INDONESIAN JOURNAL OF EDUCATIONAL RESEARCH AND REVIEW
Volume 7 Nomor 3 2024, pp 566-580
E-ISSN: 2621-8984; P-ISSN: 2621-4792
DOI: https://doi.org/10.23887/ijerr.v7i3.80790

Feasibility and Effectiveness of E-Test Through Vinesa to Improve Statistical Reasoning Ability Central Tendency Material

Hari Sugiharto Setyaedhi[1*]
[1] Teknologi Pendidikan, Universitas Negeri Surabaya, Surabaya, Indonesia
*Corresponding author: harisetyaedhi@unesa.ac.id

Abstrak

Dosen masih kurang dalam mengembangkan tes mata kuliah statistika elektronik melalui Virtual Learning Unesa (Vinesa), sehingga menyebabkan hasil belajar mahasiswa menurun. Tujuan penelitian ini adalah untuk menguji kelayakan dan keefektifan tes materi tendensi sentral elektronik. Analisis data menggunakan deskriptif kuantitatif dan uji t. Subjek penelitian adalah 1 orang ahli materi, 1 orang ahli media, 1 orang ahli bahasa, dan 28 mahasiswa. Pengumpulan data menggunakan tes, observasi, wawancara, dan angket. Model ADDIE digunakan dalam penelitian pengembangan ini. Hasil kelayakan validasi materi, media pembelajaran, bahasa, uji coba individu dan kelompok kecil masing-masing berada pada kategori sangat layak; hasil efektivitas menunjukkan bahwa kelas kontrol (0,45) berada pada kategori rendah dan kelas eksperimen (0,6) berada pada kategori cukup; dan hasil uji t pada sig. (2 ekor) = 0,000 < 0,05. Kesimpulan pengembangan elektronik tes denganmateri tendensi sentral layak digunakan dalam tes dan efektif dalam meningkatkan kemampuan penalaran mahasiswa. Implikasi dari penelitian ini adalah produk pengembangan elektronik tes yang dapat dimonitoring tingkat kelayakannya, terhadap peningkatan penalaran mahasiswa untuk mata kuliah statistik.

***Kata kunci*:** *Kelayakan, Efektivitas, Tes Elektronik, Kemampuan Bernalar*

Abstract

Lecturers are still lacking in developing electronic statistics course tests through Virtual Learning Unesa (Vinesa), which causes student learning outcomes to decline. The aim of this study was to test the feasibility and effectiveness of electronic central tendency material tests. Data analysis uses quantitative descriptive and t tests. The research subjects were 1 material expert, 1 media expert, 1 language expert, and 28 students. Data collection uses tests, observations, interviews, and questionnaires. The ADDIE model is used in this development research. The feasibility results of material validation, learning media, language, individual and small group trials are each in the very feasible category; the effectiveness results show that the control class (0.45) is in the low category and the experimental class (0.6) is in the sufficient category; and the test results t on sig. (2 tailed) = 0.000 < 0.05. Conclusion The development of electronic tests with central tendency material is suitable for use in tests and is effective in improving students' reasoning abilities. The implication of this research is the development of electronic test products that can be monitored for their level of feasibility, with the aim of improving student reasoning in statistics courses.

Keywords: Feasibility, Effectiveness, Electronics Test, Reasoning Ability

History:
Received : June 18, 2024
Accepted : October 15, 2024
Published : October 25, 2024

Publisher: Undiksha Press

1. INTRODUCTION

Evaluate the extent to which educational goals have been achieved (Nurfillaili et al., 2016; Sari et al., 2022). Evaluation is the basis for concluding what should be done in the next lesson (Muliyani & Huriaty, 2016; Setiyawan & Wijayanti, 2020). In conducting learning evaluations, lecturers measure students through tests. Evaluation tools can be used in test or non-test formats (Jumrah et al., 2023; Masitoh & Aedi, 2020; Widiyawati et al., 2019). The use of tests is important in the learning evaluation process. Test are used to measure the success of achieving a particular program (Fahrurrozi & Laili Rahmawati, 2021; Sholihah et al., 2017). Tests have the benefit of knowing the learning process effectively (Manfaat & Nurhairiyah, 2021; Setiyawan & Wijayanti, 2020). Tests are important to determine student learning outcomes during the learning process (Aisyah et al., 2021; Birch et al., 2017; Nurfillaili et al., 2016; Sa'diyyah et al., 2021). The test is prepared based on the answers to the questions and statements contained in the instrument. In this case, students are

encouraged to answer the test questions as best as possible. Tests are used for various purposes, such as new student admissions, graduation, selection, and others (Litna et al., 2021; Ma'rifah et al., 2021; Sa'diyyah et al., 2021).

One of the subjects taught in universities and used for quantitative research is statistics. Mastering statistics courses is very important for students who want to become teachers (Cahyawati et al., 2020; Subekti & Jazuli, 2022). This course is related to precision and analytical thinking. Statistics is a mandatory subject in every educational program (Afifah & Wicaksana, 2014; Robin Bastian Waruwu et al., 2022). However, even though this subject is considered important, there are some students who find it difficult to study it and are less interested. For most students, statistics is a difficult subject, so it feels boring. Because statistics is a difficult subject for most students, many students try to avoid it because it is considered boring (Inayah, 2017; Ririen & Hartika, 2021; Rizki & Fauziddin, 2021). Common mistakes made by students include not being careful, process errors, difficulty understanding questions, and transformation and notation errors (Sutrisno & Murtianto, 2016; Yuniarti, 2022).

Conceptual errors, obstacles in understanding, drawing conclusions, and testing hypotheses. It is important for lecturers to carry out error analysis to identify students' difficulties and weaknesses when solving statistical problems (Apriani, 2016; Yuniarti, 2022). In this way, teachers can improve the quality of learning and develop students' thinking abilities. Central tendency material is the basis for calculations most commonly used in statistics (Rizki & Fauziddin, 2021; Rosidah & Ikram, 2021; Yuniarti, 2022). Research shows that many students have difficulty understanding the concept of central tendency in this material (Sutrisno & Murtianto, 2016). According to a study, 78% of 26 students were unable to reason, draw conclusions, and interpret when working on statistics questions. Learning statistics requires the ability to collect, present, analyze, and draw conclusions (Fardillah et al., 2019; Subekti & Jazuli, 2022).

Advances in information and communication technology (ICT). Manual or traditional testing systems are gradually being replaced by computer-based testing systems. Online tests are part of the distance learning information system via the internet. In the final assessment by lecturers, the use of information technology is very important to obtain accurate evaluation results and an efficient process (Adiarta & Divayana, 2019; La Nani, 2020). Electronic tests (e-tests), are a modern student assessment method that is an ideal alternative to traditional paper tests. e-tests have the advantage of saving energy and time when assessing student performance. Test electronics enable test evaluation, performance measurement, and direct feedback quickly and accurately (Jasman & Sidin, 2018; Muspardi et al., 2020; Nurhasanah, 2018). E-tests are inexpensive, offer flexibility in administering multiple tests, and are easy for teachers to set up and use for most questions. There are two kinds of ways to present e-tests. (1) Synchronous outdoor face-to-face tests describe tests that are available face-to-face without changes in testing procedures, (2) Asynchronous home tests are tests at any time using high-quality proctoring software (Gamaliel & Arliyanto, 2021; Jusuf, 2019; Poonpon, 2021).

The urgency of this research is that the use of electronic test instruments via the Vinesa LMS can simplify the assessment process used by lecturers. It is easier for lecturers to combine student activities. Electronic tests via Vinesa are suitable to be used as an alternative medium to solve problems in implementing learning evaluation at Surabaya State University. The resulting test product is integrated with the internet network so that student score results can be read by lecturers, students, and parents. Apart from that, the use of electronic tests via the Vinesa LMS is practical, cheap, and easy to carry out.

This is in accordance with research findings that computer-based tests (CBT) are a viable alternative for solving the problem of implementing learning assessments. Several

important aspects that can be considered when developing CBT are practicality, cheapness, and ease of carrying out the testing process (Kamila & Prihatiningtyas, 2022; Mastroleo et al., 2020; Nurhikmah et al., 2021). Online tests have advantages: 1) teachers can prepare higher-quality test materials; 2) the test administration process is standardized; and 3) students can be supervised. The test results will be integrated into the internet network so that they can be observed by teachers, students, and parents. By utilizing technological developments, online tests can create a system that provides quick and accurate evaluations, making it easier to complete the evaluation process (Kamila & Prihatiningtyas, 2022; Mastuti, 2016).

The novelty of this research was that there is still a lack of lecturers who are able to develop electronic test properly and effectively, and there are not many researchers who study electronic test. Apart from that, most Unesa lecturers have not yet accessed the test s optimally via Vinesa (Virtual Learning Unesa). Not many Unesa lecturers have researched the feasibility and effectiveness, as well as the use of electronic test, through Unesa virtual learning. To overcome these learning problems, it is necessary to develop test that can overcome them because learning is taught not only in theory but also in practice to stimulate students' skills (Krisnanto et al., 2021; Novrianti, 2014).. The aim of this research was to assess the feasibility and effectiveness of developing electronic test. Feasibility is proven through a validation process by a team of learning media experts, learning material experts, and language experts, as well as individual and small group trials (Ririen & Hartika, 2021; Rizki & Fauziddin, 2021). The effectiveness of developing electronic test is presented through t-test statistical data processing. It is hoped that this research will produce electronic test products that are feasible and effective and can be used more widely, especially for educational technology students.

2. METHODS

Development research with a quantitative approach is aimed at producing products that are suitable for use (Rahman et al., 2021). For the purposes of testing the feasibility of the test, development research subjects for validation are required, namely: 1 material expert, 1 media expert, and 1 language expert. Individual product trials require 3 students and 6 students for small group trials. Students were selected heterogeneously, namely clever, medium, and poor. Data analysis is quantitative and descriptive. Quantitative analysis techniques are used to calculate the feasibility and effectiveness of electronic test. The electronic test is first validated by experts (expert judgment) in terms of media, material, and language. The prerequisite tests for normality and homogeneity are carried out before the hypothesis is tested t (Wingga & Saputra, 2019).

The data collection techniques used were observation, interviews, questionnaires, validation sheets, and tests. a) The observation technique is data collection by direct observation of the research object to see the activities carried out (Muchta, 2019; Pramana et al., 2020). The observation method is arranged sequentially, structured, and systematic (Khoridah et al., 2019). Data obtained through the observation method is more accurate because the researcher observes directly according to reality, b) The interview method is the process of collecting information through questions and answers, which is carried out in a structured manner, and the results are recorded in the form of a document. The interview method is by making a list of questions for students who will be researched, c) questionnaire technique: the researcher makes a list of questions given to media experts, material experts, and learning design experts to measure product suitability (Dwiqi et al., 2020; Pramana et al., 2020). Questionnaire techniques are used to obtain accurate data from the questions contained in the questionnaire, d) sheet validation technique used to determine the validity of

the test equipment that has been created by researchers (Krisnanto et al., 2021), e) Test techniques carried out during trials to obtain information about the description and results of student learning (Krisnanto et al., 2021).

The development model uses ADDIE (Analyze, Design, Development, Implementation, Evaluation). The ADDIE development model has advantages, one of which is that at each stage it goes through an evaluation stage so that it can minimize errors from the start (Artha & Putra, 2021; Pramana et al., 2020). The ADDIE model is carried out in stages, in detail and systematically. In the ADDIE model, evaluation is carried out at every stage of the activity. The ADDIE stages are: Analysis stage. Collecting various information so that product designs can be developed as planning material (Firda & Nurhadi, 2023; Wirda et al., 2017). Design stage. This stage is to design test products namely by determining basic and core competencies and compiling a grid of questions that will be developed into test (Farahiba, 2022; Firda & Nurhadi, 2023). The development stage is an activity to expand questions or statements originating from the framework or grid created at the planning stage (Ahmad, 2018; Farahiba, 2022). This stage also includes theoretical validation by expert lecturers, and testing of test (Ahmad, 2018; Firda & Nurhadi, 2023; Wirda et al., 2017). Implementation stage. Researchers conducted a limited trial of the instrument on several undergraduate students in Educational Technology to determine the level of feasibility (Firda & Nurhadi, 2023; Wirda et al., 2017). The evaluation stage, namely 1) Assessing whether the steps taken previously were appropriate and produced accurate and reliable data (Farahiba, 2022), 2) looking at the feasibility of the limited trial results. If after evaluation there are problems with the test, a revision will be carried out, this is done in order to obtain a suitable test.

The instrument grid from learning material experts, learning evaluation experts, and language, individual and small group trials, respectively, is presented in tables 3, 4, 5, and 6. This instrument aims to determine the quality of the statistics course test which has been developed. Pay attention to Table 1.

Table 1. Grid of Learning Material Experts

No	Aspect		Indicator
1	Curriculum	a	Conformity of test items with basic competencies.
		b	Correspondence of test items to indicators.
		c	Correspondence of test items to learning outcomes.
2	Material	a	The questions in the test are easy to understand.
		b	Depth of test material.
		c	Tests are supported by appropriate media.
3	Language	a	The language is easy to understand.
		b	The language used is EYD.
4	Evaluation	a	The difficulty level of the questions is determined by ability.
		b	Suitability of test items to the material.

This instrument is used to determine the quality of the statistics course test instruments that have been developed. Pay attention to Table 2.

Table 2. Grid of Learning Media Experts

No	Aspects		Indicator
1	Text Aspects	a	The text on the e-book can be read well.
		b	The text size and typeface on the e-book are appropriate.
		c	Clarity of the material description.

No	Aspects	Indicator	
		d	Clarity of instructions.
		e	Background on e-books.
		f	Colors in e-books.
2	Image Aspect	a	E-book image layout.
		b	Image quality in e-books.
		c	The attractiveness of images in e-books.
3	Accessibility Aspect	a	The ability of e-books to facilitate students in learning.
		b	The ability of e-books to facilitate teachers.
		c	Ease of accessing e-books.
		d	E-books increase competence.

This instrument is used to determine the linguistic quality of the statistics courses that have been developed. Pay attention to Table 3.

Table 3. Grid of Linguistic Experts

No	Aspect	Indicator	
1	Language	a	Simple language to communicate.
		b	Language using EYD.
		c	The language is adjusted to the student's cognitive level.
		d	Sentences are clear and easy to understand.
		e	Preparing the correct sentence structure.
		f	There are explanations for terms that are difficult to understand.

This instrument is used to determine student responses related to the use of media for developing statistical test. The instrument grid for individual groups and small groups is presented in Table 4.

Table 4. Grid of Individual and Small Group Trials

No	Aspect	Indicator	
1	Ease of reading questions	a	The language in the test is easy to understand.
		b	There are no errors in writing the test.
2	Material	a	Test items are appropriate for the material.
		b	In line with the content of the material.
		c	Clarity of the material description.
3	Test sheet display	a	Colors attract students' attention.
		b	Images and graphs are clearly visible.
		c	The text is legible.
		d	Voice clarity.

After the instrument grid and trials have been completed, the next step is to test the validity of the instrument. Experts validate the instrument items, and the validation results are in accordance with the feasibility category. Categories for making decisions are presented in Table 5.

Table 5. Eligibility Criteria for Teaching Materials with Scale 5

Score in Percent (%)	Category
< 21%	Not Worth It
21% - 40%	Not feasible

Score in Percent (%)	Category
40% - 60%	Decent Enough
60% - 80%	Worthy
80% - 100%	Very Worth It

The quasi experimental design method with a nonequivalent control group design was used to determine the effectiveness of the electronic test. In research with a quasi-experimental design, students are grouped into a control group and an experimental group, but the placement of students in a group is not carried out randomly. The research design is presented in Table 6.

Table 6. Research Design

Group	Pre-test	Treatment	Post-test
Experiment	O_1	X_1	O_1
Control	O_1		O_1

3. RESULTS AND DISCUSSION

Results

This development research aims to produce a product in the form of an electronic central tendency material test (e-test) to improve statistical reasoning as well as test the feasibility and effectiveness of an electronic central tendency material test. The feasibility test goes through several stages of validity testing, consisting of learning material expert testing, learning evaluation and language experts, individual testing, small group testing, and field testing. Meanwhile, the effectiveness test uses pre-tests and post-tests, as well as t-tests. This research uses the ADDIE development model, which includes the analysis stage, design stage, development stage, implementation stage, and evaluation stage.

The research results are as follows: Analysis stage. The purpose of needs analysis is to identify gaps between current real conditions and ideal conditions in the field. The results of observations and interviews show that there are gaps, including: a) Learning is not optimal because students do not understand central tendencies, b) lecturers have not optimally developed e-tests; c) students do not utilize the various e-tests they have to study. Design stage. Determining product specifications requires knowing the results of the needs analysis and determining the product to be developed, namely designing an e-test. The design is in the form of a simple model that contains a general description of e-test creation. The process of creating an e-test design using a computer-based application. The application used is open source and has been provided by Vinesa. This application is used to create online-based e-tests with electronic displays.

At development stage, the research team carried out product development using online applications. Explicitly, the e-test products that have been produced can be seen as follows: The results of product development that have been completed are then tested. e) product trials, which include individual group trials with 3 students and small group trials with 6 students selected heterogeneously with high, medium, and low abilities. The design of vinesa learning management system is show in Figure 1.

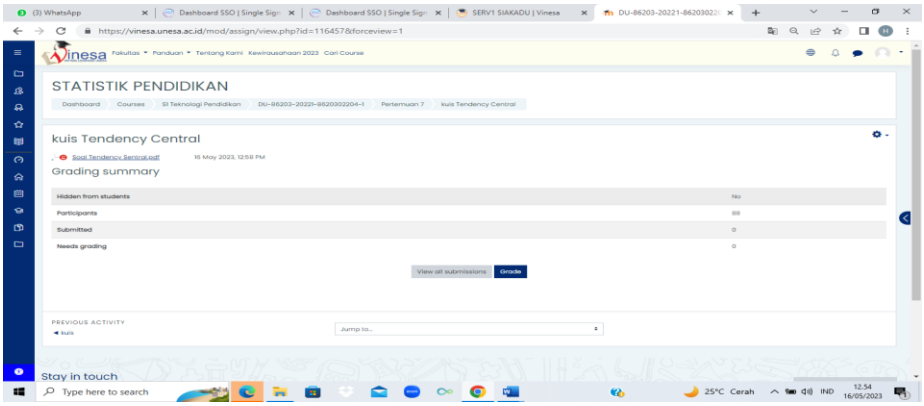

Figure. 1. Design the Vinesa Learning Management System (LMS)

In Figure 1, you can see the appearance of the tendency central quiz, which has been submitted by the lecturer in the Vinesa LMS and is ready for students to complete. Students can access these questions via Vinesa and are ready to take the test**.** Students can work on statistics questions and then upload the answers via Vinesa LMS. Lecturers can see students' answers via Vinesa. It can be seen from the answers that the students' reasoning process is in line with expectations. Implementation stage. This stage includes: 1) At this stage, experts will assess the feasibility of this e-test product. After that, the researcher carried out revisions according to the notes provided by the relevant parties so that the test instruments produced met standards and met student needs; 2) product trials through individual and small group trials. The goal is to minimize deficiencies. This stage is intended to determine the feasibility of the test instrument that has been developed as show in Table 7.

Table 7. Product Validation Assessment

Test Subjects	Validity Results	Category
Learning materials expert	96%	very worthy
Learning evaluation expert	96%	very worthy
Linguist expert	94%	very worthy
Individual trials	81%	very worthy
Small group trials	82%	very worthy

Base on Table 7, feasibility test using Calculation of All Aspects (PSA) with a scale of 5. Expert assessment of the content of the learning material obtained a result of 96% categorized as very feasible; expert assessment of the learning media obtained a result of 96% categorized as very feasible; expert assessment of the learning design obtained a result of 94% categorized as very feasible; individual trials obtained results of 81% categorized as very feasible; small group trials obtained results of 82% categorized as very feasible.

Evaluation stage, the activities carried out at the evaluation stage include the evaluation of statistical test products. The evaluation carried out in this development research was based on a feasibility test. The feasibility of statistical test is evaluated based on validation results from experts and product trials. The evaluation results are revised according to the evaluation results. The evaluation of the development of statistical test aims as follows: 1) student attitudes towards learning activities as a whole; 2) increasing student abilities; 3) institutional profits as a result of increasing student competency.

Before carrying out the t test, there are two requirements, namely: the data must be normally distributed and homogeneous. The t test was carried out to determine the overall calculation results of the research. The result of experimental group is show in Table 8.

Table 8. Experimental Group

Variable	Experimental Group	Shapiro-Wilk		
		Statistic	df	Sig.
Learning outcomes	Pretest	0.988	28	0.983
	Postest	0.987	28	0.980

Based on Table 8, it is known that the Sig value for pretest is 0.983 and the Sig value for posttest is 0.980. Sig value. both groups > 0.05, it can be concluded that the pretest and posttest experimental group student learning outcomes data are normally distributed.

Table 9. Control Group

Variable	Control Group	Shapiro-Wilk		
		Statistic	df	Sig.
Learning Outcomes	Pretest	0.978	28	0.806
	Postest	0.977	28	0.776

Based on Table 9, the sig values of pretest 0.806 and post-test 0.776 > 0.05, so it can be concluded that the pre-test and post-test control group student learning outcomes data are normally distributed. In Table 12, it can be seen that the mean value for the experimental group is 63.36, while the control group is 61.32. From descriptive statistics, it can be concluded that there is a mean difference in student learning outcomes between the experimental group and the control group. The comparison is show in Table 10.

Table 10. Comparison of Pretest Group Means

Variable	Pretest Group	N	Mean
Learning outcomes	Experiment	28	63.36
	Control	28	61.32

To prove whether the mean has a significant difference or not, we need to interpret the output of the independent samples test as show in Table 11.

Table 11. Independent Samples Test Pretest Group

Levene's Test for Equality of Variances		t-test for Equality of Means		
F	Sig.	t	df	Sig. (2-tailed)
0.000	0.986	1.426	54	0.160

In Table 11, the Sig value. Levene's 0.986 > 0.05 can be interpreted as meaning that the data variance between the experimental group and the control group is homogeneous. Based on Table 13, the Sig value. (2-tailed) is 0.160 > 0.05, so it can be concluded that there is no significant difference between the mean student learning outcomes in the experimental group and the control group. The calculated t_{value} is 1.426 < t_{table} 2.005, so there is no difference in the mean student learning outcomes between the experimental group and the control group. The aim is to prove that students who took the pretest in the experimental and control groups had the same abilities before being given treatment in the form of electronic test, so no one was superior. Comparison of posttest group means is show in Table 12.

Table 12. Comparison of Posttest Group Means

Variable	Group	N	Mean
Learning Outcomes	Experiment	28	83.36
	Control	28	73.18

Base on Table 12, to prove whether the mean difference is significant or not, it is necessary to interpret the output of the independent samples test as show in Table 13.

Table 13. Independent Samples Test Group Posttest

Levene's Test for Equality of Variances		t-test for Equality of Means		
F	**Sig.**	**t**	**df**	**Sig. (2-tailed)**
0.015	0.903	7.225	54	0.000

In Table 13, the Sig value. Levene's 0.903 > 0.05 can be interpreted as meaning that the data variance between the experimental group and the control group is homogeneous. Based on table 15 and the sig value. (2-tailed) of 0.000 < 0.05, it can be concluded that there is a significant difference between the mean student learning outcomes in the experimental group and the control group. The calculated t_{value} is 7.225 > t_{table} 2.005, so there is a difference in the mean of student learning outcomes between the experimental group and the control group. Students who use electronic test have much better student scores compared to conventional control classes, so it can be concluded that electronic test are effective in improving statistical reasoning abilities.

Discussion

Research into the development of electronic test using central tendency material Vinesa is said to be feasible after going through the stages of material validation testing, learning media validation testing, learning design validation testing, individual trials, and small group trials, all of which are said to be very feasible, as well as the effectiveness of using the t-test, which is obtained from the average value of the pretest and posttest trials (Cahyadi, 2019; Wingga & Saputra, 2019). This research is in line with research which concluded that a HOTS-based e-quiz was said to be feasible after going through a series of validation tests and field trials, especially in mathematics subjects (Rohmah & Bukhori, 2020). Other research also shows that students' response to the implementation of online physics test media was 40% and was in the appropriate category. Meanwhile, 60% are in the very feasible category. So it is suitable for use in formative physics tests (Prisuna, 2020). Moreover research on the development of science e-modules containing online tests shows that e-modules containing online tests are effective in improving learning outcomes (Lestari & Parmiti, 2020). Research on developing higher-order thinking skills e-test questions. The validation result was 94.47%, with very feasible criteria. Based on validation results, the development of HOTS e-test questions was declared suitable for use in learning activities in schools (Wulandari et al., 2023). According to research based on the results of the feasibility test of the computer-based test, it can be concluded that the use of the computer-based test is suitable for use as a learning evaluation medium in mathematics subjects (Nurhikmah et al., 2021).

Research on online tests, such as research entitled online Tests for Computer-Based Final Examination Preparation, states that online tests are effective for development (Munadi, 2015; Muspardi et al., 2020; Udiasih & Ghozali, 2019). Research entitled online exam systems to improve student learning quality states that online exam tests can improve the quality of learning (Martono et al., 2020). Research conducted previously concluded that

diagnostic tests given online can provide effectiveness and convenience (Halim et al., 2018). Research entitled Developing Web-based online Tests states that online tests can improve student learning outcomes (Rokhaniyah & Virgantara Putra, 2021). Research on the effectiveness of learning through online quizzes states that online learning is very effective (The & Usagawa, 2018). Research on formative testing in e-learning for mathematics states that formative tests using e-learning are more effective than those without (Tempelaar et al., 2012). Research on automated assessment generation in e-learning systems using combinational testing and formal concept analysis shows the feasibility of using assessment systems using e-learning (Skopljanac-Macina et al., 2021). In his research, Bellotti explained two learning objectives for computer testing: fun and challenging (Prisuna, 2020).

The limitation of this research is that the individual group trials only obtained 0.81 and the small group trials 0.82, which is close to the feasible category but still in the very feasible category to be able to improve students' statistical reasoning abilities. The implications of electronic test development products can be monitored for their level of feasibility in improving the reasoning of undergraduate students in the Educational Technology study program. The various advantages of electronic or online tests are: 1) Students can score their exams directly (Kuncahyono et al., 2020; Prisuna, 2020); 2) analysis of question items can be done directly; 3) question corrections can be done automatically (Jusuf, 2019; Kuncahyono et al., 2020; Prisuna, 2020); 4) randomization of questions and answer options so that it is difficult for students to collaborate (Jusuf, 2019; Kuncahyono et al., 2020; Mastuti, 2016).

The limitation of this research is that the individual group trials only obtained 0.81 and the small group trials 0.82, which is close to the feasible category but is still feasible for improving students' statistical reasoning abilities. The implication of this research is the development of electronic test products that can be monitored for their level of feasibility towards improving student reasoning in statistics courses (Arsanti, 2018; Muspardi et al., 2020). Apart from that, the implications of this research can be developed in other materials and courses, but training should be carried out for lecturers on how to develop and use electronic tests on the Vinesa Learning Management System (LMS) so that they can attract students' interest in active learning. It is hoped that the research contribution to the development of electronic tests through Vinesa will make students more active and creative in taking tests, and students can make optimal use of Unesa's LMS (Kamaruddin et al., 2020; Nurhasanah, 2018). Apart from that, electronic tests via Vinesa are expected to help lecturers monitor student assignments submitted via the Vinesa LMS quickly and precisely. The advantages of this research are: 1) teachers can monitor reasoning abilities that are not seen in traditional methods; 2) students can directly see their exam scores (Kuncahyono et al., 2020; Prisuna, 2020); 3) analysis of question items can be done directly; 4) question correction can be done automatically (Jusuf, 2019; Kuncahyono et al., 2020; Prisuna, 2020); 5) randomization of question items and answer options so that it is difficult for students to collaborate (Jusuf, 2019; Kuncahyono et al., 2020; Mastuti, 2016); 6) increase student motivation, concentration, and performance (Mastuti, 2016).

4. CONCLUSION

Development of e-tests to improve statistical reasoning abilities in undergraduate students' central tendency tests educational technology through Vinesa with the ADDIE development model has received qualifications from experts and students, so it can be concluded that electronic statistical test has proven feasible and effective for improving students' reasoning abilities. Electronic test can foster a culture of independent learning, so that

the development of electronic test is feasible and effective as a measurement recommendation that can be accessed anytime and anywhere.

5. REFERENCES

Adiarta, A., & Divayana, D. G. H. (2019). Pengembangan Soal Test Digital Matakuliah Asesmen Dan Evaluasi Menggunakan Aplikasi Wondershare. *Jurnal Pendidikan Teknologi Dan Kejuruan*, *16*(2), 287. https://doi.org/10.23887/jptk-undiksha.v16i2.19199.

Afifah, S. N., & Wicaksana, E. J. (2014). Persepsi Mahasiswa tentang Mata Kuliah Statistik serta Pengaruhnya terhadap Prestasi Belajar Statistik Mahasiswa IKIP PGRI Madiun. *Jurnal CARE*, *2*(1), 1–8. https://e-journal.unipma.ac.id/index.php/JPAUD/article/view/857.

Ahmad, M. (2018). Development of Student Misconception Test Instrument Using Combination of Three-Tier Test and Certainty of Response Index on Animal Kingdom Grade X High School. *Bioedu Berkala Ilmiah Pendidikan Biologi*, *7*(2), 119–128. https://ejournal.unesa.ac.id/index.php/bioedu/article/download/28756/26328.

Aisyah, S. N., Taena, L., & Ili, L. (2021). Pengembangan Tes Hasil Belajar Ekonomi Kelas X SMA melalui Google Formulir. *Jurnal Online Program Studi Pendidikan Ekonomi*, *6*(3), 106--114. http://repository.uinsu.ac.id/id/eprint/13576.

Apriani, M. S. (2016). Jenis-Jenis Kesalahan dalam Menyelesaikan Masalah Statistika Inferensia pada Mahasiswa Program Studi Pendidikan Biologi 2014 / 2015. *Seminar Nasional Matematika X Unes*, 204–212.

Arsanti, M. (2018). Pengembangan Bahan Ajar Mata Kuliah Penulisan Kreatif Bermuatan Nilai-Nilai Pendidikan Karakter Religius Bagi Mahasiswa Prodi Pbsi, Fkip, Unissula. *KREDO : Jurnal Ilmiah Bahasa Dan Sastra*, *1*(2), 71–90. https://doi.org/10.24176/kredo.v1i2.2107.

Artha, W. A., & Putra, dewa bagus ketut ngurah. (2021). Pengembangan Media Audio Visual Dengan Model Analysis Design Development Imlementation Evaluation Pada Muatan IPA Kelas IV SD Negeri 2 Pejeng Tahun Ajaran 2020/2021. *Jurnal Edutech Undiksha*, *9*(1), 149–157. https://doi.org/10.23887/jeu.v9i1.32974.

Birch, C., Lichy, J., Mulholland, G., & Kachour, M. (2017). An enquiry into potential graduate entrepreneurship: Is higher education turning off the pipeline of graduate entrepreneurs? *Journal of Management Development*, *36*(6), 743–760. https://doi.org/10.1108/JMD-03-2016-0036.

Cahyadi, R. A. H. (2019). Pengembangan Bahan Ajar Berbasis Addie Model. *Halaqa: Islamic Education Journal*, *3*(1), 35–42. https://doi.org/10.21070/halaqa.v3i1.2124.

Cahyawati, D., Wahyudin, & Prabawanto, S. (2020). Students' statistical reasoning reviewed from academic internal factors in two learning classes. *Journal of Physics: Conference Series*, *1480*(1). https://doi.org/10.1088/1742-6596/1480/1/012051.

Dwiqi, G. C. S., Sudatha, I. G. W., & Sukmana, A. I. W. I. Y. (2020). Pengembangan Multimedia Pembelajaran Interaktif Mata Pelajaran IPA Untuk Siswa SD Kelas V. *Jurnal Edutech Undiksha*, *8*(2), 33. https://doi.org/10.23887/jeu.v8i2.28934.

Fahrurrozi, M., & Laili Rahmawati, S. N. (2021). Pengembangan Model Instrumen Evaluasi Menggunakan Aplikasi Kahoot Pada Pembelajaran Ekonomi. *Jurnal PROFIT Kajian Pendidikan Ekonomi Dan Ilmu Ekonomi*, *8*(1), 1–10. https://doi.org/10.36706/jp.v8i1.13090.

Farahiba, A. S. (2022). Pengembangan Instrumen Tes Literasi Peserta Didik pada Materi Teks Anekdot. *Jurnal Dimensi Pendidikan Dan Pembelajaran (JDPP)*, *10*(2), 146–

154. https://litabmas.umpo.ac.id/index.php/dimensi/article/view/4554.
Fardillah, F., Sutaagra, O., Supriani, Y., Farlina, E., & Priatna, N. (2019). Developing Statistical Reasoning Ability of Industrial Engineering Students Through Experiential Learning. *Journal of Physics: Conference Series*, *1179*(1). https://doi.org/10.1088/1742-6596/1179/1/012068.
Firda, H., & Nurhadi, D. (2023). Penerapan Model ADDIE Dalam Pengembangan Instrumen Penilaian Diri Sendidri Peserta Didik SMA Negeri Kabupaten Mojokerto. *Jurnal Hikari*, *7*(1), 14–26. https://ejournal.unesa.ac.id/index.php/hikari/article/view/50739/41577.
Gamaliel, F., & Arliyanto, P. Y. D. (2021). Perancangan Aplikasi Ujian Online Berbasis Website. *Jurnal Manajamen Informatika Jayakarta*, *1*(4), 270. https://doi.org/10.52362/jmijayakarta.v1i4.503.
Halim, A., Mustafa, Nurulwati, Soewarno, & Nanda, N. (2018). Development of Two-Tier Diagnostic Test Based on E-Learning. *Journal of Physics: Conference Series*, *1120*(1). https://doi.org/10.1088/1742-6596/1120/1/012030.
Inayah, N. (2017). The Influence of Mathematical Reasoning and Cognitive Style Toward Student's Comunication and Connection Ability to the Statistical Topic in Class XI Exact of Public Senior High School Palu. *Jurnal Daya Matematis*, *5*(1), 120–128. https://pdfs.semanticscholar.org/db4f/22887efb1746cf1fdf6d6478cf71de7be4a2.pdf.
Jasman, J., & Sidin, M. (2018). Development of Instrument of Physics Learning Result of Students at SMAN 16 In Makassar. *Prosiding Seminar Nasional Fisik*, *1*(1), 1–9. http://journal.staihubbulwathan.id/index.php/alishlah/article/view/4717.
Jumrah, Rukli, & Sulfasyah. (2023). Pengembangan Instrumen Tes Berbasis HOTS dengan Pendekatan Pengukuran Rasch pada Pelajaran Matematika Topik Bangun Ruang untuk Siswa Sekolah Dasar. *Jurnal Basicedu*, *7*(1), 11–27. https://doi.org/10.31004/basicedu.v7i1.4207.
Jusuf, H. (2019). Pengembangan E-Test sebagai Sarana Belajar Mandiri Peserta Didik Berbasis Moodle. *Progresif: Jurnal Ilmiah Komputer*, *5*(1), 45–50. http://ojs.stmik-banjarbaru.ac.id/index.php/progresif/article/view/325.
Kamaruddin, Djabba, R., & Fajar. (2020). PKM Pelatihan Pengembangan Instrumen Tes Bagi Guru SDN Wilayah II Kecamatan Tanete Rilau Kabupaten Barru. *PENGABDI: Jurnal Hasil Pengabdian Masyarakat*, *1*(1), 77–83. http://journal.staihubbulwathan.id/index.php/alishlah/article/view/4717.
Kamila, S., & Prihatiningtyas, S. (2022). Instrumen E-Test Sebagai Inovasi Penilaian Berbasis Website. *Exact Papers in Compilation*, *4*(3), 603–608. http://ojs.unwaha.ac.id/index.php/epic/article/view/806.
Khoridah, F., Prasetiyawati, D., & Baedowi, S. (2019). Analisis Penerapan Metode SAS (Struktural Analitik Sintetik) dalam Kemampuan Menulis Permulaan. *Journal for Lesson and Learning Studies*, *2*(3), 396–403. https://doi.org/10.23887/jlls.v2i3.19899.
Krisnanto, I., Kurniawan, E. S., & Maftukhin, A. (2021). Pengembangan Instrumen Tes Hasil Belajar Fisika Berbasis STEM untuk Mengukur Keterampilan Proses Sains Peserta Didik. *Jurnal Inovasi Pendidikan Sains (JIPS)*, *2*(2), 76–81. https://doi.org/10.37729/jips.v2i2.1375.
Kuncahyono, Suwandayani, B. I., & Muzakki, A. (2020). Aplikasi E-Test "That Quiz" sebagai Digitalisasi Keterampilan Pembelajaran Abad 21 di Sekolah Indonesia Bangkok. *Lectura: Jurnal Pendidikan*, *4*(13–29), 791–792. https://journal.unilak.ac.id/index.php/lectura/article/view/4687.
La Nani, K. (2020). Peningkatan Kemampuan Penalaran Statistis Mahasiswa Melalui Pembelajaran Berbasis Proyek Berbantuan ICT. *Edukasi*, *18*(1), 234. https://doi.org/10.33387/j.edu.v18i1.1617.

Lestari, H. D., & Parmiti, D. P. (2020). Pengembangan E-Modul IPA Bermuatan Tes Online untuk Meningkatkan Hasil Belajar. *Journal of Education Technology*, *4*(1), 73–79. https://doi.org/10.23887/jet.v4i1.24095.

Litna, K. ., Mertasari, N. M. ., & Sudhirta, G. (2021). Pengembangan Instrumen Tes Higher Order Thinking Skills (HOTS) Matematika SMA Kelas X. *Jurnal Penelitian Dan Evaluasi Pendidikan Indonesia*, *11*(1), 10–20. https://repo.undiksha.ac.id/6637/.

Ma'rifah, U., Algiovan, N., & Sutarsyah, C. (2021). An Item Analysis of English Test During Online Learning. *International Journal of Multicultural and Multireligious Understanding*, *8*(12), 647–654. https://doi.org/10.18415/ijmmu.v8i12.3396.

Manfaat, B., & Nurhairiyah, S. (2021). Pengembangan Instrumen Tes untuk Mengukur Kemampuan Penalaran Statistik Mahasiswa Tadris Matematika. *Jurnal Pembelajaran Matematika*, *1*(1), 1–19. https://repository.syekhnurjati.ac.id/1668/.

Martono, A., Yulianjani, A., & Desrianti, D. I. (2020). Online Exam System to Improve Student Learning Quality in State Vocational School 5 Tangerang City. *CCIT Journal*, *13*(1), 65–74. https://doi.org/10.33050/ccit.v13i1.895.

Masitoh, L. F., & Aedi, W. G. (2020). Pengembangan Instrumen Asesmen Higher Order Thinking Skills (HOTS) Matematika di SMP Kelas VII. *Jurnal Cendekia : Jurnal Pendidikan Matematika*, *4*(2), 886–897. https://doi.org/10.31004/cendekia.v4i2.328.

Mastroleo, N. R., Humm, L., Williams, C. M., & Kiluk, B. D. (2020). Initial Testing of a Computer-Based Simulation Training Module to Support Clinicians' Acquisition of CBT Skills for Substance use Disorder Treatment. *Journal of Substance Abuse Treatment*, 114. https://doi.org/10.1016/j.jsat.2020.108014.

Mastuti, E. (2016). Pemanfaatan Teknologi dalam Memyusun Evaluasi Hasil Belajar : Kelebihan dan Kelemahan " Tes Online " untuk Mengukur Hasil Belajar Mahasiswa. *Peneltian Psikologi*, *07*(01), 10–19. https://jurnalfpk.uinsa.ac.id/index.php/JPP/article/view/93.

Muchta, A. (2019). *Teknik Pengumpulan Data Observasi : Pengertian, Metode dan Contohnya*.

Muliyani, T., & Huriaty, D. (2016). Pengembangan Instrumen Tes Geometri dan Pengukuran pada Jenjang SMP. *Math Didactic: Jurnal Pendidikan Matematika*, *2*(2), 91–98. https://doi.org/10.33654/math.v2i2.33.

Munadi, S. (2015). Analisis Validasi Kualitas Soal Tes Hasil Belajar pada Pelaksanaan Program Pembelajaran. *Jurnal Cakrawala Pendidikan*, *1*(1), 145–159. https://doi.org/10.21831/cp.v1i1.4197.

Muspardi, Yusmanila, & Desmariani, E. (2020). Problematika Dosen Dalam Menggunakan Instrumen Penilaian Pendidikan Keterampilan Umum Mahasiswa Berbasis Standar Nasional Tinggi (SN-Dikti). *Jurnal Education and Development*, *8*(4), 54–58. https://journal.ipts.ac.id/index.php/ED/article/view/2077.

Novrianti. (2014). Pengembangan Computer Based Testing (CBT) Sebagai Alternatif Teknik Penilaian Hasil Belajar. *Jurnal Ilmu Tarbiah Dan Keguruan*, *17*(1), 34–42. https://journal3.uin-alauddin.ac.id/index.php/lentera_pendidikan/article/view/514.

Nurfillaili, U., Yusuf, M., & Santih, A. (2016). Pengembangan Instrumen Tes Hasil Belajar Kognitif Mata Pelajaran Fisika pada Pokok Bahasan Usaha dan Energi SMA Negeri Khusus Jeneponto Kelas XI Semester I. *Jurnal Pendidikan Fisika*, *4*(2), 83. https://journal3.uin-alauddin.ac.id/index.php/PendidikanFisika/article/download/3709/3382.

Nurhasanah, N. (2018). Pengembangan Tes untuk Mengukur Kemampuan Penalaran Mahasiswa Mata Kuliah Geometri. *Pepatudzu : Media Pendidikan Dan Sosial Kemasyarakatan*, *14*(1), 62. https://doi.org/10.35329/fkip.v14i1.186.

Nurhikmah, Gani, H. A., Pratama, M. P., & Wijaya, H. (2021). Development of an Android-

based Computer Based Test (CBT) In Middle School. *Journal of Education Technology*, *5*(2), 272–281. https://doi.org/10.23887/jet.v5i2.33527.

Poonpon, K. (2021). Test Takers' Perceptions of Design and Implementation of an Online Language Testing System at a Thai University during the COVID-19 Pandemic. *Pasaa*, *62*(December), 1–28. https://eric.ed.gov/?id=EJ1334994.

Pramana, M. W. A., Jampel, I. N., & Pudjawan, K. (2020). Meningkatkan Hasil Belajar Biologi Melalui E-Modul Berbasis Problem Based Learning. *Jurnal Edutech Undiksha*, *8*(2), 17. https://doi.org/10.23887/jeu.v8i2.28921.

Prisuna, B. F. (2020). Media Development of online-Based Physics Test for Senior High School. *Jurnal Penelitian Ilmu Pendidikan*, *13*(1), 41–51. https://journal.uny.ac.id/index.php/jpip/article/view/30167.

Rahman, D., Suyasa, P. W. A., & Wahyuni, D. S. (2021). Pengembangan Media Pembelajaran E-Learning dengan Model Pembelajaran Flipped Classroom Berbasis Edmodo pada Mata Pelajaran Informatika. *Kumpulan Artikel Mahasiswa Pendidikan Teknik Informatika (KARMAPATI)*, *10*(1), 13. https://doi.org/10.23887/karmapati.v10i1.31092.

Ririen, D., & Hartika, D. (2021). Identifikasi Kesulitan Belajar Mahasiswa pada Mata Kuliah Statistika Selama Masa Pandemi Covid-I9. *Jurnal Ilmiah Universitas Batanghari Jambi*, *21*(1), 148–155. https://doi.org/10.33087/jiubj.v21i1.1236.

Rizki, L. M., & Fauziddin, M. (2021). Studi Kasus pada Mahasiswa yang Mengalami Kesulitan Belajar Matakuliah Statistika. *Jurnal Pendidikan Tambusai*, *5*(3), 11304–11314. https://jptam.org/index.php/jptam/article/view/3010.

Robin Bastian Waruwu, Nayanda Privanezsa Hao, & Prilia Handayani Hia. (2022). Analisis Kemampuan Pemahaman Mahasiswa Pada Mata Kuliah Statistika di STIKes Santa Elisabeth Medan Tahun 2022. *SEHATMAS: Jurnal Ilmiah Kesehatan Masyarakat*, *1*(3), 318–327. https://doi.org/10.55123/sehatmas.v1i3.653.

Rohmah, F. N., & Bukhori, I. (2020). Pengembangan Media Pembelajaran Interaktif Mata Pelajaran Korespondensi Berbasis Android Menggunakan Articulate Storyline 3. *ECOEDUCATION (Economic & Education Journal)*, *2*(2), 169–182. https://doi.org/10.33503/ecoducation.v2i2.892.

Rokhaniyah, H., & Virgantara Putra, O. (2021). Developing Web-Based Online Test System to Boost Ielts Academic Reading Score. *English Review: Journal of English Education Volume*, *9*(2), 235–244. https://journal.iaingorontalo.ac.id/index.php/al/article/view/2504.

Rosidah, R., & Ikram, F. Z. (2021). Measure of Central Tendency: Undergraduate Students' Error in Decision-Making Perspective. *International Journal of Education*, *14*(1), 39–47. https://doi.org/10.17509/ije.v14i1.29408.

Sa'diyyah, F. N., Mania, S., & Suharti. (2021). Pengembangan Instrumen Tes untuk Mengukur Kemampuan Berpikir Komputasi Siswa. *Jurnal Pembelajaran Matematika Inovatif*, *4*(1), 17–26. https://doi.org/10.22460/jpmi.v4i1.17-26.

Sari, V. N. I., Utomo, A. P. Y., & Sumarwati. (2022). Kualitas Soal Bahasa Indonesia di SMP Muhammadiyah 1 Pontianak: Analisis Butir Soal. *Jurnal Pendidikan Bahasa Dan Sastra Indonesia*, *11*(2), 112–119. https://journal.unnes.ac.id/sju/jpbsi/article/view/58091.

Setiyawan, R. A., & Wijayanti, P. S. (2020). Analisis Kualitas Instrumen untuk Mengukur Kemampuan Pemecahan Masalah Siswa Selama Pembelajaran Daring di Masa Pandemi. *Lebesgue: Jurnal Ilmiah Pendidikan Matematika, Matematika, Dan Statistika*, *1*(2), 130–139. https://doi.org/10.46306/lb.v1i2.

Sholihah, I., Fahrurrozi, & Abdul Gafur, H. (2017). Pengembangan Modul Pembelajaran Ekonomi Berbasis Guided Inquiry untuk Meningkatkan Hasil Belajar Siswa Kelas XI

MA NW Sukamulia. *Pembelajaran Ekonomi*, *1*(2), 104–117. http://eprints.hamzanwadi.ac.id/4531/.

Skopljanac-Macina, F., Zakarija, I., & Blaskovic, B. (2021). Towards Automated Assessment Generation in e-Learning Systems Using Combinatorial Testing and Formal Concept Analysis. *IEEE Access*, *9*, 52957–52976. https://doi.org/10.1109/ACCESS.2021.3070510.

Subekti, F. E., & Jazuli, A. (2022). Pengembangan Modul Aplikasi Penalitian Berbasis Penalaran Statistik. *Aksioma: Jurnal Program Studi Pendidikan Matematika*, *11*(1), 1–11. https://www.academia.edu/download/114761136/pdf.pdf.

Sutrisno, & Murtianto, Y. H. (2016). Miskonsepsi Mahasiswa Pada Mata Kuliah Statistika Deskriptif Materi Ukuran Tendensi Sentral. *Media Penelitian Pendidikan, Jurnal Penelitian Dalam Bidang Pendidikan Dan Pengajaran*, *Vol 10*(No 1), 1–14. http://journal.upgris.ac.id/index.php/mediapenelitianpendidikan/article/view/1256.

Tempelaar, D. T., Kuperus, B., Cuypers, H., van der Kooij, H., van de Vrie, E., & Heck, A. (2012). The Role of Digital, Formative Testing in e-Learning for Mathematics: A Case Study in the Netherlands. *RUSC Universities and Knowledge Society Journal*, *9*(1), 284–305. https://doi.org/10.7238/rusc.v9i1.1272.

The, M. M., & Usagawa, T. (2018). Effectiveness of e-learning experience through online quizzes: A case study of Myanmar students. *International Journal of Emerging Technologies in Learning*, *13*(12), 157–176. https://doi.org/10.3991/ijet.v13i12.9114.

Udiasih, Y., & Ghozali, I. (2019). Online Test Design for Computer-Based Final Examination Preparation. *Journal of English Language and Pedagogy*, *2*(1), 57. https://doi.org/10.36597/jelp.v2i1.3619.

Widiyawati, Y., Nurwahidah, I., & Sari, D. S. (2019). Pengembangan Instrumen Integrated Science Test Tipe Pilihan Ganda Beralasan Untuk Mengukur HOTS Peserta Didik. *Saintifika*, *21*(2), 1–14. https://core.ac.uk/download/pdf/297204372.pdf.

Wingga, P., & Saputra, N. A. H. J. (2019). Pengaruh Media Game Edukasi "Teka Teki Pengetahuan" Terhadap Minat Belajar Siswa Pada Mata Pelajaran IPA Kelas 5 SDN 03 Protomulyo. *Jurnal Profesi Keguruan*, *5*(1), 15–22. https://journal.unnes.ac.id/nju/jpk/article/view/10752.

Wirda, M. A., Berutu, N., Rahmad, R., & Rohani. (2017). Pengembangan Tes Standar Berbasis Teknologi Informasi DI Jurusan Pendidikan Geografi. *Jurnal Tunas Geografi*, *6*(2), 101–115. https://digilib.unimed.ac.id/id/eprint/51758/.

Wulandari, I., Hairida, H., Rasmawan, R., Enawati, E., & Muharini, R. (2023). Development of Problems for Higher Order Thinking Skill (HOTS) e-Test on Additive Substance Subject Material. *Hydrogen: Jurnal Kependidikan Kimia*, *11*(5), 698. https://doi.org/10.33394/hjkk.v11i5.8730.

Yuniarti, R. (2022). Kesalahan Mahasiswa Program Studi Administrasi Publik Dalam Menyelesaikan Soal Statistika Deskriptif Dan Statistika Inferensial. *Jurnal Sains Matematika Dan Statistika*, *8*(1), 46. https://doi.org/10.24014/jsms.v8i1.13312.