Measure of Central Tendency: Undergraduate Students' Error in Decision-Making Perspective

Rosidah, Fadhil Zil Ikram
[1]*Mathematics Department, Universitas Negeri Makassar, Indonesia*
[2]*Mathematics Department, Universitas Negeri Malang, Indonesia*

dhilikram@gmail.com

Abstract

The purpose of this research was to examine undergraduates' understanding of the central tendency measure from a decision-making perspective. The research adopted a qualitative method by employing an interview and test to obtain the data. It enrolled 93 undergraduate students who had previously studied basic statistics and applied statistics. Four students were selected for interviews out of the 93 participants. The analysis model used included data condensation, data visualization, and conclusion and verification. A large number of students were unable to provide explanations for their decisions. The majority of students related the test with the necessity of calculating an average or selecting a more straightforward measure. None of the students was aware of the presence and effect of outliers in the data. The undergraduate students demonstrated a lack of awareness of the factors that could influence their decision-making. The students did not consider other variables. The majority of them were unaware of the benefits and drawbacks of using mean, median, and mode to describe data.

Keywords: decision-making; errors; measures of central tendency; undergraduate students

INTRODUCTION
Statistics has become inextricably linked to our daily lives, particularly at work. We come into contact with data on a near-daily basis as the vastness of information grows. Statistics is necessary because it teaches us how to collect, organize, and analyze data and draw conclusions and make appropriate decisions (Tiro, 2008). One of the most common measures used in statistics is the measure of central tendency.

Numerous individuals have used central tendency measures to characterize data used in decision-making in various sectors, including business, politics, and education (Gravetter & Wallnau, 2017; NCTM, 2000). Each measure of central tendency has its use (Witte & Witte, 2017). For instance, polling companies used mode as a central tendency to indicate which presidential candidate received the most votes in their quick count. They cannot use the mean or median because the values of the measures convey no information. In another instance, where there are multiple modes, this is inappropriate because the optimal one only involves a single value or category (Weisberg, 1992). The

illustrations demonstrate how a particular sector applies central tendency measures in practice.

There have been numerous and growing studies on measures of central tendency in education. There has been much research on the following topics: measure comprehension and interpretation (Saidi & Siew, 2019; Santos & da Ponte, 2013); statistical reasoning and misconceptions (Rosidah et al., 2018; Maryati & Priatna, 2018; Ismail & Chan, 2015; Zaidan et al., 2012); decision-making (Roy et al., 2016; Holt & Scariano, 2009); (Manikandan, 2011). Additionally, one examined teachers' knowledge of the measure (Groth & Bergner, 2006), and another examined averages, including the concept of average and its development (Sharma, 2008; Mokros & Russell, 1995; Strauss & Bichler, 1988), as well as the history of averages (Bakker, 2003). There are also studies discussing the implementation of a particular method or learning design to teach the concept of measures of central tendencies, such as one that used realistic mathematics education (Meitrilova & Putri, 2019) or one by Kraus (2010) that employed a fictional story.

Central tendency measures is a prevalent topic of research in Indonesia. The majority of the research

discussed the implementation of a specific learning method to teach it, such as improving mathematical reasoning through a problem-posing approach (Chasanah et al., 2019) or utilizing a hypothetical learning trajectory in a game rating context (Kusumaningsih et al., 2019). However, there is a dearth of research in education that discusses the central tendency measure in the context of decision-making in problem-solving.

Decision-making is a critical aspect of our daily lives because an error in this area can have a negative impact on the outcome of our work. To illustrate, the government wishes to implement several programs to assist a village with sixteen heads of families and an average income of no more than 1.5 million rupiahs. The administration directed census officers to collect monthly income data for each head of the family in the village and describe it using central tendency measures. The following are the findings.

Table 1
Monthly income of a certain village

Head of Family	Monthly Income
A	1,000,000
B	1,000,000
C	900,000
D	1,100,000
E	1,200,000
F	1,100,000
G	800,000
H	900,000
I	700,000
J	1,000,000
K	1,000,000
L	1,300,000
M	800,000
N	1,200,000
O	16,000,000
Total	30,000,000

According to Table 1, if census officers used the mean (2,000,000) to describe the data center, the government could cancel the programs, putting the village at a disadvantage. However, upon closer examination, the officers discover a value that differs significantly. They could use the median in this case, which is 1,000,000. It is referred to as an outlier value.

Outliers are values that deviate significantly from the other observations or data. It is one of the factors causing the error when deciding to use a specific measure to explain our data. If ignored, it has a significant impact on our decision-making when solving problems. The preceding illustration demonstrates how outliers have a considerable effect on decision-making.

The presence of extreme values is not the only factor influencing how an individual decides a statistics problem involving a measure of central tendency. Our decision is also significantly affected by the type of data, mode values, the distribution, and other factors (Weisberg, 1992).

We conducted a preliminary study to ask undergraduate students about the most appropriate measure to use when describing data with outliers. The data presented were identical to those in Table 1. Nevertheless, all students agreed that the mean was the best measure because its formula encompassed all cases. They did not consider the outliers. It implied that they lacked a fundamental understanding of outliers and were unaware of their existence. This lack of comprehension may result in fatal errors in the future, as data encounters are inevitable in the working world.

To summarize, the literature and preliminary research emphasize the importance and necessity of conducting a study on undergraduate students' decision-making errors, particularly regarding the concept of measure of central tendency. The purpose of this study is to describe undergraduate students' misconceptions in selecting the most appropriate measures of central tendency when faced with a decision-making problem.

METHOD
Participants and Sites
This research's objective was to describe the phenomenon of undergraduate students' understanding of central tendency measures using a qualitative approach. The study was conducted in November, the odd semester of 2019/2020, at three universities in Indonesia in the mathematics and economics departments.

Ninety-three undergraduate students, who were in their third and fourth years, participated in our study. Students who became the participants were ones that already took the basic statistics and applied for statistics courses. Four students, consisting of two females and two males, were recruited to be interviewed to clarify and elucidate their answers. Their lecturers recommended interviewees based on our criteria, which included good communication skills, as our objective was to delve deeply into their understanding.

Data Collection
A test and an interview were used to elicit data of students' understanding of central tendency measures. The test required the participant to determine whether two statements were true or false and explain their reasoning. We conducted the test by administering one item for one day. The three statements that comprise the test are listed in Table 2.

Table 2
Statements regarding decision making

Number	Statement
1	Monthly income of 18 people in a housing estate in million rupiah are 5, 5, 8, 8, 67, 68, 69, 70, 72, 73, 75, 77, 78, 81, 88, 90, 95, and 98. Among mean, median, and mode as a measure of central tendency, the best measure to describe the data is mean
2	Test results of 14 students in a classroom are 20, 22, 88, 95, 88, 98, 100, 95, 88, 96, 97, 89, 99, and 92. Among mean, median, and mode as a measure of central tendency, the best measure to describe the data is mean
3	IQ Test results of nine high school students are 122, 124, 123, 123, 125, 125, 125, 124, 124, 126. Among mean, median, and mode as a central tendency measure, the best measure to describe the data is mean.

The data in the first statement contains outliers. As a result, the mean is not the best way to describe the data. Additionally, the mode values are outliers, indicating that the mode is not the optimal choice. Regarding statement 2, because there are extreme values, the mean is not the optimal measure. Finally, the third statement contains data that is normally distributed. As a result, the mean, median, and mode values are identical, and they are the most appropriate way to describe the data.

The interview lasted between 15 and 20 minutes. After conducting the interview, we immediately transcribed it to obtain the interview's raw data. Transcribing is necessary because analyzing data in written form was easier for us.

Data Analysis
The data analysis process consisted of three stages: condensing the data, displaying the data, and concluding (Miles et al., 2014). We coded, selected, and filtered interview transcripts and students' test responses to obtain pertinent information during data condensation. We categorized participants' responses that were similar and had the same meaning. This process occurred continuously until the final report was finished.

We summarized students' responses and organized them into tables to illustrate the variety of students' responses to each item. Part of the interview transcripts selected was then displayed in an excerpt to provide clarifications or explanations of participants' response. Data display helped us determine the next course of action; we returned to data condensation or further analysis and conclusion.

Finally, the researchers compared the data and looked for patterns or explanations before drawing conclusions about students' comprehension of central tendency measures in decision-making. Before

concluding, we verified our findings by reexamining the raw or condensed data.

This study used triangulation to ensure the data's credibility by administering a test and conducting an interview. Triangulation was used to compare the test and interview results to identify patterns or data consistency.

RESULTS AND DISCUSSION
Results of undergraduate students' responses to the test are presented in Table 3, Table 4, and Table 5.

Table 3
Summary of students' reasons in Statement 1

Reason	Freq.
Mean is also used to determine mode and median	2
Center is mean	1
Mean describes average	18
Mean is easier to calculate	2
Explaining the formula of mean	2
Mean is a grouped data	1
Median is the most appropriate because it is sorted	5
Mean could be calculated	2
All value is the most appropriate because all of them are computable	2
There is no average	1
The mode is easier to find, and the mode are 5 and 8	11
Mean is sorted	1
Mean involves all data	3
Median and mode are better because it is a sorted data	1
There is no median and mode	1
Repeating the statement	6
Does not give reasons	34

Table 4
Summary of students' reasons in statement 2

Reason	Freq.
Because mean describes the average	15
The mode is better because it appears the most, is easier to find, and the value is 88	14
Explaining the formula of mean	1
Because most of the value is around 90	1
All value is the most appropriate because all of them could be calculated	6
Mean involves all data	3
Easier	1
There is no average	3
Mean is computable	2
Repeating the statement	4
Does not give reasons	43

Table 5

Summary of students' reasons in statement 3

Reason	Freq.
Because we can calculate the mean	7
Average is mean	2
Mean is the overall value	2
Mean involves all data	3
Explaining the formula of mean	2
Median is the most appropriate to find the middle value	3
The mode is also the right one because the data has a mode	3
All measure is right because we can calculate it	1
The mode is also the most appropriate because it involves all data	1
The mode is the most appropriate because it appears the most often. Its value is 124	3
The mode is the most appropriate because we can know the average of data	1
Median and mode because the value is the same, which is 124	4
Repeating statement	7
Does not give reasons	54

Table 3, Table 4, and Table 5 show that no students correctly answered the questions. They provided insufficient justifications for their determination of the statements' correctness. Over 36% (36.6 percent, 46.2 percent, and 58.1 percent) of students could not explain their reasoning for all items. Additionally, more than half of the participants did not state why they chose a particular response.

According to Table 3, the most frequently stated responses to the statement were that the mode is easier to locate and that the mean accurately describes the average. Additionally, some (6 students) repeated the sentences to demonstrate the statement's correctness. There was no mention of outliers or extreme values by students.

Similarly, in Table 4, 15 students indicated that statement 1 was correct because the mean accurately describes the average, while 14 students indicated that statement 1 was incorrect because the mode is easier to calculate. Additionally, Table 4 demonstrates that, while some of them (3 students) understood the concept of mean, which is calculated using all data, none of the participants appeared to be aware of outliers' existence. The number of students who only rewrote the second statement is lower than the other statements (4 participants).

Table 5 shows that none of the students mentioned the term *distribution of data*. The most common answer was that the mean is the best measure because it is computable. The number of students who rewrote the statement as their reasons was the highest in this statement.

The four students, who participated in the in-depth interview, were coded as S1, S2, S3, and S4. Three sections present and discuss their interview results. They consist of the interviewees' responses to the first, second, and third statements.

Participants' Response to Statement 1

The findings of the study revealed that S1 and S3 responses were quite similar. The following excerpts show their answers.

Table 6

S1's responses to statement 1

		Responses and Questions
P	:	In this statement, why did you say that mean was the best?
S1	:	Because we want to know the monthly income of the residents
P	:	How about the median and the mode?
S1	:	It is not necessary because we want to know the average
P	:	Have you calculated the mean value?
S1	:	Not yet, wait, *calculating the value and then showing the results
P	:	Is it alright if I want to describe the data by this value?
S1	:	Yes, it is the mean

Table 7

S3's responses to statement 1

		Responses and Questions
P	:	You said that the mean is the best measure to describe the data. Why did you choose it?
S3	:	We must find the average income of the housing, so the mean is the right choice.
P	:	How about the median and the mode?
S3	:	Mean, because we cannot find the housing's average income by calculating the median and mode.

According to Tables 6 and 7, both S1 and S3 believed the statement was correct because they assumed the task required them to calculate the average. Although the instruction made no mention of it, the mean was always the best option for them without regard for any criteria. While the interviewees calculated the value in S1, the participant was unaware that outliers affect the mean value. Their responses implied that a single measure applied to all types of data.

The response given by S2 and S4 regarding statement 1 was different. The following Tables reveal the interview excerpts.

Table 8
S2's responses to statement 1

Responses and Questions	
P	: Why is it wrong?
S2	: In terms of income, we have to find the break event point in the middle. Therefore, I choose the median. As for mode, we also can use it by looking at the highest intensity of the value

Table 9
S4's responses to statement 1

Responses and Questions	
P	: Why did you choose the mode?
S4	: Because it is easy. We can directly find the mode values just by looking at the data, which are 5 and 8

As shown in Table 8, S2 connected the answer to the context of income, where the break event point is a value that must be determined. The interviewee claimed that the median was the best income data measure because it represented the break-even point (BEP). However, BEP is a state in which a business makes no profit and incurs no losses and does not involve the middle point of income. In this instance, S2 appeared to have forgotten about the concept of the break-even point. Additionally, the interviewee was unaware of the existence of outliers and their role in data mode.

As for S4, the interviewee stated that the statement was incorrect because the most appropriate one to describe the data was easier to find. Thus, S4 chose modes 5 and 8, even though these values were outliers in the data. It appeared as though the participant believed that the most appropriate measure of central tendency to use in decision-making was the one that is the simplest to calculate and locate. As a result, S4 was also unaware of the outliers.

Participants' Response to Statement 2
The results of the interview show that both S1 and S3 responded similarly. Their responses were evident in the following interview excerpts.

Table 10 – S1's Responses to Statement 2

Responses and Questions	
P	: Why did you choose to mean?
S1	: For the average, the average score
P	: Why did not you choose the median or the mode?
S1	: Mean is the most appropriate
P	: Why
S1	: To find out the average of the score

Table 11
S3's responses to statement 2

Responses and Questions	
P	: You said that the mean is the best for this statement. Why?
S3	: To obtain the average score
P	: How about the median and the mode?
S3	: To find the average score of all students, we have to calculate the mean

Both S1 and S3 associated the statement with the necessity of determining the data's average. They did not refer to alternative measures of central tendency. They frequently concentrated on the mean without considering the existence of all possible values in the data. It demonstrated how there was a tendency to choose to mean arbitrarily.

Nevertheless, in this statement, S2 and S4 gave different responses compared to S1 and S3. The following Table 12 and Table 13 show their answers.

Table 12
S2's responses to statement 2

Responses and Questions	
P	: Is this statement correct or incorrect?
S2	: *thinking for a while Wrong, because all value can be calculated

Table 13
S4's responses to statement 2

Responses and Questions	
P	: Why did not you choose the mean?
S4	: Because the mode is easier to find

S2 failed to contextualize the data in statement 2. Rather than that, the interviewee stated that all three measures were the most appropriate due to their calculable values. As with S1, the participant chose the mode because it was easier to determine for the interviewee. S4's response was very similar to those of S1 and S3. The three believed that the optimal measure of central tendency is a single measure that applies to all data types.

Participants' Response to Statement 3
The results of the interview show that both S1 and S3 responded similarly. Their responses are as follows.

Table 14
S1's responses to statement 3

	Responses and Questions
P :	In this data, why did you say pick the mean as the best measure?
S1 :	Because we want to know the average
P :	How about the median and the mode?
S1 :	It is not necessary. We need the average
P :	Have you calculated the mean, median, and mode?
S1 :	No, wait, *calculating the mean, median, and the mode and then showing the results The values are the same as the mean. If the median and mode values are the same as the mean, then all three measures are the most appropriate.
P :	Why do you always involve the average?
S1 :	Because average is the best one to describe data and mean is the average
P :	So how did you choose the best measure of central tendency to describe a set of data?
S1 :	By checking another measure whether their values are the same as the mean or not.

Table 15
S3's responses to statement 3

	Responses and Questions
P :	You said that the mean is the best measure to describe the data. Why did you choose it?
S3 :	Because what we must find is the average of the IQ, so the mean is the right choice.
P :	How about the median and the mode? Try calculating their values
S3 :	*calculating mean, median, and mode and showing the results My choices are mean, median, and mode. Providing that the median value is the same as the mean and the mode is not, then the best measures are the mean and the median. Whichever value is the same as the mean, then that value is also the right choice to describe the data
P :	So, what factors did you think could affect your decision in choosing the best measure?
S3 :	Whether the values of the other measures are the same as mean or not.

S1 and S3 considered another central tendency to describe the typical in the data based on Table 14 and Table 15. They compared the median and mode values to the mean, with the mean serving as the reference point. If the median or mode value was equal to the mean, then the median or mode might also be the best measure for describing data. Their errors stemmed from their consideration of unrelated factors and their failure to consider other factors that might affect their decision-making. When the interviewer asked them to calculate the three measures, their decision changed because their values were identical. Thus, they took into account false factors when deciding on central tendency measures and still had an inclination toward the use of mean.

S2 and S4 responded differently to statement 3. Their answers are as follows.

Table 16
S2's Responses to Statement 3

	Responses and Questions
P :	Why do you say that the statement is incorrect?
S2 :	Because all values are the same, so the most appropriate measures are mean, median, and mode.
P :	In statement 2, you said that when all values can be calculated, then mean is not the only appropriate measure, and in Statement 3, you said that if the values are all the same, then the three measure is the most appropriate. Therefore, what factors did you think affecting our decision-making in statistics problems which involves the measure of central tendency?
S2 :	*thinking for a while The first one is the context. For example, in the context of income in statement 1, the median is the most appropriate. However, there are sets of data in which their context do not need specific measures to describe the typical. In this case, when all the measures could be calculated, then mean, median and mode best describes the data.

Table 17
S4's responses to statement 3

Responses and Questions	
P	: Why did you choose the mode?
S4	: Because the best measure is one that is the easiest to find.
P	So, did you mean that the mean and median value is not the easiest one to determine?
S4	It applies to the mean because we have to calculate the sum of all values and then divide it with the number of data. It is not convenient. However, for the median, if the data is already sorted, I think the median also could be one of the right choices to describe the data

As illustrated in Table 16, S2's response to their strategy, or how the participant selects the best measure to describe the data, was still ambiguous. As a result, the researchers conducted a thorough investigation by requesting clarification from the interviewee. S2 reflected on the responses to the clarifying questions that had been posed. Finally, the participant concluded that while some sets of data with a particular context require specific measures to describe them adequately, others do not. In other instances, a particular type of data did not necessitate the application of special measures. As a result, the interviewee stated that all three values (mean, median, and mode) were acceptable, as long as the mode value existed. To summarize, S2 did not take distributional or outlier-related factors into account.

As for S4, the interviewee still picked the mode as the answer due to the ease of calculating its value. According to the participant, the best central tendency measure for describing data is the simplest to determine. However, S4 stated that the median also became a value that best describes it if the data is sorted. It is because calculating the median is much easier in sorted data than calculating the mean. As a result, S4 was unaware of the benefits and drawbacks of using each central tendency to describe the data. They considered an irrelevant factor, which is the convenience in determining the measure values.

DISCUSSION
The results indicated that none of the undergraduate students who took the test could provide the correct reasoning for the statement's correctness. The majority of them related mean to average directly without considering certain factors, such as how each value affects the mean value. It demonstrated a lack of understanding regarding outliers. They were unaware of the advantages and disadvantages of using the mean, median, and mode to describe a set of data as central tendency measures.

Numerous students believed that a single measure would be the most appropriate for all data collection. The majority of participants selected mean and mode. Some chose the former because they believed it more accurately represented the average, while others picked the latter for convenience. Its value can be determined easily by comparing the frequency of data and selecting the most frequent.

Many undergraduate students in this study were unable to calculate the mean, median, and mode values. This step is required to compare the data's measures and select the most accurate that describes the data. Even when the interviewer instructed them to compute the value, they were unaware of outliers or other conditions necessary for making a correct decision.

This finding was consistent with previous research in which a large proportion of participants were unfamiliar with the procedure for calculating the measure of central tendency. They were unaware of the existence of outliers, how sensitive the mean is to outliers, or how resistant the median is to outliers (Zawojewski & Shaughnessy, 2000; Groth & Bergner, 2006; Jacobbe, 2012; Karatoprak et al., 2015; Sharma, 2008; Ulusoy & Altay, 2016).

According to APOS theory, students' comprehension in this study is limited to the stage of process conception. At this stage, particularly in statistics, students understand only a portion of their principles or characteristics (Arnon et al., 2014). They only knew that using mean benefits is beneficial because it incorporates all of the data points. Additionally, as indicated by the results in Tables 3 and 4, some participants were still in the action conception stage. They were only aware of the mean calculation formula and how to substitute the value. These students were unaware of the benefits and drawbacks of using the mean.

The error which the students made was called *mis-logical construction*. These mistakes happened because the individual did not know several conditions that must be satisfied for a statement to be true (Subanji, 2015). In this case, the participants did not consider the distribution of data, the existence of outliers, the number of modes, and other factors.

The errors made by interviewees were also conceptual errors. It occurs when an individual is unaware of the fundamental principles or properties (Nolting, 2012). In this instance, S2 was unaware of the relative merits and demerits of mean, median, and mode. Additionally, their error implied that there was a flaw in their concept construction. Subanji (2015) asserts that students who have a construction hole have an incomplete schema. The participants in this study were unaware that selecting the most appropriate measure to describe data necessitates some conditions, including outliers.

Many undergraduate students participating in this study already knew that choosing the best measure requires several considerations. They did not know the merits and demerits of using every measure of central tendency to make a decision. Thus, they were unable to correctly answer the statements given.

Several unusual responses were discovered among the 93 participants' responses. For instance, several participants stated in statement 1 that they could use the mean value to determine the mode and median of data. Some asserted that there was no average, and others claimed that the median or mode did not exist. In statement 3, one student believed that mode encompassed all data. These are elementary errors. It is an error that should not occur at a certain level of education (Brodie, 2010). At the undergraduate level, particularly for students who have previously taken or are currently enrolled in a statistics course, the errors are unexpected. Their comprehension did not even extend to the concept of action.

CONCLUSION

To conclude, undergraduate students' understanding of factors affecting decision making is still lacking. They were not aware of many things, including outliers, the number of modes, or data distribution.

Further investigation on the topic by considering some aspects or conducting a different study is also necessary. For example, one can investigate the same theme based on their background (gender, learning styles, cognitive styles, intelligence, mathematical identity, and education level). One also can develop learning methods or designs to teach the measure of central tendency. A study exploring undergraduate students' basic errors is also necessary to determine why they made the mistakes at their current education level.

The findings of this research, hopefully, will give contribution both to students and teachers in general. This contribution could be in teaching materials preparation, development of mathematics textbooks, and statistics curriculum design. Teachers could also use the results of this study to prevent or tackle the students' misconception.

REFERENCES

Arnon, I., Cottrill, J., Dubinsky, E., Oktaç, A., Fuentes, S. R., Trigueros, M., & Weller, K. (2014). *APOS theory: A framework for research and curriculum development in mathematics education*. Springer.

Bakker, A. (2003). The early history of average values and implications for education. *Journal of Statistics Education, 11(1), 1-18*. DOI: 10.1080/10691898.2003.11910694

Brodie, K. (2010). *Teaching mathematical reasoning in secondary school classrooms*. Springer US.

Chasanah, N. A., Sisworo, S., & Dwiyana, D. (2019). Meningkatkan penalaran matematis siswa pada materi ukuran pemusatan data melalui pendekatan *problem posing*. *Jurnal Pendidikan: Teori, Penelitian, dan Pengembangan, 4*(4), 542-547.

Gravetter, F. J., & Wallnau, L. B. (2017). *Statistics for the behavioral sciences* (10th Ed). Cengage Learning.

Groth, R. E., & Bergner, J. A. (2006). Preservice elementary teachers' conceptual and procedural knowledge of mean, median, and mode. *Mathematical Thinking and Learning, 8(1), 37–63*. DOI: 10.1207/s15327833mtl0801_3

Holt, M. M., & Scariano, S. M. (2009). Mean, median, and mode from a decision perspective. *Journal of Statistics Education, 17(3), 1-16*. DOI: 10.1080/10691898.2009.11889533

Ismail, Z., & Chan, S. W. (2015). Malaysian students' misconceptions about measures of central tendency: An error analysis. In AIP Conference Proceedings, *1643*(1), 93-100. American Institute of Physics. DOI: 10.1063/1.4907430

Jacobbe, T. (2012). Elementary school teachers' understanding of the mean and median. *International Journal of Science and Mathematics Education, 10*(5), 1143–1161. DOI: 10.1007/s10763-011-9321-0.

Karatoprak, R., Gülseren, K. A., & Bengü, B. (2015). Prospective elementary and secondary school mathematics teachers' statistical reasoning. *International Electronic Journal of Elementary Education, 7*(2), 107–124.

Kraus, S. (2010). Monstrous methods for teaching central tendency concepts. *Teaching Statistics, 32*(1), 21-23. DOI: 10.1111/j.1467-9639.2009.00403.x

Kusumaningsih, W., Albab, I. U., & Angga, S. D. (2019). Desain pembelajaran ukuran pemusatan data menggunakan konteks game rating. *JIPMat, 4*(2), 182-188.

Manikandan, S. (2011). Measures of central tendency: Median and mode. *Journal of Pharmacology and Pharmacotherapeutics, 2*(3), 214-215. DOI: 10.4103/0976-500X.83300

Maryati, I., & Priatna, N. (2018). Analysis of statistical misconception in terms of statistical reasoning. *Journal of Physics: Conference Series, 1013*(1), 1-7. DOI: 10.1088/1742-6596/1013/1/012206

Meitrilova, A., & Putri, R. I. I. (2020). Learning design using PMRI to teach central tendency materials. *Journal of Physics: Conference Series, 1470*(1), 1-11. DOI:10.1088/1742-6596/1470/1/012086

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: a methods sourcebook* (3rd Ed). SAGE Publications, Inc.

Mokros, J., & Russell, S. J. (1995). Children's concepts of average and representativeness. *Journal for Research in Mathematics Education, 26(1), 20-39*. DOI: 10.2307/749226

National Council of Teachers of Mathematics. (2000). *Principles and standards for school Mathematics*. National Council of Teachers of Mathematics.

Nolting, P. D. (2012). *Math study skills workbook : Your guide to reducing text anxiety and improving study strategies* (4th ed.). Cengage Learning.

Rosidah, Budayasa, I. K., & Juniati, D. (2018). An analysis of statistical reasoning process of high

school students in solving the statistical problem. *Journal of Physics: Conference Series, 1028*(1), 1-5. DOI: 10.1088/1742-6596/1028/1/012125

Roy, G. J., Hodges, T. E., & Graul, L. (2016). How many jelly beans are in the Jar?. *Mathematics Teaching in the Middle School, 21*(7), 424-430. DOI: 10.5951/mathteacmiddscho.21.7.0424

Saidi, S. S., & Siew, N. M. (2019). Assessing students' understanding of the measures of central tendency and attitude towards statistics in rural secondary schools. *International Electronic Journal of Mathematics Education, 14*(1), 73-86. DOI: 10.12973/iejme/3968

Santos, R., & da Ponte, J. P. (2013). Prospective elementary school teachers' interpretation of central tendency measures during a statistical investigation. *In Proceedings of the Eighth Congress of the European Society for Research in Mathematics Education (CERME8).* Middle East Technical University.

Sharma, S. (2008). Exploring Fijian high school students' conceptions of averages. *Pacific-Asian Education, 20(2), 29–45.*

Strauss, S., & Bichler, E. (1988). The development of children's concepts of the arithmetic average. *Journal for Research in Mathematics Education, 19*(1), 64–80. DOI: 10.2307/749111

Subanji, S. (2015). *Teori kesalahan konstruksi konsep dan pemecahan masalah matematika.* Universitas Negeri Malang Press.

Tiro, M. A. (2008). *Dasar-dasar Statistika* (3rd Ed). Andira Publisher.

Ulusoy, C. A., & Altay, M. K. (2016). Analyzing the statistical reasoning levels of pre-service elementary school teachers in the context of a model eliciting activity. *International Journal of Research in Education and Science, 3*(1), 20–30. DOI: 10.21890/ijres.267363.

Weisberg, H. F. (1992). *Central tendency and variability.* Sage Publications.

Witte, R. S., & Witte, J. S. (2017). Statistics (11th edition). Wiley.

Zaidan, A., Ismail, Z., Yusof, Y. M., & Kashefi, H. (2012). Misconceptions in descriptive statistics among postgraduates in social sciences. *Procedia - Social and Behavioral Sciences, 46,* 3535–3540. DOI: 10.1016/j.sbspro.2012.06.100

Zawojewski, J. S., & Shaughnessy, J. M. (2000). Mean and median: Are they really so easy?. *Mathematics Teaching in the Middle School, 5*(7), 436-440. DOI: 10.2307/41180850