12. STUDENTS' MISCONCEPTIONS IN INTERPRETING THE MEAN OF THE DATA PRESENTED IN A BAR GRAPH.-halaman-1-16

by desirt76@gmail.com 1

Submission date: 07-Apr-2023 08:48AM (UTC+0300)
Submission ID: 2058202834
File name: THE_MEAN_OF_THE_DATA_PRESENTED_IN_A_BAR_GRAPH.-halaman-1-16.pdf (1.1M)
Word count: 6680
Character count: 32121

ISSN 2615-2495

International Journal of Insights for Mathematics Teaching
Volume 02, No. 1, April 2019, pp. 57-74

STUDENTS' MISCONCEPTIONS IN INTERPRETING THE MEAN OF THE DATA PRESENTED IN A BAR GRAPH

Desi Rahmatina[1], Norasykin Mohd Zaid[2]

[1]Raja Ali Haji Maritime University, Indonesia
[2]Universiti Teknologi Malaysia
desirahmatina@umrah.ac.id

Abstract

Interpreting the mean of the data presented in a bar graph constitutes a mix of two concepts highly useful for testing students' level of understanding of the way. This study aimed to describe students' misconceptions in the interpretation of the mean of data that are represented in a bar graph and the causes of such misconceptions and to examine whether misconceptions differed by gender and grade. The participants of this study consisted of 112 students (48 males, 64 females) of the Natural Science program of SMAN 1 Tanjungpinang in three grades-tenth, eleventh, twelfth. Employing a mixed method with an explanatory sequential design, this study collected and analyzed quantitative data before qualitative ones. The research identified 12 misconceptions about the mean and 8 causes of such misconceptions, and based on the chi-squared test results, neither gender- nor grade-based difference in students' misconceptions was found. These results have an implication for teachers and other educational stakeholders in considering the achievement of learning objectives and core competencies in the learning process, especially in the processing, reasoning, and presentation of the mean of data that are presented in a bar graph.

Keywords: misconception, mean, bar graph, interpretation

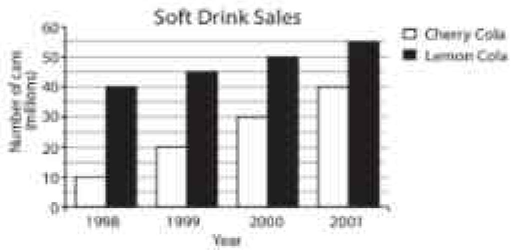
A misconception is an event in which one misinterprets a concept. A misconception is not an error, although both seem similar in terms of wrong results. While misconceptions may cause one error, others may stem from carelessness, problems reading or interpreting a chart or lack of understanding of data. As stated by Spooner (2002), "A misconception is the product of a lack of understanding or in many cases the misapplication of a 'rule' or mathematical generalization." Some researchers found students to have difficulties in learning and understanding statistics concepts (Jacqueline R. et al., 2013; Brett Berry, 2016; Gagnier J. et al., 2017). Jacqueline R. et al. (2013) found that students often make errors in answering questions related to the mean, median, and mode.

Statistics is a science of collecting, analyzing, presenting, and interpreting data as well as making a decision based on such analyses (Prem S. Mann, 2013). Statistics are commonly used in a wide range of fields, such as business (Bennett & Briggs, 2014), health (Gagnier J. J. & Morgenstern H., 2017), and education (Joan Garfiel et al., 2014; Jennifer Noll, 2012; Theodosia, 2016; Maria Meletious, 2015). In the educational curriculums applied in Indonesia, the materials on statistics are included in all educational levels, from elementary through higher education. Even the National Council of Teachers of Mathematics (2007) also incorporated materials on data analysis and chance into the mathematics curriculums of primary school, junior secondary school, and senior secondary school. A statistical content like the mean is a foundation in the learning of inferential statistics concepts, for example, the concepts of correlation test, regression test, ANOVA test, and MANOVA test. Without a correct understanding of the mean concept, it will be challenging to understand further statistical concepts as mathematics learning is hierarchical, with a topic learnable only if the prerequisite topics are well-understood. Psychologist Gagne (1997) states that no one will be able to learn a given topic if they fail to master previous topics that support that topic. For this reason, the material on data presentation in the form of a bar graph is introduced in the mathematics curriculum of grade 7, while the content on the measure of central tendency (mean, median, mode) is included in the mathematics curriculum of grade 8 (Pendidikan, Kebudayaan, & Indonesia, 2016), meaning that the students have been introduced to the concepts of data presentation in the form of graphs before they start on the mean. The National Council of Teachers of Mathematics (NCTM) (2007) recommended that students should develop understanding on the mean, median, and mode before pursuing studies at senior secondary level, and they can find, use, and interpret the measure of center and measure of variability, including the mean, at grades six through eight (p. 401).

Making the connection of the concept of data interpretation in the form of bar graphs and the mean will be of great use in testing students' level of understanding of the mean concept and graph reading ability. Computing the mean of ungrouped data is common, but despite that, many students and teachers still stumble upon difficulties in explaining which value representative is of the measure of center (Jacobbe, 2012). From interviews, it was found out that teachers defined mean by way of summing all values and divided by the number of data. They admitted that they only knew the definition of the mean and found it easier to deliver the definition through examples. As an instance, for the data 1, 2, 3, 4, and 5, the mean is 1 + 2 + 3 + 4 + 5 divided by 5. Two of the three teachers interviewed did not even understand what uses the mean can bring, and in what situation it can be used to represent the measure of center.

The mean may be interpreted when it is presented in a bar graph, but it will take reasoning skills. Cognitive reasoning of a graph is a common way to demonstrate students' thought upon the information hidden in a graph (Wang et al., 2012). Data distribution curve shape showing whether the data are normally distributed or not can be based on the graph. Designating the mean position in a graph will be considerably dependent on the graph interpreting skill. Mhlolo M. (2015) investigated students' meta-representation competence when they were constructing bar graphs.

Meanwhile, Shah P. & Freedman E. G. (2011) studied the top-down and bottom-up processes taking place in bar and line chart construction. Interpretation of the data presented in bar graphs even became an item of assessment of eight graders in the 2011 Trends in International Mathematics and Science Study (TIMSS) (Figure 1) (IAE, 2013). Based on the 2011 TIMSS assessment, only 40% of the eighth graders were able to give correct answers (see Figure 1). In Indonesia, the results of the junior and senior secondary school national exams in 2018 show that only 45.71% of junior secondary school students nationwide were able to answer questions on statistics and probability correctly, 62.51% of whom answered correctly the questions on data presented in the form of bar graphs, and out of all senior secondary school students of natural and social science programs, only 37.49% and 31.66% were able to give the correct answers to questions on statistics and chance, respectively (Puspendik, 2019). This shows that students' ability to solve statistical problems still fell into a low category, with less than 50% giving the correct answers.



The graph shows the sales of two types of soft drink over 4 years. If the sales trends continue for the next 10 years, determine the year in which the sales of Cherry Cola will be the same as the sales of Lemon Cola.

A. 2003

B. 2004

C. 2005

D. 2006

Figure 1. Interpretation of data in bar graph [International Association for the Evaluation of Educational Achievement (IEA), 2013]

Sharma S. V. (2006) presented and discussed how students gained an understanding of graph representation (table and bar graph). It was found that many of the students used experience-based strategies and intuitive ones. As in the case of grade school students, misconceptions in graph interpretation also occurred in students of higher education institutions. Lem S., Onghena P., Verschaffel L., and Van Dooren W. (2013) carried out a research study of 125 first-year students in Leuven, Belgia, related to data presented in the form of histogram and box plots and found that many students had false interpretation. Kaplan, Gabrosek, Curtiss, and Malone (2014) investigated students' understanding of the histogram and identified four misconceptions, namely those of the difference between bar graph and histogram, the difference between horizontal and vertical axes, histogram shape in relation to variability, and time component along the x-axis. Another researcher, Aoyama K. (2007), investigated the hierarchy of students' graph interpretation and identified several hurdles rendered by students' learning experience leading to them thinking narrowly over open-ended questions.

The understanding of the mean of data in graphs was once examined by Cooper L. and Shore F., (2008), who identified some misconceptions in the interpretation of the means of data in histogram and stem-and-leaf plot. From students' answers to the test questions and interview results, they identified that students had difficulties predicting the mean of the data represented when there was a skewness in the histogram. Susac, A et al. (2017) state that data representation in the form of graphs can improve students' understanding of measures, helps with data processing and data visualization, and reduce students' cognitive burdens when conducting data analysis and measurement. This makes it essential to encourage graph use by students. Additionally, box plots will also help students think about numeric values outside the box, as stated by Thomas G. Edwards et al. (2017).

Data representation in the form of graphs is not uncommon in print or electronic media. A deeper understanding of graph used in interpreting values is highly necessary as data representation in the form of graphs is deemed more effective and efficient. A further study of data representation in the form of graphs for analyzing various statistical concepts is, thus, needed. In light of that, this research aimed to 1) study in a more in-depth manner the misconceptions tenth-, eleventh-, and twelfth-graders might have in interpreting the mean of data represented in a bar graph, 2) identify the causes of such misconceptions, 3) test whether there was a significant difference in students' misconceptions based on gender, and 4) test whether there was a difference in students' misconceptions based on grade. A similar study on the misconceptions in the interpretation of the mean of data that are presented in graphs was once conducted by Cooper L. and Shore F. (2008), but it examined how students reasoned the mean when the data were presented in histogram and did not examine the cause of misconceptions nor the difference in misconceptions by gender and grade. The present study, however, examined students' misconceptions in interpreting the data presented in a bar graph and the causes of such misconceptions. It also sought to figure out whether the difference in misconceptions by gender and grade existed. Based on the research objectives above, the questions posed in this study are as follows: 1) What are the misconceptions tenth-, eleventh-, and twelfth-graders had in interpreting the mean of data that are presented in a bar graph? 2)What cause students' misconceptions in understanding the mean of data that are offered in a bar graph? 3) Are there any gender-based differences in students' misconceptions in interpreting the mean of data that are presented in a bar graph? 4) Are there any grade-based differences in students' misconceptions in interpreting the mean of data that are presented in a bar graph?

METHOD

The participants in this study consisted of 112 (48 males, 64 females) of 114 students of Natural Sciences program in three grades at SMAN 1 Tanjungpinang (tenth grade, eleventh grade, twelfth grade) in the second semester of the academic year 2018/2019. Two students from the twelfth grade did not participate in this research. The students were 15 to 18 years of age. The data of this research's respondents are presented in Table 1. The researchers examined the students' gender- and grade-wise backgrounds to investigate in a more detailed fashion, the causes of the students' misconceptions in interpreting the mean of data that are presented in a bar graph. The research method used was mixed-method. Mixed-method is a procedure for collecting, analyzing, and mixing quantitative and qualitative methods in a study or a set of studies to understand a research problem (Creswell J. W., 2012). Sequential mixed method was employed in this study. In the first stage, qualitative data were collected and analyzed to answer the first and second research questions, while in the second stage, students' misconceptions data were collected and analyzed based on the data collected and analyzed in the previous step quantitatively to answer the third and fourth questions.

The research design used was the explanatory sequential design. According to Cresswell J. W. (2012), in explanatory sequential design, a researcher collects and analyzes quantitative data before moving to qualitative data. In this research, the emphasis was placed more on qualitative data. The data were collected through a written test with one question regarding the mean of data that are presented in a bar graph. The tests were given classically to students in a 60-minute class at each grade, and the researchers conducted a direct observation when the students were working on the question to see how they answered the question. Interviews were carried out after the students finished the question.

The question used was modified from the question developed by Cooper & Shore (2010) and Shiau & Ismail (2014). Semi-structured interviews were conducted with those who answered correctly and those who did not. This was aimed to verify the students' answers and figure out what caused their misconceptions in interpreting the mean of data that are presented in a bar graph. Also, the activities taking place in the first stage were documented in the form of photographs to allow the researchers to observe how students' misconceptions developed. Afterward, the data were analyzed through tabulation to encode students' misconception by sorting the students with the correct answer and those with the wrong ones by gender and grade. Two students from each grade (one who gave correct answer and one who gave false answer) were selected randomly to be interviewed this was aimed that The researcher could generate in-depth information about students' misconceptions about the arithmetic "mean" in the form of a bar chart, through this activity the researcher was able to find out how the students could simply give the wrong answers and how it happened, this is called the 'real wrong' thinking process (Subanji, 2011), but interviews were also done to students who could determine the correct answer yet the students gave the wrong explanations, according to Subanji (2011) these students experience "pseudo right" thinking. The interviews were recorded to avoid missing some of the information conveyed by the students in relation to their misconceptions of the mean of bar graph data. The students interviewed consisted of Student 1, Student 2, Student 3, Student 4, Student 5, and Student 6.

As for the quantitative analysis, the data were analyzed using a statistical test tool to test whether there were significant differences in students' misconceptions about the mean of data that are presented in a bar graph by gender and grade. Non-Parametric statistical analysis was undertaken using a chi-squared test for hypothesis testing to declare whether there were significant differences in students' misconceptions about the mean of data that are presented in a bar graph by gender and grade with the aid of the program SPSS v.22. If the significance level (p) was $< \alpha$, it could be concluded that there were significant gender- and grade-based differences in students' misconceptions about the mean. By contrast, if the significance level was greater or equal to α, it can be concluded that there were no such significant differences.

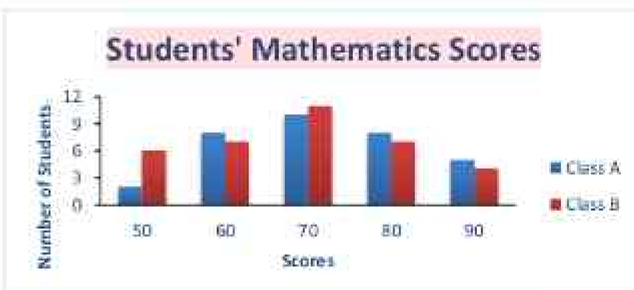
Table 1. Number of respondents

Grade	Gender			Interview
	Male	Female	Total	
X	17	25	42	2
XI	15	21	36	2
XII	16	18	34	2
Total	48	64	112	6

RESULTS AND DISCUSSION

The question as shown in Figure 1 was given to the students to find out about their misconceptions on the mean, which, in this case, was the mean of mathematics scores in two classes, namely Class A and Class B, that were presented in bar graphs. From the question, students were told to determine which class had the largest mean and what kind of reasons were given by the students for the answers.

The bar graph below shows the mathematics scores of two classes (Class A and Class B).



Based on the data of mathematics scores presented in the bar graph above,

Which class gained the highest mean?

Answers:

Explanation:

Figure 1. Question on the mean of data presented in a bar graph

Table 2. Students' answers by gender and grade

Grade	Male			Female			Total
	$\bar{x}_A > \bar{x}_B$	$\bar{x}_A < \bar{x}_B$	Did not know	$\bar{x}_A > \bar{x}_B$	$\bar{x}_A < \bar{x}_B$	Did not know	
X	14(33.3%)	3(7.2%)	0(0%)	25(59.5%)	0(0%)	0(0%)	42(100%)
XI	10(27.8%)	5(13.9%)	0(0%)	16(44.4%)	5(13.9%)	0(0%)	36(100%)
XII	11(32.3%)	4(11.8%)	1(2.9%)	16(47.1%)	2(5.9%)	0(0%)	34(100%)

Total	35(31.3%)	12(10.7%)	1(0.9%)	57(50.9%)	7(6.3%)	0(0%)	112(100%)

Based on the answers outlined in Table 2, 50.9% of female and 31.3% of male students were able to answer correctly (Class A's mean ($\bar{x}_A$) is greater than Class B's ($\bar{x}_B$)). Grade Ten had a higher percentage of the correct answer (39%) than Grade Eleven (26%) and Grade Twelve (27%). Table 3 presents the misconceptions of students with the correct answer and the causes of such misconceptions, while Table 4 students with wrong answers and the causes of such misconceptions. The researchers interviewed both students with the correct answer and those with the wrong ones to find out more about the causes of misconceptions on the mean of data that are presented in bar graphs. The details of students' misconceptions are presented in Table 3 for students who gave the right answer, and Table 4 for those who gave the wrong answers. Tables 3 and 4 show that most misconceptions occurred in eleventh graders (16 students, 14.29%), followed by twelfth (12 students, 10.71%) and tenth graders (9 students, 8.04%). It can be seen that there was hardly any gender-based students' misconception difference, with 18 female students (16.07%) and 19 male students (16.96%) having misconceptions.

This research has found 12 misconceptions on the mean of data that are presented in bar graphs (M1–M12). The percentages of misconceptions M2, M3, M5, M6, M8, and M9 were 8.04%, 1.79%, 0.89%, 6.25%, 0.89%, and 0.89%, respectively. This finding is in line with that of Ismail & Wei (2015). Besides the six misconceptions above, there were six other misconceptions found in this research, namely M1, M4, M7, M10, M11, and M12 at percentages of 4.46%, 0.89%, 2.68%, 1.79%, 3.57%, and 0.89%, respectively. This study also identified 8 causes of students' misconceptions on the mean of data that are presented in bar graphs: 1) misinterpretation of the concept of grouped data (Spooner, 2002); 2) lack of understanding of a mean value's meaning as a value that represents a set of data; 3) lack of understanding of mean value's position in a bar graph; 4) lack of familiarity of bar used for determining the mean; 5) error in mathematical computation when using the mean formula; 6) carelessness in selecting bigger and smaller numbers; 7) error in determining the number of data on the vertical axis; and 8) inability to distinguish between the use of the values on the horizontal axis and that on the vertical axis (Kaplan et al., 2014).

Table 3. Students' explanation for their answers by grade and gender

(Male: M, Female: F) correct answer ($\bar{x}_A > \bar{x}_B$)

No	Misconception	Cause	Number of Students: Grade X M	Grade X F	Grade XI M	Grade XI F	Grade XII M	Grade XII F	Total
M1	Computing the mean following the formula for the mean, but the position of the result was far from the mean value of the data in the bar graph. 1. Class A's mean = 90, Class B's mean = 68.8 2. Class A's mean = 49.0, Class B's mean = 46.2 3. Class A's mean = 72.4, Class B's mean = 68.8 4. Class A's mean = 69.3, Class B's mean = 63.6	The students were too fixated to the formula for the mean and did not recheck whether the result obtained was at the right position in the bar graph.	4	1					5 (4.46%)
M2.	The mean was determined based on the value on the horizontal axis divided by the frequency (the bar's height) on the vertical axis. The shorter the bar, the greater the mean. The number of students in Class A (n = 33) was smaller than that in Class B (n = 35). Thus, the division of the same total scores by a smaller number will yield a greater result. Class A's mean $\frac{50+60+70+80+90}{2+8+10+8+5} = \frac{350}{33} = 10.6$ Class B's mean $\frac{50+60+70+80+90}{6+7+11+7+4} = \frac{350}{35} = 10$	The misinterpretation of the concepts of the mean of grouped data and division occurred when students believed that the total scores gained by Class A and Class B were the same (350), so they concluded that Class A's mean was bigger than Class B's because the denominator for Class A was smaller than that for Class B.		1		5		3	9 (8.04%)
M3.	Determining the mean based on the frequency (bar's height) of each class divided by the number of categories on the horizontal axis. Class A's mean = $\frac{2+8+10+8+5}{5} = 6.6$ Class B's mean = $\frac{6+7+11+7+4}{5} = 7$	The students were unable to distinguish between the concept of the mean for ungrouped data and that for grouped data. They also erroneously determined which number was smaller and which was bigger, mistaking 6.6 as bigger than 7.					2		2 (1.79%)

M4.	Determining the number of students on the vertical axis (Class A = 33, Class B = 31), leading to erroneous computation.	Error in determining the number of students of each class.						1	1 (0.89%)
M5.	Determining the mean using the mean formula for grouped data, but the number of students did not correspond with the height of the bar, causing the mean obtained to be far from the mean value. Class A's mean $\frac{50\,(2) + 60\,(8) + 70\,(10) + 80\,(7) + 90\,(}{2 + 8 + 10 + 7 + 4}$ $= \frac{100 + 480 + 560 + 560 + 450}{31}$ $= 70.9$ Class B's mean $\frac{50\,(6) + 60\,(7) + 70\,(11) + 80\,(6) + 90\,(}{6 + 7 + 11 + 6 + 3}$ $= \frac{300 + 420 + 770 + 560 + 360}{33}$ $= 58.7$	The students did not know the position of the mean in the bar graph.			1				1 (0.89%)
Total			4	2	1	5	2	4	18 (16.07%)

Table 4. Students' explanation for their answers based on grade and gender

(Male: M, Female: F) wrong answer ($\bar{x}_A < \bar{x}_B$)

No	Misconception	Cause	Number of Students						Total
			Grade X		Grade XI		Grade XII		
			M	F	M	F	M	F	
M6	Determining the mean by summing the frequencies (the bar's height) of each class then dividing by the number of categories on the horizontal axis. Class A's mean $\frac{2 + 8 + 10 + 8 + 5}{5} = \frac{33}{5} = 6.6$ Class B's mean $\frac{6 + 7 + 11 + 7 + 4}{5} = \frac{35}{5} = 7$	The students were unable to distinguish between the use of the value on the horizontal axis and that on the vertical axis, and they did not know that the data presented in the bar graph were grouped data, so they computed the mean like they would do for ungrouped data by summing the scores in Class A and those	1				4	2	7 (6.25 %)

		in Class B on the vertical axis then dividing by the number of data.							
M7	The mean was positioned too far away from the mean value on the horizontal axis of the bar graph. 1. Class A's mean = 47.4, Class B's mean = 48.2 2. Class B's mean = 94 3. Class A's mean = 475, Class B's mean = 495	The students estimated the mean based on the bar graph and did not know that the mean was also the mean value of the set of data.	1		2				3 (2.68 %)
M8.	Determining the mean by summing the values on the horizontal axis and dividing by the number of data on such horizontal axis. The mean was obtained from the class with the tallest bar for the mean obtained. $\bar{x} = \frac{50+60+70+80+90}{5}$ $= \frac{350}{70} = 70$ The tallest bar was found at score 70 in Class B. They concluded that Class B had the highest mean.	The student only computed the mean of ungrouped data by summing the scores and dividing by the number of data, and they did not know that the data presented in the bar graph were ungrouped.	1						1 (0.89 %)
M9.	Computing the mean by summing the multiplication of the number of students and their respective scores then dividing by the number of students in each class, causing the mean of Class A to be positioned on the left side of the mean value. Class A's mean $= \frac{50(2)+60(8)+70(10)+80(8)+}{2+8+10+8+5}$ $= \frac{100+480+560+640+450}{33}$ $= 67.5$ Class B's mean $\frac{50(6)+60(7)+70(11)+80(7)+9}{6+7+11+7+4}$	The student made an error in the multiplication and did not recheck the correctness of their answer, mathematically or based on the mean's position in the bar graph.			1				1 (0.89 %)

	$= \frac{300 + 420 + 770 + 560 + 360}{35}$ $= 68.8$								
M10.	Seeking the mean from the bar's height on the vertical axis and finding that Class B's mean was greater than Class A's. Class A = 2 + 8 + 10 + 8 + 5 = 33 Class B = 6 + 7 + 11 + 7 + 4 = 35	The students did not know the concept of the mean.			1	1			2 (1.79)
M11.	Computing the mean based on the height of the bars in each class and dividing by the highest score on the vertical axis. Mean $= \frac{Number\ of\ data}{number\ of\ students}$ Class A's mean $\frac{2+8+10+8+5}{12} = \frac{33}{12}$ Class B's mean $\frac{6+7+11+7+4}{12} = \frac{35}{12}$	The students did not know the number of data on the vertical axis.				4			4 (3.57 %)
M12.	Determining the mean only based on estimation, so Class B's mean was found to be higher than Class A's.	The student was not used to compute the mean based on the data that were presented in a bar graph.			1				1 (0.89 %)
Total			3	0	5	5	4	2	19 (16.96%)

Table 3 and Table 4 show that 33.05% of the students (16.07% answering correctly, 16.96% wrongly) had misconceptions in interpreting the mean of data that are presented in a bar graph. Although Table 2 clearly shows that the percentage of female students who answered correctly (50.9%) was higher than that of their male counterparts (32.1%), the chi-squared value obtained based on Tables 3 and 4 ($\chi^2 = 2.179$, $df = 1$, $p = 0.140$) shows that there was no significant difference in students' misconceptions about the mean of data that are presented in bar graph between male and female students. This is consistent with the results of the research by Louis & Mistele (2012), who used TIMSS

2007 to test whether there was a significant difference between male and female students in terms of mathematical scores achieved and found that there was not any. Differently, Preckel et al. (2008) conducted a research study of 181 gifted students and 181 non-gifted students and found that male students gained test scores significantly higher than their female counterparts but no difference in mathematics grade between male and female students. Not only students, a research study once studied the gender-based difference in interpreting graphs among teachers (Patahuddin & Lowrie, 2018). The study found that there was no difference in students' understanding of graph interpretation based on gender.

Students' misconceptions about the mean of data that are presented in bar graphs differed by not only gender but also grade, but although Table 2 shows that the tenth grade had a higher percentage of students with correct answer (39%) than that of the eleventh grade (26%) and the twelfth (27%), the chi-squared value that was obtained based on Tables 3 and 4 show otherwise ($\chi^2 = 1.974$, $df = 2$, $p = 0.373$). This indicates that no significant grade-based difference in students' misconceptions was found. In other words, the misconceptions about the mean of data that are presented in a bar graph did not differ between tenth, eleventh, and twelfth graders. A research study of how to grade difference affected students' achievement in mathematics was once conducted by Garcia-mila, Marti, & Gilabert (2014). They compared the difficulties found by fifth and sixth graders (elementary school students) and those found by seventh and eighth graders (secondary school students) in developing a bar graph from raw data. Their research shows that there was a significant difference in the frequency at which the difficulties in making the bar graph was found between the two student groups (elementary and secondary school students).

To confirm students' answers, the researchers interviewed six students, two for each grade (Students 1 and 2 from grade ten, Students 3 and 4 from grade eleven, Students 5 and 6 from grade twelve), with one of the two giving correct answer (Students 2, 3, 5) and the other giving wrong answers (Students 1, 4, 6).

Researcher: What is meant by mean?

Student 1: *It's like… the sum of the scores we have divided by the number of all students.*

Researcher: How do you compute the mean?

Student 1: *Say we have scores of 70, 80, 90, 75, 80*, the mean would be *(70 + 80 + 90 + 75 + 80) / 5 = 79.*

Researcher: What is the significance of the value 79?

Student 1: *Well, it's the mean.*

Researcher: What is meant by mean?

Student 2: *Mean is the total scores divided by the number of students.*

Researcher: How do you compute the mean?

Student 2: *Let's see... hmm... oh right... suppose we are counting income. On Monday, we earn 300k, Tuesday 500k, Wednesday 150k, Thursday 100k, Friday 200k, and Saturday 400k. The mean would be 300k + 500k + 150k + 100k + 200k + 400k divided by 6 since there are six days so that it would be about 300k.*

Researcher: What is the significance of the value 300k?

Student 2: *That is the weekly income.*

Researcher: What is meant by mean?

Student 3: *The average value obtained from a set of values or the sum of the scores and to which we apply the formula for the mean.*

Researcher: How do you compute the mean?

Student 3: *For example, we have 7, 7, 7, 8, 8. So, we add all of them then divide by the number of the members, here we have 5, hmmm (the student was calculating), we get 7.4.*

Researcher: What is the significance of the value 7.4?

Student 3: *That is the mean.*

Researcher: What is meant by mean?

Student 4: *Mean is, like, there are scores from 50 to 90, then it is the score the students get on average, for example, 70. The mean then is 70.*

Researcher: How do you compute the mean?

Student 4: *Ah...(thinking...) Let's say 1 from 10... (silent).*

Researcher: What is the mean?

Student 4: *(silent)*

Researcher: How do you compute it?

Student 4: *Ehmm... I'm afraid I don't know.*

Researcher: Suppose we have 1, 2, 3, 4, 5. What is the mean?

Student 4: *It's 5.*

Researcher: Why so?

Student 4: *Because eh... (unable to answer).*

Researcher: Do you know the formula for the mean?

Student 4: *(shaking head) I don't know.*

Researcher: Did you learn the mean before?

Student 4: *I did, but I'm afraid I have forgotten it.*

Researcher: What is meant by mean?

Student 5: *Mean is the total scores divided by the frequency.*

Researcher: How do you compute the mean?

Student 5: *Hmmm, suppose we have 6, 7, 8, 8, 8, 9, 10. All of the scores are added then divided by the frequency, which is 7. The result is 8.*

Researcher: What is the significance of the value you got?

Student 5: *The total scores like... the scores that appear the most.*

Researcher: Have you ever used a bar graph to determine the mean of a set of data?

Student 5: *Yes, I have. I've encountered such questions when I was a junior secondary school student.*

Researcher: What is meant by mean?

Student 6: *It is the sum of the scores divided by the number of data.*

Researcher: How do you compute the mean?

Student 6: *The data were multiplied by the score then divided by the number of data.*

Researcher: Could you show me how you compute the mean?

Student 6: *Like the mathematics scores of a class, 2 students get a score of 8, 5 get 9, 6 get 7. The scores are summed then multiplied by the frequency.*

Researcher: Are you familiar with computing the mean using a bar graph?

Student 6: *Yes, I am.*

Researcher: Where did you encounter a question on computing the mean using a bar graph?

Student 6: *In the classroom when I was in grade 10, 11, and, currently, 12*

From the interview, it was found that all students (Students 1 through 6) were not familiar with the meaning of the mean. Even Student 4 was unable to explain how to compute the mean of a set of data nor articulate the formula for the mean because of being unable to recall it. Meanwhile, Student 5 mistaken the mean for modus, which is the value that appears most often. The students answered the question regarding the meaning of the mean using the formula for the mean (Jacobbe, 2012) and did not state the significance of the value obtained nor understand what is meant by mean and its function as a number that represents a set of data. In line with this finding, Jacobbe (2012) found that two out of three teachers had difficulty explaining the meaning of mean. They possessed the procedural knowledge, but they lacked the conceptual knowledge regarding the mean. This leads to their misconceptions in interpreting the mean of the data presented in a bar graph. Although they were able to calculate the mean, they were unable to determine where the mean was positioned in the bar graph without using the formula for the mean. It is in line with the research conducted by Cooper, L and Shore, F (2008) revealing that students experience difficulties when determining the position of the arithmetic mean on a histogram which shape is inclined to the right. This shows that students still find difficulties in determining the arithmetic mean when it is presented in graphical form. Because the representation of the data in the graph can visualize the data, students must be encouraged to leverage the graph (Ana

Susac et al. 2017), with this encouragement, students can be proficient in interpreting the graph (Wang et al., 2012), thus student's misconception in determining the arithmetic mean through diagrams or graphs can be minimized. Not only misconceptions in determining the arithmetic mean in the form of bar charts and histograms, Patahuddin & Lowrie (2018) revealed that teachers also experience difficulties in interpreting line graphs when "reading outside the data" so that in this phase, teachers need to have reliable and robust knowledge in understanding graphics.

The students worked on the question in a variety of ways. Some of the students solved the problem using the formula for the mean. Some others used their verbal ability as to how many students there were in each class was not stated for sure. They used their verbal ability to read the bar graph in the form of narration. Figures 2, 3, and 4 present students' answers. As shown in Figure 2, student 2 used the concept of the mean of grouped data and used the formula for the mean for determining which class had the highest mean. The student found the mean of Class A ($\bar{x}_A = 71.8$) higher than that of Class B ($\bar{x}_B = 68.85$). In Figure 3, student 3 used his verbal ability instead of the formula and found that the mean of Class A was higher than that of Class B on the basis that Class A had more students with scores 80 and 90 than Class B did. Meanwhile, in Figure 4, student 6 used the concept of ungrouped data in solving the question and found Class B to have higher mean than Class A. Even though in the interview the student claimed that he had been familiar with the calculation of the mean of data in a bar graph since grades ten through twelve, his written answer showed that he could not distinguish between ungrouped and grouped data. Ismail & Wei also found this misconception, (2015) where there were 10.68% of students from 412 students at grade ten in Malaysia who used a single average concept in solving group averaged questions presented in histograms by summing the height of the histogram and dividing it with the highest value on the vertical axis.

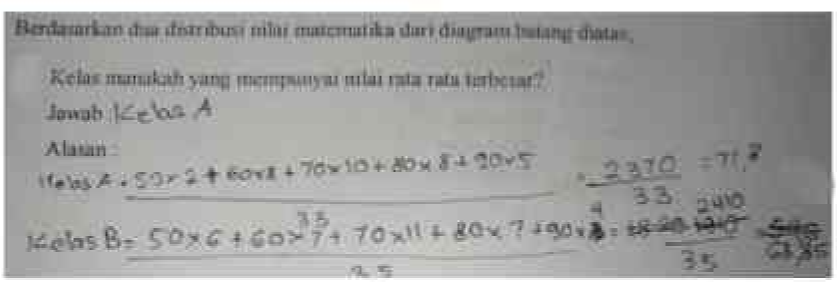


Figure 2. Student 2's answer

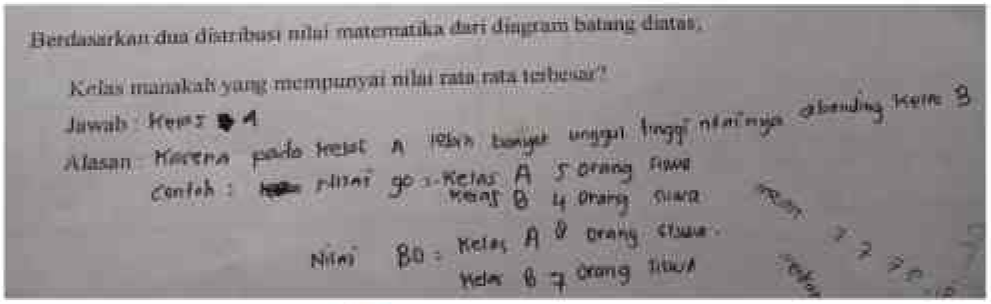


Figure 3 . Student 3's answer

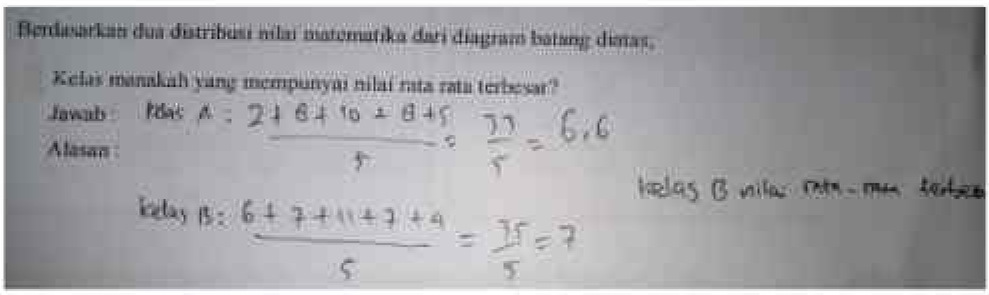


Figure 4. Student 6's answer

CONCLUSION

This research unveiled students' misconceptions about the mean of the data presented in a bar graph and the cause of such misconceptions, and to examine whether misconceptions differed by gender and grade. As many as 12 misconceptions (M1-M12), six of the misconceptions were in agreement with those found in a previous study, namely M2, M3, M5, M6, M8 and M9 and six other were freshly found in this research, namely M1, M4, M7, M10, M11, and M12 misconceptions. And 8 causes of such misconceptions were identified in this research, namely 1) misinterpretation of the concept of grouped data; 2) lack of understanding of a mean value's meaning as a value that represents a set of data; 3) lack of knowledge of mean value's position in a bar graph; 4) lack of familiarity of bar used for determining the mean; 5) error in mathematical computation when using the mean formula; 6) carelessness in selecting bigger and smaller numbers; 7) error in determining the amount of data on the vertical axis; and 8) inability to distinguish between the use of the values on the horizontal axis and that on the vertical axis.

It was also found that there was no significant gender-based difference in the students' misconceptions. Although academically, twelfth graders had more mathematics learning experiences, especially for statistical materials, and they had been preparing for the national exam, it was proven that

12. STUDENTS' MISCONCEPTIONS IN INTERPRETING THE MEAN OF THE DATA PRESENTED IN A BAR GRAPH.-halaman-1-16

ORIGINALITY REPORT

97%	97%	1%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	journal2.um.ac.id Internet Source	97%

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		