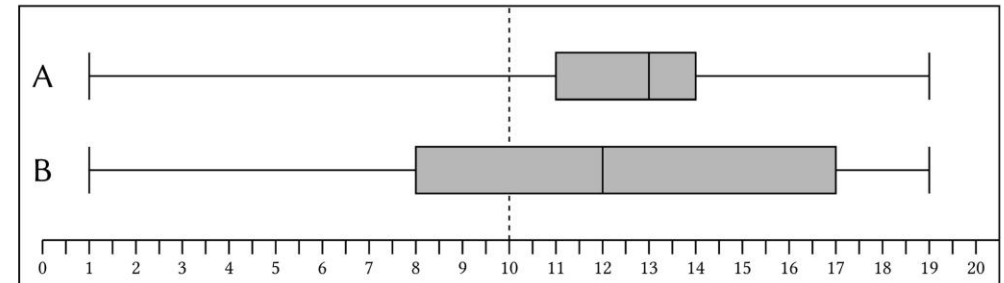
Decide in which boxplot a higher percentage of the sample is above 10.

C



Decide in which boxplot a higher percentage of the sample is above 10.

D



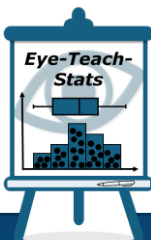
Decide in which boxplot a higher percentage of the sample is above 10.

☐ A

☐ B

☐ C

☐ D



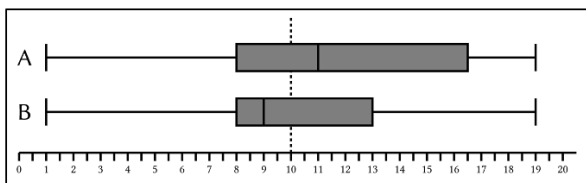
Practice II



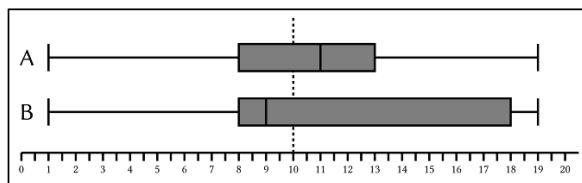
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

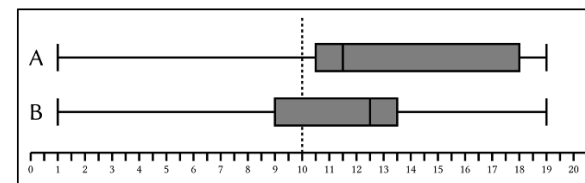
Your student Tom answered the following six items, sometimes correctly sometimes incorrectly.



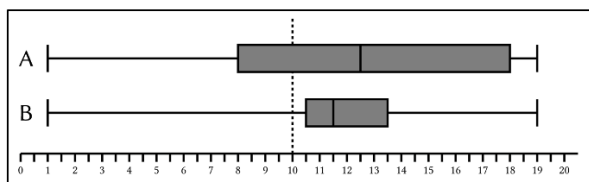
✓ Tom answered correctly (A)



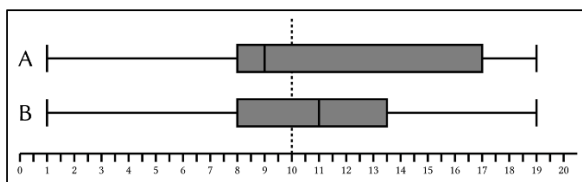
✗ Tom answered incorrectly (B)



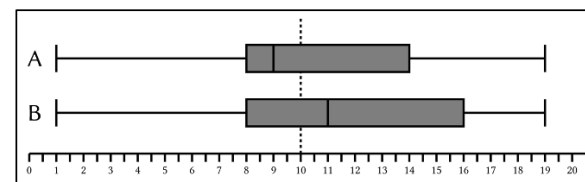
✓ Tom answered correctly (A)



✓ Tom answered correctly (B)

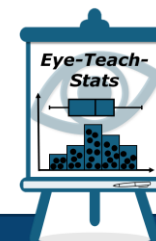
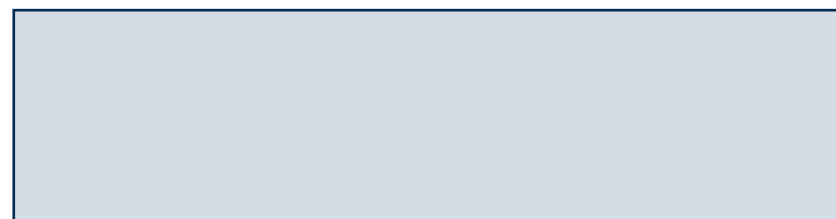


✗ Tom answered incorrectly (A)



✓ Tom answered correctly (B)

What does Tom's response suggest about his understanding of the meaning of the box (area) in boxplots, including possible conceptual difficulties that may still remain?





Eye-Teach-Stats Eye-Tracking Vignettes © 2026



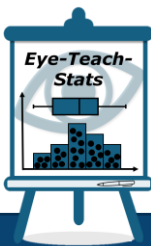
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor the granting authority can be held responsible for them.

16 Boxplot

Vignette on an error caused by median bias



Introducing situation



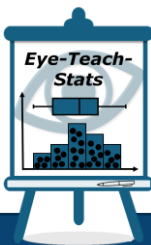
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Imagine you are teaching an 8th-grade class on data interpretation, and you introduce your students to comparing data distributions using box plots.

You notice your student Katharina is having difficulties with a task you've assigned. Katharina gives an incorrect answer.

You will now see the task Katharina is working on.



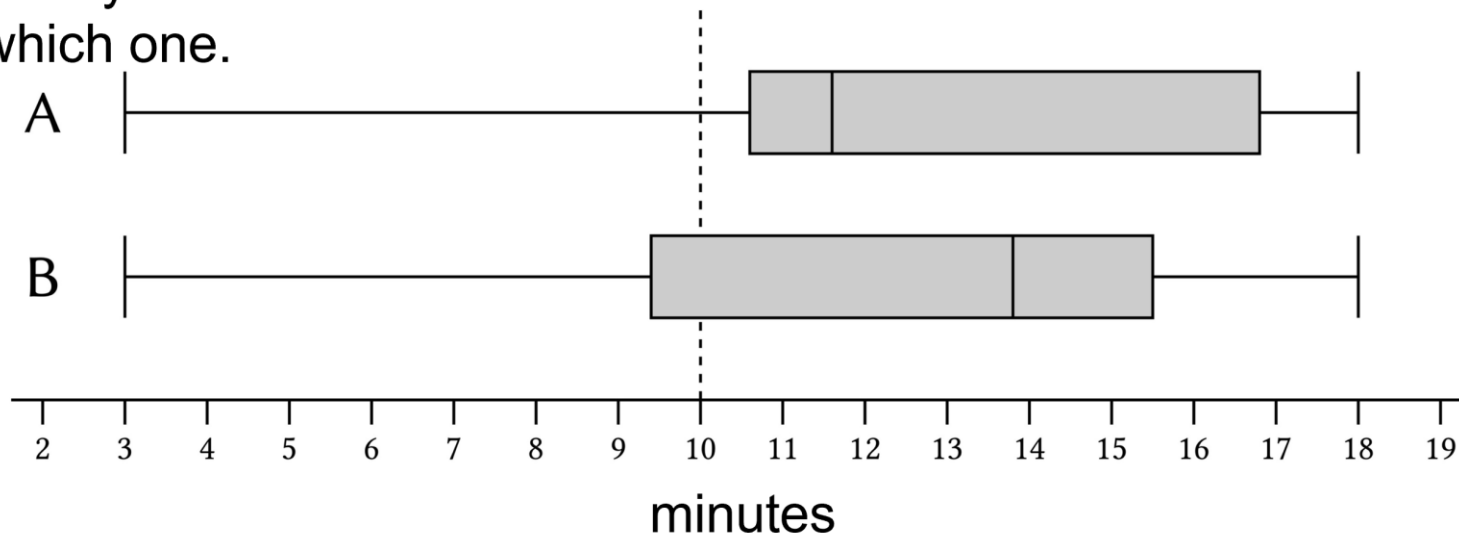
Student's task



Co-funded by
the European Union

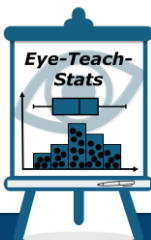
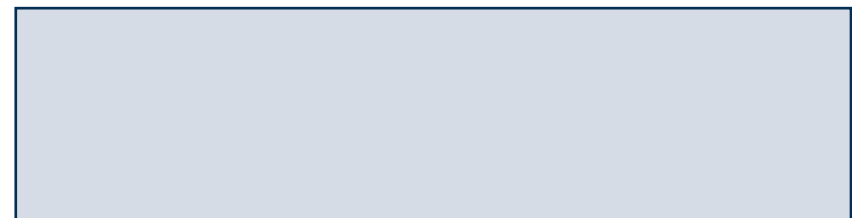
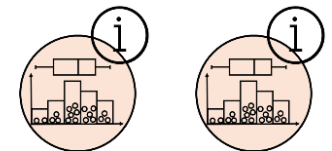
Erasmus+
Enriching lives, opening minds.

Two schools A and B are attended by the same number of students. All students have indicated in survey how many minutes they need to get to school in the morning. The results are presented by two box plots. One school (A or B) is attended by more students with a walk to school of more than 10 minutes. Decide which one.



This is the task Katharina is working on.

Please complete the task on your own first and enter your solution.



Introducing eye-tracking

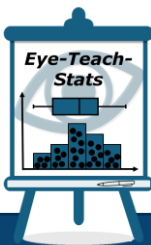


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

A group of researchers in mathematics education is pursuing the idea of supporting teachers with additional diagnostic material: To this end, eye movement data from students is recorded using eye-tracking technology and made available as supplementary material.

You will receive a gaze video that visualizes fixation points of your student Katharina that was collected while she was working on the task. This video indicates where her visual focus lies. Often it is possible to infer what your student's thinking process might have been.



First look

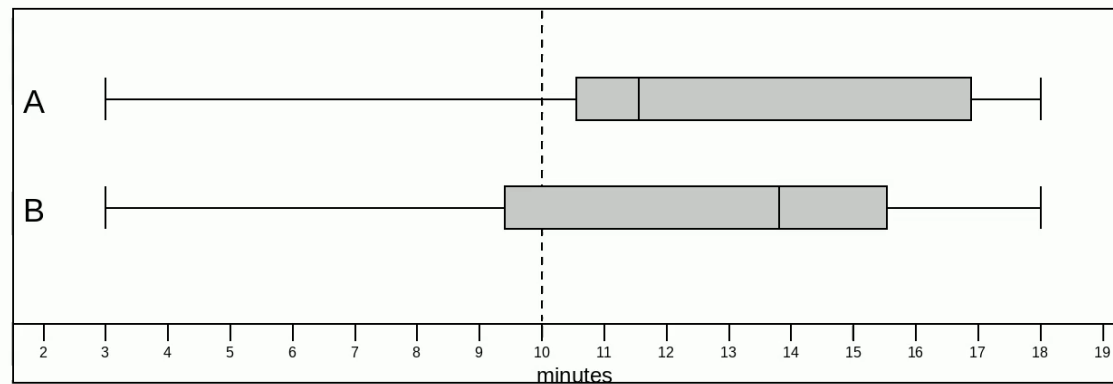


Co-funded by
the European Union

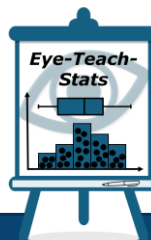
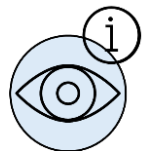
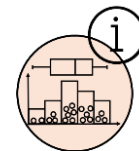
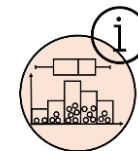
Erasmus+
Enriching lives, opening minds.

Katharina's Answer: „For me, School B is clearly the right answer. You can easily see from the boxplot B that many more pupils stated that they walked more than 10 minutes to school.“

One school has a larger number of children with a walk to school longer than 10 minutes.
Decide which one.



Have a look at the gaze video that shows your student Katharina's gaze movement and her answer.



Take perspective

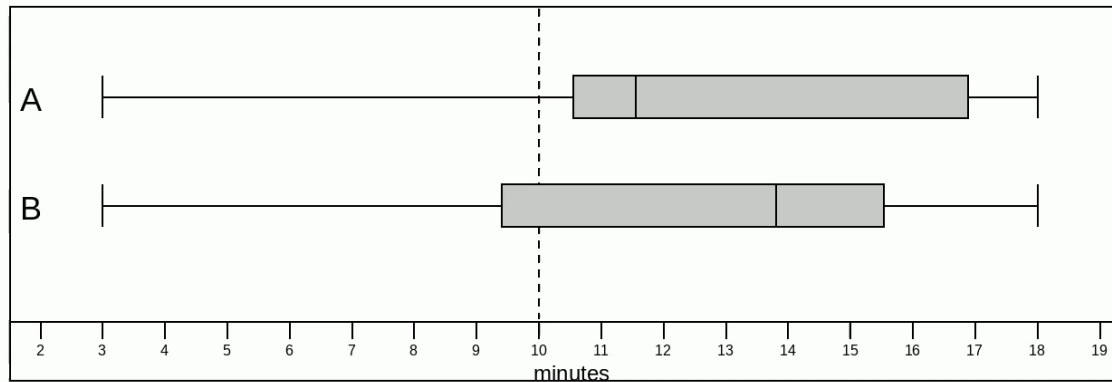


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

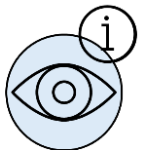
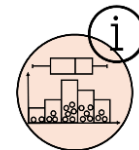
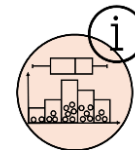
Katharina's Answer: „For me, School B is clearly the right answer. You can easily see from the boxplot B that many more pupils stated that they walked more than 10 minutes to school.“

One school has a larger number of children with a walk to school longer than 10 minutes.
Decide which one.



After the task Katharina was shown her own gaze video. She explains retrospectively what she was thinking while working on the task.

Try to take Katharina's perspective, and formulate what she may comment.



Indicate assumption

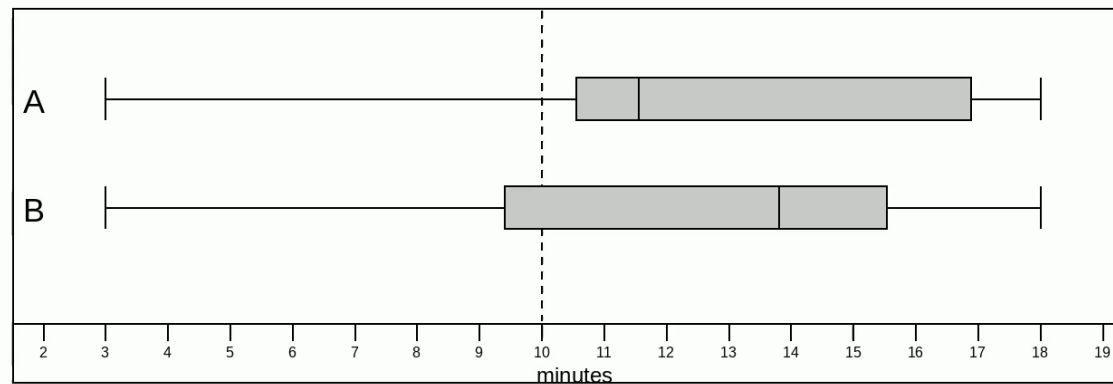


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Katharina's Comment: "The line in the middle of the box is the median, which is an average value. So I mainly compared the two medians. You can clearly see that the children from school B take much longer on average to get to school than those from school A. So I selected B."

One school has a larger number of children with a walk to school longer than 10 minutes.
Decide which one.

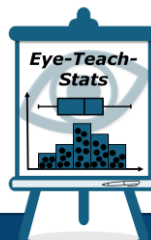
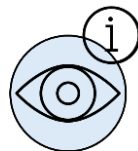
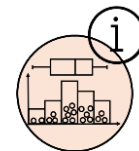
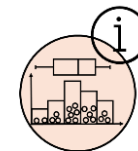


Now you can read the original comment of Katharina.

Indicate to what extent Katharina's comment agrees with what you assumed.

- ☐ No alignment at all
- ☐ Little alignment

- ☐ Quite some alignment
- ☐ Complete alignment



Evaluate solution

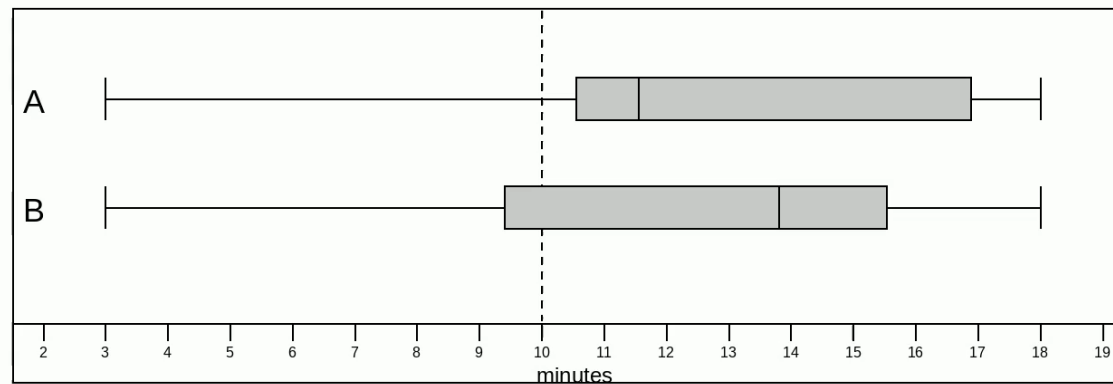


Co-funded by
the European Union

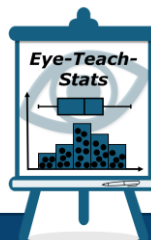
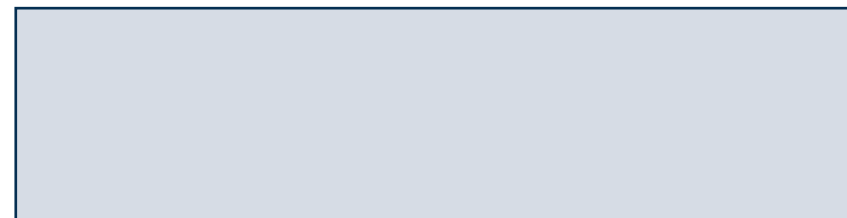
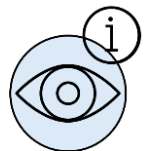
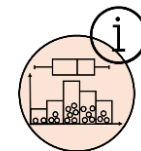
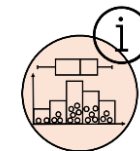
Erasmus+
Enriching lives, opening minds.

Katharina's Comment: "The line in the middle of the box is the median, which is an average value. So I mainly compared the two medians. You can clearly see that the children from school B take much longer on average to get to school than those from school A. So I selected B."

One school has a larger number of children with a walk to school longer than 10 minutes.
Decide which one.



Evaluate Katharina's solution. What can you say about her understanding?



Introducing systematic error



Co-funded by
the European Union

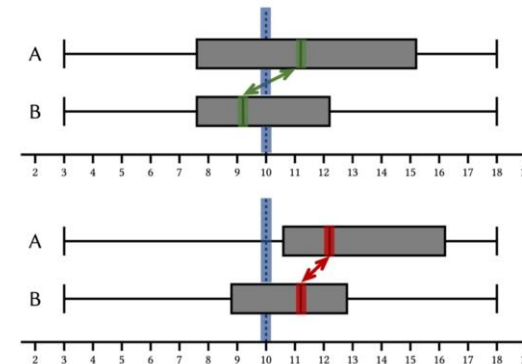
Erasmus+
Enriching lives, opening minds.

Understanding student errors:

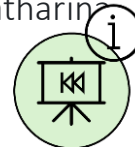
Katharina has an underlying conceptual difficulty regarding the interpretation of boxplots. This leads to an observable systematic error, called median bias.

Observable behavior: In her comment, Katharina describes that she decides the question of which of two boxplots represents more data points above a certain value based on which of the two boxplots has the higher median. Her eye movements also show a clear focus on the two medians.

Interpretation by the teacher/researcher: A plausible explanation for Katharina's approach and her eye movements is that she does not interpret the median primarily as a cut-off point that indicates where the boundary between the upper and lower half of the data is, but uses the median to decide which of the two statistical distributions tends to have the higher values and uses this central tendency for her decision. Her eye movements, which indicate that she is comparing the two medians directly with each other and not their position relative to the critical value, also fit with this assumption.



Underlying conceptual difficulty: When learners encounter box plots for the first time, they are usually already familiar with averages such as the median. In particular, mean values are not only familiar from school contexts, but also occur in different contexts outside of school. If learners identify the (well-known) median in the boxplot representation, it is plausible to assume that they use this familiar parameter to make decisions especially when they have not (yet) acquired the boxplot-specific knowledge that would be necessary for a correct solution. This overgeneralization of the median then leads to the systematic error that you have just observed with your student Katharina.



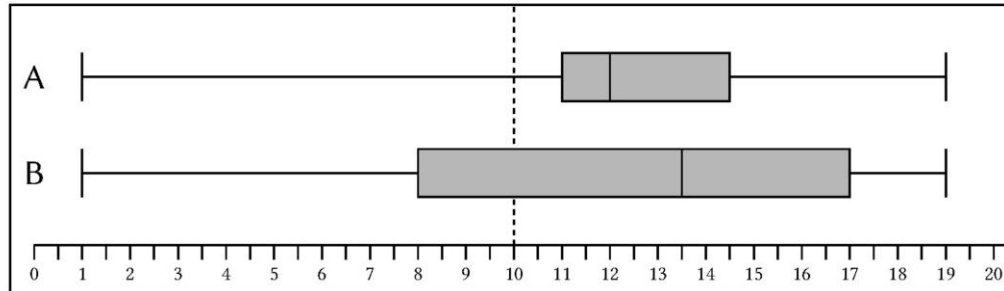
Practice



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Decide in which boxplot a higher percentage of the sample is above 10.



Three students described a different reasoning when solving the task:

Tom: “The quartiles split the data points into roughly four equal parts. You can therefore see from the position of the quartiles in which area of the data set which percentage of the data points falls.”

Lionel: “The median shows the typical value of a sample. Thus, it is a way of estimating how high the values in a sample are and you can compare the medians to estimate where the values tend to lie further to the right.”

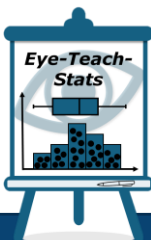
Kira: “While the range is the whole sample, the box stands for the middle half of the data. If you look at the position of the box, you can see how far to the right or left the middle half of the data lies.”

Which of these arguments leads most likely to an incorrect answer?

☐ Tom's

☐ Lionel's

☐ Kira's



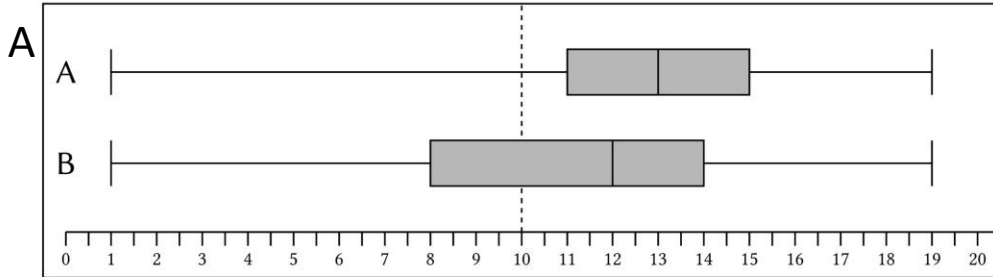
Practice II



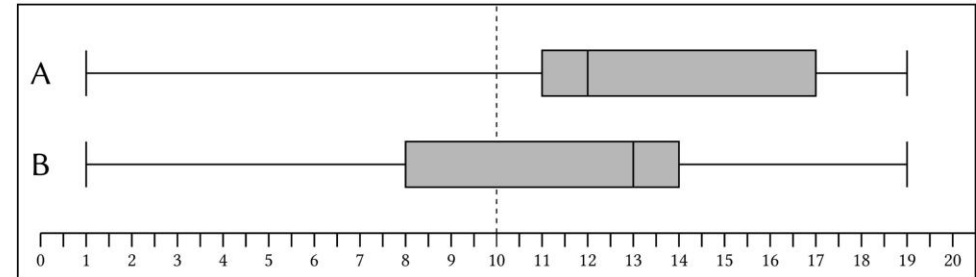
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

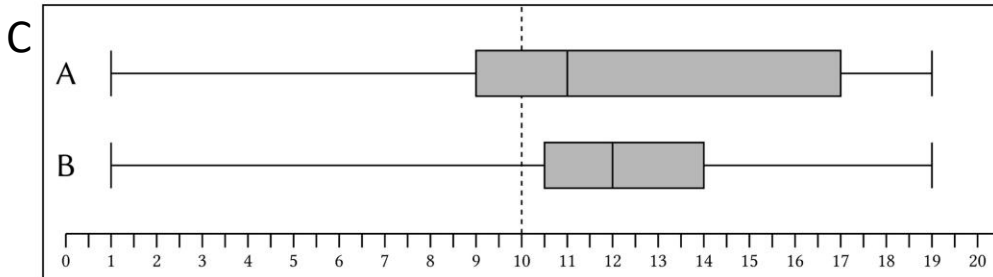
Suppose you suspect that one of your students has conceptual difficulties when interpreting the median. Choose one of the following items, that you would let your student solve to diagnose this possible conceptual difficulty?



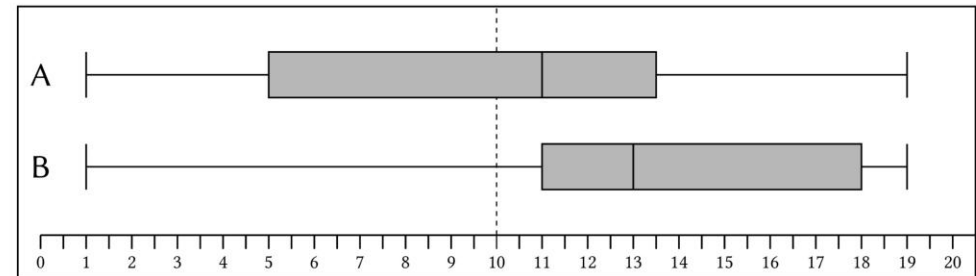
Decide in which boxplot a higher percentage of the sample is above 10.



Decide in which boxplot a higher percentage of the sample is above 10.



Decide in which boxplot a higher percentage of the sample is above 10.



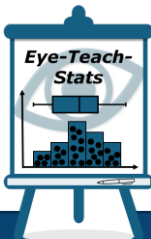
Decide in which boxplot a higher percentage of the sample is above 10.

☐ A

☐ B

☐ C

☐ D





Eye-Teach-Stats Eye-Tracking Vignettes © 2026



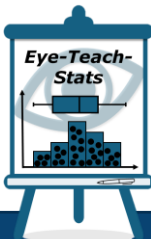
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor the granting authority can be held responsible for them.

17 Histogram

Vignette on an error caused by frequency value confusion



Introducing situation



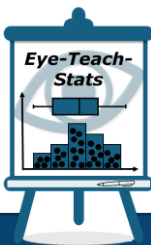
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Imagine you are teaching a 10th-grade classroom on data interpretation, and you introduce your students to find the approximate mean of weights represented in a histogram.

You notice your student Fabian is having difficulties with a task you've assigned. He gives an incorrect answer.

You will now see the task Fabian is working on.



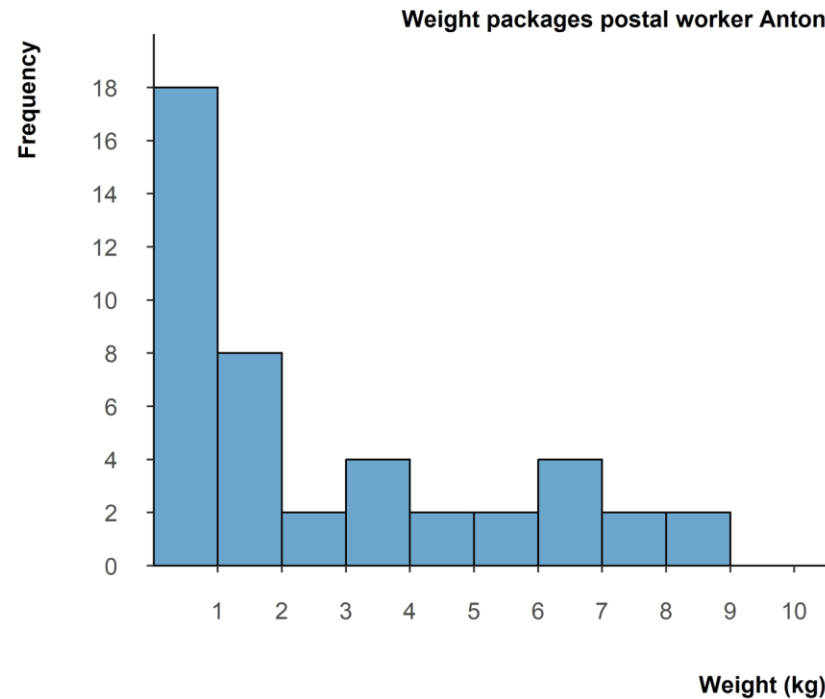
Student's task



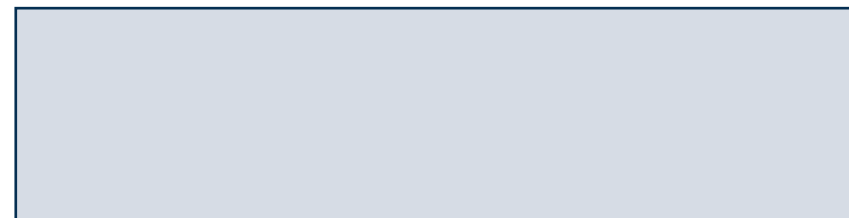
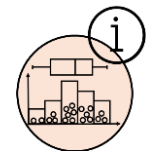
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

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



This is the task Fabian is working on. Please complete the task on your own first and enter your solution.



Introducing eye-tracking

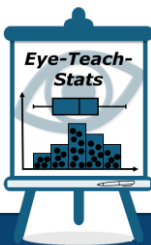


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

A group of researchers in mathematics education is pursuing the idea of supporting teachers with additional diagnostic material: To this end, eye movement data from students is recorded using eye-tracking technology and made available as supplementary material.

You will receive a gaze video that visualizes gaze data of your student Fabian that was collected while he was working on the task. The gaze data indicates where his visual focus lies. Often it is possible to infer what Fabian's thinking process might have been.



First look

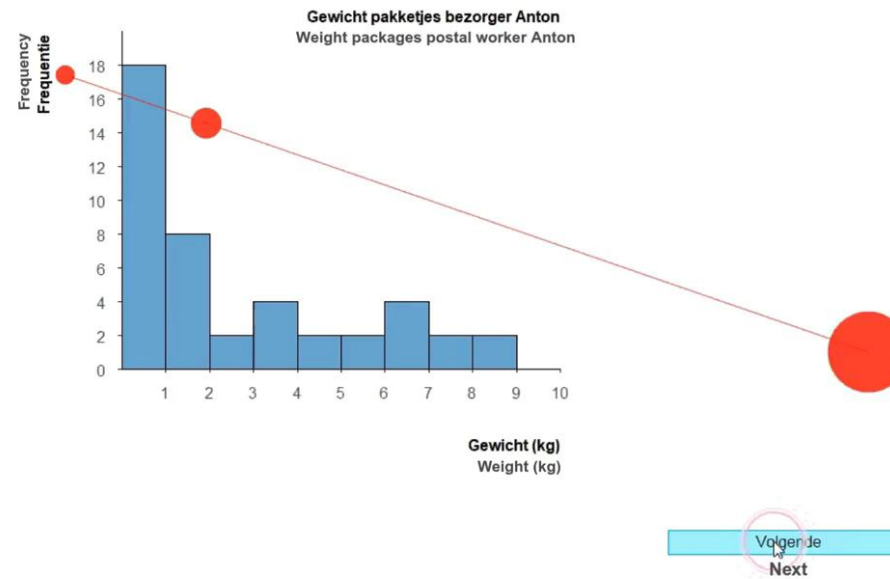


Co-funded by
the European Union

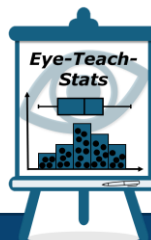
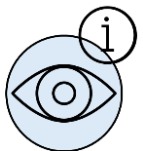
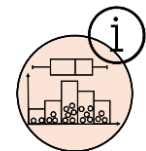
Erasmus+
Enriching lives, opening minds.

Fabian's Answer: "The arithmetic mean is about six."

What is approximately the mean weight of the packages Anton delivers?
Wat is ongeveer het gemiddelde gewicht van de pakketjes die Anton bezorgt?



Have a look at the gaze video that shows your student Fabian's gaze and his answer.



Take perspective

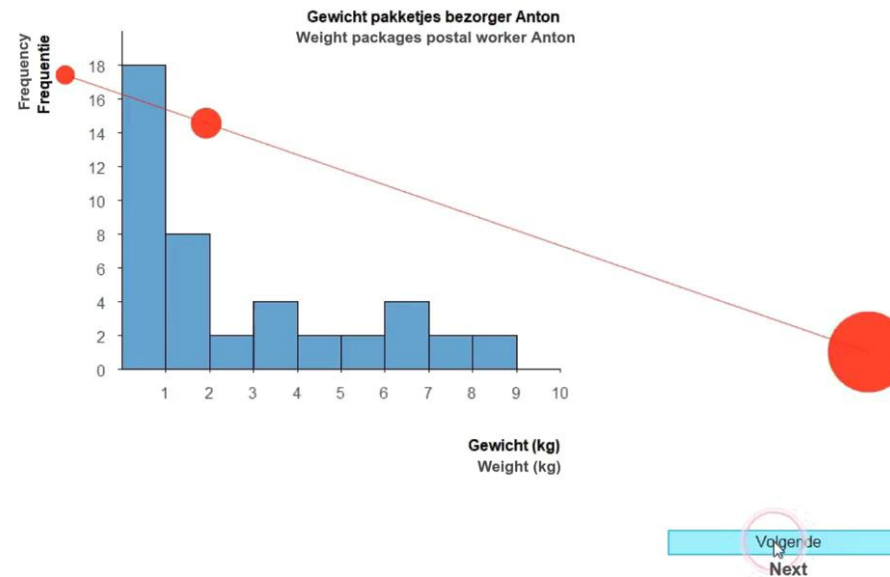


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

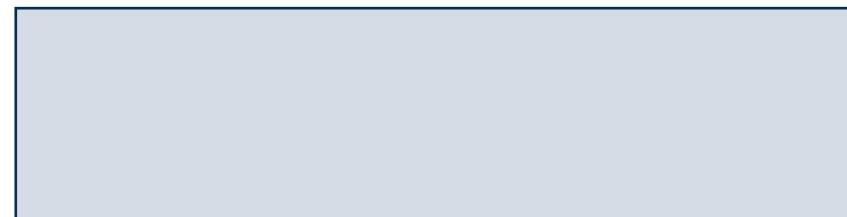
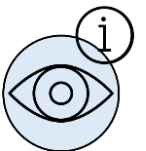
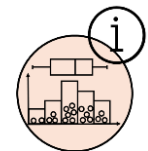
Fabian's Answer: "The arithmetic mean is about six."

What is approximately the mean weight of the packages Anton delivers?
Wat is ongeveer het gemiddelde gewicht van de pakketjes die Anton bezorgt?



After the task Fabian was shown his own gaze video. He explains retrospectively what he was thinking while working on the task.

Try to take Fabian's perspective, and formulate what he may comment.



Indicate assumption

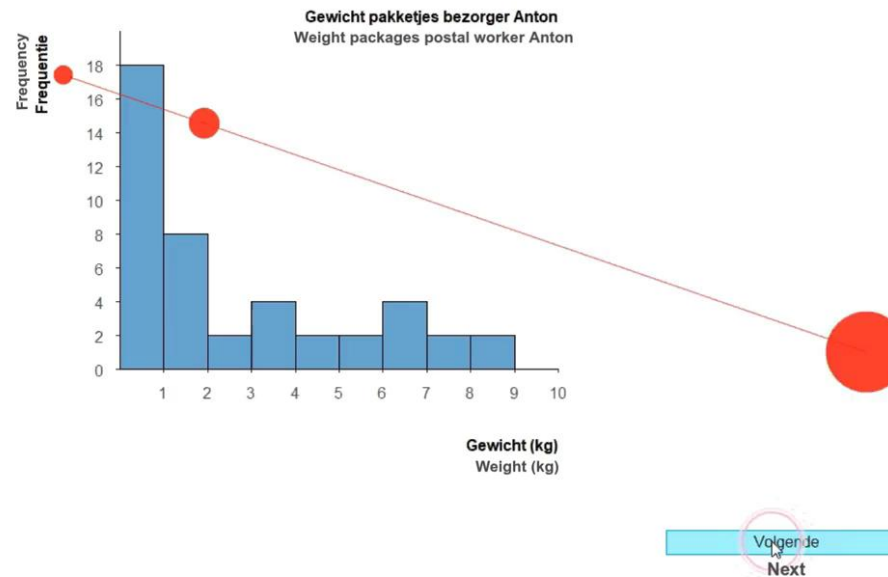


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Fabian's Comment: "I looked at the values on the axis — I saw mostly low ones and one high one, so I estimated it would be somewhere around six."

What is approximately the mean weight of the packages Anton delivers?
Wat is ongeveer het gemiddelde gewicht van de pakketjes die Anton bezorgt?

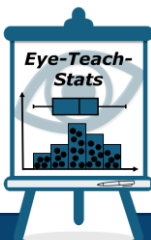
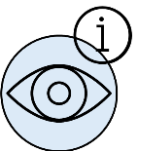
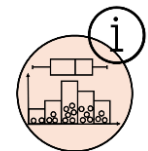


Now you can read the original comment of Fabian.

Indicate to what extent Fabian's comment agrees with what you assumed.

- ☐ No alignment at all
- ☐ Little alignment

- ☐ Quite some alignment
- ☐ Complete alignment



Evaluate solution

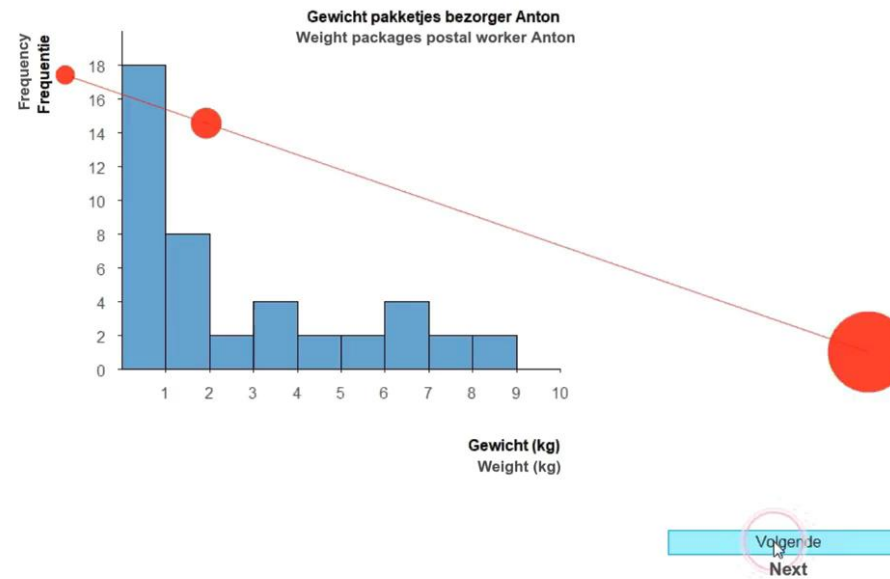


Co-funded by
the European Union

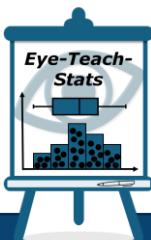
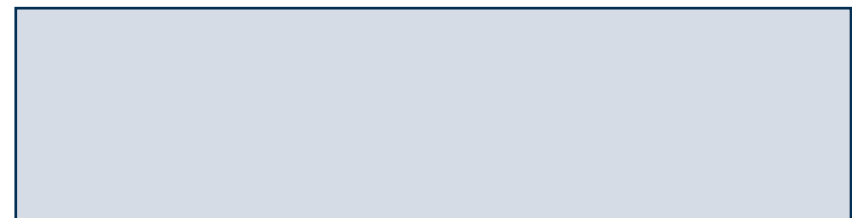
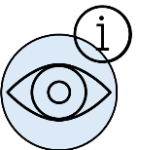
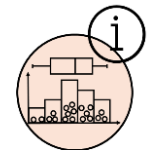
Erasmus+
Enriching lives, opening minds.

Fabian's Comment: "I looked at the values on the axis — I saw mostly low ones and one high one, so I estimated it would be somewhere around six."

What is approximately the mean weight of the packages Anton delivers?
Wat is ongeveer het gemiddelde gewicht van de pakketjes die Anton bezorgt?



Evaluate Fabian's solution. What can you say about his understanding?



Introducing systematic error



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

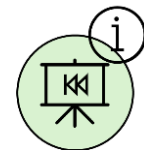
Understanding student errors:

Fabian has an underlying conceptual difficulty regarding the interpretation of histograms. This leads to an observable systematic error, called "**frequency value confusion**".

Observable behavior: Fabian said he looked at the values on the [vertical] axis, noticing mostly low ones and one high one, and estimated the mean to be around six. The gaze video substantiates this as his gaze moved along the vertical axis, fixating on the frequency values. This indicates an observable systematic error, which we term "frequency value confusion" in the interpretation of histograms.

Interpretation by the teacher/researcher: It appears that Fabian applied a strategy to the histogram that is typically used for case-value plots. From the explanation and gazes of Fabian it appears that he read the frequency values on the vertical axis as if they were the measured data values themselves, and estimated the mean from those. However, in a histogram, the heights of bars indicate frequencies, not measured values. Fabian'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 (i.e., 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 interval. Although incorrect for histograms, this strategy is correct for finding the arithmetic mean in a case-value plot.

Underlying conceptual difficulty: Fabian'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.

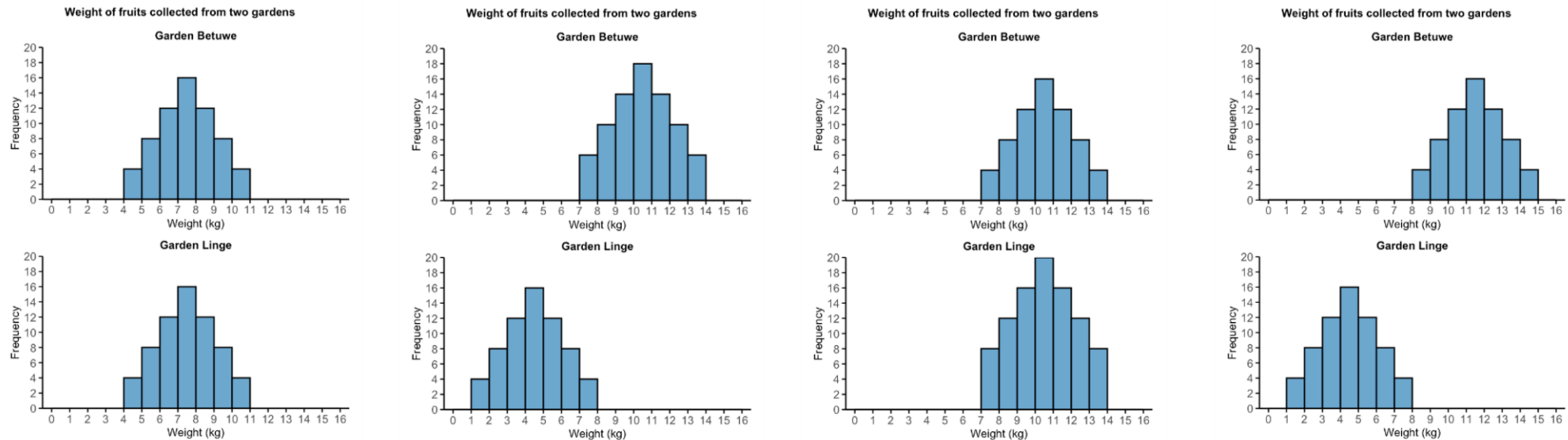


Practice



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.



A group of students participated in a science project where they measured and recorded the weight (in kilograms) of fruits collected from different gardens. Each of the five student groups collected fruit samples from two different gardens and created a pair of histograms to show the weight distributions.

Later, a student named Jamie was asked to examine these four histogram pairs and compare the mean weights between the two gardens in each pair. Jamie has an underlying conceptual difficulty regarding the interpretation of histograms. This leads to an observable systematic error, known as frequency value confusion, where Jamie tends to interpret the bar height as a data value rather than a frequency.

In which of the histogram pairs is Jamie's reasoning strategy likely to result in a wrong conclusion regarding the mean? Select all that apply.

☐ A

☐ B

☐ C

☐ D



Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor the granting authority can be held responsible for them.

18 Case-Value Plot

Vignette on an error caused by an overgeneralization of a histogram approach

Introducing situation



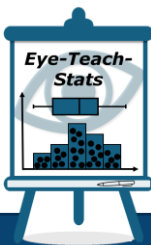
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Imagine you are teaching a 10th grade classroom on data interpretation, and you introduce your students to find the approximate mean of weights represented in a case-value plot.

You notice your student Ella is having difficulties with a task you've assigned. She gives an incorrect answer.

You will now see the task Ella is working on.

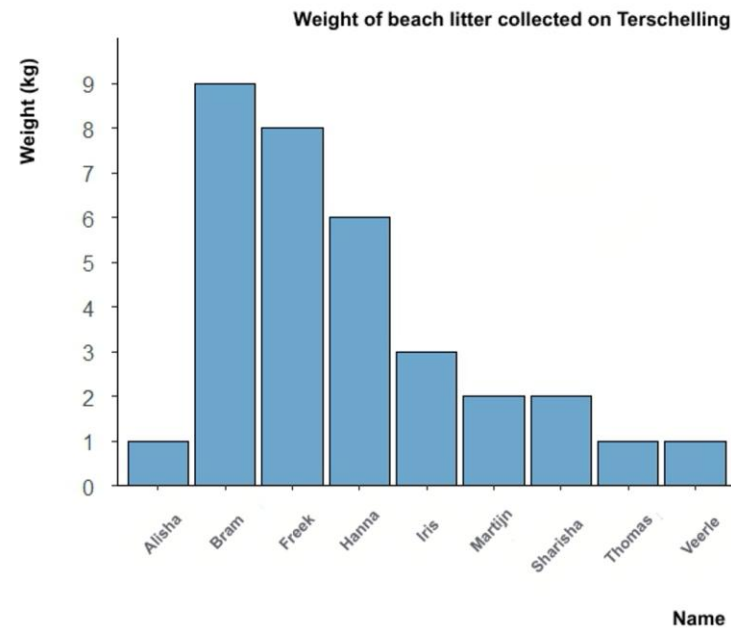


Student's task



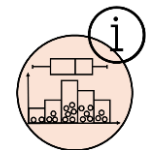
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.



What is approximately the mean weight collected per person?

This is the task Ella is working on. Please complete the task on your own first and enter your solution.



Introducing eye-tracking

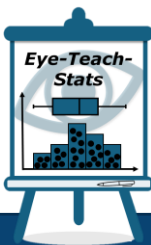


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

A group of researchers in mathematics education is pursuing the idea of supporting teachers with additional diagnostic material: To this end, eye movement data from students is recorded using eye-tracking technology and made available as supplementary material.

You will receive a gaze plot that visualizes gaze data of your student Ella that were collected while she was working on the task. The gaze plot indicates where her visual focus lies. Often it is possible to infer what Ella's thinking process might have been.



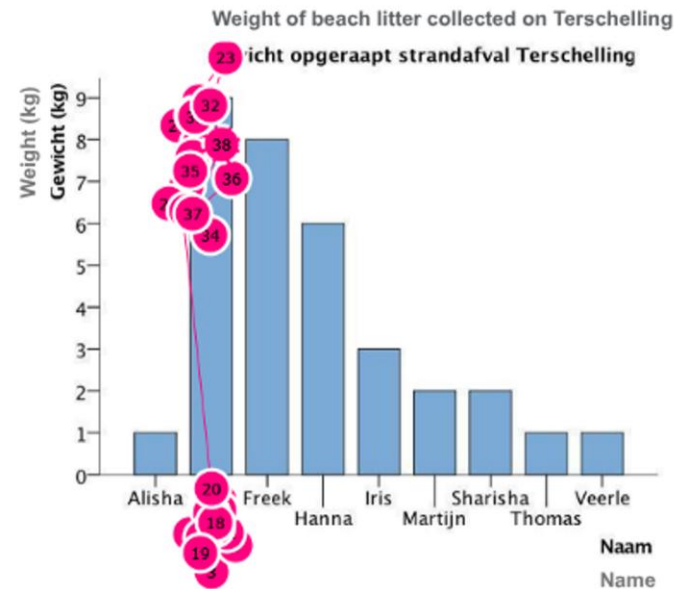
First look



Co-funded by
the European Union

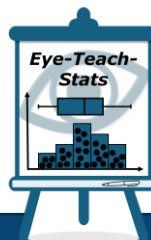
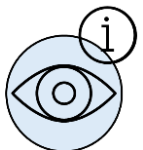
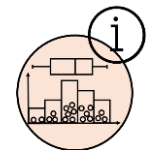
Erasmus+
Enriching lives, opening minds.

Ella's answer: "The mean weight is nine."



What is approximately the mean weight collected per person?

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



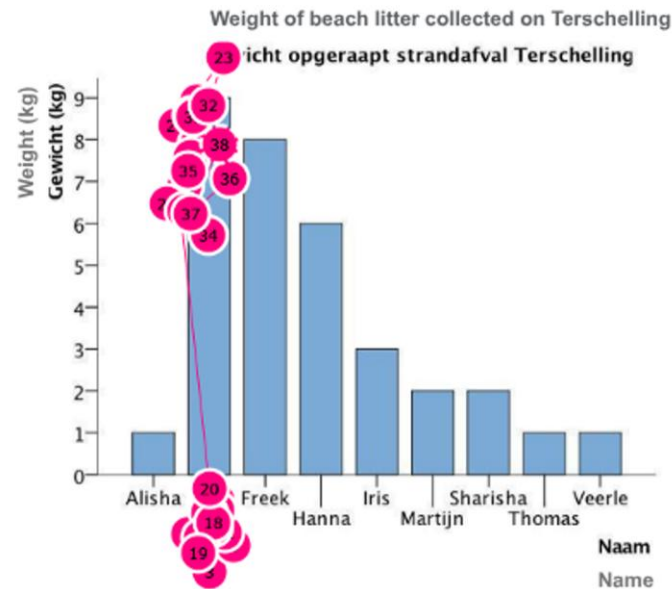
Take perspective



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

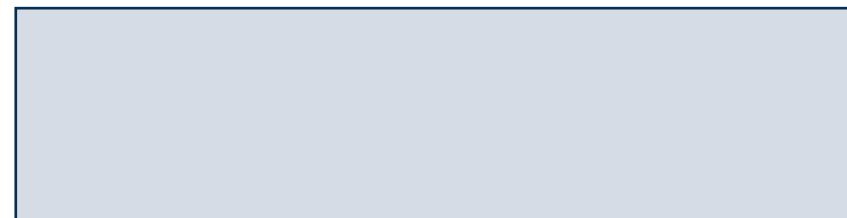
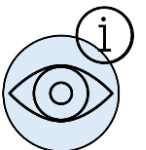
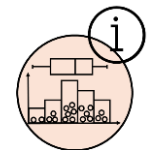
Ella's answer: "The mean weight is nine."



What is approximately the mean weight collected per person?

After the task Ella was shown her own gaze plot. She explains retrospectively what she was thinking while working on the task.

Try to take Ella's perspective and formulate what she may comment.



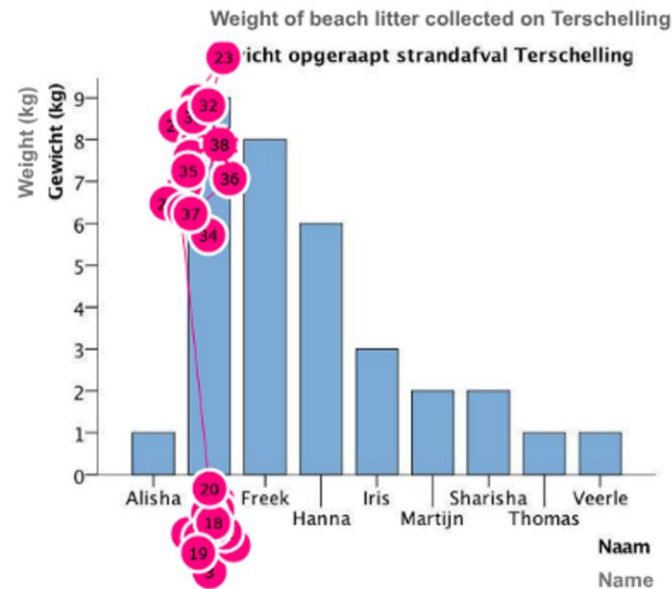
Indicate assumption



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Ella's Comment: "I looked at where the graph was kind of in balance, and that seemed to be around Bram, so I took 9 as the mean."

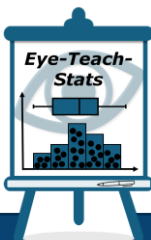
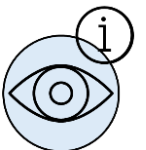
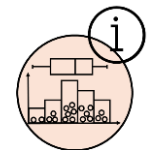


What is approximately the mean weight collected per person?

Now you can read the original comment of Ella.

Indicate to what extent Ella's comment agrees with what you assumed.

- ☐ No alignment at all
- ☐ Little alignment
- ☐ Quite some alignment
- ☐ Complete alignment



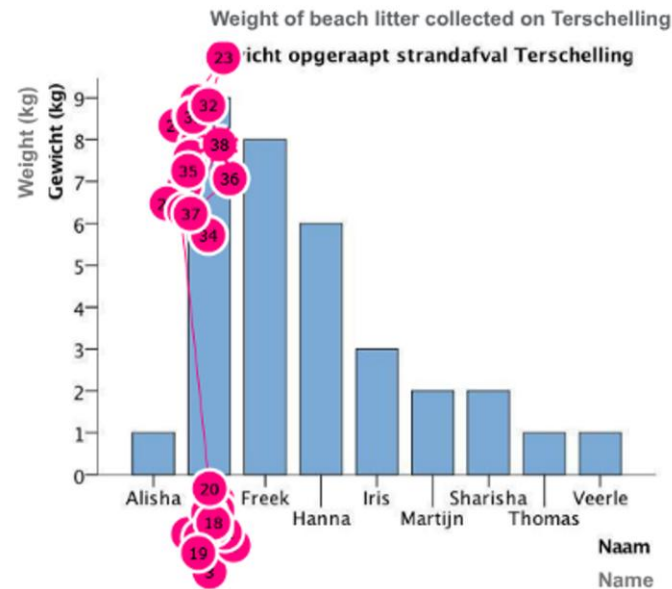
Evaluate solution



Co-funded by
the European Union

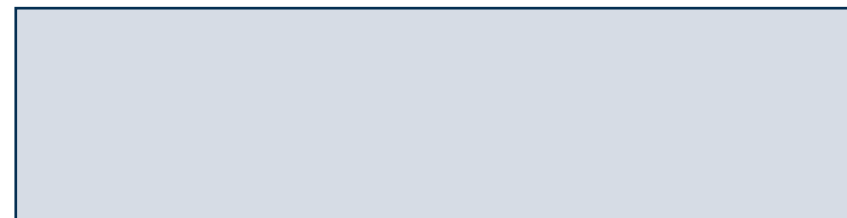
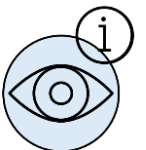
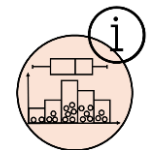
Erasmus+
Enriching lives, opening minds.

Ella's Comment: "I looked at where the graph was kind of in balance, and that seemed to be around Bram, so I took 9 as the mean."



What is approximately the mean weight collected per person?

Evaluate Ella's solution. What can you say about her understanding?



Introducing systematic error



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

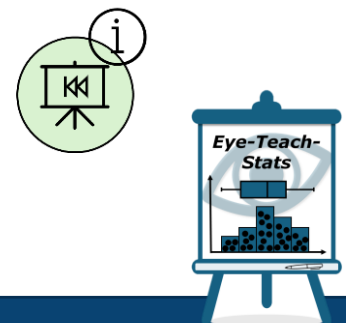
Understanding student errors:

Ella has an underlying conceptual difficulty regarding the interpretation of case-value plots. This leads to an observable systematic error, called “**overgeneralization of a histogram approach**”.

Observable behavior: Ella explained that she looked for the point where the graph was in balance, and that this seemed to be around Bram, leading her to estimate the mean as 9 kg. Her gaze data show that Ella’s fixations moved up and down between a point on the horizontal axis and the tops of the bars, creating a vertical scanning pattern around Bram’s bar. Together this observable behavior points at a systematic error, which we term overgeneralization of a histogram approach when she interprets case-value plots.

Interpretation by the teacher/researcher: It appears that Ella applied a reasoning typically used for histograms — visually searching for the balance point along the horizontal axis — and identified Bram’s position as that point. Such a strategy is appropriate for histograms, where the balance point on the horizontal axis approximates the mean of the distribution. However, it was incorrectly transferred to a case-value plot, where each bar represents one individual measured value and the mean is not found by locating a balance point on the horizontal axis.

Underlying conceptual difficulty: Ella’s systematic error is likely due to an underlying conceptual difficulty with the relation between the concept of distribution and the concept of center; therefore, in understanding how the mean is determined in case-value plots versus histograms. In a case-value plot, each bar’s height is a measured value, and the mean is estimated by finding a horizontal line that would make all bars equally high. In a histogram, bar heights represent frequencies, and the balance point of the graph along the horizontal axis approximates the mean.



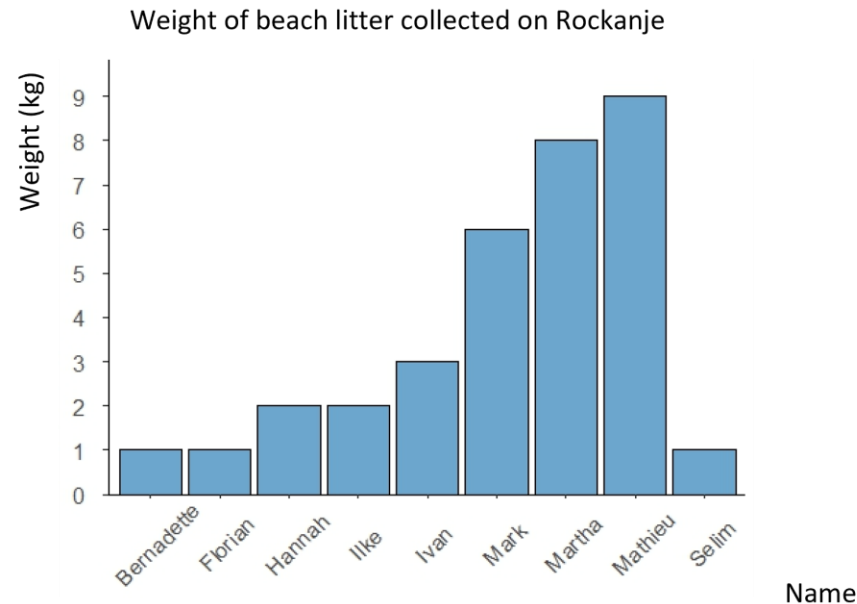
Practice



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

A group of nine people collected beach litter on the coast of Rockanje. The weight (in kg) collected by each person is shown in the case-value plot below. Compare the amounts collected and then evaluate: what was the mean weight collected per person?



Three students described different reasoning while solving the task. Which student answer fits best with the assumption that this student overgeneralises a histogram approach to a case-value plot, leading to an incorrect answer?

Robin: I added up the weights collected by all nine people and divided the total by nine. That gave me a mean of about 3.7 kg.

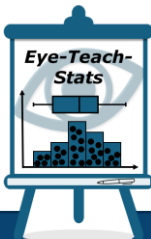
Sam: I imagined redistributing the weights so that all nine bars were the same height. That level looked like it was a little below 4 kg, so I estimated the mean to be about 4 kg.

Alex: I looked for where the bars seemed to balance, left to right, like finding the center of a distribution. That point was around Mark, so I used Mark's value, 6 kg, as the mean.

☐ Robin's

☐ Sam's

☐ Alex's



Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor the granting authority can be held responsible for them.

19 Histogram

Vignette on an error caused by overgeneralization of the sum of frequencies for mean

Introducing situation



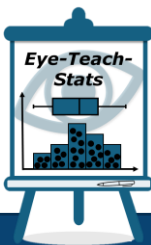
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Imagine you are teaching a 10th-grade classroom on data interpretation and you introduce your students to find the approximate mean of weights represented in a histogram.

You notice your student Martijn is having difficulties with a task you've assigned. He gives an incorrect answer.

You will now see the task Martijn is working on.



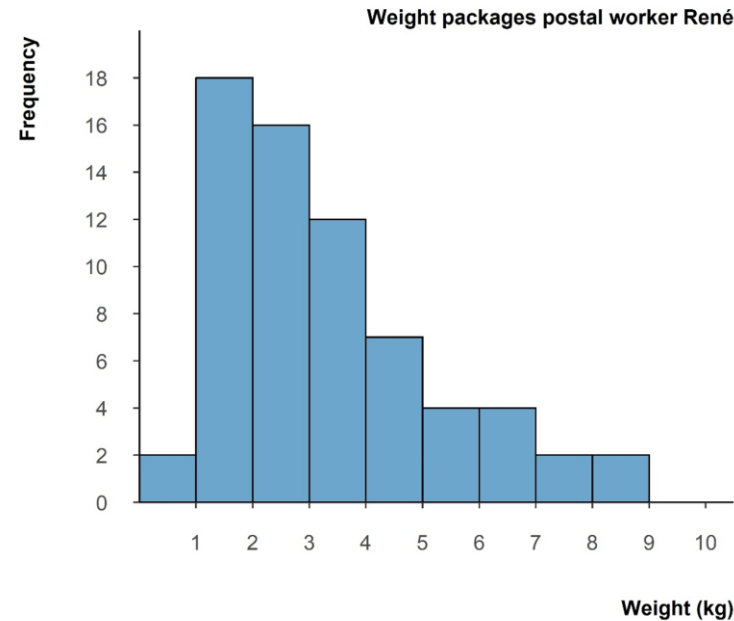
Student's task



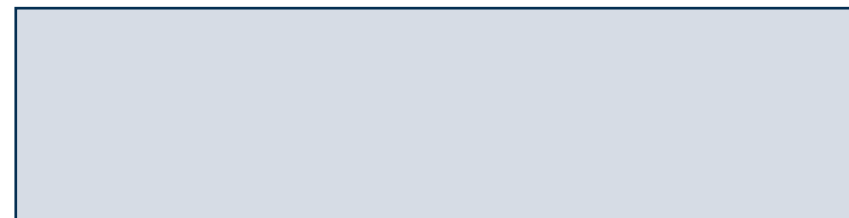
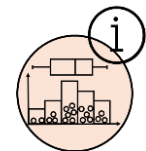
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

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



This is the task Martijn is working on. Please complete the task on your own first and enter your solution.



Introducing eye-tracking

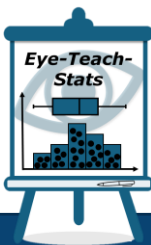


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

A group of researchers in mathematics education is pursuing the idea of supporting teachers with additional diagnostic material: To this end, eye movement data from students is recorded using eye-tracking technology and made available as supplementary material.

You will receive a gaze plot that visualizes gaze data of your student Martijn as he was working on the task. The gaze data indicates where his visual focus lies. Often it is possible to infer what Martijn's thinking process might have been.



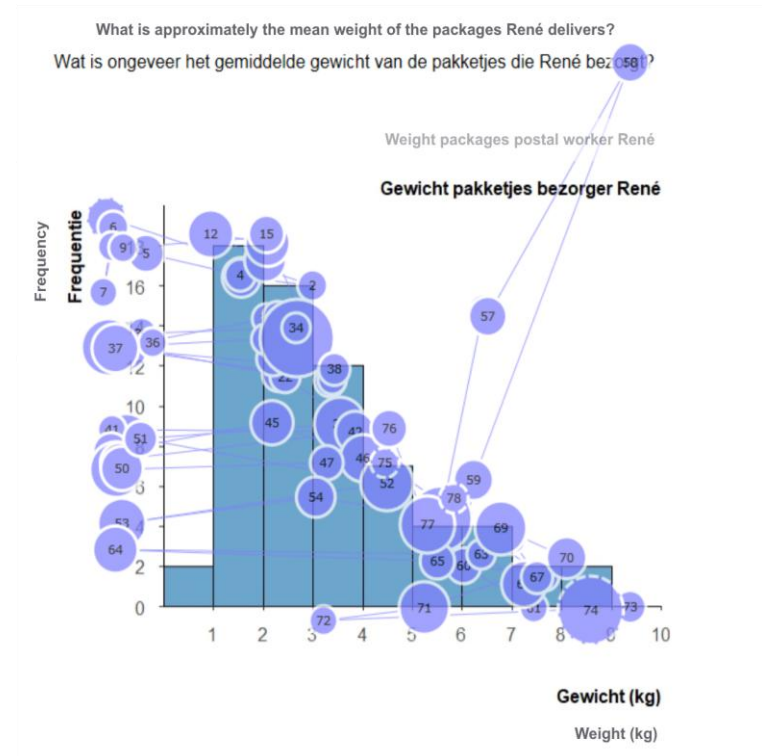
First look



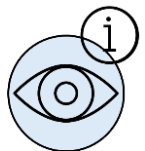
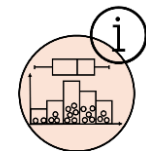
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Martijn's Answer: "The arithmetic mean is about six."



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



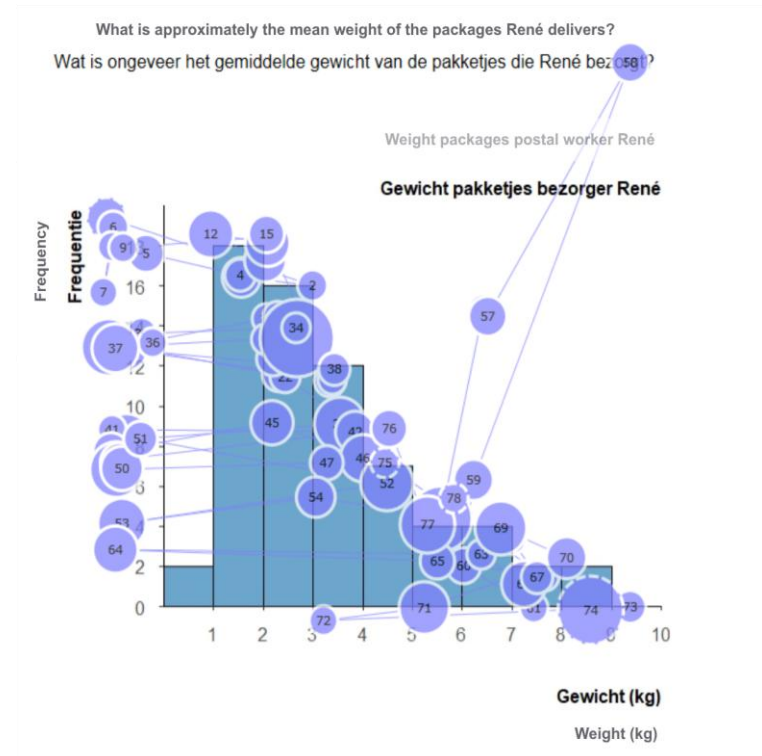
Take perspective



Co-funded by
the European Union

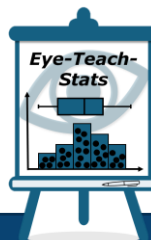
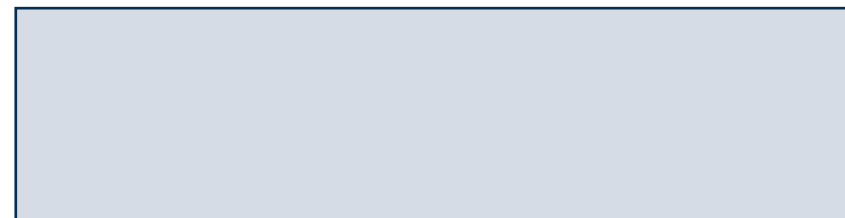
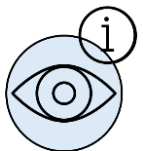
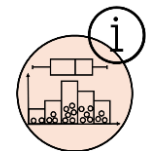
Erasmus+
Enriching lives, opening minds.

Martijn's Answer: "The arithmetic mean is about six."



After the task Martijn was shown his own gaze plot. He explains retrospectively what he was thinking while working on the task.

Try to take Martijn's perspective and formulate what he may comment.



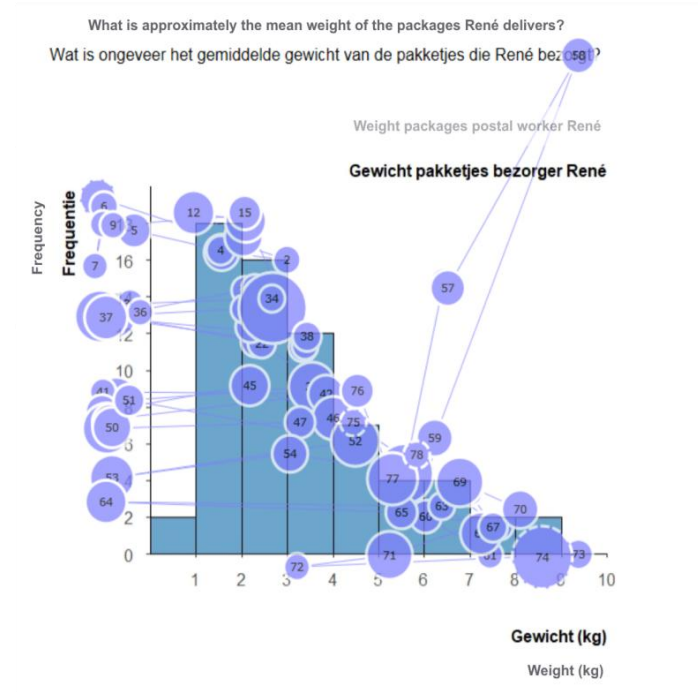
Indicate assumption



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Martijn's Comment: "Reading off how many each bar was and adding those together. So I started with two, then eighteen, then sixteen, and I kept adding that up as I looked at the bars. Then I divided by nine."

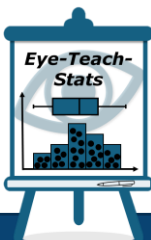
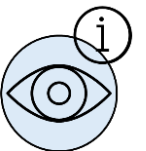
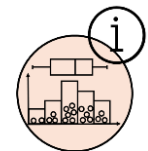


Now you can read the original comment of Martijn.

Indicate to what extent Martijn's comment agrees with what you assumed.

- ☐ No alignment at all
- ☐ Little alignment

- ☐ Quite some alignment
- ☐ Complete alignment



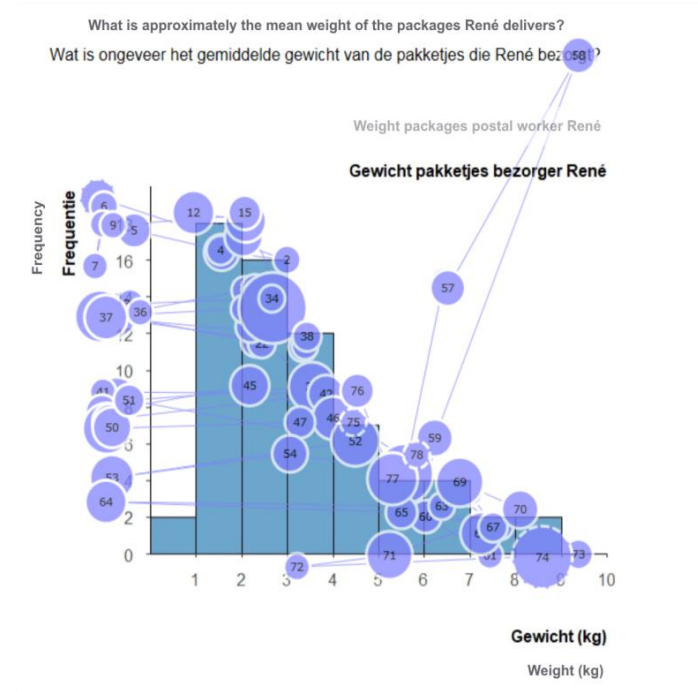
Evaluate solution



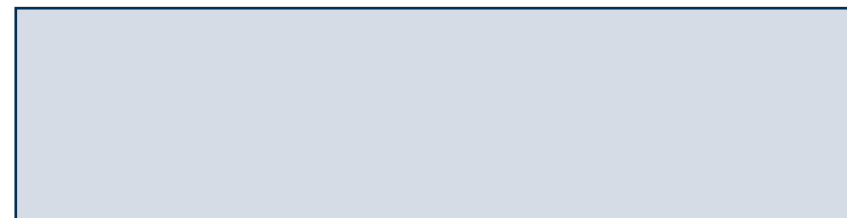
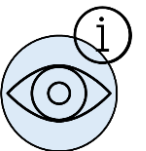
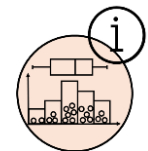
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Martijn's Comment: "Reading off how many each bar was and adding those together. So I started with two, then eighteen, then sixteen, and I kept adding that up as I looked at the bars. Then I divided by nine."



Evaluate Martijn's solution. What can you say about his understanding?



Introducing systematic error



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Understanding student errors:

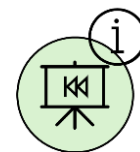
Martijn has an underlying conceptual difficulty regarding the interpretation of histograms. This leads to an observable systematic error, called **“sum for mean: sum of frequencies”**.

Observable behavior: Martijn’s verbal explanation reveals a clear procedural approach: he read off the height of each bar and kept a running total, starting with two, then adding eighteen, then sixteen, and so on. He then divided this sum by nine. His gaze supports this: fixations move systematically along the y-axis, dwelling longer at certain bar heights, suggesting he is reading off and accumulating frequency values. The gaze subsequently shifts to the x-axis, landing on the number 9, which corresponds to the total number of bars. This behavior together indicates an observable systematic error, which we term “sum for mean: sum of frequencies”.

Interpretation by the teacher/researcher: Martijn’s approach reveals that he used the bar heights — which represent frequencies — as the values to sum when calculating the mean. He added up the frequencies shown on the vertical axis and divided by the number of bars, arriving at 6. The longer fixations on the y-axis and bars likely reflect mental accumulation — he is adding up the frequencies in his head as he reads each bar. In a histogram, the vertical axis represents how many measurements fall within each bin, not the measured values themselves. Those are represented on the horizontal axis. The error occurs when students add up the frequencies and divide by the number of bars, applying a procedure that resembles calculating an arithmetic mean but on the wrong values. While adding up bar heights and dividing by the number of bars would be a valid procedure in a case-value plot — where bar heights represent measured values — it is not appropriate for a histogram, where bar heights represent frequencies.

Underlying conceptual difficulty: Martijn’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 the role of the vertical axis in a histogram and in a case-value plot. 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, and that this frequency value itself is not what should be summed to find the mean.

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 a case, their gaze would move up and down between a point on the horizontal axis and the tops of the bars, creating a scan path 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).



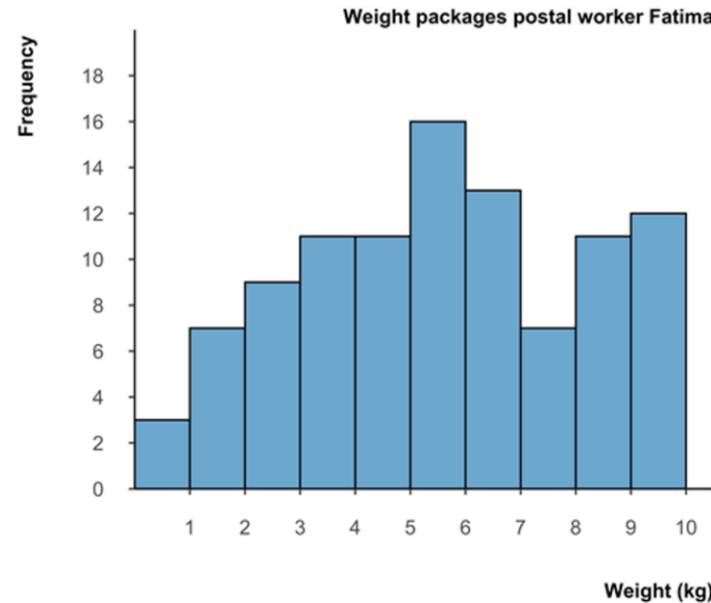
Practice



Co-funded by
the European Union

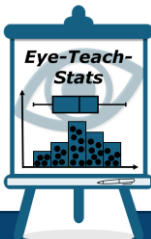
Erasmus+
Enriching lives, opening minds.

The histogram below shows the weights of packages delivered by postal worker Fatima. Elis looked at the histogram and said the mean is “around 10”. However, the actual mean based on the data is approximately 5.5.



Which of the following reasoning strategies most likely led to the incorrect solution “around 10” for the mean weight of the packages? Select one option.

- a) Elis added the frequencies and divided by the number of bars.
- b) Elis assumed the mean was at the center of the x-axis, without considering the distribution of the data.
- c) Elis focused on the bin with the highest bar and assumed its value was the mean.
- d) Elis looked at the values on the x-axis and calculated their mean, ignoring the frequencies.
- e) Elis estimated the mean by identifying a balance point in the histogram, as if placing a vertical line where the distribution appears visually balanced.





Eye-Teach-Stats Eye-Tracking Vignettes © 2026



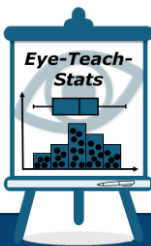
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor the granting authority can be held responsible for them.

20 Histogram

Vignette on an error caused by a count-and-compute strategy



Introducing situation



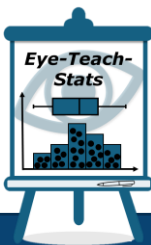
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Imagine you are teaching a 10th grade classroom on data interpretation, and you introduce your students to comparing data distributions using histograms.

You notice your student Erika is having difficulties with a task you've assigned. She gives an incorrect answer.

You will now see the task Erika is working on.

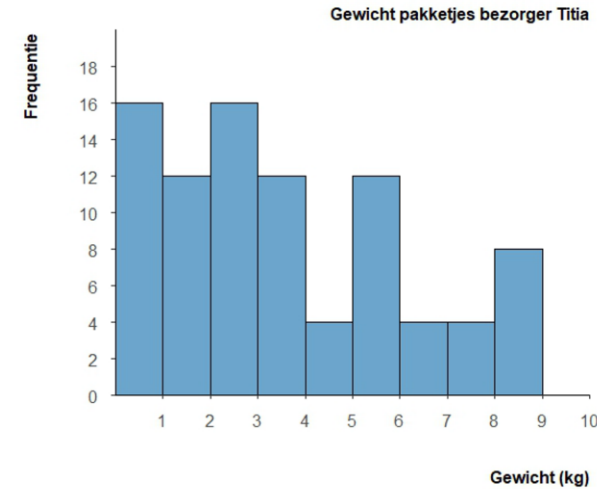
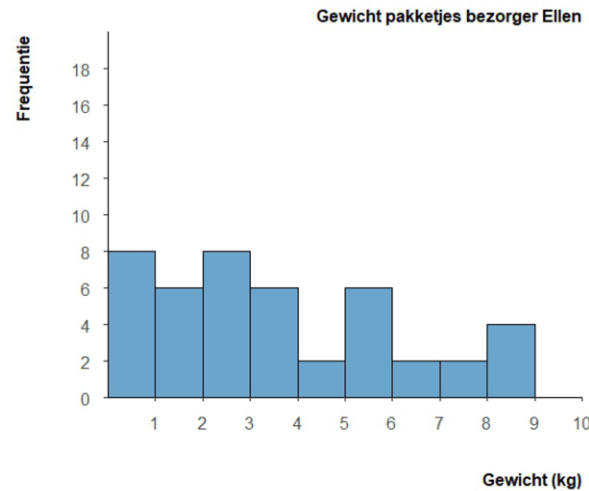


Student's task



Co-funded by
the European Union

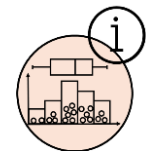
Erasmus+
Enriching lives, opening minds.



Which postal worker delivers on average heaviest packages?

- a) Ellen delivers on average the heaviest packages.
- b) Titia delivers on average the heaviest packages.
- c) The mean weight for both is approximately the same.

This is the task Erika is working on. **Please complete the task on your own first and enter your solution.**



- a) Ellen delivers on average the heaviest packages.
- b) Titia delivers on average the heaviest packages.
- c) The mean weight for both is approximately the same.



Introducing eye-tracking

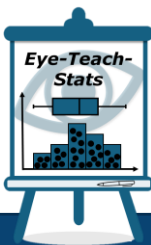


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

A group of researchers in mathematics education is pursuing the idea of supporting teachers with additional diagnostic material: To this end, eye movement data from students is recorded using eye-tracking technology and made available as supplementary material.

You will receive a short gaze video that shows the eye movements of your student Erika that were collected while she was working on the task. The gaze data indicates where her visual focus lies. Often it is possible to infer what Erika's thinking process might have been.



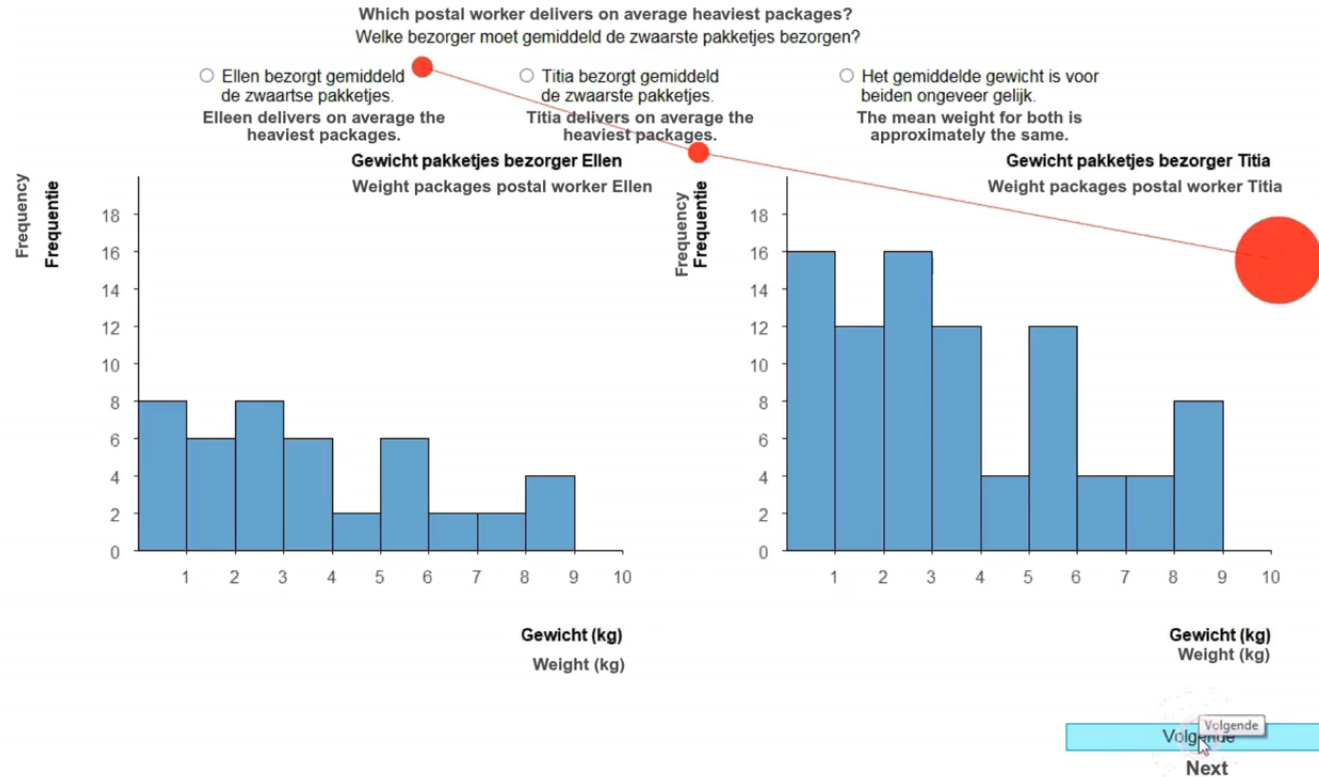
First look



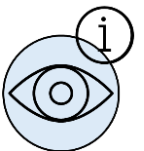
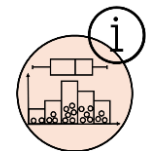
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Erika's Answer: "Titia delivers on average the heaviest packages."



Have a look at the video sequence that shows your student Erika's gaze and her answer.



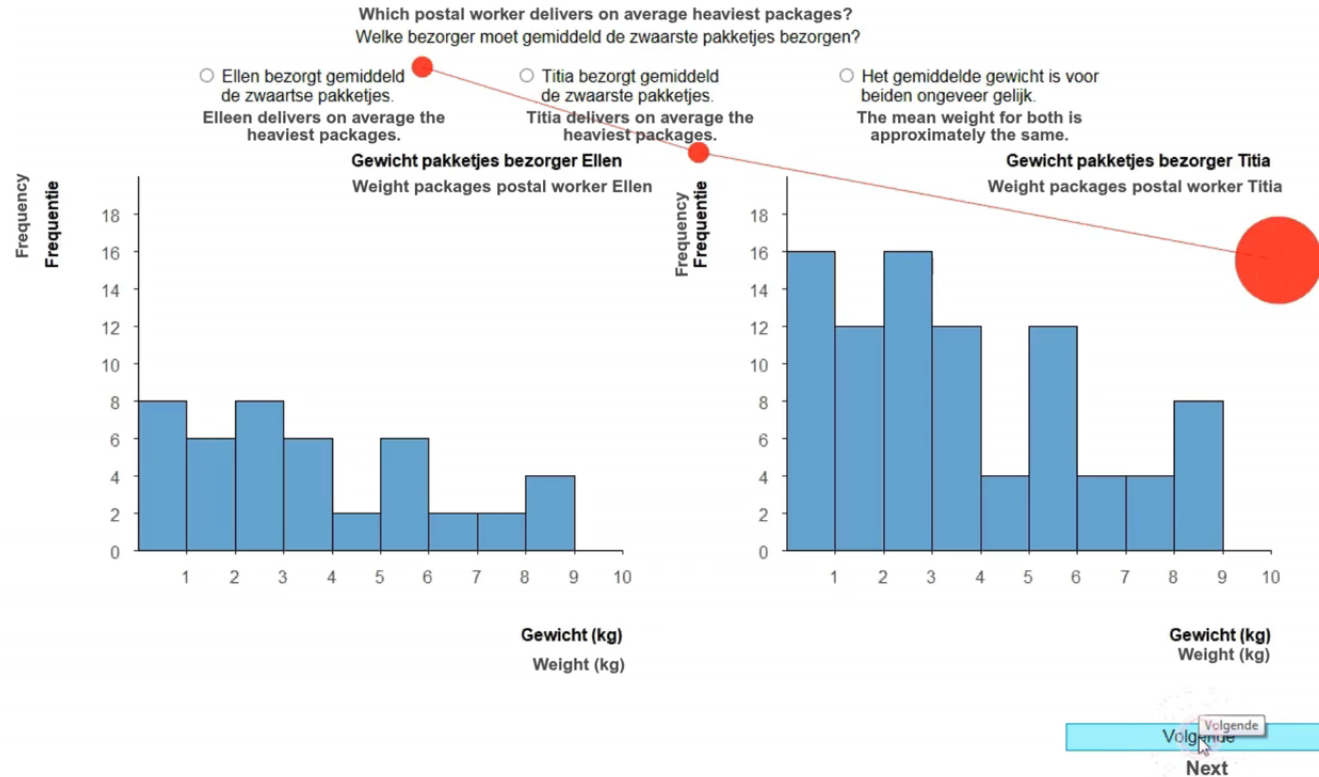
Take perspective



Co-funded by
the European Union

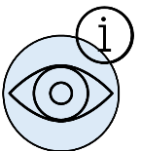
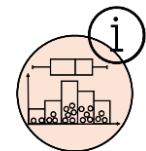
Erasmus+
Enriching lives, opening minds.

Erika's Answer: "Titia delivers on average the heaviest packages."



After the task Erika was shown her own gaze video. She explains retrospectively what she was thinking while working on the task.

Try to take Erika's perspective and formulate what she may comment.



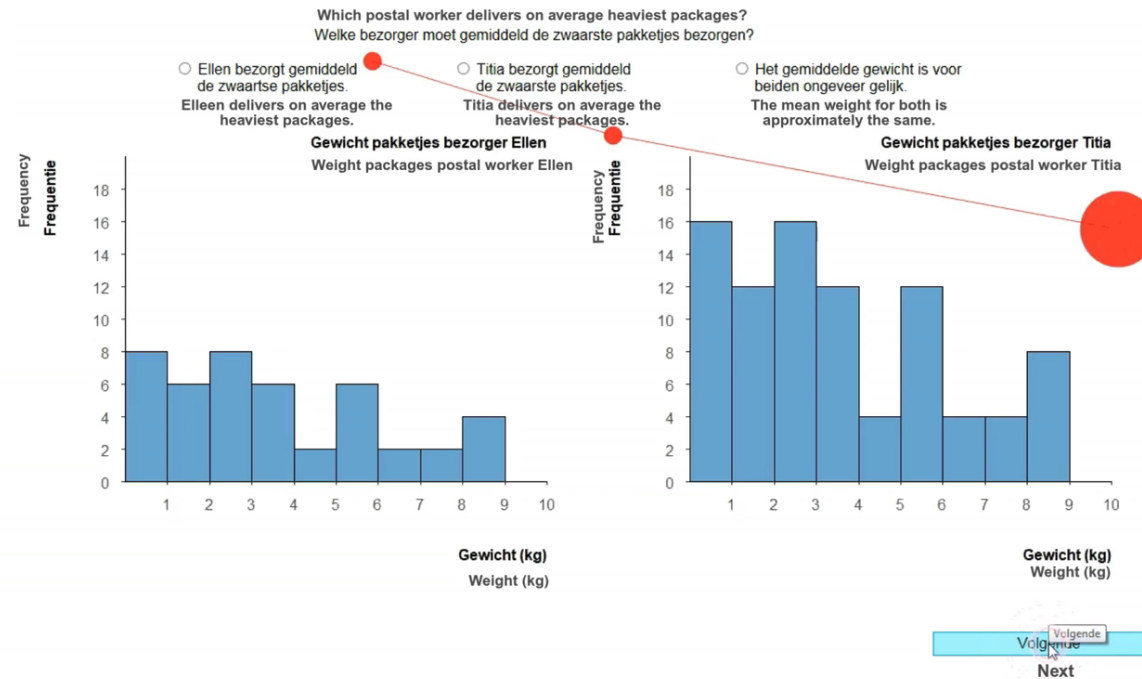
Indicate assumption



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Erika's Comment: "I wanted to find the mean weight, so I looked at each bar. For every bar, I multiplied the weight value by the frequency with which it occurred. Then I added all these results together and divided by the number of bars, because each bar represents one part of the data. I did this for both histograms and got a higher result for Titia, so I chose Titia."

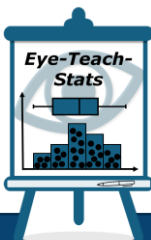
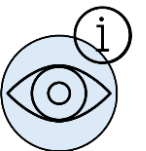
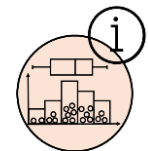


Now you can read the original comment of Erika.

Indicate to what extent Erika's comment agrees with what you assumed.

- ☐ No alignment at all
- ☐ Little alignment

- ☐ Quite some alignment
- ☐ Complete alignment



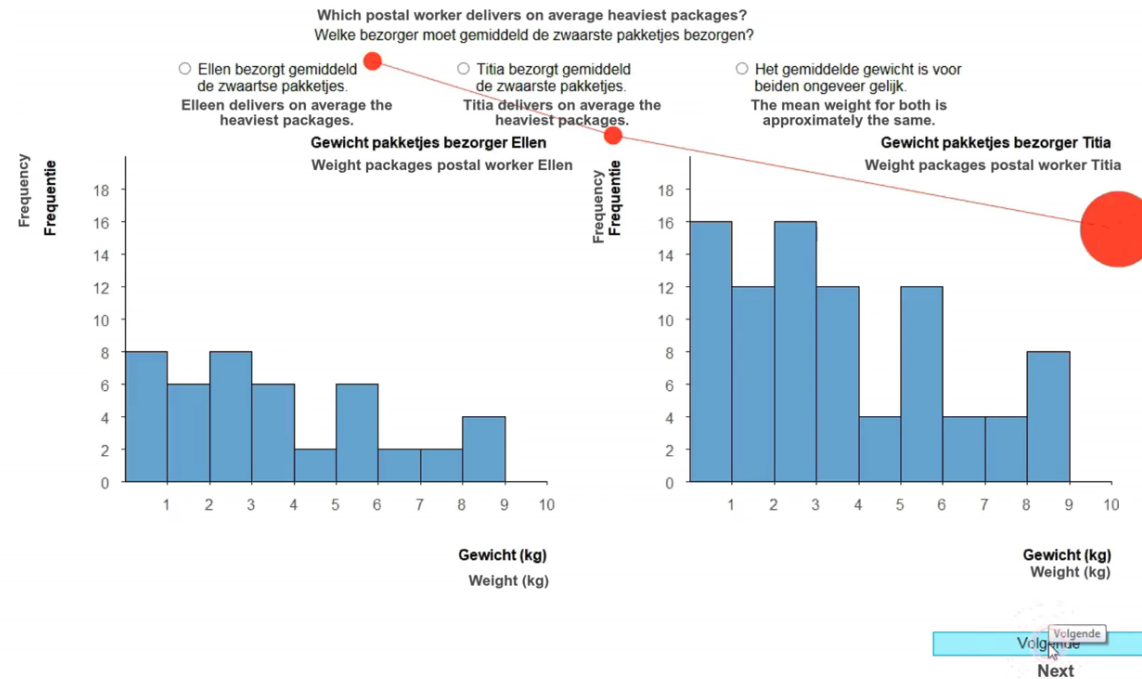
Evaluate solution



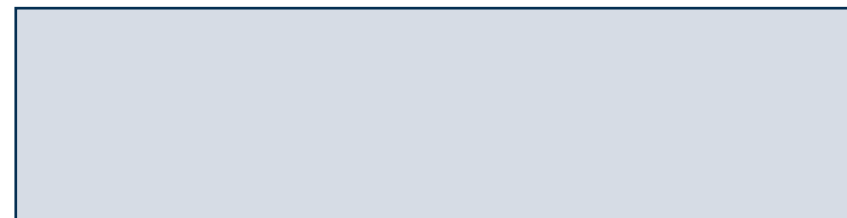
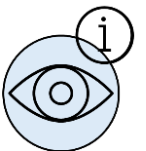
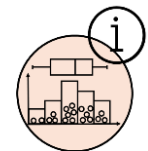
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Erika's Comment: "I wanted to find the mean weight, so I looked at each bar. For every bar, I multiplied the weight value by the frequency with which it occurred. Then I added all these results together and divided by the number of bars, because each bar represents one part of the data. I did this for both histograms and got a higher result for Titia, so I chose Titia."



Evaluate Erika's solution. What can you say about her understanding?



Introducing systematic error



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

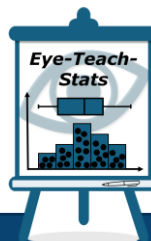
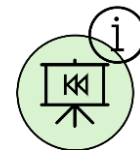
Understanding student errors

Erika has an underlying conceptual difficulty regarding the interpretation of histograms when comparing averages. This leads to an observable systematic error that is consistent with **overemphasizing local features: a count-and-compute strategy**.

Observable behavior: Erika stated that she multiplied the weight values by the frequencies and divided the result by the number of bars. She applied this procedure to both histograms and concluded that Titia delivers on average the heaviest packages. The gaze data show that Erika worked on one histogram at a time. While working on a histogram, her gaze followed a repeating right-angle pattern. She repeatedly moved her gaze from the vertical axis to the top of a bar, then down toward the horizontal axis, back to the bar, and then to the next bar. Within this pattern, she spent a long time looking at the bars.

Interpretation by the teacher/researcher: Based on her observed behavior, it appears that Erika approached the task from a local rather than a global view of the data. She focused on the data, working through the bars sequentially within one histogram before moving to the other. This suggests that she treated the bars as the main units for her reasoning, instead of comparing the two histograms as whole distributions. The long time spent looking at one bar, may indicate that she is thinking instead of actually looking at the bar (in this case, most likely mentally adding up numbers). She multiplied the values by the corresponding frequencies, added them, but then divided by the number of bars, as if each bar represented one data point. In histograms, the data are aggregated into bins, and determining and comparing the means requires considering the overall distribution and the total number of observations, which is the sum of the frequencies, not the number of bars. As a result, this local approach can lead to incorrect conclusions when comparing which postal worker has the higher mean package weight.

Underlying conceptual difficulty: The systematic error is likely due to an underlying conceptual difficulty with the concept of distribution when determining and comparing means in histograms. It suggests a procedural rather than conceptual integration of frequencies. Erika appears to apply a learned rule of value $\times$ frequency without aligning it with the corresponding denominator, dividing by the number of bars as if each bar represented one measured value. The error reflects a breakdown in distributional reasoning: Erika partially accounts for aggregation by weighting values with frequencies, but she fails to integrate this aggregation into a global view of the distribution. In addition, she does not use distributional aspects that are central for comparing means in double histograms, such as shape and shift. In this case, the two histograms have the same overall shape and the bars are located at the same positions, while the frequencies are scaled (e.g., doubled). This changes the number of observations but does not change the location of the distribution along the horizontal axis; therefore, the mean remains the same.

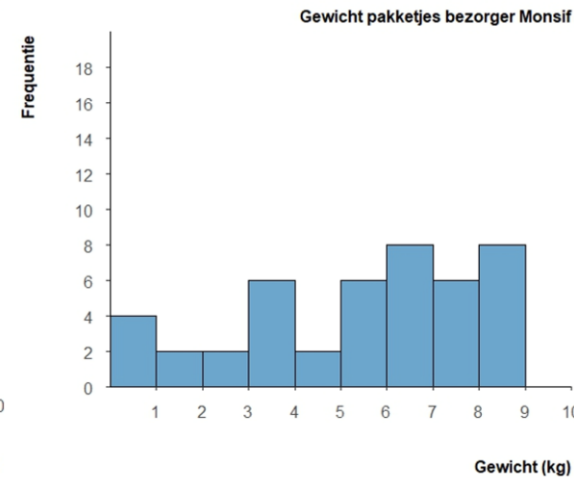
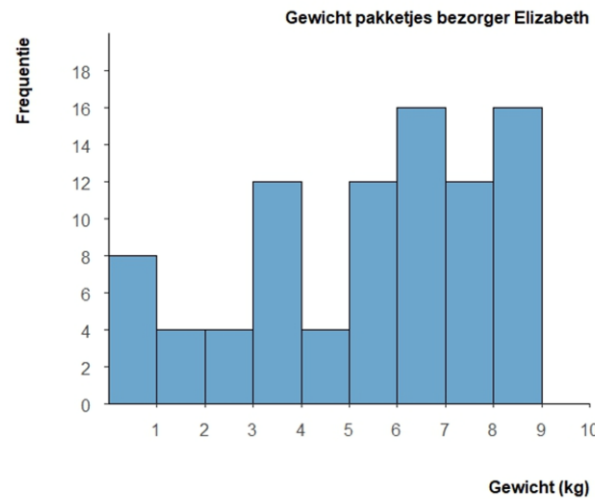


Practice



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

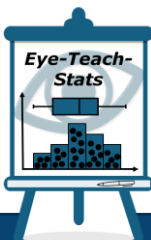
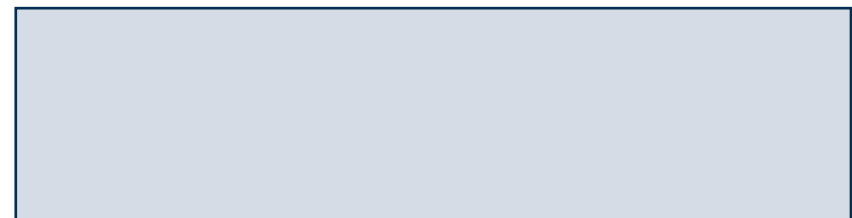


Which postal worker delivers on average the heaviest packages?

- a. Elizabeth delivers on average the heaviest packages.
- b. Monsif delivers on average the heaviest packages.
- c. The mean weight for both is approximately the same.

Imagine a student makes the same systematic error of overemphasizing local features: count-and-compute strategy.

Describe the possible gaze pattern the student may show. Do you think this error occurs here? If so, why? If not, why not?



Supportive Information

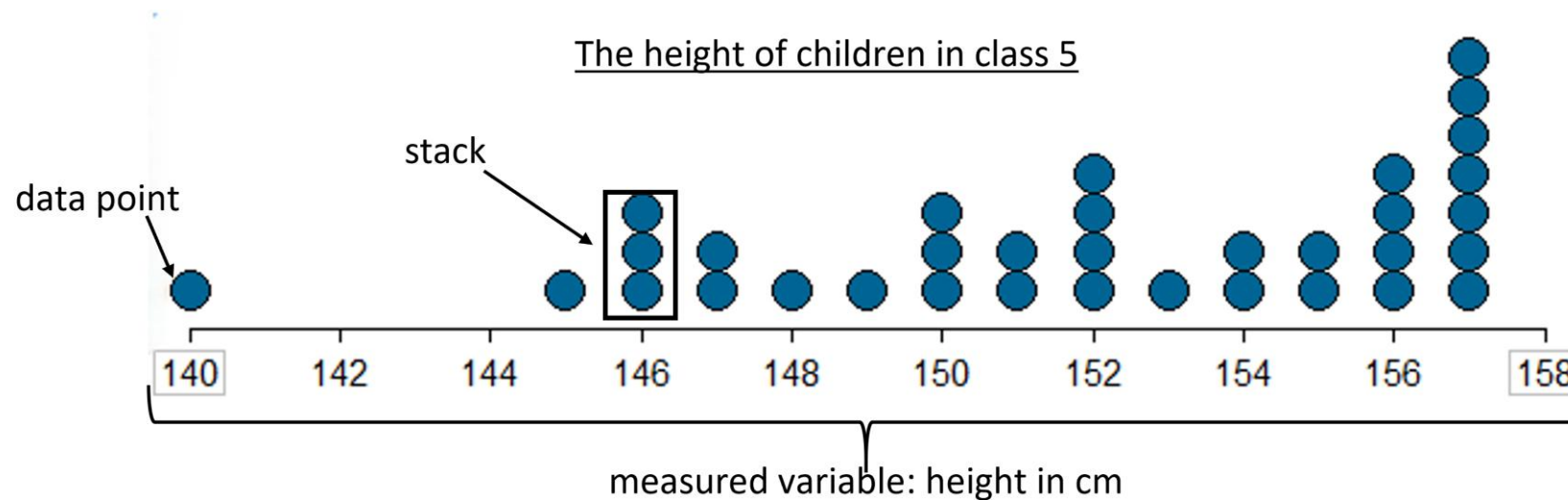
Dotplot



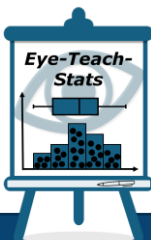
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Dotplots are graphical representations used to display the distribution of a dataset. In a dotplot, each data point is represented by a dot along a single axis, typically the horizontal axis. If multiple data points have the same value, their dots are stacked vertically. They are particularly useful for small to medium-sized datasets.



To compare distributions presented in dotplots global features as well as local details can be taken into account.



Supportive Information global – local view



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Perspectives on data distributions: Local vs. global views:

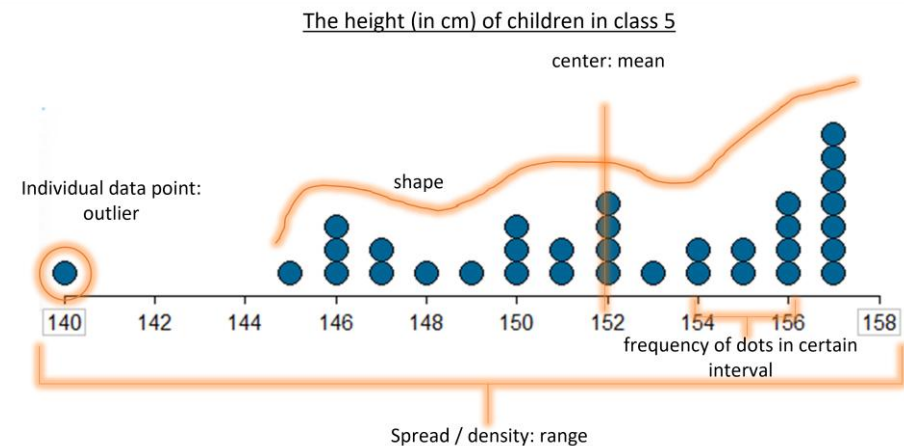
A central aim in statistics education is to guide students to also develop a “global view” of data distributions. In this perspective, students perceive distributions as conceptual entities with holistic features such as center, spread, and shape. However, students frequently perceive distributions as mere collections of individual data points, leading them to focus excessively on local details of the distribution. This is called a “local view” of data distributions.

Perspectives on data distribution: local vs. global view

distribution (conceptual entity)		
global features		
center <i>visually estimated mean ...</i>	spread/density <i>observed spread-out-ness, majority, range...</i>	shape <i>observed skewness, peaks, tails...</i>
local features		
individual data points <i>extreme values, outlier, absolute frequency of dots in certain intervals, value of certain intervals of dots, disjoint edge values...</i>		

Adapted according to Schreiter & Vogel, 2023

To compare distributions presented in dotplots, global features (e.g., center, spread/density, shape) as well as local features (e.g., outlier) can be taken into account.



Supportive Information

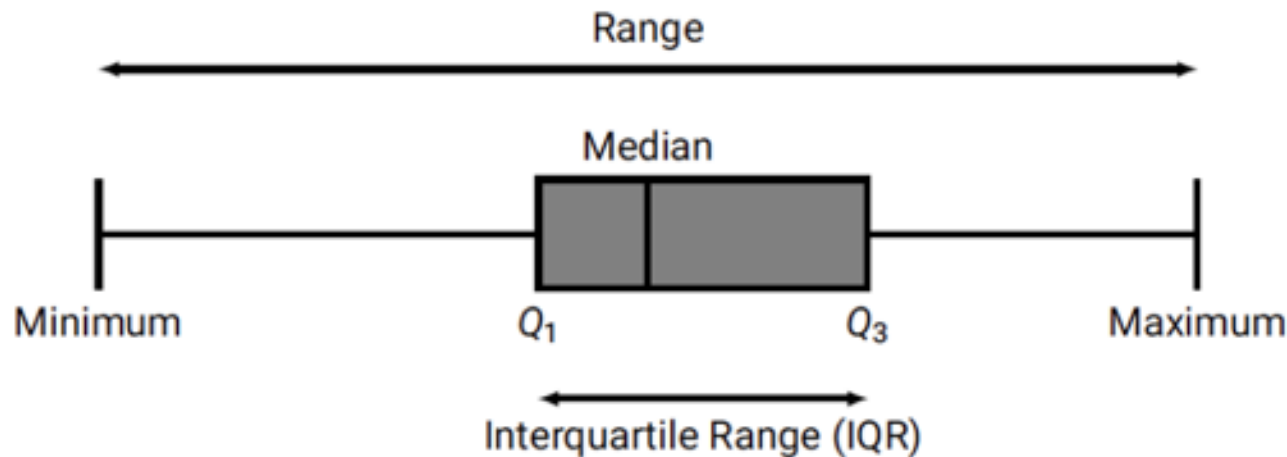
Boxplot



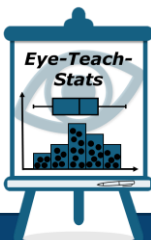
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

A boxplot, also known as a box-and-whisker plot, is a graphical representation that provides a concise summary of the distribution of a dataset. It effectively displays key statistical measures in a visually intuitive manner. By showcasing these measures, boxplots offer valuable insights into the central tendency, variability, and skewness of the data. At its core, a boxplot consists of five main components: the median, the first quartile (Q_1), the third quartile (Q_3), the extremes (maximum and minimum) and the range and the interquartile range (IQR).



By combining these components, a boxplot offers a clear and concise summary of the distribution, and it is particularly useful for comparing the distributions of multiple datasets simultaneously, facilitating quick and effective data analysis and interpretation. In summary, a boxplot serves as a powerful tool in exploratory data analysis; whether used in statistical research, data visualization, or decision-making processes, boxplots are indispensable for understanding the distributional properties of data.



Supportive Information

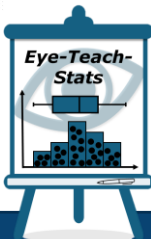
Boxplot II



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

In a class, all students indicated how many minutes they spent on homework the day before.



Supportive Information

Case-Value Plot



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

A case-value plot (CVP) is another type of graph with bars in which each bar's height represents one measured value. In a case-value plot with nominal data on the horizontal axis and numerical data on the vertical axis (two measured variables), the relevant measured value is on the vertical axis and variability can be seen as variation in the heights of the bars.

Why do we use CVP?

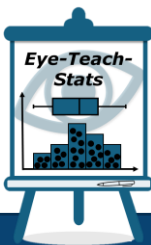
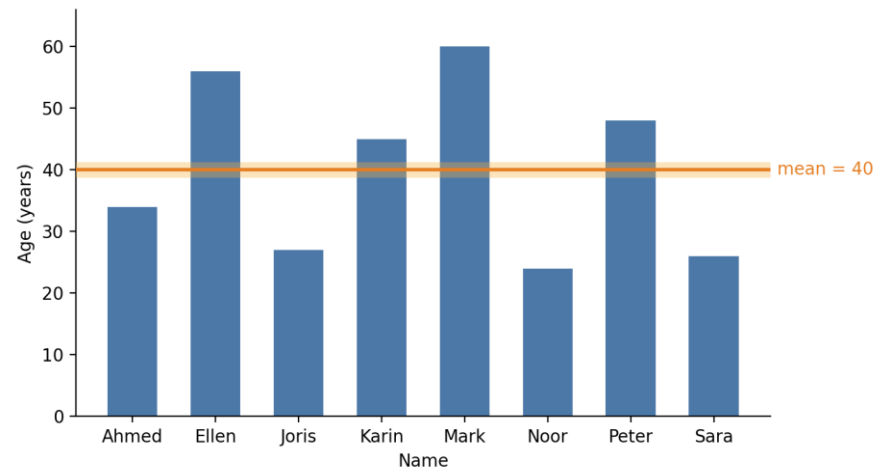
- To visually compare measurements of nominal data
- Comparing means of groups or categories (=nominal data) by case
- To see the variations between individual measurements

Example:

This case-value plot depicts the ages of the eight teachers in the mathematics department of a school. Each bar represents one teacher.

The mean age can be depicted by a horizontal line at 40 years, crossing the vertical axis at that number.

For the mean, the empty (white) space between the line and the shorter bars is exactly as large as the parts of the taller bars that rise above the line. If you moved the parts above the line into the gaps below it, all bars would become equally high.



Supportive Information Histogram



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

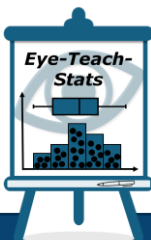
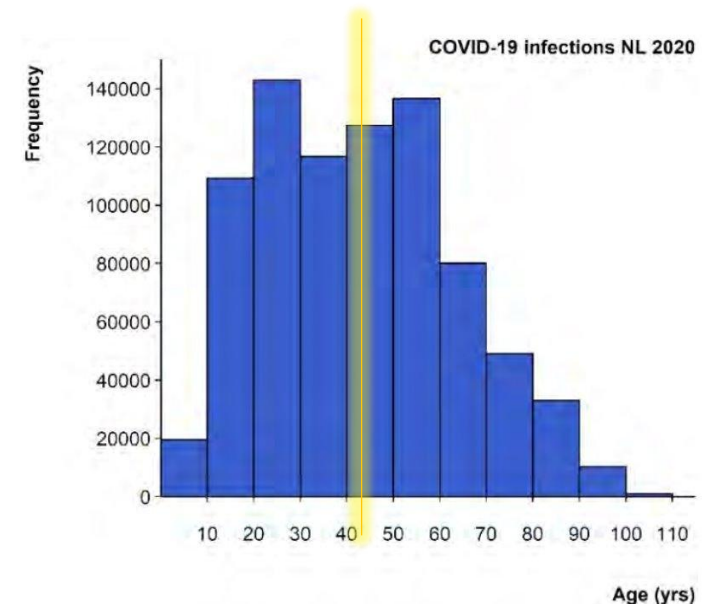
A histogram is a graph with bars that meets the following criteria:

- It consists of bars that represent grouped numerical data.
- It represents data of one statistical variable only (typically continuous).
- The statistical variable is presented conventionally on the horizontal axis.
- The statistical variable is measured at an interval or ratio measurement level.
- The vertical axis displays the class density, or—when bin widths are equal—relative or absolute frequency (counts) of the depicted statistical variable.
- The total density adds up to 1 or the total relative frequency adds up to 100%.

How to analyze and interpret a histogram:

The mean age (roughly 43 years) is along the horizontal axis and can be depicted by a vertical line that crosses the horizontal axis at 43 years; the balance point of the graph.

This can also be estimated by multiplying the midrange of each bin by the height of the bar, summing these products, and then dividing by the sum of the frequencies.



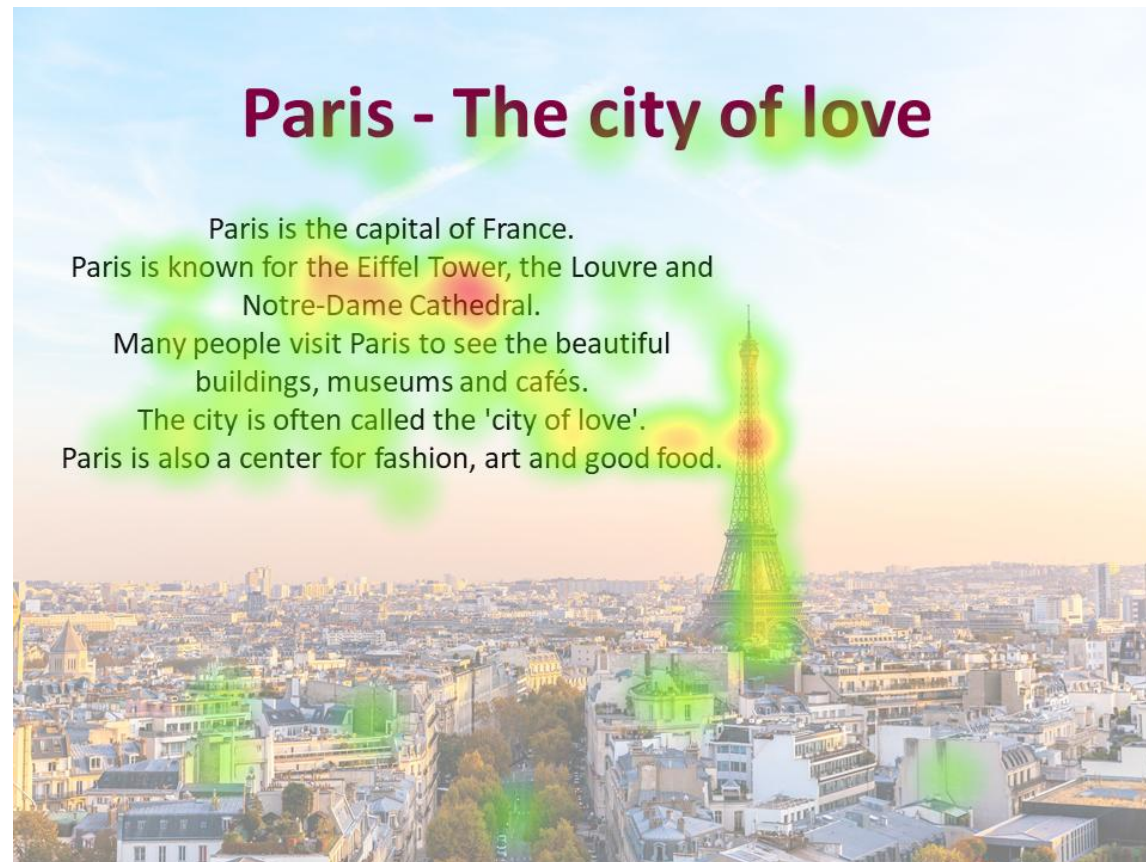
Procedural Information Heatmap



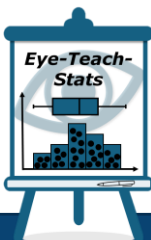
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

In heatmaps, gaze data captured by eye-tracking devices is superimposed on the original stimulus to create a combined view of both the stimulus and the participants' gaze data.



The visualization shows colored areas that represent fixations or saccades. The distribution of fixations, in which the participant remains relatively stable on an area of the stimulus, is indicated by the coloring. The redder and brighter the colors are, the longer this area is viewed.



Procedural Information

Gaze video



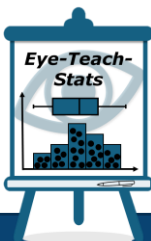
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

In a gaze video, 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.



The visualization includes dots representing fixations and lines representing saccades. Each dot indicates a moment when the participant's gaze remains relatively stable on a specific area of the stimulus. The lines indicate rapid eye movements between two points.



Procedural Information

Gaze plot



Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

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.



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 typically 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.



Amaya's solution

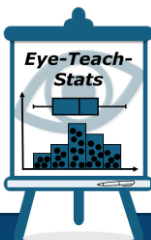
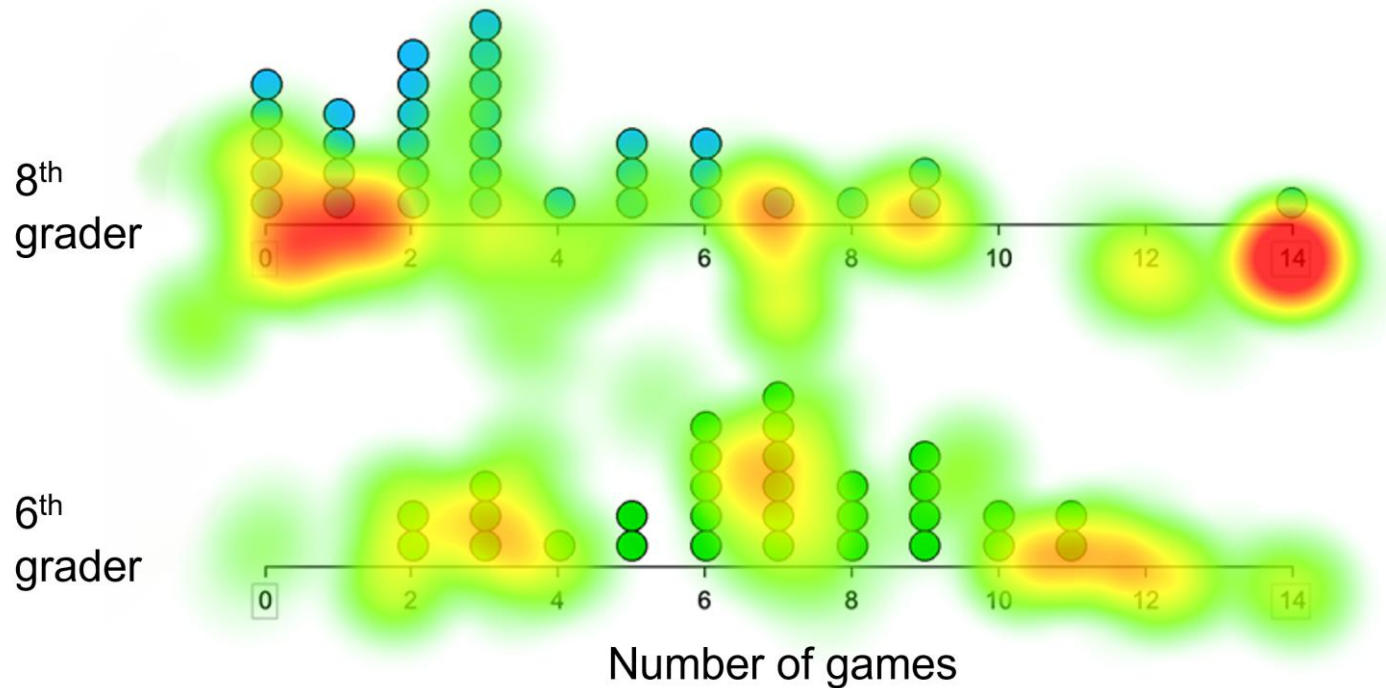


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

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

Amaya'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.”



Quinn's solution



Co-funded by
the European Union

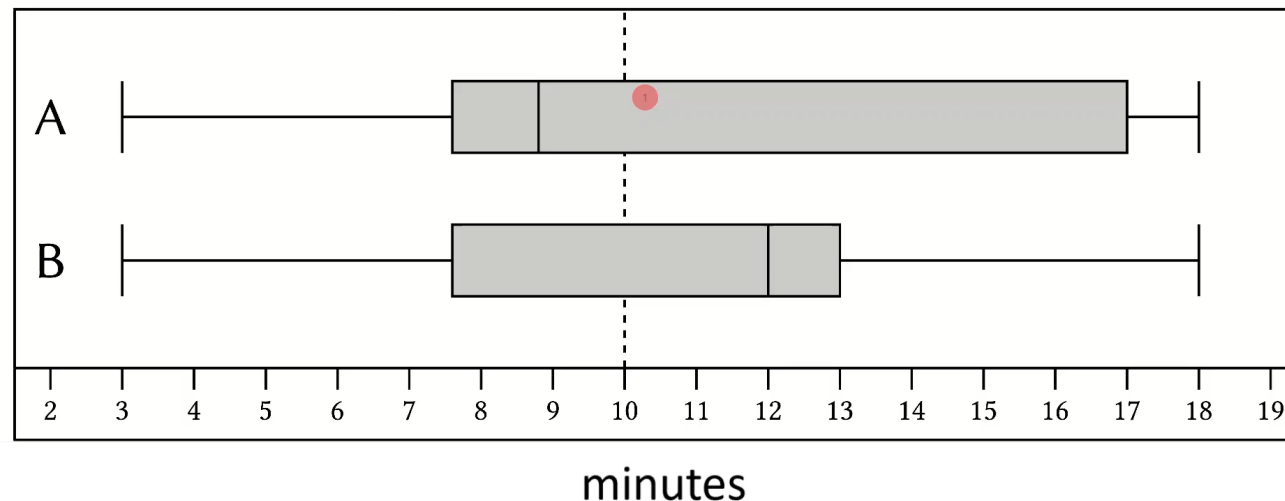
Erasmus+
Enriching lives, opening minds.

Quinn's Answer: „For me, School A is clearly the right answer. You can easily see from the box plot A that many more pupils stated that they walked more than 10 minutes to school.“

Quinn's Comment: "The box plot represents all the responses, so to speak. So I checked where more of the box is above the 10-minute limit. This is quite clearly the case with box plot A. So I also opted for school A."

One school has a larger number of children with a walk to school longer than 10 minutes.

Decide which one.



Elodie's solution

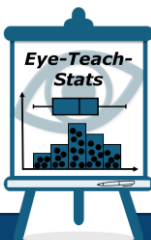
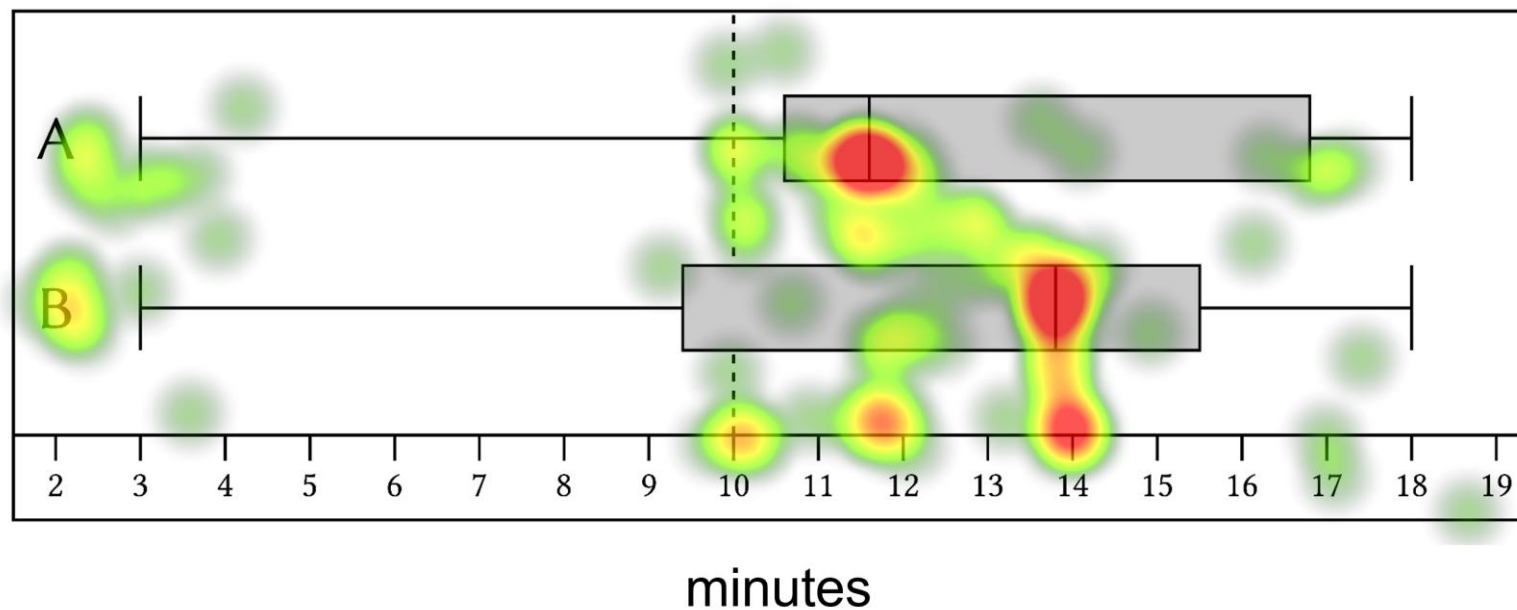


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Elodie's Answer: „For me, School B is clearly the right answer. You can easily see from the boxplot B that many more pupils stated that they walked more than 10 minutes to school.“

Elodie's Comment: “The line in the middle of the box is the median, which is an average value. So I mainly compared the two medians. You can clearly see that the children from school B take much longer on average to get to school than those from school A. So I selected B.”



Roos' solution

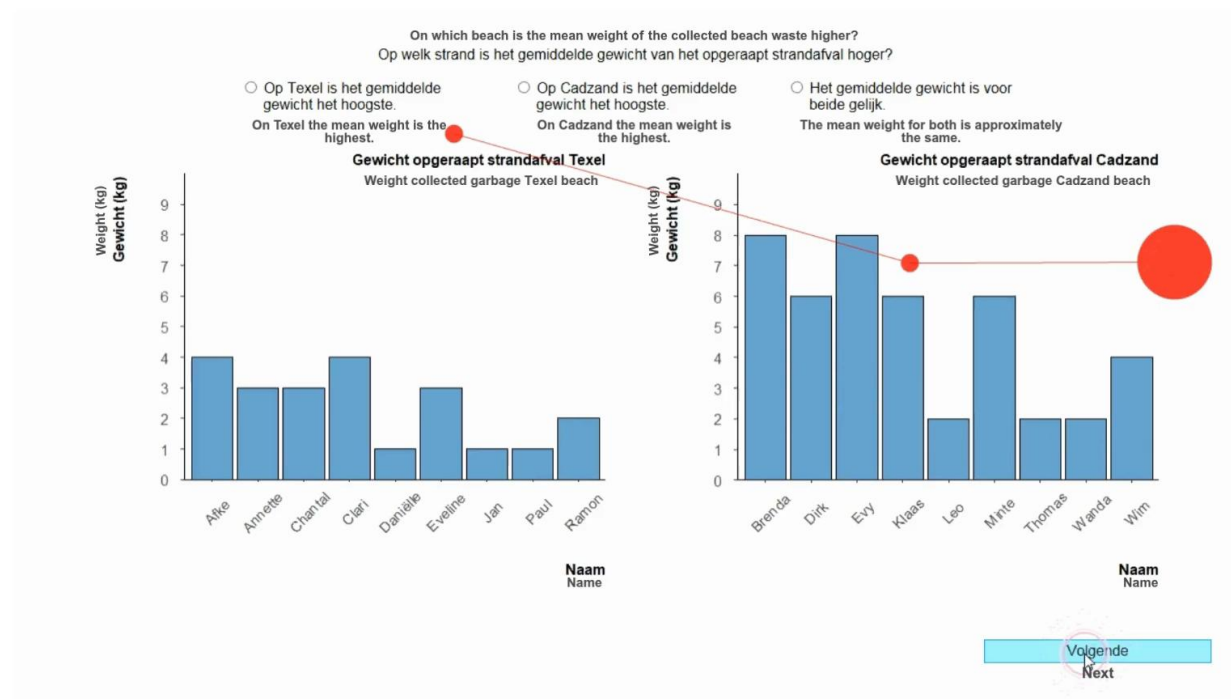


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Roos' answer: "The mean weight for both is approximately the same."

Roos' comment: "I thought the means were the same because I looked at the weights in both graphs. When I saw that the weights in one graph were twice as high, I figured it was the same pattern but with higher values. So, I thought the means would still be the same."



Aline's solution

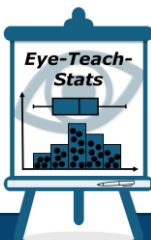
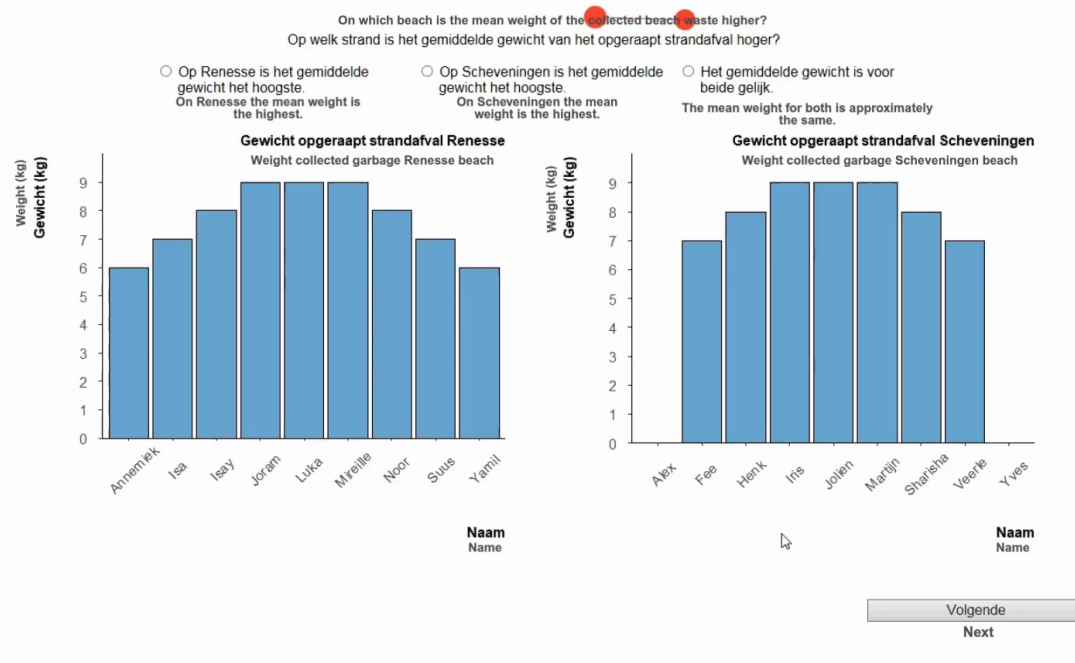


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Aline's Answer: "The mean weight for both is approximately the same."

Aline's comment: "When I compared the graphs of the weights collected from the two beaches, I thought the mean weights were pretty much the same. I looked at the graphs and tried to see if the weights balanced out on both sides. I figured that if the really heavy and really light weights looked even on each end, they would cancel each other out, so the means would be about the same. I saw that the bars on both ends of the graphs were the same distance from the middle, so I thought they balanced each other out."



Zeyd's solution

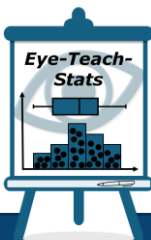
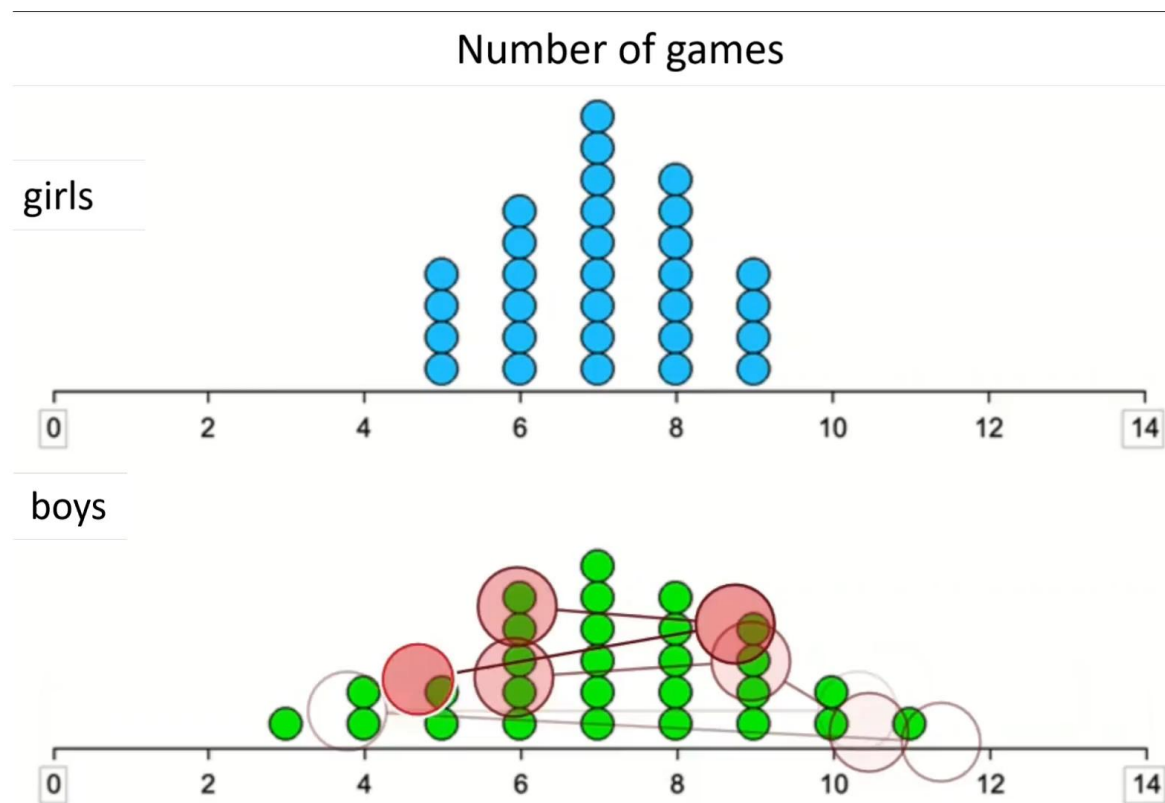


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Zeyd's answer: „I would say a). The girls have a lot more games on their phones”

Zeyd's comment: „So I looked at how many dots there were and then I compared them. For example, boys and girls have the same number of dots at nine, but the girls have more dots at seven and eight and so on. The little stacks at the outside from the boys is not so important. That's why I think the girls have more games.”



Sem's solution

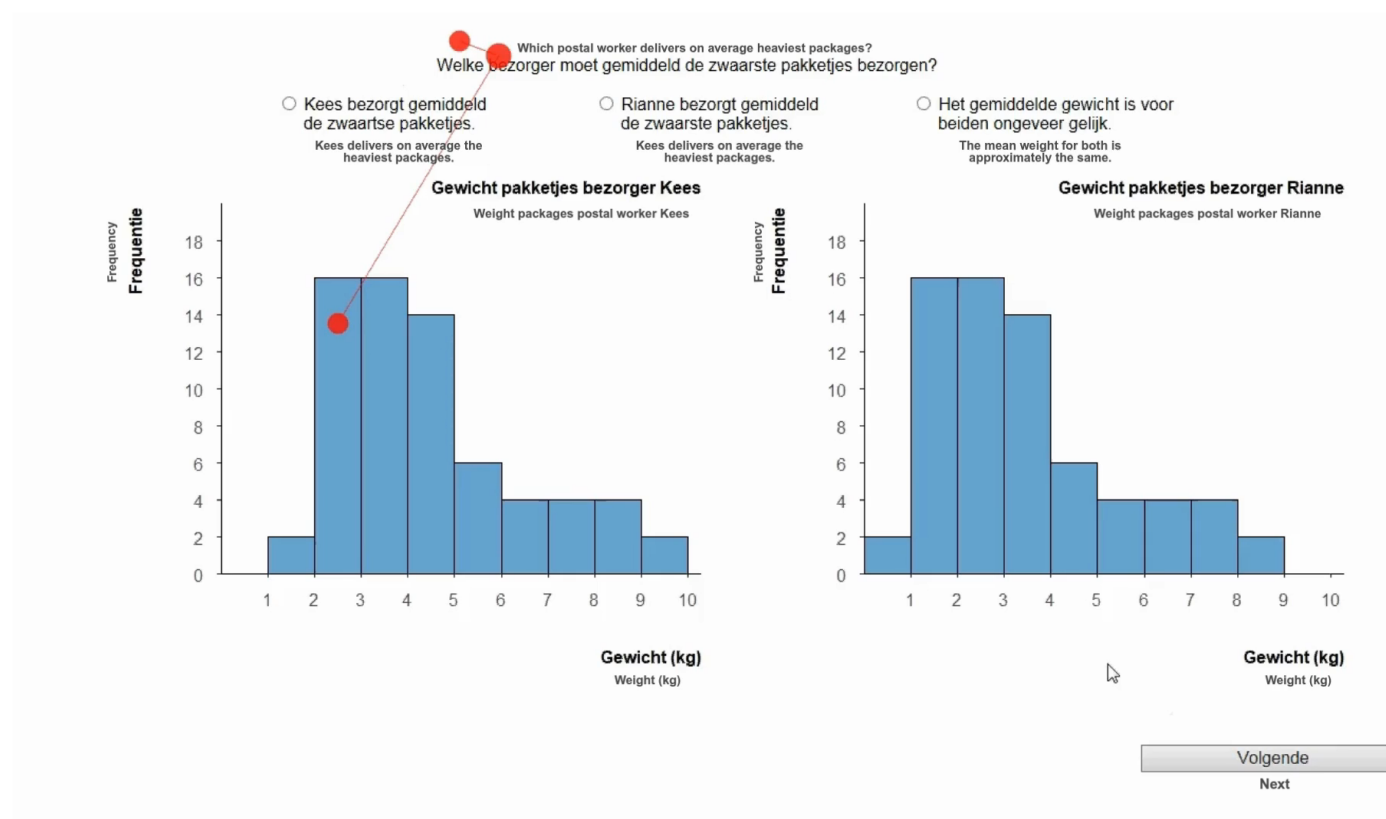


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Sem's answer: „The mean weight for both is approximately the same.”

Sem's Comment: “I saw that Kees had heavier packages with fewer light ones, and Rianne had mostly lighter packages with fewer heavy ones. But the distributions were similar. I compared each pair of stacks across the histograms, and it looked like the differences in stacks could cancel each other out, so I think the approximate means were the same.”



Eva's solution

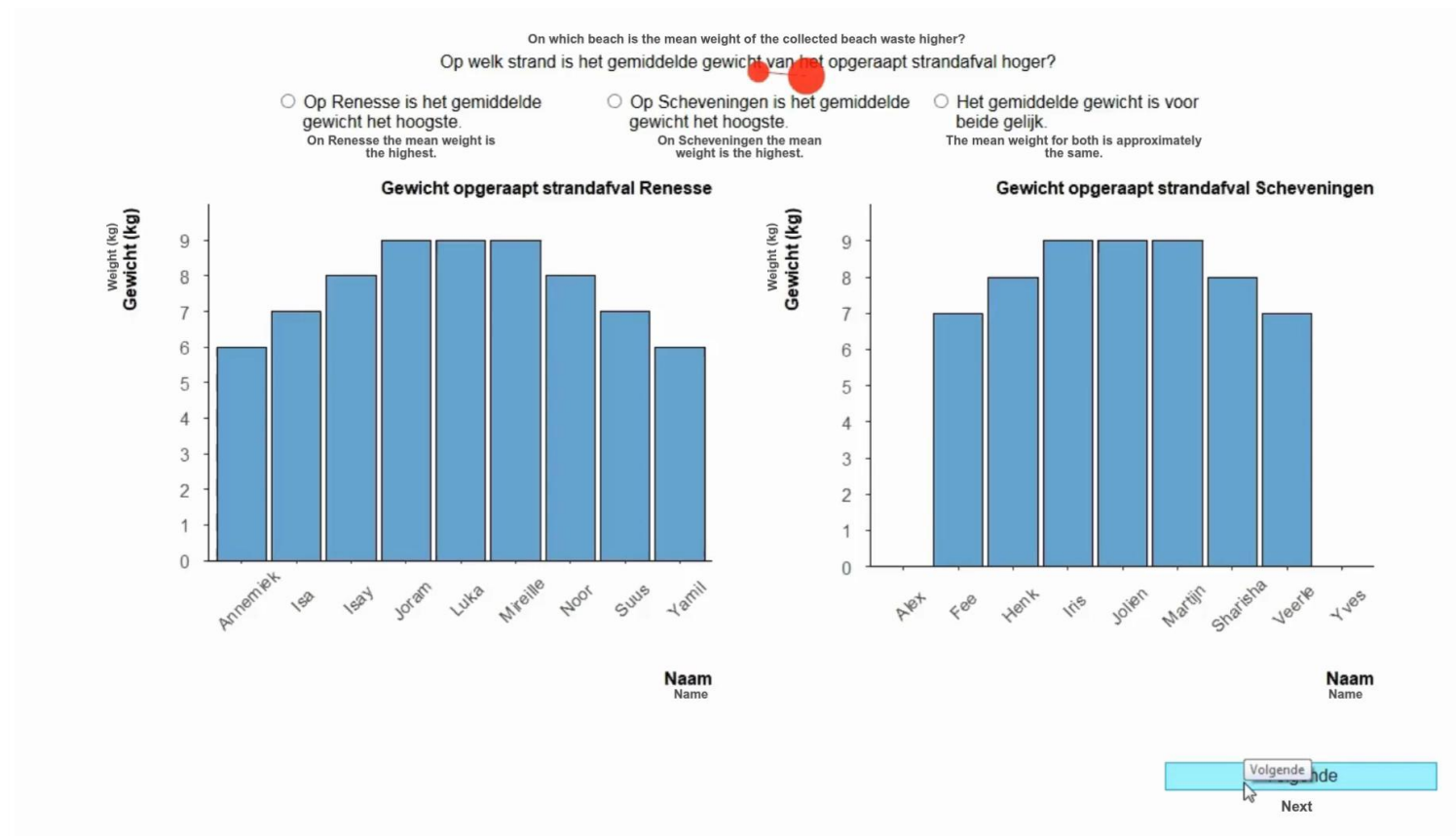


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Eva's answer: "On Scheveningen the mean weight is the highest."

Eva's comment: "I compared some bars from each graph one by one and saw that Renesse had more bars with lower values. So, I thought Scheveningen must have a higher beach waste weight."



Sophie's solution

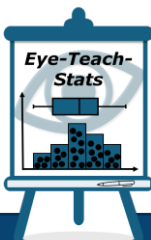
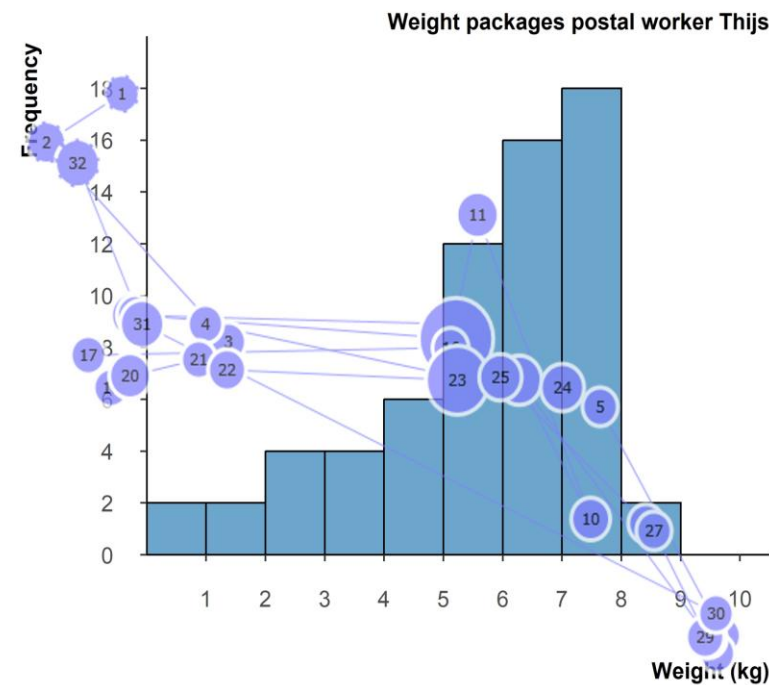


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Sophie's Answer: "The arithmetic mean is about eight."

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."



Noel's solution

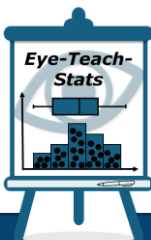
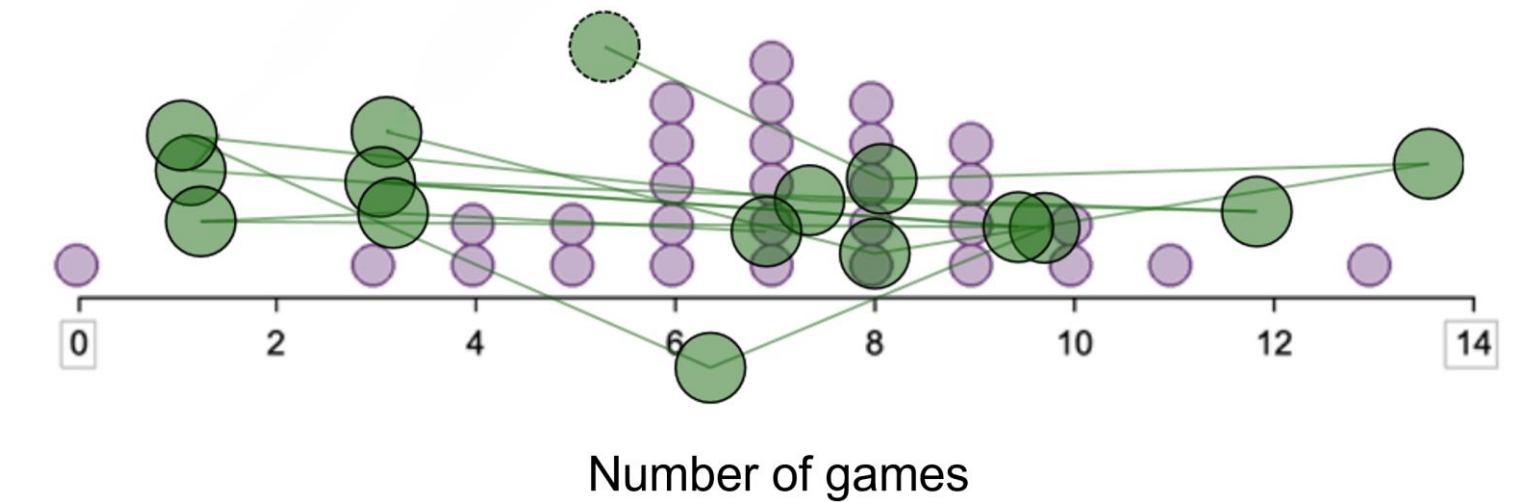


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Noel's Answer: „The arithmetic mean is about two.“

Noel's Comment: "I think it's two because when I look at it, there are some with many games and some with no games and I tried to make the stacks equally high to estimate the arithmetic mean. Then there are about two dots everywhere."



Talia's solution

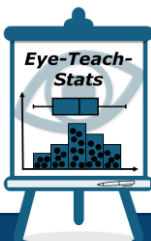
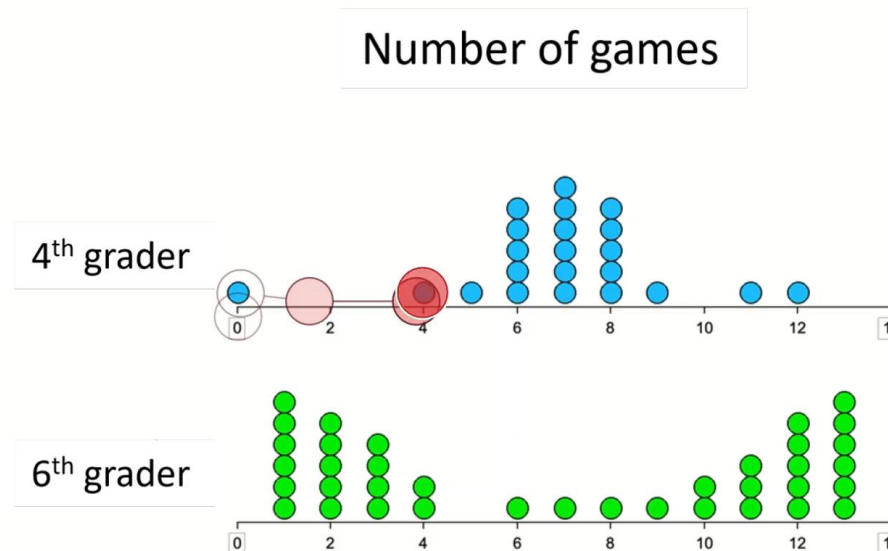


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Talia's Answer: „Therefore, I would say b) is the correct answer: 6th graders have a lot more games on their phones.”

Talia's Comment „First, I counted four plus five, which makes nine. Here, I counted five people who have six games, and calculated five times six, which makes thirty. Then I added the nine from before, making a total of 39. Then I counted... ..and here I added 9, 11, and 12, which makes a total of 134 games for the fourth graders. For the sixth graders, I just calculated 6 times 13 plus 5 times 12, and that was already more than the fourth graders have..”



Mustafa's solution

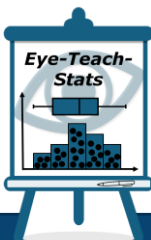
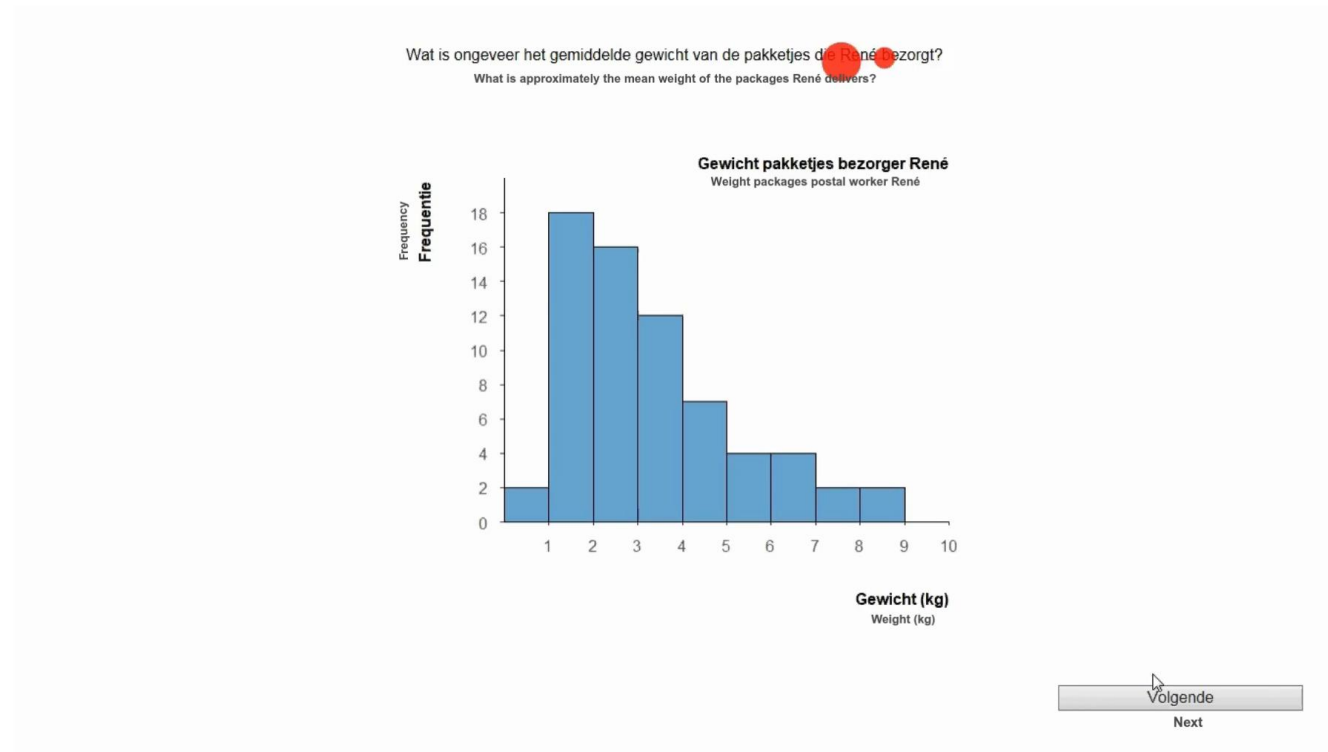


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Mustafa's Answer: "The arithmetic mean is about seven."

Mustafa's Comment: "I added all the frequencies together and then divided by nine. And then I got a number that should be about the mean."



Finn's solution

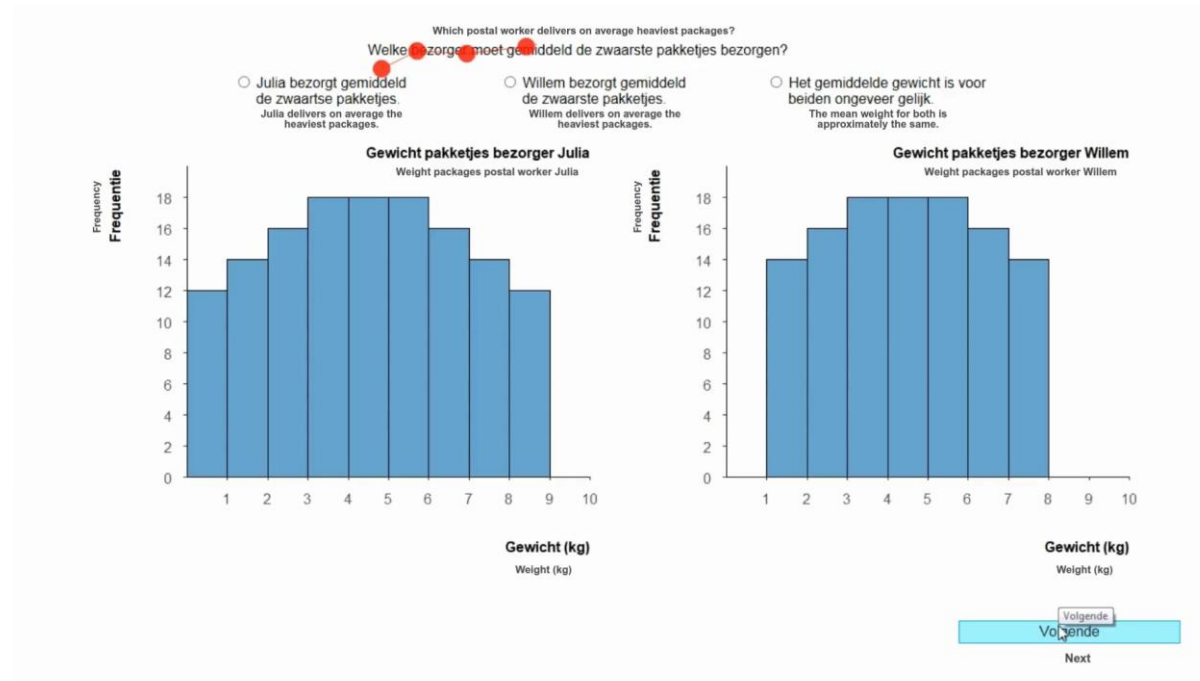


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Finn's Answer: "Julia delivers on average the heaviest packages."

Finn's Comment: "I looked at the bars and compared them. I saw that Willem has fewer bars and Julia has more bars spread across the range. Since we're talking about mean weight, I thought having more bars, like in Julia's histogram, meant a higher mean weight. Even though there are missing bars for Willem, I thought that meant there was nothing in those ranges. So, I chose Julia because I saw more bars and assumed that increased the mean weight."



Linda's solution

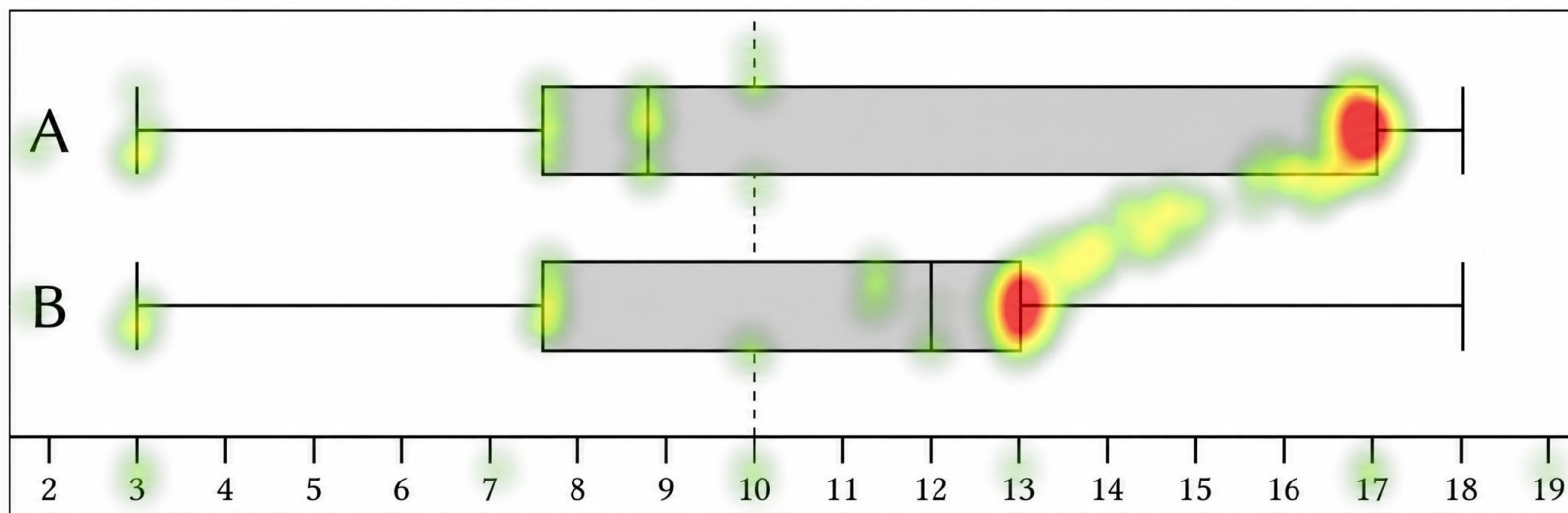


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Linda's Answer: „For me, School A is clearly the right answer. You can easily see from the box plot A that many more pupils stated that they walked more than 10 minutes to school.“

Linda's Comment: "The box plot represents all the responses, so to speak. So I checked where more of the box is above the 10-minute limit. This is quite clearly the case with box plot A. So I also opted for school A."



Paul's solution

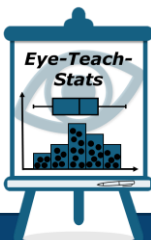
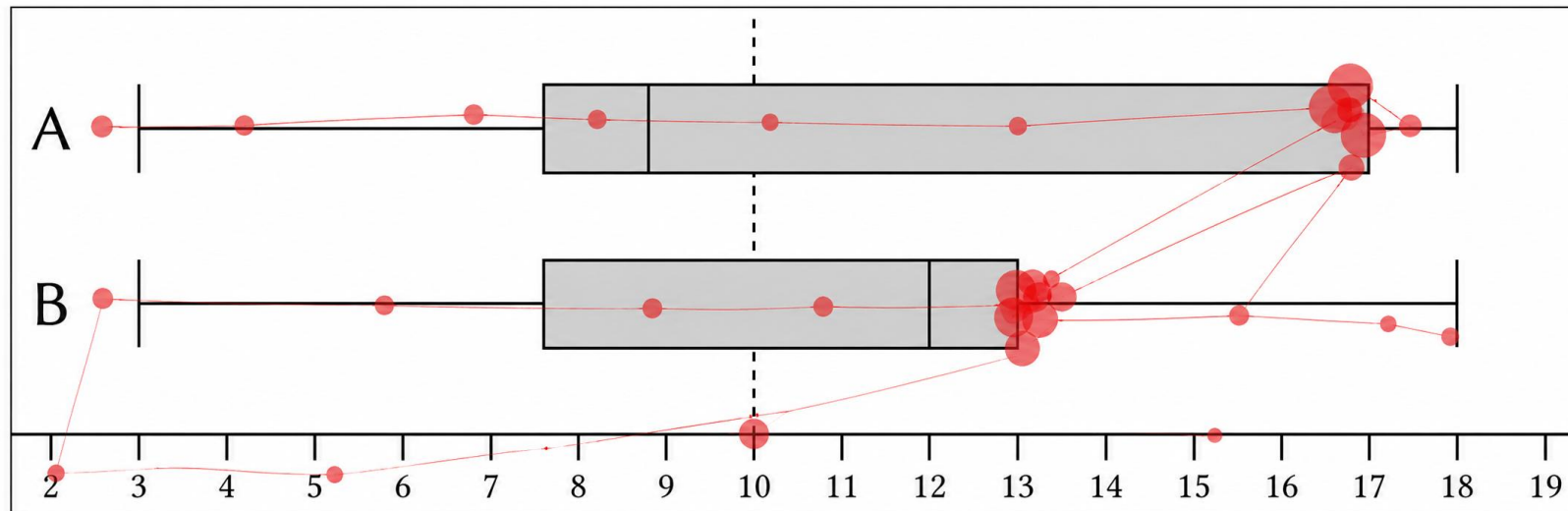


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Paul's Answer: „For me, School A is clearly the right answer. You can easily see from the box plot A that many more pupils stated that they walked more than 10 minutes to school.“

Paul's Comment: "The box plot represents all the responses, so to speak. So I checked where more of the box is above the 10-minute limit. This is quite clearly the case with box plot A. So I also opted for school A."



Katharina's solution



Co-funded by
the European Union

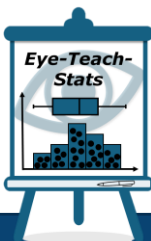
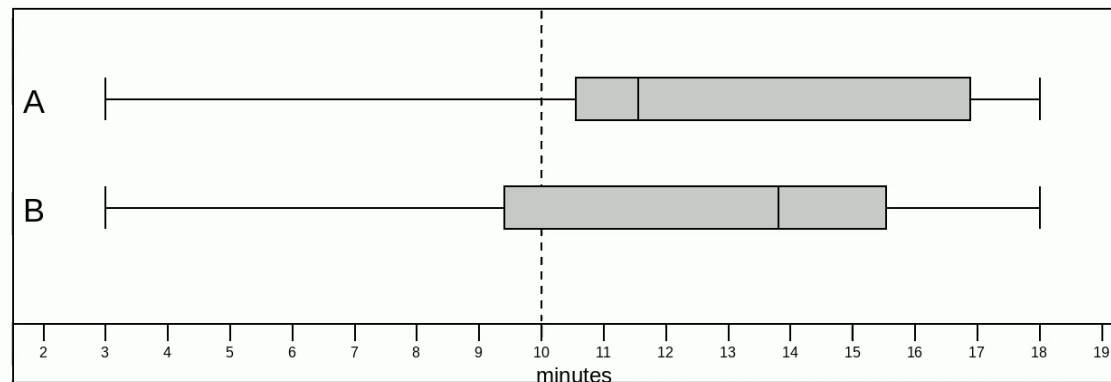
Erasmus+
Enriching lives, opening minds.

Katharina's Answer: „For me, School B is clearly the right answer. You can easily see from the boxplot B that many more pupils stated that they walked more than 10 minutes to school.“

Katharina's Comment: “The line in the middle of the box is the median, which is an average value. So I mainly compared the two medians. You can clearly see that the children from school B take much longer on average to get to school than those from school A. So I selected B.”

One school has a larger number of children with a walk to school longer than 10 minutes.

Decide which one.



Fabian's solution



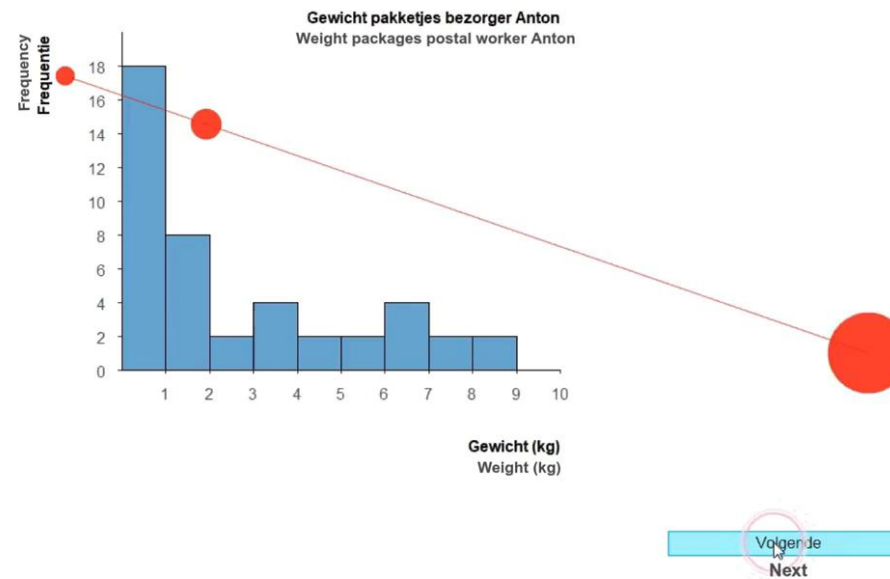
Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Fabian's Answer: "The arithmetic mean is about six."

Fabian's Comment: "I looked at the values on the axis — I saw mostly low ones and one high one, so I estimated it would be somewhere around six."

What is approximately the mean weight of the packages Anton delivers?
Wat is ongeveer het gemiddelde gewicht van de pakketjes die Anton bezorgt?



Ella's solution

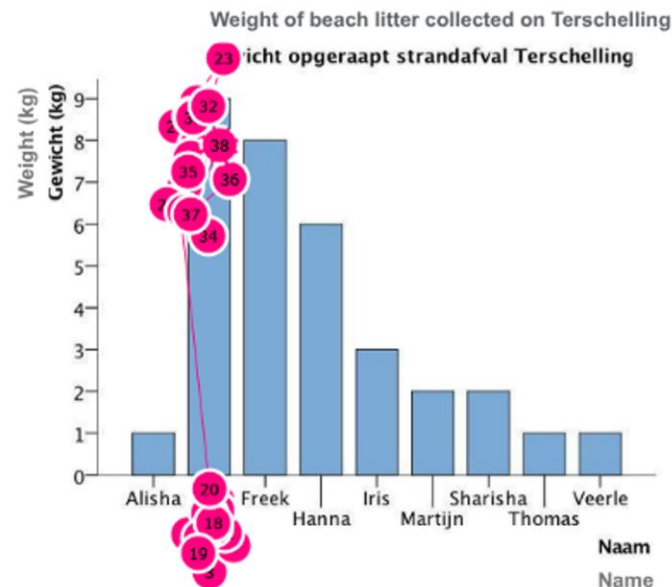


Co-funded by
the European Union

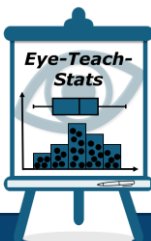
Erasmus+
Enriching lives, opening minds.

Ella's Answer: "The mean weight is nine."

Ella's Comment: "I looked at where the graph was kind of in balance, and that seemed to be around Bram, so I took 9 as the mean."



What is approximately the mean weight collected per person?



Martijn's solution

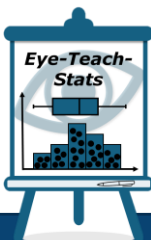
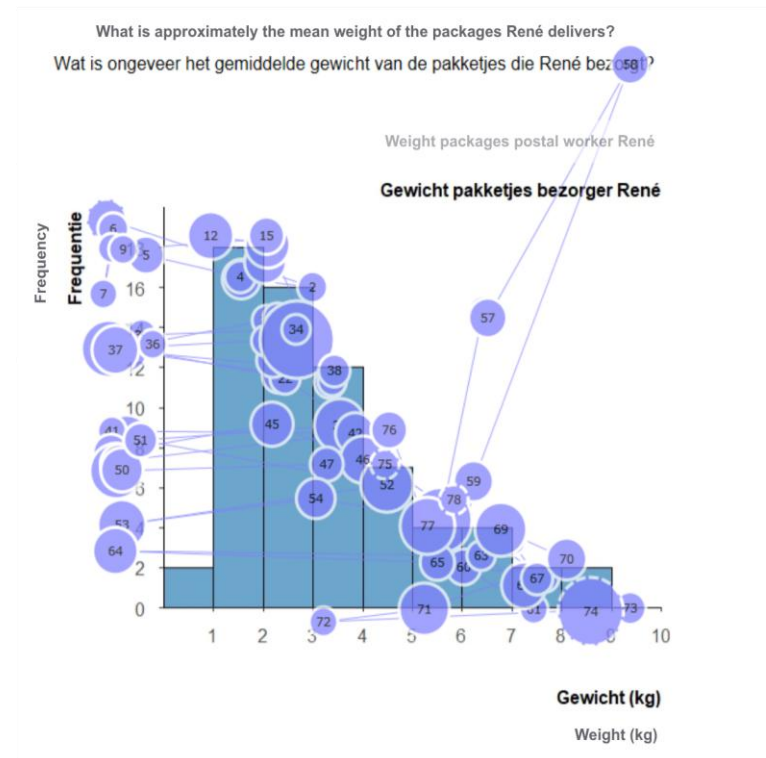


Co-funded by
the European Union

Erasmus+
Enriching lives, opening minds.

Martijn's Answer: "The arithmetic mean is about six."

Martijn's Comment: "Reading off how many each bar was and adding those together. So I started with two, then eighteen, then sixteen, and I kept adding that up as I looked at the bars. Then I divided by nine."



Erika's Answer: "Titia delivers on average the heaviest packages."

Erika's Comment: "I wanted to find the mean weight, so I looked at each bar. For every bar, I multiplied the weight value by the frequency with which it occurred. Then I added all these results together and divided by the number of bars, because each bar represents one part of the data. I did this for both histograms and got a higher result for Titia, so I chose Titia."

