Pre-service Teachers’ Statistical Reasoning based on Cognitive Style

Yusfita Yusuf1*, Y.L. Sukestiyarno2

1Mathematics Education Department, Universitas Sebelas April, Indonesia
2Mathematics Education Department, Universitas Negeri Semarang, Indonesia

*Email: yusfitayusuf87@gmail.com

Received: 26 January 2022 ; Revised: 5 April 2022 ; Accepted: 30 April 2022

Abstract. The use of statistical concepts in real-life problems has become a challenge for students. This challenge is related to statistical reasoning (SR) ability which is influenced by several factors including cognitive style. Therefore, this research aims to understand the process of statistical reasoning from the perspective of cognitive style. A qualitative method was used and the subjects were students who are pre-service mathematics teachers. The instrument used were Group Embedded Figures Test (GEFT), statistical reasoning tests, as well as interview guidelines, and the data were analyzed using the qualitative descriptive method. The results showed statistical reasoning of the field independent pre-service mathematics teacher is better than the field dependent. Also, those with the same cognitive style, which is field independent, but different genders showed different reasoning processes. It was concluded that males have better Statistical Reasoning processes than females, and cognitive style as well as gender affect the reasoning abilities.

Keywords: statistical reasoning, cognitive style, field dependent, field independent

Introduction

Statistics is one of the applied mathematics that discusses theories and methods of collecting, measuring, classifying, computing, explaining, synthesizing, analyzing, and interpreting data (Yusuf, Suyitno, Sukestiyarno & Isnarto, 2020). Furthermore, it is viewed as a tool to solve problems that occur in daily life, at work, and in science (Moore, 1997). Statistics is particularly used to describe and predict phenomena using a set of results from measurements (Yusuf, 2017). There are currently many statistical data available in daily life such as the number of participants in a debate or community action, phenomena such as crime rate, population growth, disease spread, number of productions, educational attainment, job trends, etc (NCTM, 2005; Watson & Callingham, 2003). Therefore, statistical knowledge is necessary to interpret, understand, and make good decisions for the data. This is as stated by Karatoprak, Akar and Börkan (2015) that to understand these data, statistical reasoning (SR) skills are required.

Based on the previously described needs, conceptual understanding and SR are the main goals in present statistics learning (Chan, Ismail, & Sumintono, 2016; Jin, Kim, McGhee, & Reiser, 2011; Kalobo, 2016; Türegün, 2014). Students can comprehensively understand statistics when they have a good conceptual understanding and statistical reasoning. Rumsey (2002) stated that the goal of the learning is to have an in-depth understanding to obtain information from existing data, criticize, make decisions from the given information, and
develop research skills. Therefore, statistics course has become a necessity for completing lectures (Garfield & Ben-Zvi, 2005), since the knowledge is a requirement in various research fields. Based on the description, students need to master and understand the concept.

Gal and Ginsburg (Jin, et. al, 2011) stated that students assume statistical courses to be difficult, which may hamper them in completing the research. This also occurred in a workplace where the pre-service teacher often associates statistics courses as difficult. This problem is based on students not understanding the concept, the interrelationship between them, and how to apply it in real life (Chiesi & Primi, 2010; Jin, et. al, 2011; Yusuf, et. al, 2020).

The knowledge of statistics, the interrelationship, and how to use it are closely related to the SR ability. Therefore, statistical reasoning can be defined as a way of reasoning with statistical ideas and understanding information (Gal & Garfield, 1997). This involves making interpretations, representations, and summaries of data. This form of reasoning combines ideas about data and probability and leads to the conclusion and interpretation of results. Furthermore, SR is based on important concepts of data centering, range, probability, correlation, and association, as well as sampling. Lovett (2001) interpreted this reasoning as using statistical tools and concepts to make summaries, predictions, and draw conclusions from the data. In accordance with the opinion of Ben-Zvi and Garfield (2004) that SR is a way of thinking using statistical information. Meanwhile, DelMas (2002) stated that statistical reasoning is the ability to explain the why and how of a result in production in order to draw conclusions. Chervaney et al. (Garfield, 2002) also defined SR as what students can do with statistical content and utilizing their skills in using the concepts to solve problems. They see SR as a process consisting of the following steps (1) comprehension, (2) planning and decision making, as well as (3) evaluation and interpretation. Chan and Ismail (2014) stated that there are four key constructs of statistical reasoning assessment based on the framework of Jones et al., they include (1) describe, (2) organize and reduce, (3) represent, as well as (4) analyze and interpret the data.

When pre-service teachers do reasoning, they use representations to rearrange issues in their mental imagination. This rearrangement can be attributed to the specific cognitive style of the subject, how their cognitive systems operate, and how information is obtained and processed. This is stated by Bendall, Galpin, Marrow, and Cassidi (2016), that individuals have a habitual way of approaching tasks and situations related to certain patterns in cognitive processes including decision making, problem-solving, perception, and attention. This individual style affects the way they view, remember, think, organize and solve specific problems (Kozhevnikov, Hegarty, & Mayer, 2002; Kozhevnikov, Kosslyn, & Shephard, 2005; Presmeg, 1986; Sternberg & Grigorenko, 1997). Also, when students are solving problems, they look for the right solution in their own way (Ali, Hukamdad, Akhter, & Khan, 2010; Ahghar,
According to Faiola and Matei (2005), cognitive style is a strategy of an individual to filter, receive, and process information from the environment. The cognitive styles include stable attitudes, choices, or custom strategies that distinguish individual styles from feeling, remembering, thinking, and solving problems (Saracho, 1997). Therefore, it can be concluded that the cognitive style is the way to receive different stimuli and thinking to learn.

There are several dimensions of cognitive style developed by experts who can differentiate individuals. The most important dimensions in education are field independent (FI) and dependent (FD) (Salameh, 2011). According to Farmaki, Sakkalis, Loesche, and Nisiforou (2019), field independent and dependent are broadly studied cognitive style dimensions. They are designed to measure an individual's ability to identify the embedded part of the organized visual field as a separate entity from a given field. This means students with FD cognitive style tend to look at the pattern as a whole and have difficulty separating certain aspects of a situation. Meanwhile, those with FI can see the parts that make up a large pattern (Idris, 2006). Many studies claimed that students with different cognitive styles receive information processing and problem-solving in different ways (Hassan, 2002).

It is important to know the background information of students while going to statistics lectures (Wilson & MacGillivray, 2006; Reading & Reid, 2006). Studies showed that the results of statistical learning are directly influenced by various cognitive and non-cognitive factors (Chiesi & Primi, 2010; Lai, Tanner & Stevens, 2011; Kheng, Azlan, Ahmad, Leong & Mohamed, 2016; Kheng, Idris, Mohamed & Lyn, 2016; Nasser, 2004; Tremblay, Gardner & Heipel, 2000). In fact, cognitive style is one of the factors that need to be considered in SR abilities, and it is closely related to the reasoning process at each stage. This is in accordance with Bakker (2004) that reasoning using shape as a pattern is more meaningful in understanding and solving statistical problems.

Yusuf, et. al (2020) stated that it was necessary to know the condition of the students before learning. Cognitive style is one of the conditions that need to be known in statistical learning as stated by Lailiyah, Muslimah, and Suntini (2021). Therefore, this research aims to understand the pre-service teachers' statistical reasoning based on cognitive style.

**Method**

This research used a descriptive method with a qualitative approach. Furthermore, the participants were pre-service teachers from the sixth semester of mathematics major who took statistics classes in one university in West Java, consisting of 1 class with 43 students. Research statistics is a concept that includes an explanation of data, how they are obtained, presented,
processed, and interpreted. The data collection techniques used were carried out in three stages. Firstly, participants get Group Embedded Figures Test (GEFT) to classify their cognitive styles, and secondly, they work on reasoning test questions. Lastly, 2 male and 2 female students with FI learning style, and 2 female students with FD style were selected. The SR test instruments used the reasoning stages proposed by Chervaney et al (Garfield, 2002), which are (1) comprehension to see the problem as one in the same class (2) planning and decision making, which means applying appropriate methods for problem-solving, (3) evaluation and interpretation of the result related to the original (native). Figure 1 shows one of the questions in the statistical reasoning tests.

A sample size of n = 25 has a mean = 23, median = 28, and mode = 31. When inputting the data, an error occurred, where the data that should have been 34 was entered as 43.

a. Determine the true mean!

b. Can data entry errors affect the median and mode of the data? Explain!

Figure 1. Test of statistical reasoning

For triangulation, in addition to the written test, the six participants were interviewed. Data analysis was carried out using the constant comparative method. This is because data analysis constantly compares one datum to another, then compares one category with others (Moleong, 2009). In general, the analysis process includes data reduction, categorization, synthesis, and completion with working hypotheses.

Results and Discussion

This research aims to understand statistical reasoning from the perspective of cognitive style. The GEFT data analysis showed that in one class of the sixth semester pre-service teachers who took statistics class, all males were field independent (FI) cognitive style. Meanwhile, for females, some have a field independent cognitive style and others have dependent.

Based on the results of GEFT tests on one class, it was found that all males have FI cognitive style. Suharto, Widada, Susanta, and Haji (2021) stated that there were significant differences in cognitive style between males and females. Furthermore, Onyekuru (2015) stated that males generally have a field independent cognitive style, while females have dependent. According to Witkin (Saracho, 1997), GEFT is a valid and reliable test instrument that requires the participants to put the next geometry plane in a more complex form within 20 minutes. The GEFT tests consist of spatial problems that require mental imagery in analyzing the images (Margaret, 2009). Meanwhile, spatial ability is a person's cognitive capacity to process and manipulate 2D or 3D objects both by changing their position, perception, being able to imagine
how objects look when viewed from various positions, and imagining an abstract object (Rahmatulwahidah & Zubainur, 2017). Margaret (2009) also stated that the difference in cognitive style was caused by spatial ability. The results showed the spatial abilities of males are better than females (Battista, 1990; Ganley, Vasilyeva & Dulaney, 2014). Therefore, since there are no students with FI cognitive style, FI Male, FI Female, and FD Female were selected as the participants.

The SR was analyzed using statistical reasoning stages proposed by Chervaney et al (Garfield 2002), which are (1) comprehension, (2) planning and decision making, as well as (3) evaluation and interpretation. The comprehension stage is seeing a particular matter as a problem, while planning and decision-making involve applying the appropriate method to solve the problem. Furthermore, the evaluation and interpretation stage involve interpreting the results and relating them with the initial problem. The statistical reasoning tests results of each subject for the problem are described in Figure 2.

![Figure 2.a. The result of the statistical reasoning tests for problem (a) FI female](image)

![Figure 2.b. The result of the statistical reasoning tests for problem (a) FD female](image)

![Figure 2.c. The result of the statistical reasoning tests for problem (a) FI male](image)

Based on Figure 2.a., the FI female wrote everything needed to answer the questions and the formula that would be used in performing the calculations, complete with explanations, but did not give a conclusion. Meanwhile, Figure 2.b shows FD female wrote some of the information contained in the problem using symbols (amount = n), but used the wrong concept.
in working on the given problem. In Figure 2.c., the FI male did not write the information contained in the question, did not write the formula, but the results were correct even though they were not concluded.

The interview results of female students with field independent cognitive style for problem (a) are as follows.

**Lecture** : Can you work on that question?
**FI Female** : Yes ma’am
**Lecture** : How did you do it?
**FI Female** : In part a, we are asked to determine the actual mean because there was an error in inputting the data. To find the mean, first, calculate the amount of data with a known mean. After that, the amount of data was reduced by 9 which is the difference between 43 that was entered incorrectly instead of 34. Hence, the actual number of data will be obtained. The formula for the mean is the amount of data divided by the number of data, then the actual mean will be obtained by dividing the actual number by the number of data.

From the interview results, the participants understood the concept to be used when there is an error in data entry to the average value. They can also explain the process.

The following are excerpts from interviews with female students with dependent cognitive style for problem (a).

**Lecture** : Can you work on that question?
**FD Female** : Emh, some can ma’am. But the questions are different from those that are usually given, hence, I was confused about how to do it.
**Lecture** : What is confusing?
**FD Female** : Usually, I just calculate the mean, it is different from the previous question. The question has a median with the same mode as well.
**Lecture** : Take a look at the question, what is being asked in the question?
**FD Female** : Determine the actual mean, is it not SR ma’am?
**Lecture** : What do you mean by SR?
**FD Female** : That is the one, ma’am, which is the mode minus the median then divided by the mean?
**Lecture** : Let us remember again, what is the mean?
**FD Female** : The amount of data divided by the number of data
**Lecture** : What is the median?
**FD Female** : The middle data of the data that has been sorted from smallest to largest
**Lecture** : What is that mode?
**FD Female** : Frequently appearing data
**Lecture** : The question was asked to determine the actual mean. Why were you asked to determine the actual mean?
**FD Female** : There is an incorrect data input
**Lecture** : So how should you do it?
**FD Female** : I do not know, I am confused

From the dialogue above, it can be seen that the participants did not understand the problem, and some of the information in the question was not properly understood. They are
also fixated on questions that are usually given as exercises. Therefore, when the context of the questions is different and combines several concepts in the process, they become unable to solve it, even though they understand mean, median, and mode.

The results of interviews with male students with field independent cognitive style are as follows.

Lecture : Can you work on that question?
FI Male : Yes ma’am.
Lecture : How did you do it?
FI Male : To determine the actual mean, first, calculate the amount of data from the known data and then subtract 9. After that, divide by the number of data. ...

From the interview results, the participants were able to represent the location of the wrong data input and understand the concept. In its explanation, the male students also went straight to the core of the problem.

The result of the statistical reasoning tests of each subject for problem (b) can be seen in Figure 3.

Figure 3.a. The result of the statistical reasoning tests for problem (b) FI female

Translation:
b. The data entry error can clearly affect the median and mode when it turns out that 43 is in the data which is the middle value, and will later become the median. It can also affect the mode when 43 becomes the most data, but apparently there are others whose difference in number is one below 43, while 43 should be 34

Figure 3.b. The result of the statistical reasoning tests for problem (b) FD female

Translation:
b. Yes, it affects when there is an error in data entry. It will affect the data for the median and mode, in the calculation of the data when one of the data is wrong. It will also affect the next data, hence, it is important to be careful, be more thorough, and focus when inputting data

Figure 3.c. The result of the statistical reasoning tests for problem (b) FI male

Translation:
b) The error will only affect the mean, because the location of the error is after the median location, hence, it will not affect the median value. For the mode, it will not have any effect, because the data that appears frequently (mode) remains 31

Based on Figure 3.a., the FI female assumed that the mean change is due to an error in data entry, then the other size of central tendency will also change by using the concept that is known without implementation into the given problem, which is wrong data size. The FD
female assumed that every data entry error will always affect the median and mode (Figure 3.b). The participant even advised to be careful while entering data. Meanwhile in Figure 3.c., FI male described the size of central tendency, and incorrect data was entered on a number line. For the median, the field independent male can explain precisely, but for the change in mode, there was no correct explanation.

The results of interviews of female students with field independent cognitive style are as follows.

Lecture : How about part b?
FI Female : The data error has an effect, ma'am, for both the median and the mode. For example, 43 is the middle value, when an incorrect value was inputted instead of 34, the median value will also change. The mode is the same when 43 is the most value, then there is another data that differs by one in number, then the mode will change.

Lecture : Well, now try to pay attention to the question! Is the median and mode 43?
FI Female : No ma'am.

Lecture : What is the median and mode of the question?
FI Female : Median 28 mode 31
Lecture : Now try to draw a number line and put the known values on the number line.
FI Female : (preservice teachers draw and observe). Emh, the data that changed after the median and mode ma'am.
Lecture : Did the median change?
FI Female : No ma'am
Lecture : Is the mode changed or not?
FI Female : It is possible, ma'am, when there are a lot of data, 34 and 31 are only one difference. Hence, there can be 2 modes, which are 34 and 31.

From the interview results, the student has been unable to represent the location of the wrong data input, even though they understand the concept. After some scaffolding given by the lecturer, they could relate the concepts they understand to the existing problems.

Lecture : Okay. Now if you enter the wrong data, does it affect the median and mode?
FD Female : Yes, it clearly affects ma'am.
Lecture : Why?
FD Female : Yes, the calculation will change ma'am.

FD female did not fully understand the problem, hence, cannot provide a solution. What they know is that when the data changes, the median and mode will change as well.

Lecture : How did you do it?
FI Male : .... Then for part b, the error only affects the mean. The median has no effect because of the location of the data error after the median, the mode also remains 31.
Lecture : What if for example the number 31 has 7 and 34 originally has 6?
FI Male : Yes, if that is the case, the mode has changed, hence, there are two modes. But 31 is still the mode.
From the interview described above, it can be seen that the FI male understands that median is a measure of size, therefore, the explanation was made using the concept of size. Meanwhile, for mode, an explanation was not provided. After some scaffolding given by the lecturer, they were able to explain correctly.

Based on the results of the written and interview tests on problems (a) and (b) for subjects FI Male and FI Female at the comprehension stage, the male identified the problems based on elements that are only related to the problem. This is in contrast to females who recorded all the elements regardless of which were useful or not while solving the problem. At the comprehension stage, the FI female recorded the known elements of the problem using statistical notation, while the FD recorded the elements without statistical notations, and instead tended to re-record the problems. This is in line with Morgan (Kheirzaden & Kassaian, 2011) which stated that when the data is not organized, FI individuals are likely to apply their own structure, whereas FD will accept the problems without being studied. This corresponds to the characteristics of the FI participants that they internally exhibited and processed information with their own structure (Witkin, Moore, Goodenough, & Cox, 1977). FI participants at this stage only recorded important elements, and it is already explained that gender differences affect the processing of this information. The characteristics of FI male are in line with Armstrong, Cool and Smith (2011), stated that FD individual adopts a global orientation to understand and process information, whereas FI adopts an analytical orientation.

At the planning and decision-making stage, the male took a moment to understand the question and wrote sufficiently. However, the female took longer time to understand the question and wrote the calculations in detail. FI male and FI female subjects were also able to determine the concepts to solve the given problems, but FD females were unable to determine the concepts to be used. Although the FD female knew the basic concepts, but does not understand the interrelationship between them. This is particularly evident in solving statistical reasoning tests, where the given questions are non-routine. This is in line with Johnstone and Al-Naeme (Hassan, 2002) which stated that FD participants faced difficulties in separating the 'signal' from 'noise', relevant from irrelevant, as well as what is important and what is confusing. At this stage, FI participants can easily solve existing problems, specifically the FI male that can come up with some alternative solutions. This is in line with the opinion of Hassan (2002) which stated that FI individual's way of thinking promotes a higher appearance in mathematical problem solving than FD.

In the evaluation and interpretation stage, the FI participant of both males and females can appropriately interpret the results. The FI female participant can interpret the results in several ways. Furthermore, males often utilize representations using sketches or drawings, while
females use words. The females tend to focus more on the concept and do not associate with the initial problem. However, they often make mistakes while associating the results with the initial problem. The FD female participant cannot interpret at all due to not understanding the given problems. This is similar to Amstrong, et. al (2011), that FD individual adopts a global orientation to understand and process information, whereas FI adopts an analytical orientation.

In the three stages of the statistical reasoning process, either based on the results of interviews or test results, it was stated that males only write or say things that are considered important. In doing calculations, not all stages are written or explained, and general things that everyone thinks they know are not written or explained. This condition is as stated by Geary, Sault, Liu, and Hoard (2000) and Leder, Fergas, and Jackson (2014) that males are better than females in numeracy, while females are better in reading, writing, spelling, and grammar. Males tend to make interpretations in the form of drawings or sketches, because they use logic more in solving problems such as statistical reasoning problems. The SR questions are non-routine or are seldom given in learning, where the solution requires several concepts and logical thinking. In statistical reasoning, males are better than females, according to Gallagher, Delisi, Holst, McGillicuddy-DeLisi, Morely, and Cahalan (2000) who stated that male preservice teachers are more likely to work correctly in solving non-conventional problems using logical estimation. In addition, males have high self-confidence to master statistics due to their reasoning skills as stated by Yusuf, Suyitno, Sukestiyarno, and Isnarto (2019).

Both participants (FI male and FI female) have the same field independent cognitive style, however, male statistical reasoning process is better. This is in line with Rosidah (2016), Liu and Garfield (2002), and Martin (2013) which suggested that males are better at performing more complex cognitive processes such as statistical reasoning. In terms of mathematical ability, Araiku, Sidabutar, and Mairing (2019) stated that based on Bloom's taxonomy, males are better than females. However, Liu and Garfield (2002) suggested that gender differences in reasoning abilities do not apply to students in the United States. Liu and Garfield (2002), as well as Ghasemi and Burley (2019) stated that socio-cultural factors are one of the reasons for differences in statistical reasoning in gender.

**Conclusion**

This research found that cognitive style affects the statistical reasoning process. Furthermore, the SR of the pre-service teacher with FI cognitive style is better than the FD. The characteristics of FI and FD are different in solving the given problems, and those with FD in the comprehension stage expressed the elements in the problem by using words that copy the questions. At the planning and decision-making stage, the participants cannot determine what to
use in solving the problems, even when they know the concept. Also, at the evaluation and interpretation stage, they cannot interpret and relate to the initial problem because they do not understand the given problems. FI male and FI female have the same good statistical reasoning, but the process is different. The SR process shown by male participants with FI cognitive style is as follows: In the comprehension stage, the subjects can write the elements needed to solve the problem using statistical symbols; At the planning and decision-making stage, they can determine and calculate effectively and efficiently; In terms of evaluation and interpretation, they can represent correctly in the form of schemes or drawings and correlate with the initial problems. Meanwhile, the process of statistical reasoning shown by female subjects with field independent cognitive style is as follows; In the comprehension stage, they can state all the problem elements by using statistical notations; At the planning and decision-making stage, they can determine the concept used correctly, write all the calculation phases, and present them in several ways; At the evaluation and interpretation stage, they can make interpretations in some ways, but they often do not relate to the initial problem.

Despite this research focusing on only two students, it was argued that the findings provide an insightful lens for understanding students’ SR based on their cognitive styles. This is expected to be a reference in designing learning because cognitive style affects statistical reasoning. Also, this research was conducted on preservice mathematics teachers, hence, there was no attention to their numeracy skills. As long as it will be implemented on different participants, it is necessary to pay attention to their numeracy skills because it is one of the factors that influence statistical reasoning ability.

References


Lailiyah, S., Muslimah, N., & Sutini. (2021). Do students with different cognitive styles have similar levels of statistical thinking?. *Beta: Jurnal Tadris Matematika, 14*(1), 15-33. https://doi.org/10.20414/betajtm.v14i1.438


Rahmatulwahidah, N., & Zubainur, C. M. (2017). The analysis of students’ spatial ability at senior high school in banda aceh. *Proceedings of The 7th Annual International Conference (AIC) Syiah Kuala University and The 6th International Conference on Multidisciplinary Research (ICMR) in conjunction with the International Conference on Electrical Engineering and Informatics (ICELTICs) (pp. 18-20).* Banda Aceh, Indonesia.


Rosidah. (2016). Analysis of statistical reasoning process of senior high school students on the size of central tendency (the case study for student's low math ability). *International conference on research, implementation and education of mathematics and science.*


