PAPER • OPEN ACCESS

Statistical reasoning levels and error analysis of prospective mathematics teacher students

To cite this article: A Rozak and Nurwiani 2020 *J. Phys.: Conf. Ser.* **1464** 012033

View the article online for updates and enhancements.

Statistical reasoning levels and error analysis of prospective mathematics teacher students

A Rozak*, and Nurwiani

Departement of Mathematics Education, STKIP PGRI Jombang, East Java, Indonesia.

*abd.rozak76@yahoo.co.id

Abstract. This paper describes the level of statistical reasoning and the difficulties and errors that occur in statistical problem-solving in prospective mathematics teacher students at STKIP PGRI Jombang. This research is a qualitative study with a case study strategy, the case study in the process of solving statistical problems carried out by the subject. The research subjects were the 69 students at third-semester of Mathematics Education Program STKIP PGRI Jombang who had taken descriptive statistics. The results of this study indicate that students are not found at the idiosyncratic level, most of them are at the transitional and quantitative level, while the analytical level is still a minority

1. Introduction
As students who will deal with a variety of data and information are required to be able to manage, analyze, and conclude the data correctly. Therefore students must be able to develop statistical reasoning during lectures. Statistical reasoning defined as the way people reason with statistical ideas and make sense of statistical information [1]. Statistical reasoning is ability in making interpretations based on sets of data, graphical representations, statistical summaries, combines ideas about data and chance, making inferences and interpreting statistical results. Statistical reasoning is a conceptual understanding of important ideas, such as distribution, center, spread, association, uncertainty, randomness, and sampling [2]. The existence of statistical subjects (descriptive statistics and inferential statistics) in the college curriculum is expected to be able to train and develop students' statistical reasoning skills. The benefits of these capabilities in the short term can be used in preparing the final project and in the long term can be used for other studies as well as the basis in making decisions based on data. [2] explained the level of statistical reasoning, 1) Idiosyncratic; 2) Transitional; 3) Quantitative, and 4) Analytical. Several studies related to statistical reasoning ability have been done, [3] stating that the level statistical reasoning of educational study program students and mathematics non-educational study programs is still low in the transitional and quantitative level, as well as junior/high school mathematics teachers who follow PPM Lecturer Lecturer UPI activities in Subang Regency and some participants of the junior high school mathematics teacher certification certification in BMI Lembang obtained information that the average ability of junior/high school teacher statistical Reasoning has only reached 32.15%. [4] stated that students who have low statistical reasoning have a lack of understanding of the notions of statistical objects, an understanding of the benefits or the usefulness of statistics that are quite good. But the attention to statistical subjects is still lacking.

Relating to the low ability of statistical reasoning occurs because of the difficulties and errors of students in understanding the concept of statistics and resulting in errors in solving statistics problems. [5] explains some difficulties or obstacles experienced by someone in solving problems can occur in each stage of problem-solving. When referring to the stages of problem-solving according to Polya, it is explained that: 1) lack of comprehension, 2) lack of strategy knowledge. 3) inability to translate the problem into mathematical form 4) inability to use the correct mathematics, namely the inability to use correct mathematics or difficulty in assessing whether what is done is correct.

Difficulties faced by students have an impact on the occurrence of errors in solving statistical problems. Newman [6] [7] [8] explain several errors that can occur in the stages of the problem-solving process, namely: 1) reading error, which is an error in reading the problem, errors occur in reading words or sentences or symbols contained in the problem. 2) comprehension error, which occurs in the stage of understanding the problem, where someone misunderstands what is needed/known and is asked in the problem. 3) transformation error, occurs in if wrong in interpreting and identifying the appropriate mathematical operations to solve the problem. 4) process skill error, occurs in the stage of implementing the procedure where there is an error in implementing the work procedure or in the calculation. 5) encoding error, occurs when students are wrong in writing the final answer. Previous research only outlined the level of statistical reasoning [3] [9] and errors in solving problems [10] [6] but have not analyzed errors that occur at each level of statistical reasoning ability. Based on the above description, this study aims to identify the level of statistical reasoning and the difficulties and errors that occur in solving statistical problems in prospective mathematics teacher students at STKIP PGRI Jombang.

2. Method

Research is a qualitative study with a case study strategy, in which the researcher carefully investigates a program, event, activity, process, or group of individuals [11]. The case study in question is in the process of solving mathematical problems carried out by the subject, then the researcher explores the process to obtain the research objectives. The research subjects were the third-semester students of Mathematics Education Program STKIP PGRI Jombang who had taken descriptive statistics, as many as 69 students as research subjects. Then the results of the analysis work based on the level of statistical reasoning (**Table 1**) and the error indicator (**Table 2**) are then confirmed through interviews to determine whether students experience difficulties or errors in solving problems.

Table 1. Indicator of the levels statistical reasoning [9].

Level	Indicator	Indicator based problem
Idiosyncratic	Support responses by providing examples from your own experience	Determine the importance of whether or not the time of delay is based on your own experience.
Transitional	Focusing on only one aspect of the data and limited reasoning	Decide based on the average delay or no delay
Quantitative	Making more than one comparison, difficulty to establish relationship	Don't think according to the context of the problem. There is a calculation of the mode, median, average, or standard deviation, but cannot establish a relationship between these measures
Analytical	Establishing relationship	Decide which steps are most suitable for

between the context and the data	interpreting the data and are able to interpret the meaning of this step in the context of the problem.

Table 2. Indicator of the error by Newman [11]

Error type	Indicator
Reading error	Wrong in reading terms, symbols, words or important information in a problem.
Comprehension error	1. Don't know what is actually being asked about the question. 2. Incorrect in capturing information in the problem so that it cannot complete the next process.
Transformation error	1. Failure to change to the correct form of a mathematical model. 2. Incorrect use of calculation operations 3. Wrong in applying procedures
Process Skill error	1. Wrong in calculation or computation. 2. Do not continue the settlement procedure.
Encoding error	1. Do not write the final answer that is asked for the question. 2. Cannot conclude answers according to mathematical sentences. 3. Experiencing errors due to carelessness or inaccuracy

The instruments in this study consisted of the main instruments and supporting instruments. The main instrument is the researcher because it is directly related to the subject of research and cannot be represented, as the main instrument of the role of the researcher as a planner to determine the focus of research, select research subjects, collect data, analyze data, interpret data to report research results. While the supporting instruments consist of questions of statistical problems by adapting the MEA "On-time Arrival" with some adjustments related to the context of the problem to be contextual. While the interview guideline instrument is used to confirm events or events on a particular subject that indicated difficulties and errors in solving the problem. Data analysis techniques used are data reduction, data presentation, and conclusion drawing.

3. Result and Discussion

The work results of students are analyzed and grouped based on the level of statistical reasoning, obtained as the table below:

Table 3. Statistical reasoning levels analysis of prospective mathematics teacher students

Level	Amount	%
Idiosyncratic	0	0
Transitional	45	65
Quantitative	20	29
Analytical	4	6
Total	69	100

3.1. Idiosyncratic Level

Based on the data in Table 3 above, there were no students with the Idiosyncratic level.

3.2. Transitional Level

In the study it was found that students at the transitional level dominated as many as 45 students, most students only used one statistical measure in decision making, the average time of delay, or the number of delay. The following are examples of student work presented in Figure 1 below:
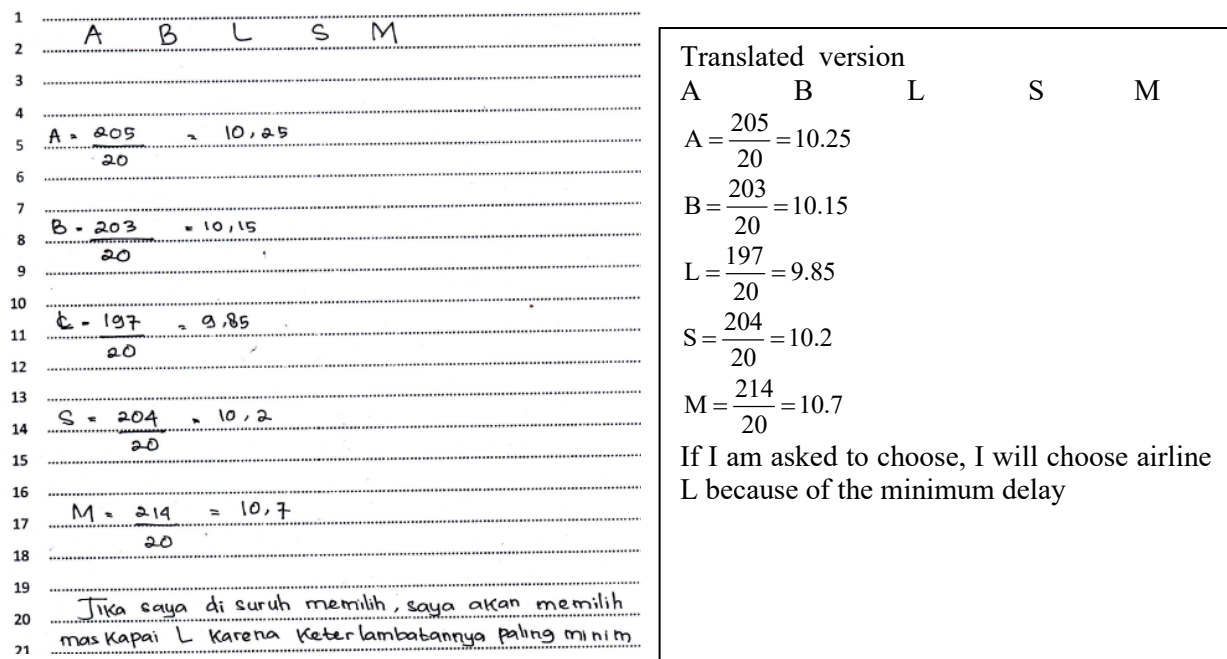

Figure 1. Examples of student work at the transitional level

Based on the results of student work, at the transitional level students focus on only one measure of the data, namely the average time of delay (**Figure 1** steps 5-17), then decide on the selected airline based only on the smallest average time delay and not linking to other statistical measures. According to [2], students on this level of start to use quantitative information in statistical problems, therefore starting to show quantitative reasoning; however, they do this in a limited. This study mentions that students have a belief that is only suitable for a central tendency. The measure for a data set is the arithmetic mean. At this transitional level, there is also a fundamental error in understanding the information or data contained in the problem, where students only understand the problem and relate it to just one statistical measure.

3.3. Quantitative Level

At the quantitative level, students use more than one statistical measure in decision making, for example, the average time delay and the number of delays, the average delay time and the maximum/minimum delay, the average time delay and the standard deviation. But students cannot connect the two statistical measures in decision making, along with examples of student work at the quantitative level.

Figure 2. Examples of student work at the quantitative level

Students in the quantitative level can associate various measures in statistics, in this case, the average delay (**Figure 2** step 3-8) and the number of flight delays (**Figure 2** steps 11-16) that occur within 20 days. Students can use more than one statistical measure and combine it in decision making (**Figure 2** steps 19-25). However, errors occur in the form of Transformation errors, where errors occur in implementing decision-making procedures, namely when multiplying between the average value of delay with the number of delays. The concept of this calculation is not statistical procedures (**Figure 2** steps 19-25). Judging from the difficulties experienced by students in the form of lack of strategy knowledge, where students have difficulty in choosing the strategies used in decision making.

3.4. Analytical Level

At this level students use more than one statistical measure, for example the average time delay and the number of delays, the average delay time and the maximum / minimum delay, the average time delay and the standard deviation and connect the two or more statistical measures in the retrieval decision, here is an example of student work at the analytical level.

Figure 3. Examples of student work at the Analytical level

Translated version

Look for the airline that is the longest delay, airline A = 90 minutes.
Look for airlines that are rarely delayed, i.e. airline A and airline S = 8 times.

$$xA = \frac{10+5+...+0}{20} = \frac{205}{20} = 10.25$$

$$xB = \frac{16+10+...+0}{20} = \frac{203}{20} = 10.25$$

$$xL = \frac{10+15+...+0}{20} = \frac{197}{20} = 9.85$$

$$xS = \frac{0+25+...+5}{20} = \frac{204}{20} = 10.2$$

$$xM = \frac{10+5+...+0}{20} = \frac{224}{20} = 11.20$$

the description above is a step to find the average delay, it was found that the airline that was the longest delay was airline M for 11.2 minutes, then the least was airline L for 9.85 minutes.

when referring to the average delay, I will choose n Airlines, but on day 4 I have experienced a delay of at most 50 minutes. My conclusion is based on a small delay even though it is often late, namely airline B, because the average delay is not too long and is the longest delay of 25 minutes.

Based on the work in Figure 3, it shows that students can link problem-solving with 3 statistical measures at once, the longest delay (**Figure 3** Step 1), the most delay (**Figure 3** Step 2) and the average delay (**Figure 3** Step 3-11). Students make decisions with the most appropriate steps in interpreting the data, namely by the combination of the smallest possible odds of delay. In this Analytical level, there are no difficulties and errors in solving statistical problems. According to Jones et al [2], students on this level are the best measure to interpret the data and are able to interpret the meaning of this measure in the context of the problem. This level, the student uses inference based on data and uses both quantitative and contextual knowledge.

[9] states that statistical reasoning process consists of 1) Describing Data Displays; 2) Organizing and Reducing Data; 3) Data Representing; and 4) Analyzing and Interpreting data, the level of statistical reasoning is in line with the statistical reasoning process that occurs. Students who are at the highest level can be interpreted as not experiencing difficulties and errors

4. Conclusion

Based on the description above, the researcher concluded that in the third semester students the prospective mathematics teacher had a tendency to better understand statistical measures related to statistical problems, but the skills in building relationships between existing measures needed to be improved. Students are still accustomed to routine statistical problems, but are not familiar with non-routine or complex problems in the form of Model Eliciting Activities. Achievement of the level of statistical reasoning with difficulties and errors that occur in each process.

References

[1]   Garfield J B, Ben-Zvi D, Chance B, Medina E, Roseth C and Zieffler A 2008 Developing students' statistical reasoning: Connecting research and teaching practice *Dev. Students' Stat. Reason. Connect. Res. Teach. Pract.* 1–408

[2]   Conway B, Gary Martin W, Strutchens M, Kraska M and Huang H 2019 The Statistical Reasoning Learning Environment: A Comparison of Students' Statistical Reasoning Ability *J. Stat. Educ.* **0** 1–34

[3]   priatna martadiputra B 2002 Kajian Tentang Kemampuan Melek Statistis (Statistical Literacy), Penalaran Statistical (Statistical Reasoning), dan Berpikir Statistis (Statistical Thinking) Guru SMP/SMA *Kemamp. Melek Stat. Penal. Stat. Dan Berpikir Stat.* 1–7

[4]   Ilham Minggi M V W 2015 Deskripsi Persepsi Tentang Statistika Ditinjau Dari Tingkat Kemampuan Berpikir Statistis (Statistical Thinking) Pada Mahasiswa Jurusan Pendidikan Ips Terpadu Fakultas Ilmu Sosial Universitas Negeri Makassar *J. Daya Mat.* **3** 70

[5]   Nikmah I L, Juandi D and Prabawanto S 2019 Students' difficulties on solving mathematical problem based on ESD objectives *J. Phys. Conf. Ser.* **1157**

[6]   Singh P, Rahman A A and Hoon T S 2010 The Newman procedure for analyzing Primary Four pupils errors on written mathematical tasks: A Malaysian perspective *Procedia - Soc. Behav. Sci.* **8** 264–71

[7]   Prakitipong N and Nakamura S 2006 Analysis of Mathematics Performance of Grade Five Students in Thailand Using Newman Procedure *J. Int. Coop. Educ.* **9** 111–22

[8]   Abdullah A H, Abidin N L Z and Ali M 2015 Analysis of students' errors in solving Higher Order Thinking Skills (HOTS) problems for the topic of fraction *Asian Soc. Sci.* **11** 133–42

[9]   Saputra D S A 2017 Инновационные подходы к обеспечению качества в здравоохраненииNo Title *Вестник Росздравнадзора* **6** 5–9

[10]   Oktaviana D 2018 Analisis Tipe Kesalahan Berdasarkan Teori Newman Dalam Menyelesaikan Soal Cerita Pada Mata Kuliah Matematika Diskrit *Edu Sains J. Pendidik. Sains Mat.* **5** 22

[11]   Creswell 2015 Perencanaan, Pelaksanaan, dan Evaluasi Riset Kualitatif & Kuantitaif