IOP Conf. Series: Materials Science and Engineering **296** (2018) 012037 doi:10.1088/1757-899X/296/1/012037

Pre-service mathematics teachers' statistical reasoning about mean

Y D Kristanto

Universitas Sanata Dharma, Yogyakarta

yosepdwikristanto@usd.ac.id

Abstract. This article offers a descriptive qualitative analysis of 3 second-year pre-service teachers' statistical reasoning about the mean. Twenty-six pre-service teachers were tested using an open-ended problem where they were expected to analyze a method in finding the mean of a data. Three of their test results are selected to be analyzed. The results suggest that the pre-service teachers did not use context to develop the interpretation of mean. Therefore, this article also offers strategies to promote statistical reasoning about mean that use various contexts.

1. Introduction

What do the pre-service teachers mean? It is worth to submit this question to find out the pre-service teachers' understanding about the concept of the mean. The answer to that question will also help the author to know about the context of the students, so we can provide appropriate treatment in the next statistical learning.

Reasoning the students towards the concept of the mean is very important [1]. This is because they will teach the concept at the junior or high school level. The following concept is the basic competency of junior high school which contains the mean concept. Analyze data based on data distribution, mean, median, mode, and data distribution to draw conclusions, make decisions, and make predictions [2]. The following concepts are the basic competencies of the senior high school that contain the mean concepts. Determine and analyze the centralized and distributed size of the data presented in the form of frequency distribution tables and histograms [2,3].

Based on two basic competencies in the field of mathematics, the competence which is able to analyze the mean that measures the center is a necessary competence for students. Therefore, pre-service teachers are required to be able to understand and interpret the mean concepts before they actually teach students in school.

However, some research indicates that there is still a lack of understanding of students about the concept of the mean [4–6]. They are still focused on the procedure for determining the mean, which is the ratio of the sum of all the data and the number of data. When they are asked to interpret the mean value, they are still unable to explain it clearly.

Interpretation of the mean can be done by connecting the mean with the context. There are four contexts that can be used to understand the mean concepts such as data reduction, fair share, a typical value, and signal in noise. The explanation of these four contexts will be further explained in the following discussion.

Formally, the mean is the central tendency of a data. However, the meaning of the mean can be understood from various contexts [7], at least for the purposes of teaching. There are four contexts that

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

can be used to understand the mean, i.e. data reduction, fair share, a typical value, and signal in noise [6].

In the context of data reduction, the mean value can be used to simplify the data set. A collection of values in a sample or population can be simplified into a value, i.e. its measure of center. Mean is one measure of the data center. The simplification of such data is intended to make easy to memorize them.

The second context is a fair share. In this context, the values in a data are shared from one value to another in order to obtain the same value. This context is suitable for elementary school students because, for the first time, they will be introduced to the concept of sharing in division operation.

Mean as a typical value, assumes that the mean is a value that often appears in the distribution or value that is located in the center of data after the data are sorted. The mean here is considered to be the same as the mode or median. Although mean, median, and mode are generally three different values, they are all used to determine the center of data.

The last context that can be used to understand the mean is a signal in noise. Based on this point of view, each random observation has an estimated value of a specific value, but it is unknown. Each observation has an average that has a certain deviation to the true mean. The deviation is called as an error. The mean of each observed average is considered to be the actual data mean. This kind of context is perfect for building an understanding of distribution.

In addition to statistical reasoning, there are two other concepts that become the focus of current statistical learning called statistical literacy and thinking [8,9]. Therefore, the three concepts are first discussed in this paper and then will be explained by focusing on what statistical reasoning is.

Statistical literacy is the essential and basic skills used to understand statistical information or research results [10]. This skill consists of the skills to organize data, create and display tables, and represent data from multiple points of view. In addition, this skill also contains an understanding of the concepts, terms, and symbols in statistics, as well as an understanding of opportunities as a measure of uncertainty.

Statistical thinking is a person's understanding of the reasons why a statistical investigation is conducted, and how to do it [11]. With this understanding, one can choose the right method of analyzing data. In addition, with statistical thinking one can criticize and evaluate the results of statistical research.

Statistical reasoning is a way for a person to interpret statistical information and statistical ideas [12]. Reasoning stands for a person's understanding of the statistical processes and the person's interpretation of the statistical results.

From that sense, it is apparent that the concept of statistical thinking and reasoning are overlapped. But these two concepts are quite distinguishable. For example, someone who has knowledge of when and how to apply certain statistical procedures, then it can be said that the person has statistical thinking. And if the person is able to interpret the results of the statistical procedures he or she performs, then the person has statistical reasoning.

Based on the previous explanation, statistical reasoning is an ability that must be owned by school students. Students should not only be able to perform certain statistical procedures, but they must also be able to interpret the results of the statistical investigations they have undertaken.

To promote statistical reasoning, teachers should make students deal with real data around them, process and present the data, and ultimately draw conclusions from the data. Thus, they are able to give the reason why their conclusions are meaningful.

2. Methodology

To find out the statistical reasoning of students about the mean, the author analyzed data collected from 26 students. Data obtained in the form of student work result to the problem of mean. These problems can be shown as follows.

When asked to find the mean of the numbers 17, 43, 22, and 38, Joko turned in the computations illustrated below to show the mean was 30. Explain her method and illustrate it to find the mean of a set of different numbers.

17	43	22	38
20	40	22	38
20	30	32	38
28	30	32	30
30	30	30	30

(Adapted from Bennett, Jr., Burton, Nelson [13])

The problem is in addition to testing the students' understanding of the mean concepts and also for testing student pedagogies. In addition, the problem wants to show the mean concept as a fair share context.

After the data has been obtained, the authors analyze the types of student work. From the analysis, obtained three level of answers. The three types of answers are further described in the next section.

3. Results

The students' work on the mean problem given in outline can be divided into three categories, described as follows.

3.1 Fail in Understanding the Problem

Polya split the steps in solving the problem into four steps, such as understand the problem, devise a plan, carry out the plan, and look back. By using these Polya steps, it can be said that most students make mistakes in the first step, i.e. understand the problem.

Most students make mistakes when reading the numbers on the table. They assume all numbers are numbers to be searched for. So, they tend to add up all the numbers, then divide by 20, that is, the number of these values. Answers in this category are represented by student 1:

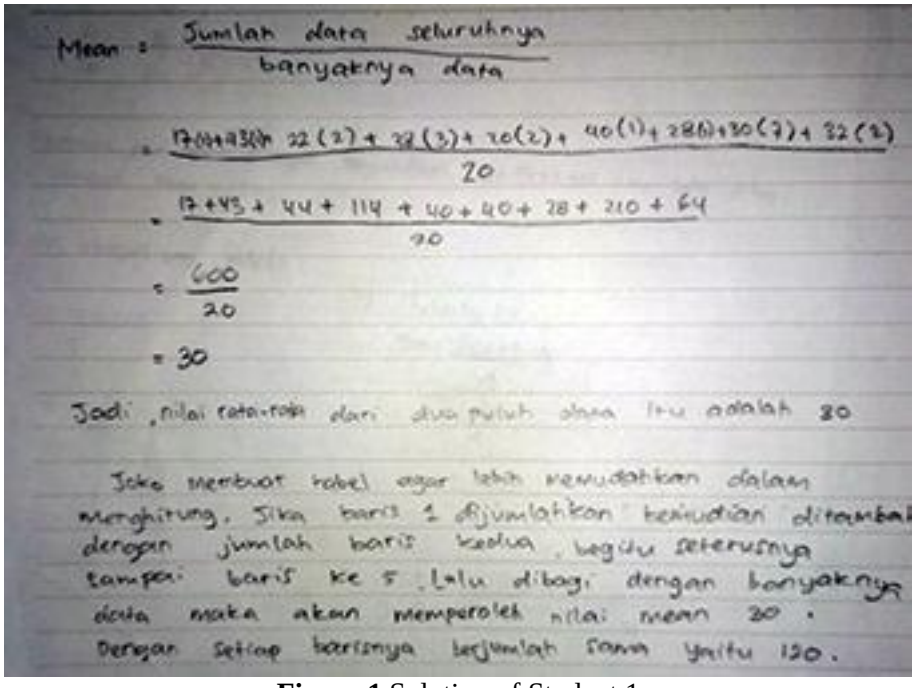


Figure 1 Solution of Student 1

The answer illustrates how students have not understood the problem as a whole. They have not analyzed in depth what relationship between the first, second, third, fourth, and fifth rows in the table. They are still fixated on the numbers that appear, and immediately assume that all of these numbers are numbers to be determined by the mean.

3.2 Success in Looking Pattern, Fail in Interpreting Method

The second category is the category where students are able to find patterns in the method of determining the mean, but they cannot interpret the method. This category is represented by the following 2 student work.

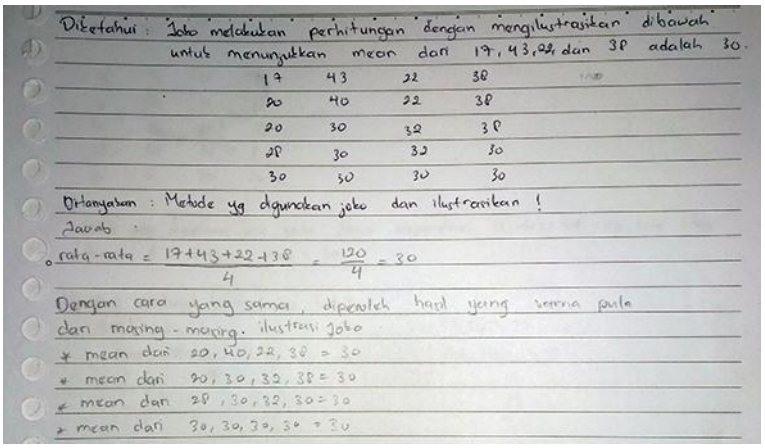


Figure 2 Solution of Student 2

The student's answer gives insight that they can already see the method of determining the mean in depth, i.e. by looking at the pattern of each of the steps taken, but they are still not able to interpret the method.

In the answer, it has been shown that each row has the same mean, but it has not been shown how the relationship between rows. In other words, the student looks at each line separately. He has not explained the relationship of the numbers in each line.

3.3 Success in Interpreting the Method

The third category is the category where the student can explain the steps or procedure of the method of determining the mean. The answers in this category are represented by the following 3 student responses.

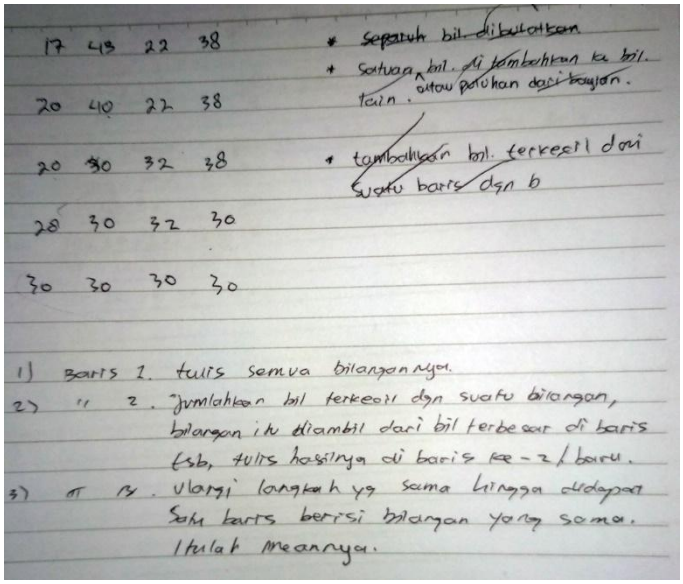


Figure 3. Solution of Student 3

From the answer, it can be seen that the student has explained the method used in detail. He has been able to find relationships between rows in the given steps. And the steps are rewritten in detail in order to be used to determine the mean in different cases.

4. Discussion

This paper wants to answer how students understand the concept of mean. Therefore, based on the data that has been collected the answers to these questions can be categorized as follows.

4.1 Ignoring Problem Context

From the data obtained, it can be said that most students still consider the mean is simply as the ratio between the sum of the overall value of the number of these values. Assuming this, they will be very sensitive to the numbers that appear, but instead, ignore the essence of the problem.

Students still cannot find the context of the given mean problem. Thus, they jump directly into the values given and look for what is asked. Once they find the given values and the ones being asked are the means, they instantly add up all those values and divide them by the number of the values. So, they do not realize that the numbers in each line are a representation of the method in the determination of the mean.

As a result, students still do not have statistical reasoning in interpreting the other method in finding mean. They have not been able to criticize the method and test its truth. Instead, they show the method they already know and used it on the given problem, even with a misconception of the context.

4.2 Not Seeing Relation Between Steps

The answer of the second category of students has shown that they have been able to see that the mean of each line is the same. Using the mean formula that they already know, they can show that the mean of each line is 30.

However, they can't find the relationship between lines. They think of one line with another separately. So, they still cannot find the "big idea" in the given method of determining mean.

If viewed from the statistical reasoning perspective, students in this category still cannot interpret the mean value. They are still fixated on the formula they've been using. Using this formula, they determine the mean of each row and find that the mean of all rows is the same. They have not found where the numbers in the second, third, fourth, and fifth rows come from.

4.3 Explaining the Method for Finding Mean

Students in this category can already find patterns in each row in the given method of determining the mean. They explain that the smallest number must be added with a certain constant, and the constant is "taken" from the larger number on the same line. In other words, they already associate the mean with a fair share context.

Based on data from this category, it can be said that some students already have enough reason statistical reasoning. They were able to analyze other methods of finding mean, by not gluing only to the methods they knew.

5. Conclusion

Based on the results of this study, it can be concluded that most students still have low statistical reasoning that in understanding the concept of mean. Most students still have not been able to analyze other methods in determining the data mean. Nevertheless, there are some students who have been able to interpret the steps in determining the mean by relating it to the fair share context.

Based on these conclusions, the authors provide recommendations in learning statistics at school or at the university level. Statistical learning should prioritize the achievement of students' statistical reasoning. This can be done by designing learning that requires students to collect data from their everyday life, then organize the data, and in finally they can make conclusions based on their own interpretations. This kind of learning can help students to connect the real data representation into a more abstract representation.

References

[1] Dani B Z and Joan G 2004 *Statistical Literacy, Reasoning, and Thinking: Goals, Definitions, and Challenges* (Netherlands: Springer)
[2] Kemendikbud 2016 *Kompetensi Inti dan Kompetensi Dasar Sekolah Menengah Pertama/Madrasah Tsanawiyah (SMA/MA)* (Jakarta: Kementerian Pendidikan dan Kebudayaan)
[3] Kemendikbud 2016 *Kompetensi Inti dan Kompetensi Dasar Sekolah Menengah Pertama/Madrasah Tsanawiyah (SMP/MTs)* (Jakarta: Kementerian Pendidikan dan Kebudayaan)
[4] Jacobbe T and Carvalho C 2011 *Teachers' Understanding of Averages* (Netherlands: Springer)
[5] Leavy A and O'Loughlin N 2006 Pre-service Teachers Understanding of the Mean: Moving Beyond the Arithmetic Average *J. Math. Teach. Educ.* **9** 53–90
[6] Pollatsek A, Lima S and Well A D 1981 Concept or computation: Students' understanding of the mean *Educ. Stud. Math.* **12** 191–204
[7] Watson J M and Moritz J B 2000 The Longitudinal Development of Understanding of Average *Math. Think. Learn.* **2** 11–50
[8] Garfield J and Ben-Zvi D 2009 Helping Students Develop Statistical Reasoning: Implementing a Statistical Reasoning Learning Environment *Teach. Stat.* **31** 72–77
[9] Garfield J and Ben-Zvi D 2004 *Research on Statistical Literacy, Reasoning, and Thinking: Issues, Challenges, and Implications* (Netherlands: Springer)
[10] Gal I 2002 Adults' Statistical Literacy: Meanings, Components, Responsibilities *Int. Stat. Rev.* **70** 1–25
[11] Greer B 2000 Statistical Thinking and Learning *Math. Think. Learn.* **2** 1–9
[12] Del Mas R C 2004 *[3] . (2004). A Comparison of Mathematical and Statistical Reasoning* (Netherlands: Springer)
[13] Bennet J and Burton N 2012 *Mathematics for Elementary Teachers: A Conceptual Approach* (New York: McGraw-Hill Companies, Inc.)