International Journal of Instruction *October 2023 ● Vol.16, No.4*
e-ISSN: 1308-1470 ● www.e-iji.net *p-ISSN: 1694-609X*
pp. 1081-1098

Article submission code: 20221231163328
Received: 31/12/2022 Revision: 29/05/2023
Accepted: 20/06/2023 OnlineFirst: 15/09/2023

Students' Reasoning about Variability When Comparing Two Sets of Data from the Perspective of Commognitive

Desi Rahmatina
Asst. Prof., Mathematics Education Department, Universitas Maritim Raja Ali Haji, Tanjungpinang, Indonesia, *desirahmatina@umrah.ac.id*

Toto Nusantara
Prof., Mathematics Department, Faculty of Mathematics and Science, Universitas Negeri Malang, Indonesia, *toto.nusantara.fmipa@um.ac.id*

I Nengah Parta
Assoc. Prof., Mathematics Department, Faculty of Mathematics and Science, Universitas Negeri Malang, Indonesia, *nengah.parta.fmipa@um.ac.id*

Hery Susanto
Assoc. Prof., Mathematics Department, Faculty of Mathematics and Science, Universitas Negeri Malang, Indonesia, *hery.susanto.fmipa@um.ac.id*

This study aims to explore students' reasoning about variability when they compare two groups of data from the perspective of commognitive. The task contains two questions, namely comparing the distribution of the patient's recovery time in the two treatments and choosing one of the two treatments that the patient should use. Data on the assignment were given in the form of ogive graphs. Interviews were conducted to get in-depth information about students' reasoning on variability. Fifty-nine mathematics education students were involved in this study (16 male and 43 female). The students have studied material about measures of dispersion. Eight of them were chosen as the subjects of the study. The assignment was given to the students to be completed by thinking aloud. Data on the assignment were given in the form of ogive graphs. The finding of this study was that there are four categories of students' reasoning when they interpret variability, namely quantitative, confirmation, single center point, and informal categories. The educators can introduce students to various types of data displays that can foster the development of students' reasoning about variability. Further research can be carried out by assessing students' reasoning about the variability of various levels of education.

Keywords: reasoning, variability, commognitive, categories, two sets of data, students' reasoning

Citation: Rahmatina, D., Nusantara, T., Parta, I. N., & Susanto, H. (2023). Students' reasoning about variability when comparing two sets of data from the perspective of commognitive. *International Journal of Instruction*, *16*(4), 1081-1098. https://doi.org/10.29333/iji.2023.16459a

INTRODUCTION

Variability is a fundamental component in statistical thinking (Pfannkuch & Wild, 2004), a key in statistical reasoning (Garfield & Ben-zvi, 2008), and the heart of statistical reasoning (Wells, 2018). Without variability, statistics as a discipline would not even exist (Cobb & Moore, 1997). Therefore, variability is very essential in statistical reasoning and thinking. It is even associated with all aspects of statistical investigations (Garfield & Ben-zvi, 2008; Watson, 2006). Therefore, Garfield & Ben-zvi (2005) suggest that "the centrality of variability needs to be given an emphasis in subjects from the lowest grades in high school to introductory statistics courses in college". However, despite the fact that it has been emphasized in every educational level, students still have misconceptions about it. Cooper (2018) studies undergraduate students' conceptions and misconceptions in interpreting variability in several groups of data presented in the form of histograms, bar graphs, and value bars. In this case, the understanding of the concept of variability contributes to the development of students' statistical thinking (Reid & Reading, 2008). Students who do not understand the concept of variability are thought to have difficulty in making correct reasoning about variability.

Reasoning about variability can be used to make decisions in choosing one among several sets of data (Amaro & Sánchez, 2019; Whitaker & Jacobbe, 2017). Rabab'h et al.(2019) state that "decisions can be based on *quantitative"* technique. Relevant to this, Mullineaux (2009) argues that the measure of variability can be determined quantitatively by using formulas such as the standard deviation and variance formulas. He also argues that determining the measure of variability quantitatively can provide an objective assessment. Hence, the students will have more confidence in the decisions they make. Consequently, students generally do not need to verify their final decisions because the decisions have a measure of certainty (Pugalee, 2004). However, if one is not sure about the decision he makes, he will tend to need the evidence for the validity of the decision. Tien et al.(2020) found that students identify the final results of their problem solution using diagrams, formulas, and patterns to confirm if the solutions taken are correct. By using good reasoning skills, students can answer complex problems (Purnomo et al., 2022).

There are several methods students can choose for doing reasoning about variability. Shaughnessy et al.(2004) categorize student's responses to reasoning about variability into six categories, namely *Specific Data Points*, *Centers*, *Variation*, *Distribution*, *Informal Inferences*, and *Context*. The Center category takes place when students make the reasoning about variability based on the mean or median, or both. However, in itself mean is not enough to make the decision about choosing one of the two sets of data. What if the mean of several sets of data is the same? Therefore, reasoning for variability can be used to make decisions in choosing one among several sets of data (Amaro & Sánchez, 2019; Kramer et al., 2017). In this case, mean can be used as the initial step to determine the measure of variability.

When someone compares two data sets and makes a decision to choose one of the two data sets presented in a graph, he/she is communicating with himself about what he

thinks of the information on the graph. This shows that thinking is an individual communication (Sfard, 2008b). Sfard (2008b) combines the word communication and cognitive and created the term *commognitive*. Lampen (2015) analyzed Sfard's (2008b) commognitive theory in the teacher's narrative in making sense of the algorithm of average. Pratiwi et al.(2022) used commognitive to describe the errors of communication throughout the minds of students in order of answers to problems solved. However, research that used commognitive as a lens for investigating reasoning about variability in the comparison of two sets of data is very rare.

This study explores students' reasoning about variability in comparing two groups of data presented as an ogive graph using Sfard's commognitive theory. Students reasoning about variability can be seen from the way students compare data distribution in two groups of data and make decisions about which data group to choose using the concept of variability.

Review of Literature

Variability can be visualized in form of graphs (Cooper & Shore, 2010) and can be used to compare several data sets (Amaro & Sánchez, 2019; Cooper, 2018; Whitaker & Jacobbe, 2017). Comparing data sets is a fundamental and crucial activity in statistics a (Biehler et al., 2018; Frischemeier, 2019; Shin, 2021), and is a productive means of motivating students to reason in statistical manner (Ben-Zvi, 2004). Futhermore, comparing different sets of data will enable students to think about various key statistical ideas (Biehler et al., 2018), such as thinking about ideas about measures of central tendency, measures of variability, and the shape of the distribution of these data sets.

Many studies have been conducted on the reasoning of variability in comparing different sets of data that are presented in graphs, such as bar graphs (Amaro & Sánchez, 2019; Cooper & Shore, 2010; Whitaker & Jacobbe, 2017), histogram (Cooper, 2018; Cooper & Shore, 2010; Whitaker & Jacobbe, 2017), dot plot (Kramer et al., 2017; Padmi, 2015; Wells, 2018). Amaro & Sanchez (2019) gave two problems to 87 ninth-grade students in two different classes. The problems contain risk context about the variation. The findings show that there are three levels of students' reasoning about variability, namely circular argumentation (Level 1), data consideration (Level 2), and data combination: Risk (Level 3).

Commognitive can be used as a lens for research (Presmeg, 2016), including research in statistical reasoning (Lampen, 2015), and reasoning of variability. Commognitive framework, according to Sfard (2008b), is made up of *word use*, *visual mediator*, *narrative*, and *routine*. Word use is the meaning of a word or keyword. Word use that involves the use of mathematical terms (e.g., "topology, polygon, prism") is categorized as literate word use, whereas word use that involves daily vocabulary that has specific meanings in mathematics, such as "limit and group" is classified as colloquial word use (Mpofu & Pournara, 2018). Visual mediators involve real objects, such as diagrams, tables, graphs, and equation (Mpofu & Pournara, 2018). Numbers, algebraic formulas, algebraic notation, graphs, pictures, and diagrams are examples of visual mediators that

are most widely used in mathematics (Rahmatina et al., 2022; Sfard, 2007, 2008a). Mpofu & Pournara (2018) categorizes visual mediators into two, namely iconic visual mediators and symbolic visual mediators. In statistics, ogive graphs and tables are iconic visual mediators, and statistical symbols such as $\bar{x}$ symbol and s symbol are symbolic visual mediators. Whereas narrative is a series of expressions made to describe objects, processes and relationships between objects (Nardi et al., 2014; Sfard, 2008b). Zayyadi et al.(2019) give an example of students explaining the area of a rectangle theorem, using the concepts of addition and subtraction. Justification and reasons for certain actions are called substantiation narrative, whereas narratives that involve formulas or rules are called memorisation narrative (Mpofu & Pournara, 2018).

Routines involve habits that are applied regularly and distinctly that are used in specific ways, such as defining, guessing, proving, estimating, generalizing, and abstracting (Nardi et al., 2014). Mpofu & Pournara (2018) categorize routines into two, namely type-based routines and property-based routines. Type-based routines are further classified into routine ritualized and exploratory routine. A ritualized routine occurs when students can perform the required procedure but are unable to explain how the procedure was obtained. An exploratory routine is the result of reflective imitation. It guides students to learn why certain steps are completed. One of the property-based routines is the applicability routine. Applicability routine can be observed in a) the ability to solve equations, b) the ability to make graphs from equations/tables, c) the ability to use tables to identify the main features of a discourse, d) the ability to use main features to create graphs.

METHOD

This study uses Sfard's (2008b) commognitive theory to explore students' reasoning about variability when comparing two groups of data presented in the ogive. This study was qualitative research of exploratory type—fifty-nine students (16 male and 43 female). The students involved in this study were third-semester students (16 students), fifth-semester (26 students), and seventh-semester (17 students) of the Mathematics Education Study Program at the Raja Ali Haji Maritime University, Indonesia. Eight of the 59 students were chosen as the research subjects, namely S1, S2, S3, S4, S5, S6, S7, and S8. They were chosen based on the following considerations: 1) they can make reason about variability, and 2) they can communicate their thoughts well in writing and verbally.

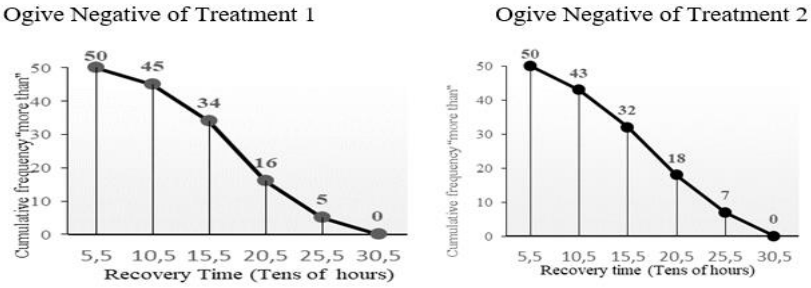
The following instruments were used for the study: 1) the researcher as the main instrument, 2) assignments presented in ogive graphs, and 3) interview guidelines. The assignments require students to make reasoning about variability. The assignments were validated by two validators, a mathematics education expert and a mathematics expert. The assignments' readability test was conducted by two mathematics education students who had completed a statistical method course. Expert validation and empirical validation tests were conducted to ensure that the assignments were worth using for this research. The assignments given to the students are presented in Figure 1. A semi-structured interview was used for this research to make it more adaptable to the actual

condition met during the interview, which may develop beyond the questions that had been prepared. The interview guidelines were validated by 2 expert validators, one from mathematics education background and the other from mathematics.

Students were given a simple statistical question to answer by thinking aloud as an exercise before they did the assignments for this study. The exercise was given to accustom the students to think aloud activity when doing the assignments for this study. The students did the think-aloud tasks through the Zoom virtual platform. Subsequently, the researcher interviewed the students through WhatsApp voice messaging, chat, and Zoom to obtain in-depth information on how they completed the assignments.

The following procedure was used to analyze the data: 1) The think-aloud interview data were transcripted, 2) the written answer, think-aloud, and interview data were studied, 3) the data were reduced by sorting the essential elements to focus on the most important things and leave out the unnecessary, 4) the data were coded, 5) students' reasoning about variability was described from the perspective of commognitive, 6) conclusions were made. Triangulation was performed by checking the validity of the assignment's answer sheets against the subject's think-aloud data and interview data.

The spread of a dangerous virus is increasing in an area, as indicated by the number of people who are positively infected with the virus. With specific treatments, people infected with the virus can recover. There are two kinds of treatment given to expedite the patient's recovery. The effectiveness of each treatment to 50 patients is presented in the form of negative ogive below.



a. Which treatment has more extended time distribution for a patient's recovery? Why?
b. Which treatment is better for the patient? Why?

Figure 1
Assignments for this study

FINDINGS AND DISCUSSION

The distribution of the students' responses in completing the task is presented in Table 1. These 35 students (59%) could not perform reasoning about variability. It means they did not use the concept of variability when comparing the distribution of the patient's recovery time. They could not choose one of the two groups of data based on variability.

Table 1
Distribution of responses of prospective subjects and research subjects

Categories of students' responses	Number of students
Quantitative	10 (17%)
Confirmation	3 (5%)
Single center point	4 (7%)
Informal	7 (12%)
Did not make reasoning about variability	35 (59%)
Total	59 (100%)

Quantitative Category

There are 10 students in this category. Two of the 10 students were chosen as the subjects of the research, namely S1 and S2. In the perspective of commognitive, subjects use literate and colloquial word use when making reasoning about variability. The examples of literate word use used by the subjects when comparing the distribution of patients' recovery time and making the decision about choosing between the two treatments include "frequency, mean value, absolute value, sigma, mean, mean deviation, standard deviation, variance, table, more than," and the examples of colloquial word use they used include "deviated, spread, center." Word use was visualized in the form of symbolic mediator visual. For example, the word use "absolute value, mean, mean deviation, standard deviation, and variance" are respectively symbolized by | |, $\bar{x}$, SR, s, and s^2 .The value table's iconic visual mediator was used by the students to help them make the measure of variability calculations. Sfard (2008b) argues that visual mediator plays a pivotal role in building effective communication. In this case, the numerical information of the graph helps students in making the calculation with the average and the measure of variability formulas. Mueller & Brand (2018) argues that numerical information can be useful for comparing various options, evaluating opportunities and risks, or for estimating expected outcomes before a decision is made.

The substantiation narrative was noticeable when students gave reasons for the distribution of patient's recovery time and their treatment of choice. To compare the distribution of patients' recovery time, S1 and S2 chose a treatment that had a considerable measure of variability as a treatment that had a more extended distribution of patient recovery time. In this case, S1 and S2 chose treatment two, which has a longer distribution of patients' recovery time compared to treatment 1. Whereas in determining which of the two treatments should be used by the patient, the subjects made different choices. S1 chose a treatment that has a small variability, whereas S2 chose a treatment that has a large variability as a better treatment for the patient, i.e. S1 chose treatment 1 and S2 chose treatment 2 as the better treatment for the patient. Memorisation narrative was found when the subject used variability relevant rules. For example, S1 said that variability is a measure of data distribution to see how wide data are distributed or deviated from the data center. In contrast, S2 said that variance is a measure that shows

how far the values are spread from the average, so that if the variance value is large, the data set is spread far from the average.

Applicability routine was found when students completed the average deviation formula and created a value table to determine the average deviation. Lavie et al.(2018) called routines like this a "numeric comparison" routine. In this procedure, students work on a formula for the measure of variability to be able to compare two groups of data and make decisions about which group of data to choose. Procedures like this help them obtain a reliable conclusion so that they no longer need to verify the conclusions obtained. For example, the large deviation standard obtained by the students also indicates that the variability is also large. This is in line with Pugalee (2004) who found that students generally do not verify their final decision because the decision has high degree of confidence. Exploratory routine was found when the subject chose the values in the table to be substituted into the average deviation formula. The commognitive components used by students are shown in Figure 2 and in the interview excerpts between the researcher (R) and the subject.

R : What do you mean by the center in your answer and your audio recording? (Researcher showed the answer sheet and played the audio recording of S1)
S1: The average value, ma'am.
Why is it that the treatment which has a smaller mean deviation is said to have less data spread from the data center? (Researcher showed the answer sheet and played
R : S1's think-aloud recording)
Because if the value of the mean deviation is small, then the distribution of the data is narrower or not far from the average. The smaller the mean deviation value, the
S1: faster the patient's recovery time is.
Why did you use the mean, standard deviation, and variance in answering questions?
P : So that I can determine which sample data distribution is better and closer to
S2: the mean.

From Figure 2 and the interview excerpt, it is clear that emphasis was given by the subject on applicability routine by comparing the distribution of patient's recovery times and choosing between the two treatments based on the measure of variability. Both S1 and S2 used the measures of variability formulas, such as mean deviation, deviation standard, and variance formulas. Furthermore, they compared the values obtained from the calculations of variability measures to get the distribution of the patient's recovery time and make the decision about which treatment is better.

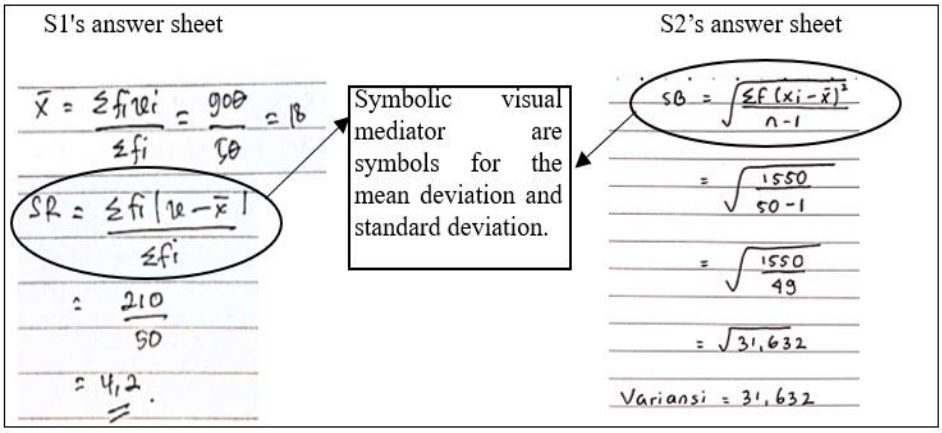


Figure 2
Applicability routine by S1 and S2

Confirmation Category

There are three students in this category. Two of the three students (S3 and S4) were chosen as subjects for this study. Both subjects represent the students in the confirmation category. In this category, subjects used two methods to compare the distribution of patients' recovery time in the two treatments and to make the decision about which treatment is better. First, the subjects identified the information in ogive. Then, they calculated the measure of variability to confirm the decision made with the first method. The identification of the information in ogive made by the subject can be seen in the interview excerpt between the researcher (R) and the subjects below.

P : What did you notice when you saw the ogive?
S3 : I looked at the values in the X and Y axes. Then the information in the ogive tells me that there are points of distance that are inconsistent in treatment 1, some are far, and some are too close, whereas in treatment 2, they are more consistent, and the fracture is not too obvious (S3 draws in the ogive).
S4 : I noticed the graphs and tried to see the change in the data and the differences between the two graphs. More patients recovered in treatment 1 than in treatment 2 in the recovery time range of 14.5 and 20.4 hours.

Then the subject performed applicability and exploratory routines procedures. S3 and S4 used performed applicability routine in completing the formula for the mean, mean deviation, and standard deviation, as well as in making the table of values. The subjects used the average value to calculate the variability measures. An exploratory routine was found when the subject chose the values in the table to determine the average value, mean deviation, standard deviation, and variance. Applicability routine performed by the subject was also apparent in the think aloud transcript below.

S3: To get the standard deviation, first we have to find ……. the variance ….the formula is frequency times square absolute ….what is the formula for square ....just multiply it with the frequency, the total of which gets the variance, the results of which is divided by 49, which is 50 minus 1.

S4: The formula for standard deviation is the root of the sum of fx squared minus the mean squared. We should first find the mean, which is equal to the sum of fx divided by the sum of f. The sum of fx is 5, the x is 5 point 5 plus 10 point 5, 15 16 divided by 2 equals eight. Therefore, the x is 8.

Based the interview and the think aloud excerpt, it is show that the subjects were focusing on using the iconic visual mediator and applicability routine to make the reasoning about variability. Based on the ogive's iconic visual mediator, both S3 and S4 expressed their substantiation narrative for the reasons for their choice of treatment that has longer recovery time distribution. The subjects performed applicability routine in working out the formula on the measure of variability such as the formula of deviation standard. They did this to confirm the assumptions they made about variability based on the iconic visual moderator. This is because the conclusion obtained based on the observation of the iconic visual mediator is only a conjecture and has a low degree of confidence. To increase confidence, they need to perform an applicability routine by working on the formula of measure of variability. Tien et al. (2020) found that students identified their final result in solving the problem to confirm that the solution was correct. In this case, students use the applicability routine to confirm the conclusions that have been obtained based on information on the iconic visual mediator. S3 and S4 chose treatment with a larger measure of variability as the treatment that has the longer distribution of recovery time.

In this case, S3 and S4 chose treatment 2, which has a longer distribution of patients' recovery time compared to treatment 1. The decision was based on the calculation results of the measure of the variability of patients' recovery time of treatment 2, which is larger than the measure of the variability of treatment 1. S3 and S4 believed that the treatment with smaller recovery time distribution is a better choice for the patient's recovery from the dangerous virus. S3 used a substantiation narrative that a smaller measure of variability means that the distribution of recovery time is also smaller. Hence the recovery time is faster, and it is better for the patients infected with the dangerous virus. Whereas S4 used a memorization narrative that a smaller standard deviation means that the patient's recovery time is close to average. He assumes that if the recovery time is close to the average, then the recovery time of the treatment is more accurate.

Single Center Point Category

There are four students in this category. Two of the four students (namely, S5 and S6) were chosen as subjects for this study. Both subjects represent the students in the single center point category. From the perspective of commognitive, the words that subjects use when reasoning about variability, such as "frequency, table, negative ogive," are categorized as literate word use, and words such as "center, homogeneous, and heterogeneous" are categorized as colloquial word use. The iconic visual mediator is the

table that makes it easier for the subjects to calculate the average value and to identify the frequency around the average, and the ogive determines the position of the data center point of the ogive. Whereas symbolic visual mediators are symbols that define the average values. Substantiation narrative is the reason that the subject gives in determining the treatment that has the more extended distribution of patients' recovery time and in choosing the best treatment for the patients. In this case, S5 believes that the treatment with a smaller number of recovered patients around the data center is the treatment that has the more extended distribution of patients' recovery time. At the same time, S6 believes that the treatment with more recovered patients outside the mean is the treatment that has longer distribution of patients' recovery time. Both S5 and S6 choose treatment 1 as the treatment of choice for patient's recovery. Their reason being that there are more recovered patients in the data center or in terms of recovery time average in treatment 1 than treatment 2.

Memorisation narrative can be found in variability related rules when the subject thinks that the size of the distribution of data can be seen from the amount of data gathered in the data center. For example, S5 believes that variability or the size of data distribution can be used to see how far the data spread from the data center, the larger the size of the data spread, the more heterogeneous the data are, and the smaller the size of the spread, the more homogeneous they are. Below is the interview excerpt between the researcher (R) and S5:

R : Are you familiar with the concept of data measure spread?
S5: I do, ma'am.
R : What is data measure spread?
S5: The measure used to see the extent to which the data spread from the data center, the larger the size of the data spread, the more heterogeneous the data are. If the data distribution is small, the data is homogeneous.
R : Which one is better, homogeneous or heterogeneous recovery time distribution?
S5: Homogeneous recovery time distribution is better.

Then the subject performed applicability routine and exploratory routine procedures. S5 and S6 performed applicability routine in completing the average formula and created a value table to make it easier for them to perform calculations on the average formula. An exploratory routine was found when the subject chose the values in the table to determine the average value and identify the number of recovered patients around the ogive's mean. In this category, students used mostly applicability routine, exploratory routine, and iconic visual moderator to make reasoning about variability. The commognitive elements performed by the subjects are shown in Figure 3 and the interview transcript below.

R : Why didn't you use the data spread measure formula to find the longer recovery time distribution?
S6: Because from the average value we can see how much data is not on the average, So it can be assumed that the data that is not around the average must be spread away from the average.

From Figure 3 and the interview transcript, it is evident that the subjects placed emphasis on applicability routine, exploratory routines, and visual mediators. Applicability routine was performed by completing the average formula. From the average value obtained, the subjects performed an exploratory routine by identifying the frequency of the patient's recovery time within and outside the average. The identification of recovery time frequency was made with the help of the iconic visual mediator. In this case, the subjects chose treatment 2 that has a longer distribution of patients' recovery time compared to treatment 1. The choice was made based on the subjects' declared substantiation narrative that treatment 2 has more numbers of recovered patients outside the mean than treatment 1, i.e. 36 patients recovered outside the mean in treatment 2 compared to only 32 patients recovered outside the mean in treatment 1.

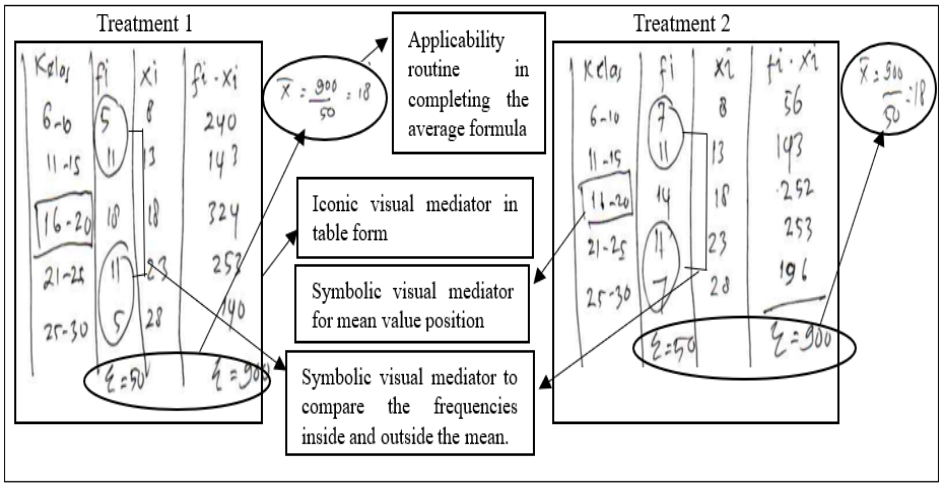


Figure 3
Variability comparison of two groups of data by one of the subjects in the single center point category

In the category, students use one numerical value, the average value, to determine variability. Students in this category are similar to those in the center's category in Shaughnessy et al.(2004) study. However, the center category in Shaughnessy et al. (2004) study is aimed at students who choose one of two data sets based on the average or median value or both. In this study, students use the word "average," where the average value is not the final decision in the decision-making process. Instead, the average value obtained from the applicability routine results is explored by identifying the frequencies around the average in the iconic visual mediator to make the decision.

Informal Category

There are 7 students in this category. Two of the seven students (namely, S7 and S8) were chosen as subjects for this study. Both subjects represent the students in the

informal category. In the commognitive perspective, the subject's use of word when making reasoning about variability include the word use literate of "frequency, data center, negative ogive, data range". To compare two groups of data based on variability, the subjects used iconic visual mediator in the form of table and ogive to determine the position of the average value. S7 used the ogive-based table, while S8 used ogive to determine the average value.

The substantiation narrative expressed by the subject in determining the distribution of longer recovery times is based on the number of recovered patients in the average range of patient recovery time. S7 thinks that the fewer patients are in the data center the more patients are outside the data center. S7 assumes that the treatment with more recovered patients outside the data center means that the recovery time of the treatment is faster. Whereas S8 assumes that, at 15.5-20.5 recovery time intervals (in tens of hours), fewer patients recovered in treatment 2 than in treatment 1. Memorisation narrative is a rule related to variability, about which the subject believes that the size of the data spread shows how far the data spreads from the average.

The subject's exploratory routine is evident when the subject performed a procedure or steps in proving the distribution of a patient's long recovery time. The ritualized routine occurs when the subject determines the position of the average recovery time as a basis for comparing the distribution of recovery time. The subjects could do the procedure to determine the distribution of recovery time; however, the fact that the average position is the same as the mode position cannot be explained by S7. Similarly, S8 could not explain why the average value is the same as the median value. In this case, the subjects did not use the calculation of the average formula or the measure of variability, as shown in the interview snippet below

Think aloud	Interview
S7: Which treatment should the patient use then? Why? I think treatment 2 should be the patient's choice because, from the spread of data, the frequency of treatment 2 is faster. We can see in the negative opinion of treatment 2, which is the data center, it is between 16 and 20.	P: What makes you think treatment 2 has a better recovery rate?
	S7: Because it has more recovered patients outside the average, that's why it has a better recovery rate.
S8:Treatment 2 is better for the patient because the number of patients who recovered outside the average interval was more than the number of patients who recovered in treatment 1.	P: Why didn't you use a formula to determine the measure of variability?
	S7: The answer to the question was easy to guess, so I didn't need a formula. I learned the formula to solve this problem once, but I was not sure I really understood it, so I didn't use the formula.

The students in the category only observed iconic visual mediator to reason about the variability. This shows that the role of visual mediators is very crucial in the discourse (Sfard, 2008a). Shaughnessy et al.(2004) categorize reasoning about variability as informal inferences when students make conclusions based on predictions, consistency, reliability, advantages, or disadvantages. In fact, by making the decision informally, students were trying to find the structure on their own (Rufiana et al., 2018). In this study, the student in the informal category did not perform an applicability routine to obtain average value or measure of variability but to observe an iconic visual mediator to get the position of average value in it and use it to identify every frequency around the average. About the memorization narrative, they maintained that the size of the data spread could be used to see how far the data is spread from the average. In this case, the iconic visual mediator used by students did not prevent them from making a visual mediator-based memorisation narrative. However, the visual mediator functions only to support the substantiation narrative. Tabach & Nachlieli (2011) argue that visual mediators used in communication often influence a person's ideas about what is being discussed and the chosen discursive action. Furthermore, Ripardo (2017) suggests that visual mediator is a tool and routine a process, whereas endorsed narrative is the outcome of a discourse.

The students in the informal category need sufficient knowledge and experience in making decisions based on an iconic visual mediator. It is because data interpretation is a complex process that involves cognitive and technical aspects (Queiroz et al., 2017). Lem et al.(2013) found that chart interpretation is not always easy for students. To reason correctly, one needs to be able to interpret the distribution graph correctly. For example, when students know that the median, mode, and mean values are the same, they are expected to be aware of the involvement of the concept of the normal distribution. The concept of normal distribution and the concept of central tendency are the basis for the substantiation narrative used in decision-making. That is why students must have statistical reasoning ability, of which an understanding of main concepts such as distribution, center, dispersion, association, uncertainty, randomness, and sampling is necessary (Garfield, 2002).

CONCLUSION AND SUGGESTIONS

Variability can be used to compare groups of data and to make the decision about which data group to choose. Relevant to this, commognitive is a lens that is used to examine students' reasoning about variability. There are many ways students can choose to do reasoning about variability. Some students emphasize applicability routine so they can make decisions based on the calculation of measures of variability. Some were able to emphasize the iconic visual mediator. The visual mediator helped them identify every value in the chart or table to make the decision. Some even emphasized both applicability routine and iconic visual mediator that enable them to make the decision based on both the measures of variability and identification of every value in the visual moderator. Accordingly, the existing commognitive in the students' memory can be used to make reasoning about variability.

This study has identified four categories of reasoning about variability used by students when comparing two groups of data and making the decision about which one to choose. There are the quantitative, confirmation, single center point, and informal categories. Before making the decision about which set of data to select, first, they had to describe the appearance of the data. The ability to read the appearance of the data is the foundation for the ability to predict and identify trends Jones et al.(2004). Therefore, the way they describe the data affects their approach to decision-making about which set of data to choose. Graphic data display of iconic visual mediator plays a significant role in building the substantiation narrative.

Understanding the various statistical ideas or concepts is essential for the student's ability to do reasoning about variability, especially those of the informal category. Students also need to be introduced to different types of data displays with various data conditions. Accordingly, educators play a significant role in developing students' ability to do reasoning about variability. One of the things that educators can do about it is to introduce students to various types of data displays that can foster the development of students' reasoning about variability.

This research contributes to the educator that there are a few ways to be used to assess students when they solve the problem of variability. There are at least four categories of student reasoning about variability in comparing two sets of data from the perspective of commognitive that have been found in this study. The reasoning categories explored in this study do not represent all levels of education, as the study only involved university students. The researchers believe that more student's reasoning categories may be found if more students are involved in the research. Therefore, other studies involving students of different education levels are needed.

REFERENCES

Amaro, J. A. O., & Sánchez, E. A. (2019). Students Reasoning About Variation in Risk Context. In G. Burrill and D. Ben-Zvi (Ed.), *Topics and Trends in Current Statistics Education Research* (pp. 51–69). Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-030-03472-6_3

Ben-Zvi, D. (2004). Reasoning about variability in comparing distributions. *Statistics Education Research Journal*, *3*(2), 42–63.

Biehler, R., Frischemeier, D., Reading, C., & Shaughnessy, J. M. (2018). Reasoning About Data. In J. G. Dani Ben-Zvi , Katie Makar (Ed.), *International Handbook of Research in Statistics Education* (pp. 139–192). Springer International Publishing. https://doi.org/10.1007/978-3-319-66195-7

Cobb, G. W., & Moore, D. S. (1997). Mathematics , Statistics , and Teaching. *The American Mathematical Monthly*, *104*(9), 801–823.

Cooper, L. L. (2018). Assessing Students' Understanding of Variability in Graphical Representations that Share the Common Attribute of Bars. *Journal of Statistics Education*, *26*(2), 110–124. https://doi.org/10.1080/10691898.2018.1473060

Cooper, L. L., & Shore, F. S. (2010). The effects of data and graph type on concepts and visualizations of variability. *Journal of Statistics Education*, *18*(2), 1–16. https://doi.org/10.1080/10691898.2010.11889487

Frischemeier, D. (2019). Statistical Reasoning When Comparing Groups with Software—Frameworks and Their Application to Qualitative Video Data. In G. Burrill & D. Ben-Zvi (Eds.), *Topics and Trends in Current Statistics Education Research: International Perspectives* (p. 425). Springer International Publishing. https://doi.org/10.1007/978-3-030-03472-6

Garfield, J. (2002). The challenge of developing statistical reasoning. *Journal of Statistics Education*, *10*(3), 1–12. https://doi.org/10.1080/10691898.2002.11910676

Garfield, J. B., & Ben-zvi, D. (2008). Developing Students' Statistical Reasoning. In *Developing Students' Statistical Reasoning*. Springer Science+Business Media B.V.

Garfield, J., & Ben-zvi, D. (2005). A Framework for Teaching and Assessing Reasoning about Variability. *Statistics Education Research Journal*, *4*(1), 92–99.

Hashim, S., & Nik Pa, N. A. (2014). Meaning of an Ogive by Students of Diploma in Accountancy in an Institute of Higher Education. *Open Journal of Statistics*, *4*, 137–143.

Jones, G. A., Langrall, C. W., Mooney, E. S., & Thornton, C. A. (2004). Models of Development in Statistical Reasoning. In D. Ben-zvi & J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp. 97–117). Springer Science+Business Media Dordrecht All. https://doi.org/10.1007/1-4020-2278-6

Kramer, R. S. S., Telfer, C. G. R., & Towler, A. (2017). Visual comparison of two data sets: Do people use the means and the variability? *Journal of Numerical Cognition*, *3*(1), 97–111. https://doi.org/10.5964/jnc.v3i1.100

Lampen, E. (2015). Teacher narratives in making sense of the statistical mean algorithm. *Pythagoras*, *36*(1), 1–12. https://doi.org/10.4102/pythagoras.v36i1.281

Lavie, I., Steiner, A., & Sfard, A. (2018). Routines we live by: From ritual to exploration. *Educational Studies in Mathematics*, *101*, 153–176.

Lem, S., Onghena, P., Verschaffel, L., & Van Dooren, W. (2013). On the misinterpretation of histograms and box plots. *Educational Psychology*, *33*(2), 155–174. https://doi.org/10.1080/01443410.2012.674006

Mpofu, S., & Pournara, C. (2018). Learner Participation in the Functions Discourse : A Focus on Asymptotes of the Hyperbola. *African Journal of Research in Mathematics, Science and Technology Education*, *22*(1), 2–13. https://doi.org/10.1080/18117295.2017.1409170

Mueller, S. M., & Brand, M. (2018). Approximate number processing skills contribute

to decision making under objective risk: Interactions with executive functions and objective numeracy. *Frontiers in Psychology*, *9*(1202), 1–16. https://doi.org/10.3389/fpsyg.2018.01202

Mullineaux, D. (2009). Methods for Quantifying the Variability in Data. *ISBS - Conference Proceedings Archive*, 1–5.

Nardi, E., Ryve, A., Stadler, E., & Viirman, O. (2014). Commognitive analyses of the learning and teaching of mathematics at university level: The case of discursive shifts in the study of Calculus. *Research in Mathematics Education*, *16*(2), 182–198. https://doi.org/10.1080/14794802.2014.918338

Pfannkuch, M., & Wild, C. (2004). Towards an Understanding of Statistical Thinking. In D. Ben-Zvi & J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp. 17–46). Springer, Dordrecht. https://doi.org/10.1007/1-4020-2278-6_2

Pratiwi, E., Nusantara, T., Susiswo, S., & Muksar, M. (2022). Routines ' errors when solving mathematics problems cause cognitive conflict. *International Journal of Evaluation and Research in Education*, *11*(2), 773–779. https://doi.org/10.11591/ijere.v11i2.21911

Presmeg, N. (2016). Commognition as a lens for research. *Educational Studies in Mathematics*, *91*(3), 423–430. https://doi.org/10.1007/s10649-015-9676-1

Pugalee, D. K. (2004). A Comparison of Verbal and Written Descriptions of Students' Problem Solving Processes. *Educational Studies in Mathematics*, *55*(1/3), 27–47. https://doi.org/10.1023/B

Purnomo, H., Sa'dijah, C., Hidayanto, E., Sisworo, Permadi, H., & Anwar, L. (2022). Development of Instrument Numeracy Skills Test of Minimum Competency Assessment (MCA) in Indonesia. *International Journal of Instruction*, *15*(3), 635–648. https://doi.org/10.29333/iji.2022.15335a

Queiroz, T., Monteiro, C., Carvalho, L., & François, K. (2017). Interpretation of Statistical Data : The Importance of Affective Expressions. *Statistics Education Research Journal*, *16*(1), 163–180.

Rabab'h, B. S., Omar, K. M., & Alzyoud, A. A. Y. (2019). Literature review of the Impact of the Use of Quantitative techniques in administrative Decision Making: Study (Public and private sector institutions). *International Journal of Scientific and Research Publications (IJSRP)*, *9*(7), 515–521. https://doi.org/10.29322/ijsrp.9.07.2019.p9166

Rahmatina, D., Nusantara, T., Prata, I. N., & Susanto, H. (2022). Statistical Reasoning Process of Students in Decision Making Using Commognitive Framework. *Acta Scientiae*, *24*(3), 63–88.

Reid, J., & Reading, C. (2008). Measuring the Development of Students' Consideration of Variation. *Statistics Education Research Journal*, *7*(1), 40–59.

Ripardo, R. B. (2017). Teaching mathematics from the perspective of Mathematics as a discourse. *Ciência & Educação*, *23*(4), 899–915. https://doi.org/10.1590/1516-731320170040014

Rufiana, I. S., Sa'dijah, C., Subanji, S., Susanto, H., & Abdur Rahman Asari. (2018). Informal Statistical Reasoning Of Students Taken Formal Statistics Learning Related To Distribution. *International Journal of Insight for Mathematics Teaching*, *01*(2), 130–140.

Sfard, A. (2007). When the rules of discourse change, but nobody tells you: Making sense of mathematics learning from a commognitive standpoint. *Journal of the Learning Sciences*, *16*(4), 565–613. https://doi.org/10.1080/10508400701525253

Sfard, A. (2008a). Introduction to Thinking as comunication. *The Mathematics Enthusiast*, *5*(2&3), 429–436. https://doi.org/10.1017/S1041610212001652

Sfard, A. (2008b). *Thinking as Communicating: Human development, the growth of discourses, and mathematizing*. Cambridge University Press.

Shaughnessy, J. M., Ciancetta, M., & Best, K. (2004). Students' Attention to Variability when Comparing Distributions. In S. Starkings (Ed.), *The Research Presession of the 82nd Annual Meeting of the National Council of Teachers of Mathematics* (pp. 1–44).

Shin, D. (2021). A framework for understanding how preservice teachers notice students ' statistical reasoning about comparing groups. *International Journal of Mathematical Education in Science and Technology*, *52*(5), 699–720. https://doi.org/10.1080/0020739X.2019.1699968

Tabach, M., & Nachlieli, T. (2011). Combining Theories to Analyze Classroom Discourse : a Method to Study Learning Process. *Proceedings of the Seventh Congress of the European Society for Research in Mathematics Education*, 2524–2538.

Tien, L. van, Tong, D. H., & Vy, N. T. T. (2020). Verification and control of solving pseudo-real problems related to the systems of equations. *Universal Journal of Educational Research*, *8*(8), 3551–3561. https://doi.org/10.13189/ujer.2020.080831

Watson, J. M. (2006). Statistical literacy at school: Growth and goals. In *Statistical Literacy at School*. Lawrence Erlbaum Associates, Inc. https://doi.org/10.4324/9780203053898

Wells, J. F. (2018). Dot plots and hat plots : supporting young students emerging understandings of distribution , center and variability through modeling. *ZDM*, *50*(7), 1125–1138. https://doi.org/10.1007/s11858-018-0961-1

Whitaker, D., & Jacobbe, T. (2017). Students' understanding of bar graphs and histograms: Results from the LOCUS assessments. *Journal of Statistics Education*, *25*(2), 90–102. https://doi.org/10.1080/10691898.2017.1321974

Zayyadi, M., Nusantara, T., Subanji, Hidayanto, E., & Sulandra, I. M. (2019). A

commognitive framework: The process of solving mathematical problems of middle school students. *International Journal of Learning, Teaching and Educational Research*, *18*(2), 89–102. https://doi.org/10.26803/ijlter.18.2.7