Advances in Social Science, Education and Humanities Research, volume 504
Proceedings of the 2nd International Conference Innovation in Education (ICoIE 2020)

A Local Instructional Theory to Learn the Concept of Test Criteria in Hypothesis Testing Based on Realistic Mathematics Education

Syafriandi[1,*] A Fauzan[2] Lufri[3] Armiati[4]

[1,2,3,4]*Doctoral Program of Education, Universitas Negeri Padang*
[*]*Corresponding author. Email: syafriandi_math@fmipa.unp.ac.id*

ABSTRACT
Determining the test criteria in hypothesis testing is a stage to determine whether the null hypothesis is accepted or rejected. In order for students to understand the concept of test criteria properly, it is necessary to design a learning trajectory that can facilitate students to reinvent the concept of test criteria in hypothesis testing. The purpose of this study was to produce a local instructional theory (LIT) for the topic of test criteria in hypothesis testing based on the Realistic Mathematics Education (RME) approach that is valid, practice, and effective. The research method used is a design of research with the Gravemeijer & Cobb models. This model is a cyclic process of preparing for the experiment, conducting the experiment, and retrospective analysis. Research data were collected through observation, interviews, checklists, analysis of student work results, and test. The result of this study was a learning trajectory for the topic of test criteria in hypothesis testing, which is valid because it meets the criteria for the state of the art RME, content, and construction. Practice, because the learning trajectory was useable, easy to use, and has the best appeal. The result of the learning trajectory was also effective in improving students' statistical reasoning abilities.

Keywords: *Local instructional theory, Learning trajectory, Realistic mathematics education, Statistical reasoning.*

1. INTRODUCTION

Hypothesis testing is one of the important topics in inferential statistics. Inferential statistics is the study of how to conclude conclusions about a population based on information from a sample [1]. One of the important concepts in hypothesis testing is the concept of test criteria, namely determining the critical value, the area of rejection, and the area of acceptance of the null hypothesis. This test criterion can be used to make a decision whether the null hypothesis can be accepted or rejected based on the sample data that owned. As a mathematics student, studying the test criteria in the testing hypothesis is not only about learning how to determine the test criteria, but the most important thing is to learn the theoretical concept of the test criteria. It is in accordance with the demands of the Indonesian National Qualifications Framework (KKNI), that undergraduate education is at Level 6 qualification level, which must meet the qualifications, including mastering the theoretical concepts of certain areas of knowledge in general and the theoretical concepts of a special section in the field of knowledge in depth, and being able to formulate procedural problem solving [2]. The Guidelines for Assessment and Instruction in Statistics Education (GAISE) in [3] describe the purpose of an introductory statistics course, which one of the topics is the test criteria in hypothesis testing, that is to focus students more on conceptual understanding and attainment of literacy and statistical thinking, rather than just learn a set of tools and procedures. [4] also emphasized that studying hypothesis testing requires two important things at once, namely the theoretical level and the applicative level.

The reality in the class shows that learning hypothesis testing topics are focused on application procedures more, while basic concepts at the theoretical level used to ignored. Learning the topic of hypothesys testing has been conventional in which focuses on calculation procedures without strengthening conceptual knowledge simultaneously. Students fail to provide reasons for the answers they get. Not only that, there are many misunderstandings in interpreting the test results obtained

[5], [6]. Various difficulties and misunderstandings experienced by students in studying the topic of hypothesis testing are difficulties identifying null and alternative hypotheses, misinterpreting the level of significance and p-value, and failing to understand the role of the sampling distribution in hypothesis testing [7], [8], [9]. The official statement of the American Statistical Association (ASA) regarding the misinterpretation of p-values, including: (1) the p-value can indicate how incompatible the data is with the specified statistical model; (2) the p-value does not measure the probability that the hypothesis under study is correct or the probability that the resulting data is a coincidence; and (3) p-value does not provide a good measure of evidence regarding the hypothesis [10]. The researcher's experience of teaching the topic of test criteria in hypothesis testing also found that students only memorized standard procedures which founded in many introductory statistics books without understanding the meaning of what was obtained. Students cannot give reasons for the answers given.

Learning that only focuses on calculation procedures without properly understanding the theoretical concepts that following it, is not able to train students to think statistically. Statistical thinking is a form of reasoning which a kind of ability that students must have. Students who have good reasoning skills will be able to think logic, analytic, systematic, critic, and creative to produce accountable conclusions. This reasoning according to [11] is called statistical reasoning, namely the way people think using statistical ideas and make statistical information more meaningful. To practice statistical reasoning [12] suggests statistics teachers do it in stages starting with simple contextual problems that students can understand to get to formal mathematical concepts.

Responding to the problems stated above, it is believed that the Realistic Mathematics Education (RME) approach has the potential to overcome these problems. Realistic Mathematics Education is a learning approach that emphasizes that mathematics is a human activity and learning mathematics is doing mathematics [13], [14], [15]. The main characteristic of the RME approach is learning mathematics begins by examining "realistic" contextual problems (can be understood or imagined by students) and gradually students are facilitated to develop formal mathematical concepts, namely doing horizontal and vertical mathematization. Furthermore, [15], [16] revealed three main principles of the RME approach, namely: guided reinvention and progressive mathematizing, didactical phenomenology, and self-development models. In order for the learning design to meet the RME principles, [15], [16] reveal five characteristics of the RME approach, namely: (1) phenomenological exploration, students explore a real phenomenon that will be manipulated in the lesson; (2) bridging by vertical instrument, bridging from a concrete level to a more formal level using models and symbols; (3) student contributions, contributions of students in learning activities based on their production and construction; (4) interactivity, the learning process is not only an individual activity but also a social process through class discussions and group work that provides opportunities for students to share their strategies and findings with others, to improve their strategies in finding concepts; and (5) intertwining, integrating various mathematics topics in one learning activity.

Utilizing the RME approach, students are facilitated to reinvent the concept of test criteria in hypothesis testing as previously discovered by experts. For this reason, it is necessary to design a learning trajectory that can facilitate students to carry out horizontal and vertical mathematization which is a characteristic of RME. The learning trajectory designed is packaged in the form of a hypothetical learning trajectory (HLT). HLT is useful for guiding the learning experiment process to fit the material specifications and learning hypotheses that have been designed. [17] revealed that HLT consists of three components, namely learning goals, which define the direction of learning, learning activities, and hypothetical learning processes to predict how students' thoughts and understanding will develop in the context of learning activities. Through a cyclic process, HLT was tried out and finally, a local instructional theory (LIT) was obtained for the topic of test criteria in hypothesis testing. The resulting local instructional theory describes how the topic of test criteria in hypothesis testing can be taught based on the RME principle so that it is by student characteristics and learning objectives [18]. The purpose of this study is to produce a local instructional theory (LIT) for the topic of testing criteria in hypothesis testing based on the Realistic Mathematics Education (RME) approach that is valid, practice, and effective.

2. METHOD

The type of research used in this research is design research type validation study, which is a type of design research in the form of educational intervention studies (such as learning processes, learning environments, and the like) to develop or validate theories about the process and how to design it [19] . The resulting development product is a local instructional theory (LIT) for the topic of test criteria in hypothesis testing, which begins with designing a hypothetical learning trajectory (HLT). The development model used is the Gravemeijer and Cobb models, which consist of three phases, namely preparing for the experiment, experimenting in the classroom, and retrospective analysis [20].

The development of HLT begins with a thought experiment, which is thinking about the learning trajectory that students will go through, then reflecting on the results of experimenting in the classroom. If the goal has not been achieved, then proceed with the next thought experiment. In the long term, thought experiments are always side by side to obtain LIT [20]. This relationship can be seen in Figure 1.

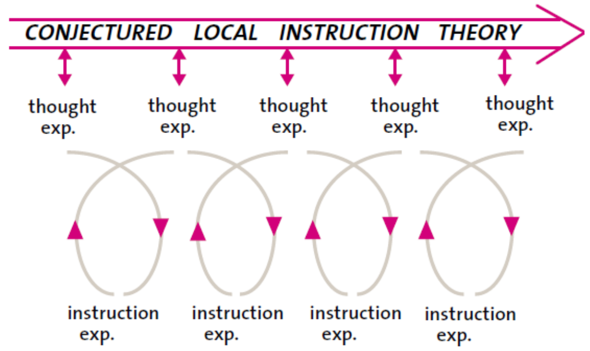

Figure 1. Reflexive relation between theory and experiments

HLT was tried out to the students of the Mathematics Education Study Program of Universitas Negeri Padang who took Elementary Statistics courses in July- December 2019 Semester. During the trial, the practicality and effectiveness of the products being developed were determined. The practicality of the product is determined based on the analysis of the learning observation sheet and questionnaire, and the effectiveness of the product is determined by providing a statistical reasoning test to students.

3. RESULT AND DISCUSSIONS

3.1. Preparing for the Experiment Phase

The purpose of this first phase is to design ILT for the topic of test criteria in hypothesis testing. There are three stages carried out in this phase, namely needs and context analysis, initial product design in the form of HLT, and formative evaluation. The results of the needs and context analysis show that the topic of the test criteria in hypothesis testing is one of the important topics in hypothesis testing which serves to conclude whether the null hypothesis formulated can be accepted or rejected. Learning outcomes on this topic were that students able to design test criteria in hypothesis testing and use them to conclude. Students' expectation is studying these test criteria were not just memorizing.

The test criteria contained in various reference books, for example, students know that if the z value calculated from the sample is more than the z value in the standard normal distribution table, then the null hypothesis is rejected. However, if a student is asked to provide a reason why this is so, the student cannot explain it. For this reason, it is necessary to understand the concept of studying the topic of the test criteria in testing the hypothesis by utilizing contextual problems.

The initial results of the product in the form of HLT contain learning goals, learning activities, and hypothetical learning processes. So that students can carry out the horizontal and vertical mathematical process, first, an iceberg is designed for the topic of test criteria in hypothesis testing as shown in Figure 2. In Figure 2 it can be seen that the contextual problem used is the expired date and reference value of the blood test results in Figure 2. laboratory. This was chosen because the two problems are related to time limits or normal limits so that students can form their models of critical values, critical areas, and areas of acceptance in hypothesis testing. Furthermore, utilizing the standard normal distribution curve students find their own concept of test criteria in hypothesis testing. HLT is designed in five stages of learning activities which can be seen in Table 1.

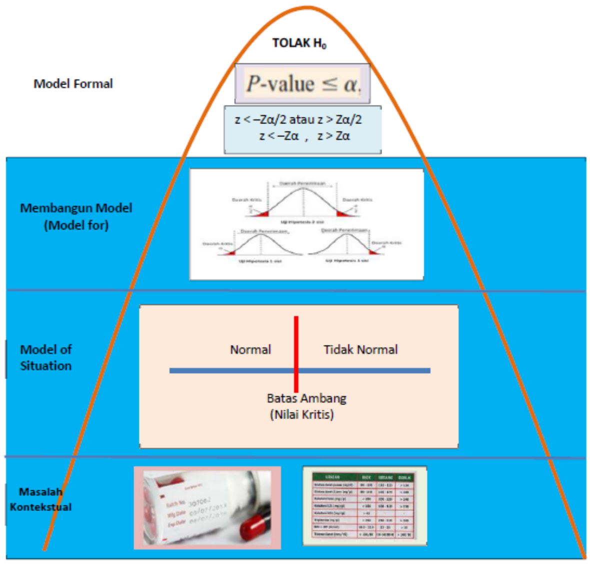

Figure 2. Iceberg of the test criteria in hypothesis testing

Furthermore, a formative evaluation is carried out on the initial HLT designs that have been produced. Formative evaluation in the form of self-evaluation and expert review aims to determine the validity of the designed product. There are three aspects evaluated, namely the content aspects, the language aspects, and the graphic aspects.

Table 1. The initial design of a hypothetical learning trajectory for test criteria in hypothesis testing

Learning goal	Stage of learning activities	Hypothetical learning process	
		Predictions of student activities	Anticipation of lecturer
Students are able to reinvent the concept of test criteria in hypothesis testing	Students examine contextual problems such as the expired date of a product and the reference value of blood tests. Next, examine the benefits of the expired date and the reference value of the blood test	Examining the contextual problem provided, and interpreting several possibilities, namely: make own model in the form of a horizontal line to interpret the values obtained. to declare the value as a threshold. declare that the value divides the area into two parts, namely the normal and the abnormal. cannot interpret at all.	The lecturer asks probing questions, such as: what does the expired date mean? can food or medicine be used beyond the expired date? a. what if the food or medicine is used in the time before the expired date? what about the reference value of blood tests in the laboratory? is this case related to the concept of critical value, critical area, and H_0 acceptance area?
	Students define the meaning of critical values, rejection region, and acceptance region in hypothesis testing using their own sentences	Associating cases with critical values, rejection region, and acceptance region using sketches. Furthermore, defining with their own language. Confused to express in their own language the definition of critical values, rejection region, and acceptance region.	The lecturer asks probing questions, such as: have you ever heard of someone being critical? What is the meaning? what does it mean that someone has passed the critical period? what is the meaning of the critical value related to the expired date and the reference value of blood tests in the case above how can the ideas above be used to define the critical value, rejection region, and acceptance region for H_0?
	Students examine the standard normal distribution curve and find the critical value, rejection region, and acceptance region for H_0 in hypothesis testing.	Utilizing the standard normal distribution curve by interpreting α to express the critical value, rejection region, and acceptance region for H_0? Can determine the critical value, but cannot determine the rejection region, and acceptance region for H_0?	The lecturer asks probing questions, such as: look again at the standard normal distribution curve. If we know $z\alpha$, where is the area of α? what is the meaning of α in the concept of testing this hypothesis? remember that $z\alpha$ divides the region into two parts, namely the area of α and $1-\alpha$. Can you understand the broad meaning of this area? Try to relate to contextual problems that have already been exploited.
	Examines the various possibilities of alternative hypotheses, and determines the critical region and rejection region H_0 of each of these possibilities.	Presents the various possibilities of the alternative hypothesis, and determines the rejection region, and acceptance region for H_0 for each of these possibilities. Presents the various possibilities of the alternative hypothesis, but cannot determine the rejection region, and acceptance region for H_0 from for each of these possibilities is not correct. Confused in determining it.	The lecturer asks probing questions, such as: recall the three possible formulations of H_1. Can the standard normal distribution curve be used to illustrate the three possible formulations of H_1? where is the area whose size is α? if H_1: $\mu \neq \mu_0$, how to determine the region whose area is α?
	Students determine the p-value and use the p- value to decide whether to accept or reject H0.	Make use of the distribution curve to present the p-value and can determine the test criteria. Make use of the distribution curve to present the p-value, but cannot determine the test criteria. Confused about determining p-value so that they cannot determine the test criteria.	The lecturer asks probing questions, such as: do you understand the meaning of p-value? can the p-value be determined if the test statistic value has not been determined? how about using the normal distribution curve to illustrate the p- value region?
One group presented the results of their discussion in front of the class and the other group give their responses.			

The results of the HLT self-evaluation show that based on the content aspect, there are not many problems. Three components that must be present in HLT, namely learning goals, learning activities, and hypothetical learning process have been presented in HLT. In the description of learning activities already characterized by RME characteristics. In the language aspect, the researcher did not find any writing errors or typos. However, the rules for writing upper or lowercase letters, and the use of punctuation marks still need to be considered. The use of the word "like" is also inappropriate, because the lecturers' anticipation must be written down everything. In the case of the word "like" the meaning is just a few examples. In the appearance and graphic aspects of HLT, it needs to be colored to make it look attractive. The results of the expert review show that the content aspect has met the state of art criteria of RME, content, and construct. In the language aspect, there are still errors in using punctuation marks. Suppose (a) the sentence does not end with a period; (b) writing the word "Purpose:" it is advisable not to use a colon; and (3) in the lecturer anticipation section, it is suggested not to start with the word "Lecturer", in order to be consistent with the alleged part of student activities that do not begin with the subject "Student". The results of the expert review validation using a questionnaire, it was found that HLT had a score of 0.89 with a very valid category.

3.2. Experimenting in the Classroom Phase

Students are facilitated to construct the concept of critical value, critical area, and area of H_0 rejection by utilizing the expired date and reference value of blood test results in the laboratory. Examining the meaning of the expired date of food and medicinal products, as well as the reference value of blood tests in the laboratory, students define in their own sentences the meaning of critical values, critical areas, and areas of acceptance of H_0 in hypothesis testing. Students have been able to make their own models of these contextual problems. All students understand the meaning of the expired date, which is the limit that separates two time periods, namely the period before the expired date and the period after the expired date.

Figure 3 presents one of the student jobs to construct test criteria in hypothesis testing. In the beginning, the students examined the contextual problems given. Students understand the meaning of the expired date and the reference value of blood tests in the laboratory. Taking advantage of this idea, students make use of the normal distribution curve with a certain α to determine the critical value, critical area, and acceptance area of H_0 in hypothesis testing. The results of student work in Figure 3 also show that students have been able to take advantage of the concepts that have been constructed on the problem of test criteria in hypothesis testing, and have been able to conclude on whether H_0 is accepted or rejected.

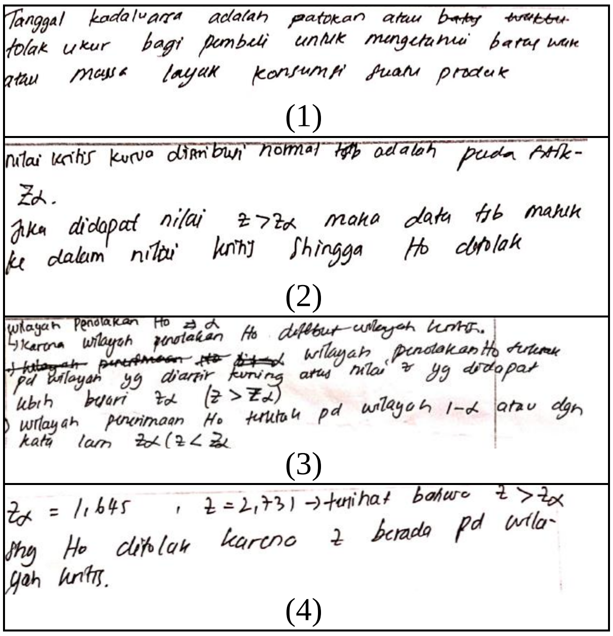

Figure 3. One Student's Work Result

3.3. Retrospective Analysis Phase

The results of experimenting in the classroom showed that the learning design in the form of a learning trajectory that was designed to run well. Prepared expectations and anticipations emerge in learning. The contextual problems provided become the motivation for students to think about various possible solutions to problems. Students are facilitated to carry out horizontal mathematical and vertical mathematical processes. Not telling something right or wrong about the response given by students has encouraged students to think for themselves constructing the concept of test criteria in hypothesis testing. This is evidenced by the various student work results data. Probing questions given by lecturers are an important step that can help students understand problems in a more meaningful way. Furthermore, each response given by students is followed by the reasons why and how it can be obtained. This is an effort to help improve students' statistical reasoning abilities.

The practicality test results show that the learning design developed is categorized as very practical. This shows that the resulting learning design can be used, easy to use, and has very good appeal. This is in accordance with what was stated [21], namely that an intervention is said to be practical if the user views that the intervention that is designed is useful and easy to use. In addition, [22] added that an intervention is said to be practical if it is attractive and useful under normal conditions. The characteristics of a very practical learning design are shown by all levels of students being able to carry out designed activities, it can be used to facilitate students to explore the theoretical concepts of the material being studied, can develop students' statistical reasoning abilities, the time available is sufficient to achieve learning objectives, the language used is easy to understand, and content display that begins with contextual problems attracts students' interest to carry out the activities provided.

To determine the effectiveness of the learning design of the hypothesis testing topic, the results of the student's statistical reasoning ability test were used in two groups. The first group is an experimental group whose learning uses a developed learning design, and the second group is a control group whose learning uses a learning design that has been applied in studying the topic of hypothesis testing. Based on the results of the students 'reasoning ability test in both groups, it showed that the learning design developed was effective in improving students' statistical reasoning abilities. This is supported by the results of research [23] that HLT can develop students' statistical thinking skills and can assist educators in creating models of students' thinking and restructuring educators' understanding of mathematics and reasoning.

The effectiveness of the design results in improving students' statistical reasoning abilities can be seen from the statistical measure of the statistical reasoning ability test results in the two groups being tested. The result is that the measure of the central tendency of the experimental group is bigger when compared to the control group. While the measure of the variation of the experimental group was smaller than the control group. This indicates that the experimental group students' statistical reasoning ability spread more uniformly at higher scores when compared to the control group. To recognize these findings, a hypothesis test was carried out on the difference between the two population mean values using the t-test. The t-test result shows that the p-value obtained is 0.011 which is smaller than the specified significance value. This shows that H_0 is rejected, it means the mean scores of students' statistical reasoning abilities in the experimental group differed significantly when compared to the average scores of students' statistical reasoning abilities in the control group. So, in general, it can be concluded that the learning design of the test criteria topic in hypothesis testing in the form of LIT which was developed effectively improves students' statistical reasoning abilities.

4. CONCLUSIONS

The result of this study is a local instructional theory for the topic of testing criteria in testing hypothesis a valid, because it meets the state of art criteria of RME, content, and construct. Practice, because the learning trajectory is usable, easy to use, and has a very good appeal. The results also show that the average value of the statistical reasoning ability of group students whose learning uses LIT is better than the average value of statistical reasoning ability of group students whose learning does not use LIT. Thus, the resulting LIT is effective in improving student’s statistical reasoning abilities.

REFERENCES

[1] Brase, C. H., & Brase, C. P. *Understandable statistics: Concepts and methods* (10th ed.), Boston, MA, Brooks/Cole Cengage Learning, 2012, pp. 10.

[2] Sekretaris Kabinet, *Lampiran Peraturan Presiden Republik Indonesia Nomor 8 Tahun 2012,* tantang Kerangka Kualifikasi Nasional Indonesia. 2012, pp. 2-3.

[3] [Online:https://sipuu.setkab.go.id/PUUdoc/17403/Pe rpre s0082012.pdf]

[4] American Statistical Association, *Guidelines for assessment and instruction in statistics education: College report*. Alexandria, VA: Author, 2016, pp. 1-6

[5] [Online:https://www.amstat.org/asa/files/pdfs/GAIS E/Ga iseCollege_Full.pdf]

[6] Kula, F., & Rüya Gökhan Koçer, R. G, Why is it difficult to understand statistical inference? Reflections on the opposing directions of construction and application of inference framework. *Teaching Mathematics and its Applications*, **00**, 2020, pp. 1–18. DOI: https://doi:10.1093/teamat/hrz014

[7] Abdulah, M. A. & Ismail, Z., Malaysian college students misconceptions in inferential statistics. *International Journal of Pure and Applied Mathematics*, 118 (24), 2018, pp. 1-14. http://www.acadpubl.eu/hub/Special Issue. ISSN: 1314-3395 (on-line version)

[8] rishnan, S. & Idris, N., Students’ misconceptions about hypothesis test. REDIMAT, *Journal of Research in Mathematics Education*, *Vol 3*(3), 2014, pp. 276-293.DOI: https://doi:10.4471/redimat.2014.54

[9] Batanero, C., Controversies around the Role of Statistical Tests in Experimental Research. *Mathematical Thinking and Learning*, *2* (1-2), 2000, pp. 75–98. DOI: https://doi: 10.1207/S15327833MTL0202_4

[10] Castro, S. A. E., Vanhoof, S., Noortgate, W. V. den, & Onghena, P., Students’ misconceptions of statistical inference: A review of the empirical evidence from research on statistics education. *Educational Research Review*, *2(2)*, 2007, pp. 98 – 113.DOI: https://doi:10.1016/j.edurev.2007.04.001

[11] Thompson, B., Effect Sizes, Confidence Intervals, And Confidence Intervals For Effect Sizes.

Psychology in the Schools, Vol. 44(5), 2007, pp. 423-432. DOI: https://doi: 10.1002/pits.20234

[12] Gagnier, J. & Morgenstern, H., Misconceptions, misuses, and misinterpretations of p-values and significance testing. *Journal Of Bone & Joint Surgery*, Vol 99 (18). 2017, pp. 1598-1603. DOI: https://doi: 10.2106/JBJS.16.01314

[13] Garfield, J. & Ben-Zvi, D., *Developing students' statistical reasoning: connecting research and practice*, California: Springer, 2008.

[14] Krishnan, S. & Idris, N., Students' misconceptions about hypothesis test. REDIMAT, *Journal of Reseach in Mathematics Education*, *Vol 3*(3), 2014, pp. 276-293.DOI: https://doi: 10.4471/redimat.2014.54

[15] Gravemeijer, K., & Van Eerde, D., Design research as a means for building a knowledge base for teachers and teaching in mathematics education. *Elementary School Journal, 109*(5), 2009, pp. 510-524. http://www.jstor.org/stable/10.1086/596999

[16] Moore, D., The predictive approach to teaching statistics. *International Statistical Review* 65(2), 1997, pp. 123-165. DOI: https://doi: 10.22004/ag.econ.267383

[17] De Lange, J., *Using and Applying Mathematics in Education. In International Handbook of Mathematics Education*, A.J. Bishop, et. al. (eds.). The Netherlands: Kluwer Academic Publishers, 1996.

[18] Gravemeijer, K. P. E., *Developing realistic mathematics education*. Utrecht: CD-β Press, 1994.

[19] Clements, D., & Sarama, J., Learning trajectories in mathematics education. *Mathematical Thinking and Learning*, *6*(2), 2004, pp. 81-89. DOI: https://doi: 10.1207/ s15327833mtl0602_1

[20] Larsen, S. P., A local instructional theory for the guided reinvention of the group and isomorphism concepts. *The Journal of Mathematical Behavior*, *32*(4), 2013, pp. 712- 725.DOI: https://doi.org/10.1016/j.jmathb.2013.04.006

[21] Plomp, T., Educational design research: An introduction", In Tjeerd Plomp dan Nienke Nieven (Eds.), *Educational Design Research*. Netherlands: Institute for Curriculum Development (SLO), 2013

[22] Gravemeijer, K., & Cobb, P., Design research from a learning design perspective. In J. van den Akker,

[23] K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational design research*, London: Routledge, 2006, pp. 17-51,

[24] Akker, J.V.D., Bannan, B., Kelly, A.E., Nieveen, N., & Plomp, T. E*ducational Research Design*. The

[25] Netherlands: Netherlands Institute for Curriculum Development (SLO), Enschede, 2013.

[26] Fauzan, A., Plomp, T., & Gravemeijer, K., *The development of an rme-based geometry course for Indonesian primary schools*. In T. Plomp, & N. Nieveen (Eds.), Educational design research – Part B: Illustrative cases, Enschede, the Netherlands: SLO, 2013, pp. 159-178.

[27] Wilson, P. H., Mojica, G. F., & Confrey, J., Learning trajectories in teacher education: Supporting teachers' understandings of students' mathematical thinking. *The Journal of Mathematical Behavior*, 32(2), 2013, pp. 103-121. DOI: 10.1016/j.jmathb.2012.12.003.