

Students' Statistical Reasoning in Statistics Method Course

Rohana¹, Yunika Lestaria Ningsih²

^{1, 2}Mathematics Education Study Program, Universitas PGRI Palembang, Jl. Jend. A. Yani Jl. Jend. 9/10 Ulu Palembang
Email : yunikalestari@univpgri-palembang.ac.id

Abstract

The role of statistics is wide and crucial in daily life, making statistics important. Many students have difficulty understanding statistics. This study aims to determine students' statistical reasoning about inference statistics, which is limited to the subject matter of the testing hypotheses about two-sample hypotheses testing. This study used descriptive research method. The subjects were 25 students of third-year Mathematics Education Department at Universitas PGRI Palembang in the academic year 2018/2019. Data were collected through tests and interviews. Data were analyzed through descriptive quantitative. The results of data analysis showed that 32% of students had level 1 statistical reasoning (the lowest level), 20% were at level 2, 28% at level 3, 12% at level 4 and 8% at level 5 (highest level). Based on the result, it can conclude that students' statistical reasoning ability in learning statistical method is not satisfactory, students are still very lacking in reasoning.

Keywords: Inference Statistics, Statistical Reasoning, Testing Hypotheses

Abstrak

Peranan statistika yang luas dan nyata dalam kehidupan sehari-hari, menjadikan statistika penting untuk dipelajari secara bermakna oleh mahasiswa. Namun untuk dapat memahami konsep statistika bukan merupakan hal yang mudah bagi mahasiswa. Banyak mahasiswa yang kesulitan dalam memahami statistika. Penelitian ini bertujuan untuk mengetahui penalaran statistis mahasiswa dalam statistika inferensia, yang dibatasi pada pokok bahasan uji hipotesis dua rata-rata untuk sampel bebas. Penelitian ini menggunakan metode penelitian deskriptif. Subjek penelitian adalah mahasiswa tahun ketiga Program Studi Pendidikan Matematika FKIP Universitas PGRI Palembang tahun akademik 2018/2019 yang berjumlah 25 orang. Data penelitian dikumpulkan melalui tes dan wawancara. Data dianalisis secara deskriptif kuantitatif. Hasil analisis data menunjukkan bahwa 32% mahasiswa memiliki penalaran statistis level 1 (level terendah), 20% berada pada level 2, 28% pada level 3, 12% pada level 4 dan 8% pada level 5 (level tertinggi). Berdasarkan hasil penelitian dapat disimpulkan bahwa penalaran mahasiswa dalam pembelajaran metode statistika belum memuaskan, banyak mahasiswa yang kesulitan dalam melakukan penalaran statistik.

Kata kunci: Penalaran Statistis, Statistik Inferensi, Uji Hipotesis 2 Rata-Rata

Cara Menulis Sitasi: Rohana, & Ningsih, Y. L. (2020). Students' statistical reasoning in statistics method course. *Jurnal Pendidikan Matematika*, 14(1), 81-90. <https://doi.org/10.22342/jpm.14.1.6732.81-90>.

INTRODUCTION

One of the abilities that must be possessed by students to compete against the challenges nowadays is reasoning ability. Muller & Maher (Rahmawati, Mardiyana, & Triyanto, 2018) stated that reasoning as a process for making a conclusion base on evidence or other assumptions. This reasoning is useful for a person in the process of building and comparing ideas from various situations faced, so that he can make the right decision in solving his life's problems. The learning process in higher education can develop students' reasoning abilities.

Statistics Method is a subject that must be attended by students in higher education. Coladarci, Cobb, Minium, & Clarke (2011) define that "statistics merely formalizes what humans do every day". Statistics can be seen as a knowledge that provides a means to be able to provide solutions to

phenomena or problems that occur in life, in the work environment, and in science itself. Therefore, based on the important role of statistics both in real life and research, students must learn statistics meaningfully (Lanani, 2014).

Students have difficulty understanding statistics (Garfield & Ben-Zvi, 2008; Chan & Ismail, 2012). In line with this, Whitaker, Foti, & Jacobbe (2015), revealed that the most difficulty faced by students was in interpreting statistical results. To overcome these difficulties, statistical capabilities such as the ability to read information, conduct analysis, draw conclusions and connect the results obtained with the problem are needed. This ability is known as statistical reasoning.

Statistical reasoning is how a person uses his mind to reason and understand the information in statistics (Garfield & Ben-Zvi, 2008). It involves many statistical concepts such as combining data concepts and opportunities. Martin (Chan & Ismail, 2013) defines it as forming conclusions and decision making based on data obtained from observations, experiments and surveys. Furthermore, Régniera & Kuznetsova, (2014) consider it one of the essential purposes of learning statistics.

Statistical reasoning according to Garfield (2002) is divided into 5 levels, namely: (1) idiosyncratic reasoning, this reasoning is level 1 reasoning, students can use some symbols in statistics but cannot comprehend them in full and relate them into information provided, (2) verbal reasoning, is level 2 reasoning, students can know the definitions and meanings of some statistical symbols but still fail to apply them, (3) transitional reasoning, is level 3 reasoning, students can understand several aspects of the statistical process, but they fail to apply the concept to find the answers, (4) procedural reasoning, is level 4 reasoning, students can identify statistical processes accurately, but they cannot interpret and understand them, and (5) integrated process reasoning is level 5 reasoning, students can understand the statistical process correctly and can explain the process.

According to the explanation above, it is known that statistical reasoning is a crucial thing. The higher the level of students' reasoning, the faster students can achieve the goal learning (Hasanah, Tafrilyanto, & Aini, 2018). Lanani (2014) added that the formation of good statistical reasoning skills in students makes students understand the concept of statistics well so that they can solve statistical problems and appreciate statistics in daily life. Therefore, it is important for paying attention to what level of statistical reasoning achieved by the students.

This study purposes to determine the level of statistical reasoning of third-year students of the Mathematics Education Department at Universitas PGRI Palembang. The level of students' statistical reasoning is evaluated based on Garfield (2002). The subject matter is limited in statistical inference: testing the hypotheses about two independent means.

METHODS

This research is a descriptive study, which aims to describe the level of student's statistical reasoning in the statistic method course. The subjects were 25 students from the third-year Mathematics

Education Department at Universitas PGRI Palembang in the academic year 2018/2019. Data collection techniques in this study were tests and interviews. The test was conducted in January 2019. The test submitted can be seen in Figure 1.

<p>Mahasiswa Prodi Pendidikan Matematika yang berjumlah 10 orang mengikuti mata kuliah Aljabar (X) dan Kalkulus (Y). Apabila hasil tes akhir perkuliahan sebagai berikut :</p> <p>X : 70 38 46 57 68 34 80 57 42 65</p> <p>Y : 75 48 95 41 52 69 70 85 64 76</p> <p>Ujilah apakah reratanya berbeda secara signifikan, gunakan $\alpha=0.05$.</p>
<p>10 Mathematics Education students attend Algebra (X) and Calculus (Y) courses. If the final test results are as follows:</p> <p>X : 70 38 46 57 68 34 80 57 42 65</p> <p>Y : 75 48 95 41 52 69 70 85 64 76</p> <p>Test whether the mean is significantly different, use $\alpha=0.05$</p>

Figure 1. The problem submitted of testing the hypotheses two independent means

The steps for solving the problem are: 1) formulate the statistical hypotheses and select a level of significance; 2) determine the desired sample size and select the sample; 3) calculate the necessary sample statistics; 4) identify the region(s) of rejection; 5) make the statistical decision and form conclusion (Coladarci, Cobb, Minium, & Clarke, 2011). The test was analyzed descriptively. Students who answer completely and correctly are categorized into level 5 statistical reasoning (integrated process reasoning), meaning that students can understand well and correctly the statistical process of testing the hypotheses about two independent means, and can explain the process. Meanwhile, the interview is conducted after the written test. Researches chose the students from every level to follow the interview. It aims to explore further the students' statistical reasoning, including also knowing the difficulties and the factors causing them.

RESULTS AND DISCUSSION

Data obtained during the study in the form of written test results about students' statistical reasoning abilities and interview results. Data were analyzed to determine the level of statistical reasoning ability based on Garfield (2002) and identify students' misconceptions in statistical reasoning.

Levels and Indicators of Statistical Reasoning

Levels and its indicators which are used in this study can be seen in Table 1.

Table 1. Levels and indicators of statistical reasoning

Level of statistical reasoning	Indicators
Level 1: Idiosyncratic reasoning	Students use several statistical symbols such as mean and standard deviation, but cannot comprehend completely and cannot relate it to the information provided to solve the problem.

Level of statistical reasoning	Indicators
Level 2 : Verbal reasoning	Students can find out the definitions and meanings of some statistical symbols but still fail to apply them.
Level 3 : Transitional reasoning	Students understand several aspects of the statistical process, but they fail to apply the concept to find the solution.
Level 4 : Procedural reasoning	Students can identify statistical processes accurately, but they cannot interpret and understand them.
Level 5 : Integrated process reasoning	Students can understand well and correctly the statistical process and can explain the process.

The Students' Statistical Reasoning

The level of students' statistical reasoning based on test results can be seen in Table 2.

Table 2. The distribution of students' level of statistical reasoning

Level of statistical reasoning	Frequency	Percentage
Level 1	8	32%
Level 2	5	20%
Level 3	7	28%
Level 4	3	12%
Level 5	2	8%
Total	25	100%

Table 2 shows that there are 8 people or 32% students in the lowest level of statistical reasoning, 5 people or 20% students in level 2 (verbal reasoning), 7 people or equal to 28% students in level 3 (transitional reasoning), 3 people or 12% students in level 4 (procedural reasoning), 2 people or 8% students are in level 5 (integrated process reasoning) which is the highest level.

The student statistical reasoning based on Garfield's level is described as follows:

Idiosyncratic Reasoning

Idiosyncratic reasoning is indicated by the ability of students to use some statistical symbols but cannot comprehend completely and cannot relate it to the information provided. The first step used to conduct a hypotheses test is to determine the hypotheses formula based on the problem. Students at this level, only pay attention to the statistical calculations needed in answering questions. They did not state the null-hypotheses and the alternative hypotheses. They directly display a calculation table to find the mean value and standard deviation. In other words, it can be concluded that students cannot understand and associate statistical symbols with the information provided.

There are 8 students with statistical reasoning at this level. After conducting interviews with some students at this level, errors and weaknesses in statistical reasoning are caused by students' lack of understanding of the testing hypotheses. Students also did not understand the purpose of the problem (Diniyah, Akbar, Akbar, Nurjaman, & Bernard, 2018). Furthermore, Link (Krishnan & Idris, 2015) said students have trouble when stating the hypotheses. Students did not understand the

definitions of hypotheses (Sotos, Vanhoof, Noortgate, & Onghena, 2009). Based on test results in this study, students did not understand how to formulate the hypotheses; they only remember how to count the mean value and standard deviation. This is in line with Batanero & Diaz (Krishnan & Idris, 2014; Dolor & Noll, 2015) who states that students have difficulty learning hypothesis testing because this test involves many statistical concepts.

Verbal reasoning: At this level, students can find out the definitions and meanings of some statistical symbols but still fail to apply them. Students at this level already understand the meaning of the mean symbol, but students still fail to apply the mean image for the two groups. Students write symbols that state the mean data of group X with the symbol $\bar{X}x$ (should be written \bar{X}) and the symbol $\bar{X}y$ which states the mean data of group Y (should be written \bar{Y}). The concept of means is part of a measure of centralization of data, the difficulty of students in understanding this concept was also put forward by Chan, Ismail, & Sumintono (2016). The mean concept and other measures of centralization are important concepts that must be understood by students to mastering the testing hypothesis (Reaburn, 2011).

Transitional reasoning

This level of reasoning is indicated by the ability of students to understand several aspects of the statistical process, but they fail to apply the concept to find answers. Students at this level already understand the statistical symbols used in determining the mean of the two groups, the data is also described correctly. The student test result to determine the mean of two groups can be seen in Figure 2.

2) Menentukan Nilai Rata-rata

$$\bar{X} = \frac{\sum x}{N} = \frac{557}{10} = 55,7$$

$$\bar{Y} = \frac{\sum Y}{N} = \frac{675}{10} = 67,5$$

Determine the mean value:

$$\bar{X} = \frac{\sum x}{N} = \frac{557}{10} = 55,7$$

$$\bar{Y} = \frac{\sum Y}{N} = \frac{675}{10} = 67,5$$

Figure 2. Students' test result to determine the mean of two groups

They use these means in testing the hypotheses. It means that students understood some of the statistical processes. However, the steps that they made are incomplete. They didn't make the step 4 and 5. The analysis results show that there are 7 out of 25 students or 28% at this level. Based on the results of interviews it is known that the difficulty of students in making conclusions from testing

hypotheses is due to students' lack to interpret statistical test results in real life. This finding is in line with Rosidah (Nisa, Zulkardi, & Susanti, 2019). Most of them difficult to relating the test result to reality (Canadas, Batanero, Diaz, & Roa, 2012). This can be interpreted that the students' statistical reasoning in understanding the problem has been running but not completely (Yusuf, 2017).

Procedural reasoning: Reasoning at this level is indicated by the ability of students to identify statistical processes accurately, but they cannot interpret and understand them. The steps at this level were close to true. Students make hypotheses formulation under the problem. It can be seen in Figure 3.

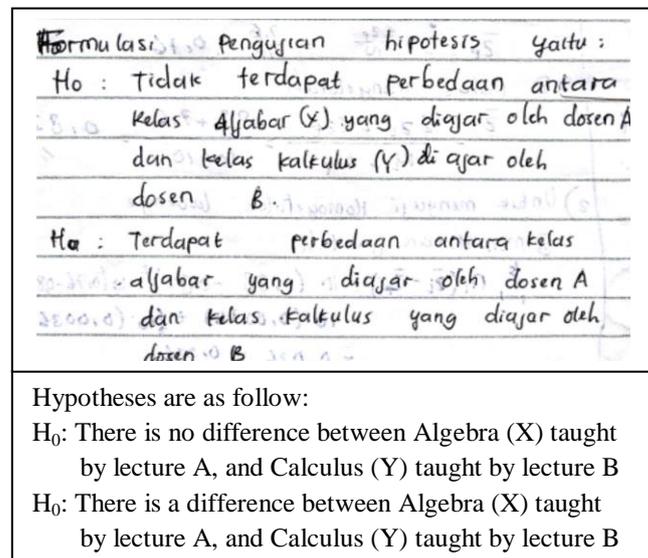


Figure 3. Student's test result to determine the hypotheses

From Figure 3, students made a mistake in determining the region of rejection. This shows that students can identify statistical processes, but students' understanding of this problem is not comprehensive (Garfield, 2002). There are 3 out of 25 students or 12% have entered level 4. After further interviews, it is known that they already understood the steps for testing hypotheses but confused in determining the region of rejection, especially in identifying the critical t value. They made an incomplete step for testing the hypotheses (Krishnan & Idris, 2014). In other words according to (Yusuf, 2017), students at this level know a concept to solve problems but not yet fully integrated.

Integrated Process Reasoning

Students can understand and explain the statistical process in testing hypotheses about two independent means. They could solve the problems correctly and completely. The students' statistical reasoning has running completely. Student's worksheet for step 1-5 can be seen in Figure 4.

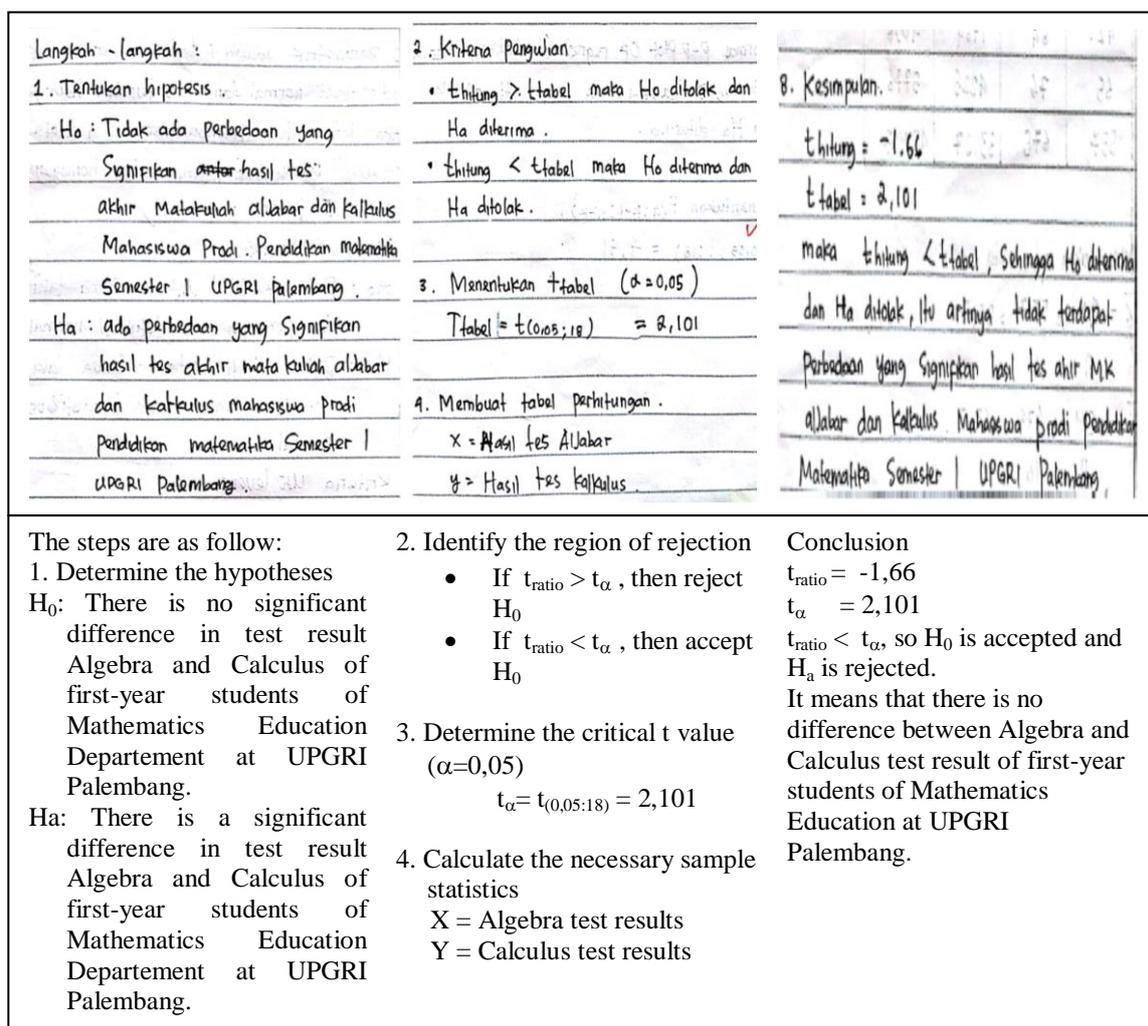


Figure 4. Student's worksheet for step 1-5

Based on the Figure 4, students wrote a hypothesis formulation as an initial step of testing hypothesis. Make test criteria, determine t-table values, perform t-test statistical calculations, and draw conclusions. Every step made by students in testing this hypothesis is correct; this shows that students can understand the statistical process in solving the problem about two-sample hypothesis testing. Based on further interviewed students at this level also could explain every step of hypothesis testing. There are 2 out of 25 students or 8% at this level, it's means that the hypotheses testing is difficult for students to understand. It is in line with the previous result of Stalveya, et al., (2019). There are only 2 students have entered the highest level of statistical reasoning in this study. At this level, they use the statistics concept to analyze data and solve problems (Yusuf, 2017).

CONCLUSION

Based on the results, it can be concluded that the students' statistical reasoning in learning the statistical method is 8 students or 32% have level 1 statistical reasoning (idiosyncratic reasoning)

which is the lowest level. The students at this level know some statistical words and symbols, such as mean and standard deviation, but they incorrectly in using these symbols. The next level is verbal reasoning, 5 students or equal to 20% have entered this level. Students at this level understood the statistical symbols such as mean and standard deviation, but they fail to apply these symbols for two groups. Level 3 of statistical reasoning is transitional reasoning. There are 7 students or 28% at this level. Students in this level show that they understood the meaning of symbols in statistics such as mean and standard deviation and can use it correctly, but they made mistake in determining the region of rejection. There are 3 students or 12% have entered level 4, procedural reasoning. Students at this level could solve the problem of testing hypotheses. However, they made incomplete step for testing hypotheses. The highest level of statistical reasoning is integrated process reasoning. There are 2 students or 8% at this level. Students at this level could solve and explain the steps for testing hypotheses testing correctly. The results of this data analysis show that the students' statistical reasoning ability in learning statistical method is not satisfactory. Students are still very lacking in reasoning. Therefore, a suggestion that can be submitted for the next researcher is an action to improve students' statistical reasoning abilities.

ACKNOWLEDGMENTS

Thank to the Rector of Universitas PGRI Palembang, the Head of LPPKM of Universitas PGRI Palembang, Dean of FKIP of Universitas PGRI Palembang, and Chair of Mathematics Education Department of Universitas PGRI Palembang. We are also grateful to our student participants, without whom this research would not be possible.

REFERENCES

- Diniyah, A.N., Akbar, G.A., Akbar, P., Nurjaman, A., & Bernard, M. (2018). Analysis of senior high school students' reasoning ability and self confidence in probability [in Bahasa]. *Journal on Education, 1*(1), 14-21.
- Canadas, G.R., Batanero, C., Diaz, C., & Roa, R. (2012). Psychology students' understanding of the chi-squared test. *Statistique et Enseignement, 3*(1), 3-18.
- Chan, S.W., & Ismail, Z. (2012). The role of information technology in developing students' statistical reasoning. *Procedia-Social and Behavioral Science, 46*, 3660-3664. <https://doi.org/10.1016/j.sbspro.2012.06.123>.
- Chan, S.W., & Ismail, Z. (2013). Developing statistical reasoning assessment instrument for high school students in descriptive statistics. *Procedia-Social and Behavioral Sciences, 116*, 4338-4343. <https://doi.org/10.1016/j.sbspro.2014.01.943>.
- Chan, S.W., Ismail, Z., & Sumintono, B. (2016). A framework for assessing high school students' statistical reasoning. *PLoS ONE, 11*(11), 1-32. <https://doi.org/10.1371/journal.pone.0163846>.

- Coladarci, T., Cobb, C.D., Minium, E.W., & Clarke, R.B. (2011). *Fundamentals of Statistical Reasoning in Education*. USA: John Wiley & Son, Inc.
- Dolor, J., & Noll, J. (2015). Using guided reinvention to develop teachers' understanding of hypotheses testing concepts. *Statistics Education Research Journal*, 14(1), 60-89.
- Garfield, J. B. (2002). The challenge of developing statistical reasoning. *Journal of Statistics Education*, 10(3). <https://doi.org/10.1080/10691898.2002.11910676>.
- Garfield, J. B., & Ben-Zvi, D. (2008). *Developing Students' Statistical Reasoning*. USA: Springer.
- Hasanah, S.I., Tafrilyanto, C.F., & Aini, Y. (2018). Mathematical reasoning: The characteristics of students' mathematical abilities in problem solving. *Journal of Physics: Conference Series*, 1188, 1-8. <https://doi.org/10.1088/1742-6596/1188/1/012057>.
- Krishnan, S., & Idris, N. (2014). Students' misconceptions about hypothesis test. *Journal of Research in Mathematics Education*, 3(3), 276-293. <https://doi.org/10.17583/redimat.2014.1140>.
- Krishnan, S., & Idris, N. (2015). An overview of students' learning problems in hypothesis testing. *Jurnal Pendidikan Malaysia*, 40(2), 193-196. <https://doi.org/10.17576/JPEN-2015-4002-12>.
- Lanani, K. (2014). Developing of ICT project-based teaching material and research instrument to enhance the students' statistical reasoning-ability, statistical communication and academic help-seeking [in Bahasa]. *Delta-pi: Jurnal Matematika dan Pendidikan Matematika*, 3(2), 1-12.
- Nisa, S., Zulkardi, & Susanti, E. (2019). Students' statistical reasoning ability on histogram data presentation materials through PMRI learning [in Bahasa]. *Jurnal Pendidikan Matematika*, 13(1), 21-40. <https://doi.org/10.22342/jpm.13.1.5460.21-40>.
- Rahmawati, R., Mardiyana, & Triyanto. (2018). Analysis of students' mathematical reasoning ability in solving mathematics problem. *Proc. The International Conference on Teacher Training Educaton (ICTTE) 2018*, 262, 311-314. <https://dx.doi.org/10.2991/ictte-18.2018.57>.
- Reaburn, R. (2011). Students' Understanding of Statistical Inference: Implications for Teaching. *Dissertation*. Australia: University of Tasmania.
- Régniera, J.-C., & Kuznetsova, E. (2014). Teaching of statistics: Formation of statistical reasoning. *Procedia-Social and Behavioral Sciences*, 154, 99-103. <https://doi.org/10.1016/j.sbspro.2014.10.119>.
- Sotos, A.E.C, Vanhoof, S., Noortgate, W.V., & Onghena, P. (2009). How confident are students in their misconceptions about hypotheses tests?. *Journal of Statistics Education*, 17(2). <https://doi.org/10.1080/10691898.2009.11889514>.
- Stalveya, H.E., Burns-Childersb, A., Chamberlain, D., Kemp, A., Meadows, L.J., & Vidakovic, D. (2019). Students' understanding of the concepts involved in one-sample hypothesis testing. *The Journal of Mathematical Behavior*, 53, 42-64. <https://doi.org/10.1016/j.jmathb.2018.03.011>.
- Whitaker, D., Foti, S., & Jacobbe, T. (2015). The levels of conceptual understanding in statistics (LOCUS) project: Results of the pilot study. *Numeracy: Advancing Education in Quantitative Literacy*, 8(2), 1-16. <http://dx.doi.org/10.5038/1936-4660.8.2.3>.

Yusuf, Y. (2017). Construction of statistical reasoning in research statistics [in Bahasa]. *Scholaria: Jurnal Pendidikan dan Kebudayaan*, 7(1), 60-69.
<https://doi.org/10.24246/j.scholaria.2017.v7.i1.p60-69>.