

Routledge
Taylor & Francis Group

Educational Psychology

An International Journal of Experimental Educational Psychology

ISSN: 0144-3410 (Print) 1469-5820 (Online) Journal homepage: http://www.tandfonline.com/loi/cedp20

Refutational text and multiple external representations as a method to remediate the misinterpretation of box plots

Stephanie Lem, Kathy Baert, Eva Ceulemans, Patrick Onghena, Lieven Verschaffel & Wim Van Dooren

To cite this article: Stephanie Lem, Kathy Baert, Eva Ceulemans, Patrick Onghena, Lieven Verschaffel & Wim Van Dooren (2017) Refutational text and multiple external representations as a method to remediate the misinterpretation of box plots, Educational Psychology, 37:10, 1281-1300, DOI: 10.1080/01443410.2017.1333574

To link to this article: https://doi.org/10.1080/01443410.2017.1333574

Published online: 02 Jun 2017.

Submit your article to this journal

Article views: 93

View related articles

View Crossmark data

Citing articles: 1 View citing articles

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=cedp20

EDUCATIONAL PSYCHOLOGY, 2017
VOL. 37, NO. 10, 1281–1300
https://doi.org/10.1080/01443410.2017.1333574

Routledge
Taylor & Francis Group

Check for updates

Refutational text and multiple external representations as a method to remediate the misinterpretation of box plots

Stephanie Lem[a], Kathy Baert[a], Eva Ceulemans[b], Patrick Onghena[b], Lieven Verschaffel[a] and Wim Van Dooren[a]

[a]Centre for Instructional Psychology and Technology, KU Leuven, Leuven, Belgium; [b]Methodology of Educational Sciences Research Group, KU Leuven, Leuven, Belgium

ABSTRACT

The ability to interpret graphs is highly important in modern society, but has proven to be a challenge for many people. In this paper, two teaching methods were used to remediate one specific misinterpretation: the area misinterpretation of box plots. First, we used refutational text to explicitly state and invalidate the area misinterpretation of box plots. Second, we used multiple external representations (MERs): Histograms were used as an overlay on box plots in order to give students a better insight in the way box plots represent data distributions. Third, we combined refutational text and MERs. We found that refutational text was successful in improving students' interpretation of box plots, but that the use of MERs did not improve students' interpretation of box plots. The addition of MERs also did not increase the effect of refutational text.

ARTICLE HISTORY
Received 11 August 2016
Accepted 18 May 2017

KEYWORDS
Graph interpretation; refutational text; multiple external representations; conceptual change; intervention

Introduction

The ability to read and interpret graphs is highly important in modern society, but this is more difficult than it might seem (Friel, Curcio, & Bright, 2001). In this article we report a study on one specific graphical misinterpretation in box plots and test interventions that can help students to overcome this misinterpretation. More specifically, we studied how students are able to overcome the area misinterpretation in box plots by using refutational text and multiple external representations (MERs), both separately and in combination.

Before we present this study, we will first discuss graphical misinterpretations in general, the importance of the data distribution concept in statistics, and the 'area misinterpretation' of box plots, which is the focus of this study. We will then summarise various teaching techniques that have already been tested to correct this misinterpretation, after which we zoom in on refutational text and MERs and their possible effects on the area misinterpretation.

Graphical misinterpretations

Graphicacy refers to the ability to understand quantitative phenomena that are presented graphically (Wainer, 1992). It is a competency that is becoming more and more important

CONTACT Stephanie Lem Stephanie.Lem@kuleuven.be

© 2017 Informa UK Limited, trading as Taylor & Francis Group

in our society (Friel et al., 2001). Numbers are all around and are often presented in a graphical manner: bar charts in the newspaper, a line graph in the television news or a pie chart in a magazine. On top of that, graphs are very important in various school subjects besides mathematics, like history, geography and biology (McKenzie & Padilla, 1986). Despite this omnipresence, interpreting graphs is not as easy as it might seem and many misinterpretations have been reported (e.g. Friel et al., 2001). Many difficulties are based on the fact that people try to interpret graphs as literal representations of the real world. Bell and Janvier (1981), for example, describe how secondary school pupils interpret the time/distance graph of a racing car as a representation of the shape of the racing track. A similar misinterpretation entails thinking that a horizontal line in a velocity-time graph means that the object is at rest (Ainsworth, 2008).

Also in statistics a lot of graphs are used, for example to visualise data distributions. In what follows we will discuss some of these misinterpretations and focus on the misinterpretation at hand for this study: the area misinterpretation of the box plot.

Data distribution and its graphical representation

Distribution is one of the key concepts of statistics as it forms the basis for statistical decision-making. The distribution of one variable is the most basic type of distribution. Its understanding can be considered a prerequisite for understanding more complex distributions, such as sampling distributions (e.g. Meletiou & Lee, 2002). The distribution of a set of observations of one variable – for instance body length – can be represented in several ways. One can list all observations in the sample one by one, but certainly with a large number of observations this representation does not allow to assess the distribution as a whole and judge its characteristics such as the centre, spread and shape. Another option is to summarise the distribution by means of a number of descriptive statistics that refer to these characteristics, such as the mean, median, standard deviation and skewness. Research has shown that students have many misinterpretations of these concepts (e.g. Garfield & Ben-Zvi, 2008; Wild, 2006).

Another way to make data distributions accessible is through one or more graphical representations, such as dot plots, histograms or box plots. Also for such graphical representations of data distributions, various misinterpretations have been reported. In histograms, for example, students use the height differences between the bars instead of the range and shape of the bars as an indicator of variation (Baker, Corbett, & Koedinger, 2002), or they interpret the histogram as representing an evolution in time (delMas, Garfield, & Ooms, 2005). Box plots are also often misinterpreted. Students think, for example, that the whiskers are confidence intervals and that they do not represent any data (Lem, Onghena, Verschaffel, & Van Dooren, 2013a). They also sometimes interpret the area of the box plot as representing the amount of data in that interval (Lem, Onghena, Verschaffel, & Van Dooren, 2013b). This last misinterpretation is the focus of the present article. Box plots have become more and more popular for displaying data distributions since their introduction in the 1970s (Tukey, 1977). Box plots are part of most statistics curricula at the university level, but are also taught at the secondary level (National Council of Teachers of Mathematics, 2000). Reasons for their popularity include the fact that they provide an overview of a distribution that can be viewed, facilitate the comparison of multiple distributions using little space, and are easy to draw by hand (Bakker & Hoffmann, 2005). However, the

interpretation of box plots is not as straightforward as it might seem (Lem et al., 2013a), making it difficult for students to interpret given data-sets, but also makes it more difficult to learn about the underlying concepts of distribution and different descriptive statistics. In what follows, we will explain what the area misinterpretation is and discuss related research.

The area misinterpretation

Box plots are frequently used to represent data distributions, but it has recently been shown that this representation is often misinterpreted. One misinterpretation, the area misinterpretation, has recently been studied systematically (Lem, Onghena, Verschaffel, & Van Dooren, 2014; Lem et al., 2013a, 2013b) and is the misinterpretation under investigation in this study. In order to explain this misinterpretation, we will use Figure 1 to explain how box plots represent data-sets. A box plot represents the data-set using five summary numbers: the minimum (represented by the end of the left horizontal line), the first quartile (Q1; represented by the most left vertical line), the median (represented by the middle vertical line), the third quartile (Q3; represented by the most right vertical line) and the maximum (represented by the end of the right horizontal line). Although the part between 9 and 18 is much larger than the part between 3 and 5, approximately 25% of the data-set is represented in each of these two parts. This is a counter-intuitive interpretation, as a larger part in most graphs stands for 'more' and it is this discrepancy between the size of an interval and the number of data points it represents that is difficult to understand for many people, including experts (Lem, Onghena, et al., 2014).

By using accuracy and reaction time data on different item types, Lem et al. (2013b) concluded that the area misinterpretation is caused by fast and intuitive reasoning, based on superficial visual cues. Moreover, they showed that it is difficult to overrule this intuitive reasoning: The explicit instruction to think each answer over carefully led to longer reaction times, but to equally inaccurate answers as with the instruction to answer as quickly as possible (Lem et al., 2013b). Also after a 10-min one-on-one instruction about box plots, students still showed signs of this incorrect intuitive reasoning (Lem et al., 2013b). Finally, Lem, Onghena, et al. (2014) showed that this misinterpretation affected even the reasoning of expert users of box plots. It hence seems important to find teaching techniques that can help students to interpret box plots correctly. We will discuss various options in the next section.

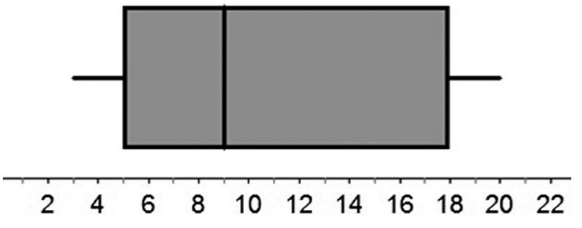


Figure 1. Example of a box plot.

Remediating misinterpretations

Because the area heuristic has been shown to be very persistent and difficult to overcome, we looked for ways of remediating this misinterpretation through instruction. We found two interesting instructional techniques for the remediation of the area misinterpretation that are associated with two different literatures: the conceptual change literature and the instructional design literature. We will discuss both techniques below.

Refutational text

The conceptual change literature focuses on ways to transform misconceptions into conceptions that are in line with current scientific concepts. Although the area misinterpretation is caused by incorrect intuitive reasoning (Lem et al., 2013b) and hence strictly speaking cannot be considered to be a misconception (while a misconception is formed *before* the correct knowledge is acquired or is the result of incorrect knowledge acquisition (Méheut, 2012) the area of a box plot is misinterpreted *despite* people having the correct knowledge), it seems likely that methods that are designed to remediate misconceptions may also have an impact on the eradication of incorrect intuitions as both misconceptions and heuristics are very persistent and students are not aware of the fact that they hold these incorrect intuitions or misconceptions.

In a refutational text, a common (but incorrect) conception is explicitly stated, after which it is refuted and replaced by an alternative (correct) theory (Hynd, 2001). For instance, (Tippett, 2004), used the following text to teach pupils that ostriches do not really bury their head in the sand:

> Many people believe that an ostrich will bury its head in the sand when it is in danger. This is not true, however. If ostriches buried their heads, they would not be able to breathe! Ostrich chicks may hide from danger by lying with their necks stretched out along the ground. Adults may listen for sound with their heads near the ground, or they might run away. (p. 202)

Refutational texts consist of two essential elements (Tippett, 2010). The first element is the explicit naming of the incorrect interpretation that is being addressed. This is very uncommon in most traditional textbooks, most probably out of fear that students would remember this incorrect interpretation as being correct, as it was part of the text. The second element consists of the explicit statement that this interpretation is incorrect and why, and the provision of a new, correct interpretation.

A positive effect of refutational text has been found in various areas already, but mostly in the area of science and mathematics education, in which many misinterpretations and misconceptions have been reported and studied (Tippett, 2010). Refutational text can lead to conceptual change by letting readers simultaneously activate incorrect prior knowledge and correct conceptions (van den Broek & Kendeou, 2008). This way, refutational text can help readers to find inconsistencies between both. This co-activation of conceptions is seen as a crucial step in the remediation of misconceptions and misinterpretations.

Multiple external representations

Another way of improving the interpretation of representations can be found in the instructional design literature. By using the simultaneous presentation of a second, more familiar, representation the interpretation of the first representation is 'constrained' by the interpretation of the second presentation. This constraining function is one of the three functions

the use of MERs can have, according to the classification of (Ainsworth, 1999) and can take two forms: constrain by familiarity or constrain by inherent properties. When using a dot plot or histogram to constrain the interpretation of box plots, both of these forms can help in the interpretation of the box plot. First, dot plots and histograms are more familiar to students and can hence scaffold the interpretation of box plots when the student goes back and forth between both representations, checking their interpretation of the box plot with the information they get from the dot plot or histogram. Second, dot plots and histograms have the property of providing a more fine-grained representation of the data distributions than the box plot does. This way their interpretation can constrain the misinterpretation of the area of the box plot.

It has already been shown in various domains that MERs can help students (Corradi, Elen, & Clarebout, 2012; Tsui & Treagust, 2003; Van Dooren, De Bock, Hessels, Janssens, & Verschaffel, 2004; Waldrip & Prain, 2012). Also for the use of MERs in teaching box plots, some empirical evidence exists. Bakker (2004) used interactive applets in which box plots were built on top of dot plots in order to teach secondary school students about box plots (see Figure 2). Although there were no pre- and post-tests and no control group, he concluded that the area misinterpretation did not occur frequently in his participants, suggesting that the use of this combination of representations helped the students. Lem et al. (2013b) based their intervention on the study of Bakker (2004) by also building box plots on top of dot plots. This intervention of Bakker seemed to be successful in most participants, although some participants' reasoning did not improve at all. On top of that, it was shown that even the students who did improve their interpretation were still affected by the area misinterpretation, based on reaction time patterns.

Combining refutational text and MERs

A recently conducted study (Lem, Kempen, et al., 2014) already suggested that it is possible to use refutational text and MERs in combination to remediate the area misinterpretation of box plots. However, as will be explained below, there were some limitations to that study that warranted a new, more systematic experiment. Lem et al. evaluated the effect of four different versions of a home-study text on university students' box plot understanding: a control text (regular explanation about histograms and box plots, without a connection being made between both representations), a refutational text (refutational text was used to explicitly refute the area misinterpretation), a MERs text (box plots were explained by overlaying them on histograms), or a combination text (MERs and refutational text were combined). The conclusion of this study was that both refutational text and MERs help to remediate the area misinterpretation in box plots, but with different effects on different types of items.

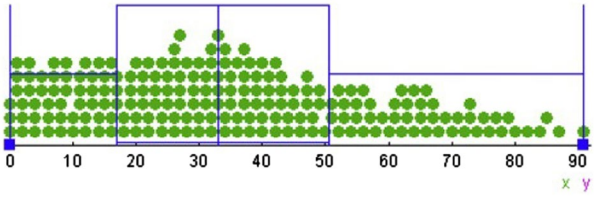


Figure 2. Box plot built on top of a dot plot. Source: Cobb, Gravemeijer, Bowers, and Doorman (1997).

Goal of the present study

The goal of this study was, first, to test whether refutational text and MERs would improve students' interpretation of box plots, and, second, whether the combination of refutational text and MERs would be even more effective.

A few changes were made to the study in order to obviate some of the shortcomings in the study of Lem, Kempen, et al. (2014). First, we included a pre-test in order to control for the possible differences between participants concerning their prior knowledge on box plots. Second, we opted to implement the experimental manipulations in actual classroom instruction instead of creating variants of home-study texts. This not only prevented students from being exposed to more than one condition (by uncontrolled interactions with each other before the post-test took place), but also assured that other variables such as time-on-task were kept the same for all participants. Moreover, according to (Tippett, 2010), frontal instruction enlarges the effect of refutational text. Third, the post-test was administered immediately after the intervention, making the time between learning and testing the same for all participants. Fourth, we expanded the test instrument by adding more items on the interpretation of box plots and on the matching of box plots and histograms. This was done to allow a better investigation of the differential effects of the different conditions on these two item types.

The current study is not only a replication of the study of Lem, Kempen, et al. (2014), but the improvements in the study design will also lead to more information concerning the way refutational text and MERs can improve the interpretation of box plots. First, the setting was more ecologically valid, as now the intervention was conducted in a real teaching setting. This makes the results more easily transferrable to real-life teaching. Second, by using an extended post-test, we will be able to get a better view on the individual merits of refutational text and MERS on the remediation of the area misinterpretation.

Method

Participants

Participants were 288 undergraduate students enrolled in the introductory statistics course taught by the third author. Of these 288 students, 199 students (194 women, 5 men; the overrepresentation of female participants is due to student enrollment patterns) participated in both the pre-test and the post-test. These 199 students were enrolled in the bachelor of educational sciences (58.3%) or speech therapy and audiological sciences (41.7%), in which this course was a compulsory course. The average age of the participants was 18.14 years and 78.4% stated that they already learned about box plots during their secondary school mathematics classes. Participation was voluntary, but it was stressed that all material taught during the intervention was part of the regular curriculum.

Pre- and post-test

The pre- and post-test were two parallel tests. Both tests consisted of two parts: four knowledge questions (exactly the same in both the pre- and post-test), and eight items in which we specifically aimed at revealing the area misinterpretation (only the numbers and contexts were different between both tests). The knowledge items asked students to name the

different elements of a box plot, to explain what Q1 means and to state which percentage of observations is represented in the highlighted parts of two given box plots. In five of the area items (box plot only items) one or more box plots had to be interpreted and in the other three area items (matching items) participants had to match a histogram to a box plot (or vice versa). In order to get a better insight in their reasoning, participants were asked to provide an explanation for their response. The 12 items of the pre-test can be found in Appendix.

Intervention

The intervention was done during the practical sessions of the introductory statistics course, about one month after the pre-test. This timing corresponds to the timing at which box plots are introduced in the normal curriculum. Before the intervention, students were already taught about descriptive statistics and histograms, but not yet about box plots. In the beginning of the academic year, students chose which of the eight practicum groups they wanted to enrol in. Based on the average scores on the pre-test per group, the eight existing practicum groups were assigned to the four conditions, thereby creating groups in which the prior knowledge was as equal as possible. A 2×2 design was used, varying whether refutational text and MERs were used in the intervention or not. This led to the following four conditions: control condition ($n = 46$), refutational text condition ($n = 58$), MERs condition ($n = 49$) and the combination condition ($n = 46$).

The intervention consisted of a short lecture of about 20 min by the second author, and about 25 students were present per session. The complete lessons were written out to ensure that all lessons would be as similar as possible, except for the experimental teaching techniques. We also asked students to keep possible questions for the end of the class, as we did not want some sessions to have a larger time-on-task than others. A PowerPoint presentation was used to teach the students about histograms and box plots. The three experimental lectures were based on the *control condition*, in which both histograms and box plots were explained, without making a connection between both representations and without explicitly saying anything about possible misinterpretations. At the end of each session four simple exercises were presented in reading and drawing histograms in box plots. These exercises were briefly discussed, which was the end of the short lecture. In the *refutation* condition, again no connection between histograms and box plots was made, but the area misinterpretation was explicitly stated and refuted, and the correct interpretation was subsequently offered. The refutational elements were verbalised by the teacher, but were also written on the slides in red text in order to emphasise the importance of this information. This is an example of a refutational text as used in the refutation condition:

> When you look at a box plot you might think that a larger part of the box represents more results than a smaller part of the box. This is incorrect! In each of the four parts of a box plot approximately the same number of observations is represented. The size of the parts of the box does hence not give an indication of the number of represented observations. A larger part of the box means that the observations here are more spread out. In a smaller part of the box plot the observations are hence less spread out. In each of the four parts of a box plot approximately 25% of the observations is represented.

In the *MERs* condition, histograms were used as a basis to explain box plots. This was done by adding lines to a histogram step-by-step, until a box plot was formed (see Figure 3 for

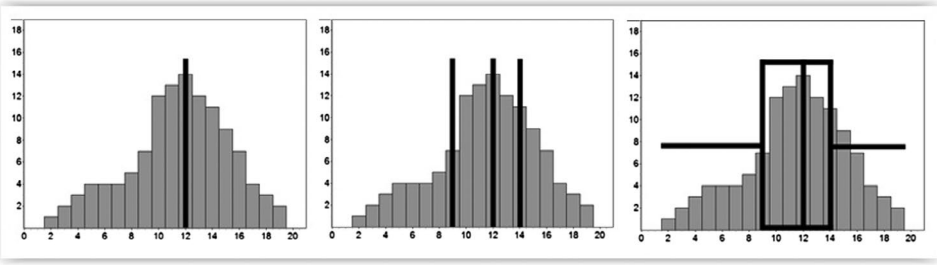


Figure 3. Example of how a box plot was built on top of a histogram in the MERs and combination conditions.

some example images). In the *combination condition*, finally, both MERs and refutational text were combined. This means that box plots were explained by using histograms as a basis, but also that the refutational text used histograms to show how box plots should be interpreted. Directly after the intervention, the post-test was administered.

Analysis

All explanations given by the participants were coded for the demonstration of the area misinterpretation. The second author coded all explanations and the first author did the same for 12% of the explanations. There was an agreement of 91.3% and Cohen's kappa was 0.89, indicating an almost perfect inter-rater reliability.

Results

Box plot knowledge

The box plot knowledge of the participants is presented in Table 1. A generalised linear mixed model with test moment, refutational text, MERs, and their interactions as independent variables, and accuracy as the dependent variable was fitted. Table 2 provides the results of this analysis. We found a significant effect of test moment, with higher accuracy in the post-test (73.49%) than in the pre-test (39.45%). This was a large effect with an odds ratio of 4.25. None of the other tested effects were significant. This suggests that box plot knowledge was the same in all conditions, not only before the intervention, but also afterwards, showing that all conditions were equally successful in improving participants' box plot knowledge. This also means that the differences found between conditions on the area misinterpretations are due to the effect of the intervention, rather than to differences in box plot knowledge.

Accuracy area misinterpretation items

The accuracy rates for the area misinterpretation items are presented in Table 3. A generalised linear mixed model with test moment, refutational text, MERs, and their interactions as independent variables, and accuracy on the area misinterpretation items as the dependent variable was fitted. The results of this analysis are presented in Table 4. First, we found a main

Table 1. Accuracy rates for box plot knowledge items in the pre- and post-test, per condition.

	Pre-test			Post-test		
	No ref	Ref		No ref	Ref	
No MERs	36.96	40.52	38.94	71.20	74.14	72.84
MERs	42.86	36.96	40.00	72.96	75.44	74.21
	40.00	38.94	39.45	72.11	74.76	73.49

Note: MERs = multiple external representations. Ref = refutational text.

Table 2. Results of the generalised linear mixed analysis with accuracy on the box plot knowledge items as dependent variable.

	Num DF	Den DF	*F*	*p*
Test moment	1	1389	178.51	<.001*
Ref	1	1389	0.18	.667
MERs	1	1389	0.35	.553
Test moment × Ref	1	1389	0.76	.383
Test moment × MERs	1	1389	0.02	.880
Ref × MERs	1	1389	0.88	.347
Test moment × Ref × MERs	1	1389	0.78	.379

Note: Ref = refutational text, MERs = multiple external representations.
*Effects that are significant at the .05-level.

Table 3. Accuracy rates for area misinterpretation items in the pre- and post-test, per condition.

	Pre-test			Post-test		
	No ref	Ref		No ref	Ref	
No MERs	40.76	46.12	43.75	65.67	76.94	71.96
MERs	44.64	49.18	46.84	63.68	77.45	70.36
	42.76	47.48	45.23	64.64	77.16	71.19

Note: MERs = multiple external representations. Ref = refutational text.

Table 4. Results of the generalised linear mixed analysis with accuracy on the area misinterpretation items as dependent variable.

	Num DF	Den DF	*F*	*p*
Test moment	1	2979	228.46	<.001*
Ref	1	2979	10.40	<.001*
MERs	1	2979	0.27	.600
Test moment × Ref	1	2979	7.98	.005*
Test moment × MERs	1	2979	1.08	.300
Ref × MERs	1	2979	0.02	.893
Test moment × Ref × MERs	1	2979	0.26	.609

Note: Ref = refutational text, MERs = multiple external representations.
*Effects that are significant at the .05-level.

effect of test moment. This effect shows that accuracy was higher in the post-test (71.19%) than in the pre-test (45.23%). Second, we found a main effect of refutational text, showing that participants exposed to refutational text solved more area interpretation items correctly (75.38%) than participants who were not exposed to refutational text (52.95%). Finally, we found an interaction effect of test moment and refutational text. This effect shows that participants exposed to refutational text showed a larger improvement (from 47.48% to 77.16%, *OR* = 3.73) than participants who were not exposed to refutational text (from 42.67% to 64.64 %, *OR* = 2.55).

We did not find a significant effect of the exposure to MERs, and, moreover, the lack of a significant interaction effect of refutational text and MERs shows that the addition of MERs to refutational text did not significantly add to the positive effect of refutational text.

To get a better understanding of the effect of the interventions, we did some additional analyses on the post-test data. A first analysis aimed to find out whether there was a different effect of the interventions for the two item types: box plot only items and matching items. An overview of the percentages of correct responses per item type is provided in Table 5. For the five box plot only items, we found a main effect of refutational text, $F(1,795) = 26.65$, $p < .001$, $OR = 2.74$, but not of MERs, $F(1,795) = 0.02$, $p = .875$, and also no interaction effect between both teaching techniques, $F(1,795) = 0.87$, $p = .352$. For the three matching items, on the other hand, accuracy rates were much more similar between conditions, resulting in no main effect of refutational text, $F(1,397) = 1.71$, $p = .191$, nor of MERs, $F(1,397) = 0.33$, $p = .567$, and also no interaction effect between both, $F(1,397) = 1.26$, $p = .263$. These results suggest that while refutational text had a positive effect on the solving of box plot only items, none of the tested teaching techniques had an effect on the solving of the matching items.

Furthermore, we looked at the number of participants in each condition that answered every area misinterpretation item in the post-test correctly. This way we could get an insight into the number of students who probably fully understood box plots after the intervention. In the control condition, there were only four such participants (8.7%), in the refutation condition 17 (29.3%), in the MERs condition 9 (18.4%) and in the combination condition 12 (28.1%). Although a χ^2-test on these numbers was not statistically significant, $\chi^2(1) = 2.785$, $p = 0.095$, $V = 0.258$, the effect size is small to medium, suggesting that there tend to be differences between the conditions concerning the number of students who achieved a full understanding of box plots. It is hence interesting to note that more than one-fourth of the students in the conditions with refutational text had a perfect score, while this was less than one-fifth of the students in the MERs condition, and one-tenth of the students in the control condition.

Occurrence of the area misinterpretation

We did not only look at the accuracy rates of the area items, but also at the percentage of items that, in the explanation of their response, showed signs of the area misinterpretation. The occurrence of the area misinterpretation is presented in Table 6. Explanations participants wrote that demonstrate the area misinterpretation are, for example: 'because that box plot is bigger' or 'the part above 10 is bigger, so there are more observations there'. A generalised linear mixed model with test moment, refutational text, MERs, and their interactions

Table 5. Accuracy rates for the box plot only items and the matching items in the post-test, per condition.

	Box plot only items			Matching items		
	No ref	Ref		No ref	Ref	
No MERs	55.02	73.45	65.32	83.33	82.76	83.01
MERs	51.02	77.83	64.00	84.93	76.81	80.99
	52.95	75.38	64.69	84.15	80.13	82.05

Note: MERs = multiple external representations. Ref = refutational text.

Table 6. Occurrence of the area misinterpretation in the pre- and post-test, per condition.

	Pre-test			Post-test		
	No ref	Ref		No ref	Ref	
No MERs	18.48	21.55	20.19	13.04	4.53	8.29
MERs	28.57	20.11	24.47	17.60	4.08	11.05
	23.68	20.91	22.24	15.39	4.33	9.61

Note: MERs = multiple external representations. Ref = refutational text.

Table 7. Results of the generalised linear mixed analysis with occurrence of the area misinterpretation as dependent variable.

	Num DF	Den DF	*F*	*p*
Test moment	1	2981	99.53	<.001*
Ref	1	2981	20.51	<.001*
MERs	1	2981	1.09	.296
Test moment × Ref	1	2981	28.83	<.001*
Test moment × MERs	1	2981	0.32	.570
Ref × MERs	1	2981	3.03	.082
Test moment × Ref × MERs	1	2981	0.18	.674

Note: Ref = refutational text, MERs = multiple external representations.
*Effects that are significant at the .05-level.

as independent variables, and occurrence of the area misinterpretation as the dependent variable was fitted. The results of this analysis are presented in Table 7. First, we found a main effect of test moment, showing that the area misinterpretation occurred more often before the intervention (22.24%) than after the intervention (9.61%). Second, we found a main effect of refutational text, meaning that participants exposed to refutational text showed the area misinterpretation less frequently (23.58%) than participants not exposed to refutational text (6.92%). Finally, we found an interaction effect of test moment and refutational text. This effect shows participants exposed to refutational text showed a larger decrease in the presentation of the area misinterpretation (from 20.91% to 4.33, *OR* = 5.84) than participants who were not exposed to refutational text (from 23.68% to 15.39%, *OR* = 1.71).

We did not find an effect of MERs and also no interaction effect of refutational text and MERs, meaning that MERs did not have an effect on the occurrence of the area misinterpretation and that the use of MERs together with refutational text also did not enlarge the effect of refutational text.

Discussion and conclusion

In this study we used two teaching techniques, refutational text and MERs, both separately and in combination, in order to eradicate the area misinterpretation of box plots in students. In what follows we will discuss the results, their implications for both educational theory and practice, and suggestions for further research.

Our results show that the use of refutational text improves the accuracy on items on the area misinterpretation significantly more than the use of MERs. The same holds for the decrease in occurrence of the area misinterpretation: participants exposed to refutational text showed a larger decline in incorrect responses based on the area misinterpretation than participants not exposed to refutational text. We did not find a similar effect of MERs, and

as we also did not find an added effect of the use of MERs next to refutational text, we could conclude that the use of MERs did not improve participants' interpretation of box plots. When looking at the two different item types, box plot only items and matching items, all conditions scored equally well on the matching items, in which both histograms and box plots had to be interpreted simultaneously, while refutational text clearly led to higher accuracy in the items in which only box plots had to be interpreted. This is slightly different from the results of Lem, Kempen, et al. (2014), who found a trend showing that matching tasks were solved better when MERs were used as an intervention. A possible interpretation of our results is that MERs do not help students in *learning* about box plots as such, but that they are more effective in constraining the incorrect interpretation of box plots during *problem-solving*.

When looking at the percentage of students in each condition that answered all area items correctly, there were remarkably more students who were exposed to refutational text who scored consistently well on all area items than when not exposed to refutational text. This suggests that refutational text is more effective in providing students with a correct mental model of the box plot. For MERs no such effect was found. However, we must note that refutational text was no panacea: Almost one-fourth of the items solved by participants exposed to refutational text still were answered incorrectly.

Although these results are promising, we must take some limitations into account. First, the design of our study did not allow to collect extensive qualitative data. By only using multiple-choice questions with a short written explanation, it is difficult to get a full understanding of the way participants interpret box plots. Second, because of curricular restrictions, we were not able to perform a delayed post-test. Because of this, we do not know whether the effects of the intervention lasted after the post-test. It is hence possible that the interventions only led to short-term effects or that for example the addition of MERs to refutational text led to longer lasting effects than the use of refutational text alone. Future research could study the long-term effects of (combining) refutational text and MERs. Third, we did not study the relation with general mathematical ability, which possibly varied a lot between participants. This variable could be a predictor for the ability to interpret box plots correctly and the added value of refutational text and MERs in this learning process. However, it was not possible to take too much time of the classes for testing these students, so we had to make choices. By having a heterogeneous group of participants concerning mathematics background, we hope to have eliminated the possibility that the effects we found are only applicable to students with a certain mathematical background, making the results more transferable to other student populations.

The results of this study have several implications for theory. First, we have united two teaching techniques from different fields, namely the conceptual change literature and the instructional design literature. Both have fundamentally different points of departure. While refutational text specifically focuses on what students interpret incorrectly, MERs depart from a representation that is already interpreted correctly by students to teach another representation. This difference makes that combining them could make a strong combination of teaching techniques. Although we could not find that this combination was more successful than the use of refutational text alone, Lem et al. (2014) did find a trend in this direction, warranting further research. A possible explanation for not finding this effect in the current study is that the instruction time might have been too short to fully process the combination of refutational text and MERs. In the study of Lem et al. the processing time

was much longer, given that the intervention was presented as a home study text and participants had one week to process the text. Second, our results suggest that, in the interpretation of box plots, MERs were more successful in improving students' *problem-solving* than in improving their *learning*. More specifically, while participants who were exposed to refutational text had a clear advantage in solving items in which only box plots had to be interpreted, all participants scored equally well when box plots and histograms had to be matched to each other. This suggests that the constraining function of MERs (Ainsworth, 1999) had its effect for all conditions during the solving of the matching tasks, while this function was less successful in teaching participants about box plots.

The results of this study are also important for educational practice. We have shown that refutational text helps students in understanding box plots and overcoming their misinterpretation. As this teaching technique is relatively easy to implement and does not require much extra teaching time, it can be recommended for teachers to implement it for the teaching of box plots. Furthermore, although one could say based on these and earlier results that box plots are too complicated for students to learn, there are still reasons to use box plots. Bakker and Hoffmann (2005) argue that box plots provide a visual overview of a distribution, facilitate the comparison of multiple distributions using little space, and are easy to draw by hand.

Further research could focus on improving the interventions we used and testing their effect, both for the interpretation of box plots and for other topics in which misconceptions occur. Furthermore, it would be interesting to collect and study process data in order to get a deeper insight in the way students interpret box plots after having been exposed to refutational text and/or MERs. Finally, it could be studied whether refutational text, MERs, and their combination can benefit learning and problem-solving in other content fields.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Stephanie Lem holds a postdoctoral fellowship of the Research Foundation – Flanders (FWO). This research was partially supported by 'Number sense: Analysis and Improvement' of the KU Leuven [grant number GOA/12/010].

References

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education, 33*, 131–152. doi:10.1016/S0360-1315(99)00029-9

Ainsworth, S. (2008). The educational value of multiple-representations when learning complex scientific concepts. In J. K. Gilbert, M. Reiner, & M. Nakhleh (Eds.), *Visualization: Theory and practice in science education* (pp. 191–208). Springer Netherlands. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4020-5267-5_9

Baker, R. S., Corbett, A. T., & Koedinger, K. R. (2002). *The resilience of overgeneralization of knowledge about data representations*. New Orleans, LA. Retrieved from http://www-2.cs.cmu.edu/~rsbaker/BCKAERA2002.pdf

Bakker, A. (2004, May). *Design research in statistics education: On symbolizing and computer tools*. Retrieved from http://igitur-archive.library.uu.nl/dissertations/2004-0513-153943/UUindex.html

Bakker, A., & Hoffmann, M. H. G. (2005). Diagrammatic reasoning as the basis for developing concepts: A semiotic analysis of students' learning about statistical distribution. *Educational Studies in Mathematics, 60*, 333–358. doi:10.1007/s10649-005-5536-8

Bell, A., & Janvier, C. (1981). The interpretation of graphs representing situations. *For the Learning of Mathematics, 2*, 34–42.

Cobb, P., Gravemeijer, K. P. E., Bowers, J., & Doorman, M. (1997). *Statistical minitools* [Computer software]. Nashville & Utrecht: Vanderbult University, TN & Freudenthal Institute, Utrecht University. Retrieved from www.wisweb.nl

Corradi, D., Elen, J., & Clarebout, G. (2012). Understanding and enhancing the use of multiple external representations in chemistry education. *Journal of Science Education and Technology, 21*, 780–795. doi:10.1007/s10956-012-9366-z

delMas, R., Garfield, J., & Ooms, A. (2005). *Using assessment items to study students' difficulty with reading and interpreting graphical representations of distributions*. Presented at the Fourth Forum on Statistical Reasoning, Thinking, and Literacy (SRTL-4), July 6, 2005, Auckland, New Zealand.

Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education, 32*, 124–158. Retrieved from http://filebox.vt.edu/s/sboyce/CurriculumSummer03/StatArticles/JRME2001-03-124a.pdf

Garfield, J. B., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting research and teaching practice*. Berlin: Kluwer Academic.

Hynd, C. R. (2001). Refutational texts and the change process. *International Journal of Educational Research, 35*, 699–714. doi:10.1016/S0883-0355(02)00010-1

Lem, S., Kempen, G., Ceulemans, E., Onghena, P., Verschaffel, L., & Van Dooren, W. (2014). Combining multiple external representations and refutational text: An intervention on learning to interpret box plots. *International Journal of Science and Mathematics Education, 13*, 909–926. doi:10.1007/s10763-014-9604-3

Lem, S., Onghena, P., Verschaffel, L., & Van Dooren, W. (2013a). On the misinterpretation of histograms and box plots. *Educational Psychology, 33*, 155–174. doi:10.1080/01443410.2012.674006

Lem, S., Onghena, P., Verschaffel, L., & Van Dooren, W. (2013b). The heuristic interpretation of box plots. *Learning and Instruction, 26*, 22–35. doi:10.1016/j.learninstruc.2013.01.001

Lem, S., Onghena, P., Verschaffel, L., & Van Dooren, W. (2014). Experts' misinterpretation of box plots – A dual processing approach. *Psychologica Belgica, 54*, 395–405. doi:10.5334/pb.az

McKenzie, D. L., & Padilla, M. J. (1986). The construction and validation of the test of graphing in science (togs). *Journal of Research in Science Teaching, 23*, 571–579. doi:10.1002/tea.3660230702

Méheut, M. (2012). Preconceptions and learning. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 2662–2664). Boston, MA: Springer. doi:10.1007/978-1-4419-1428-6

Meletiou, M., & Lee, C. (2002). *Student understanding of histograms: A stumbling stone to the development of intuitions about variation*. Retrieved from http://www.stat.auckland.ac.nz/~iase/publications/1/10_19_me.pdf

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

Tippett, C. D. (2004). *Conceptual change: The power of refutation text* (Master's thesis). University of Victoria, Canada. Retrieved from https://dspace.library.uvic.ca/handle/1828/537?show=full

Tippett, C. D. (2010). Refutation text in science education: A review of two decades of research. *International Journal of Science and Mathematics Education, 8*, 951–970. doi:10.1007/s10763-010-9203-x

Tsui, C.-Y., & Treagust, D. F. (2003). Genetics reasoning with multiple external representations. *Research in Science Education, 33*, 111–135. Retrieved from http://link.springer.com/article/10.1023/A:1023685706290

Tukey, J. W. (1977). *Exploratory data analysis*. Reading: Addison-Wesley.

van den Broek, P., & Kendeou, P. (2008). Cognitive processes in comprehension of science texts: The role of co-activation in confronting misconceptions. *Applied Cognitive Psychology, 22*, 335–351. doi:10.1002/acp.1418

Van Dooren, W., De Bock, D., Hessels, A., Janssens, D., & Verschaffel, L. (2004). Remedying secondary school students' illusion of linearity: a teaching experiment aiming at conceptual change. *Learning and Instruction, 14*, 485–501. doi:10.1016/j.learninstruc.2004.06.019

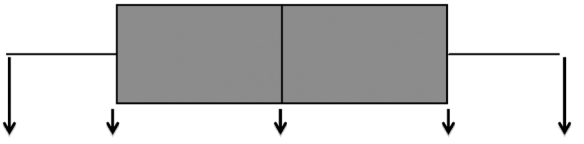
Wainer, H. (1992). Understanding graphs and tables. *Educational Researcher, 21*, 14–23. doi:10.3102/0013189X021001014

Waldrip, B., & Prain, V. (2012). Developing an understanding of ions in junior secondary school chemistry. *International Journal of Science and Mathematics Education, 10*, 1191–1213. doi:10.1007/s10763-011-9327-7

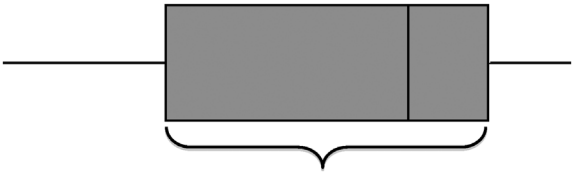
Wild, C. (2006). The concept of distribution. *Statistics Education Research Journal, 5*, 10–26. Retrieved from http://www.stat.auckland.ac.nz/~iase/serj/SERJ5(2).pdf#page=13

Appendix

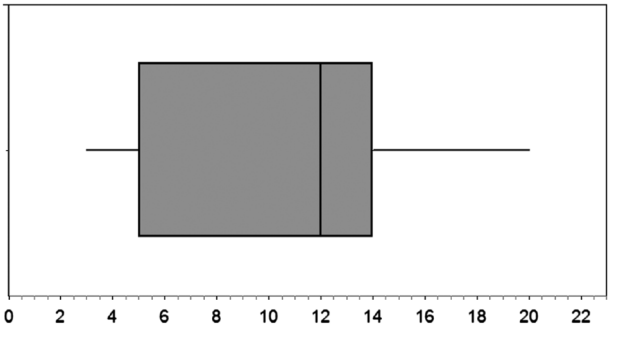
Pre-test (post-test was parallel to the pre-test and is hence not included in this appendix).
1. Name the different parts of the box plot that are indicated with an arrow.



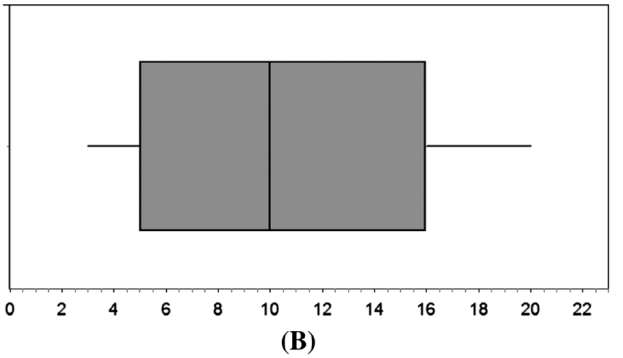
2. Define what percentile 25 means (= first quartile = Q1).
3. What is the maximum percentage of the data that is situated above percentile 75?
4. What percentage of the data is represented by the accolade?



5. Below you can see the flowering period of two types of flowers, type A and type B.



(A)

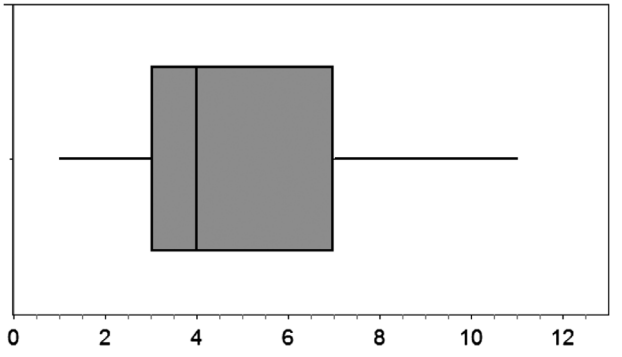


(B)

Which type of flowers would you choose in order to have the highest chance on a flowering period of more than 10 days?
A – B – You cannot tell from these box plots – I don't know

Explain how you found your answer

6. In this representation, the number of glass wearing children per class is presented.

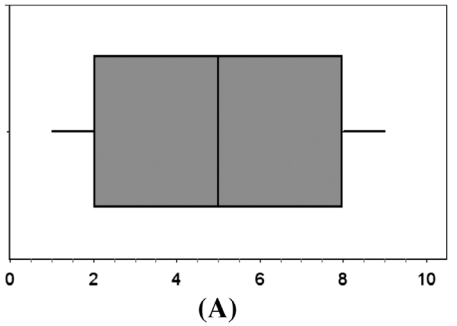


Which statement is correct?

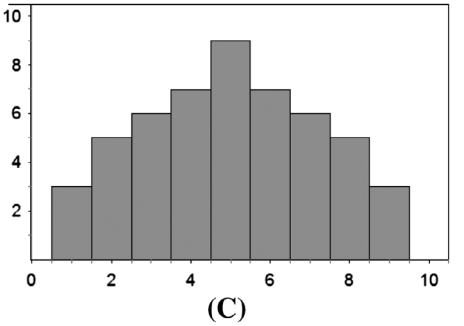
(a) There are more classes with more than four glass-wearing children than there are classes with less than four glass-wearing children.
(b) There are less classes with more than four glass-wearing children than there are classes with less than four glass-wearing children.
(c) There are as many classes with more than four glass-wearing children as there are classes with less than four glass-wearing children.
(d) You cannot be sure about any of the above statements.
(e) I don't know.

Explain how you found your answer

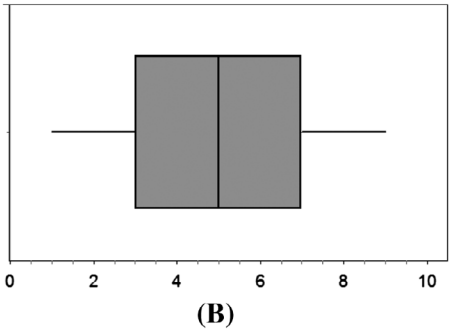
7. Link the box plots to the histograms that represent the same data-set



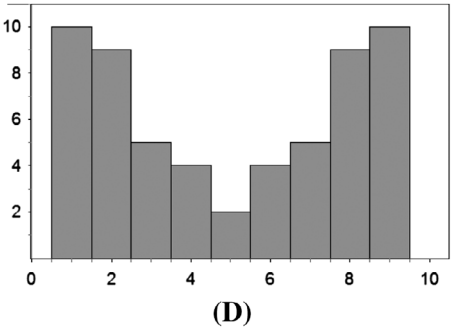
(A)



(C)



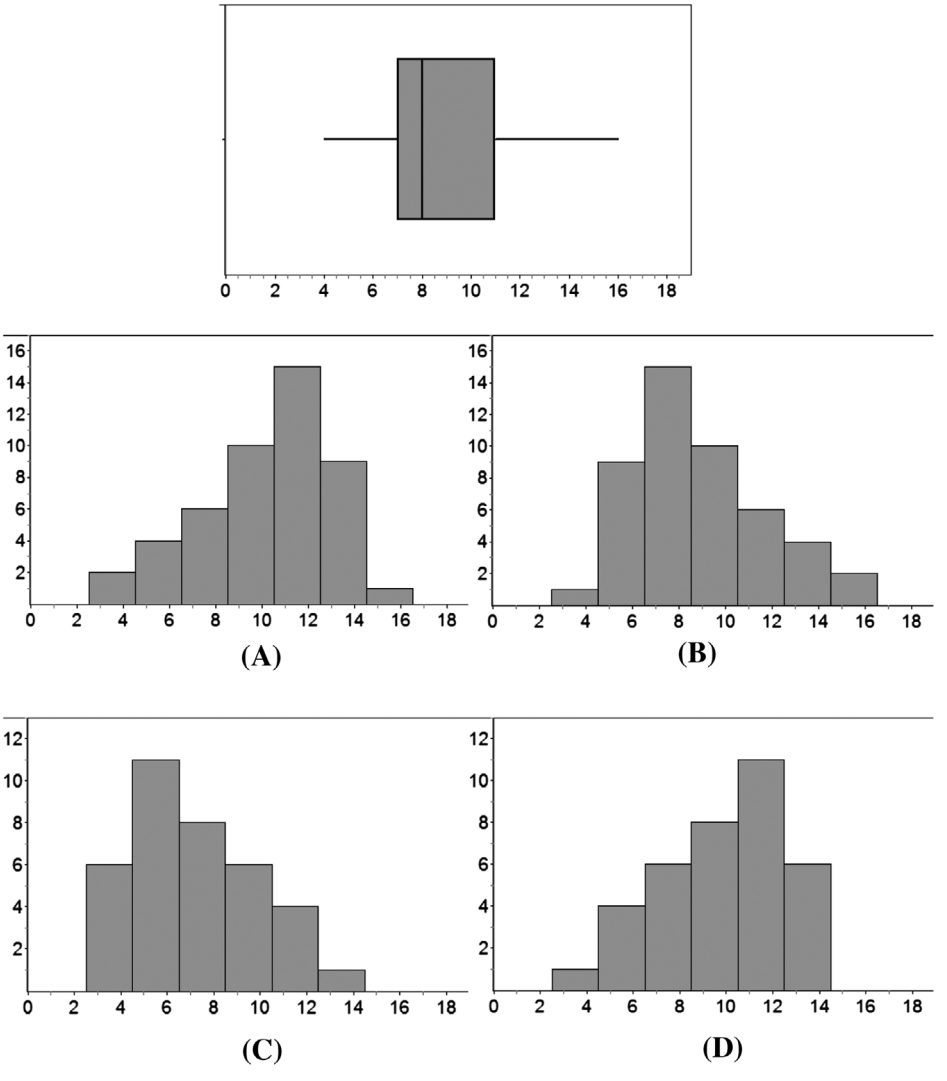
(B)



(D)

Box plot A represents the same data as histogram _____
Box plot B represents the same data as histogram _____
Explain how you found your answer

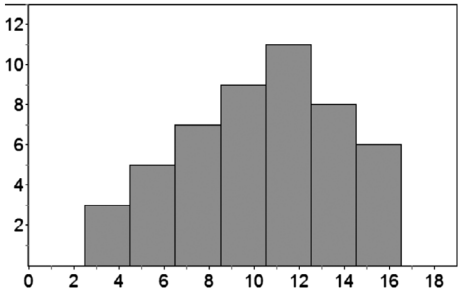
8. Link the box plot to the histogram that represents the same data

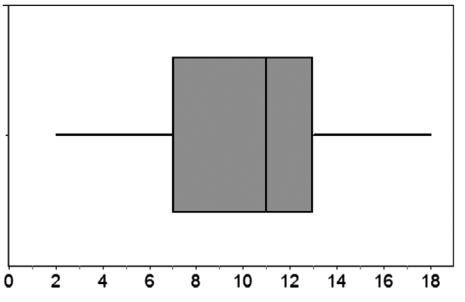


The box plot represents the same data set as histogram _____

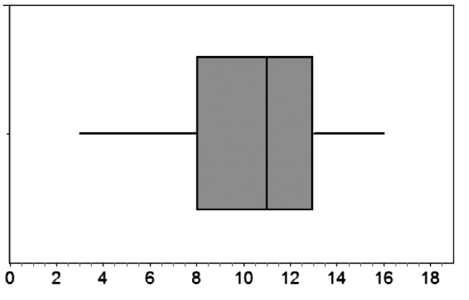
Explain how you found your answer

9. Link the box plot to the histogram that represents the same data

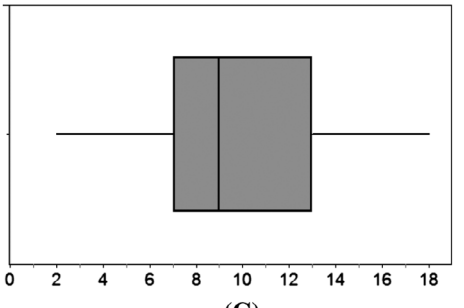




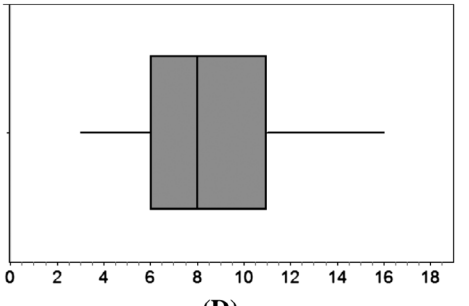
(A)



(B)



(C)

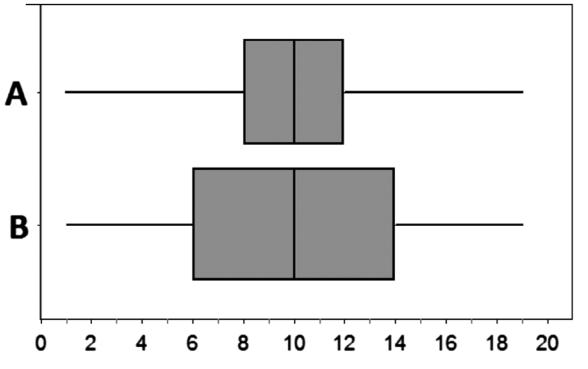


(D)

The box plot represents the same data set as histogram _____

Explain how you found your answer

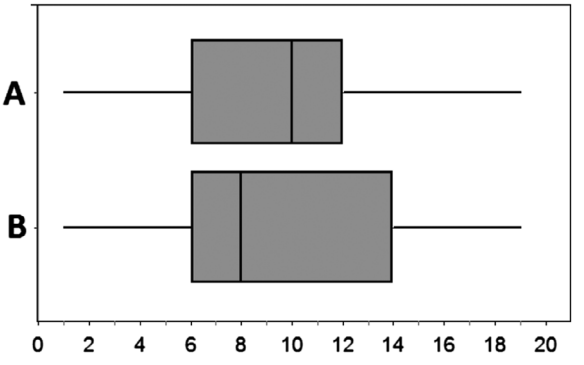
10. Below the exam results of two classes of students are represented. In which class were more students with a score of 10 or higher?



A - B - Equal

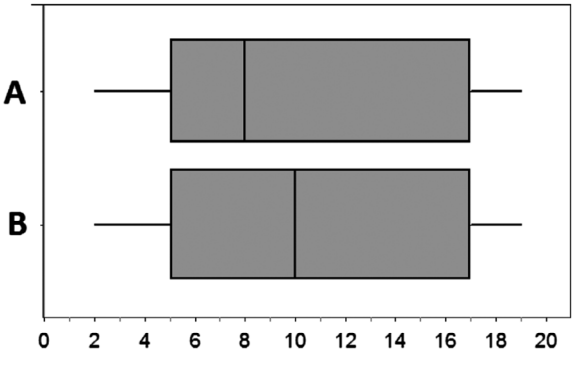
Explain how you found your answer

11. Below the exam results of two classes of students are represented. In which class were more students with a score of 10 or higher?



A - B - Equal

12. Below the exam results of two classes of students are represented. In which class were more students with a score of 10 or higher?



A - B - Equal

Explain how you found your answer