Nurhasanah, S., Alghadari, F., Nuary, R.H.
Development of Student Statistical Thinking Test Instruments Oriented to the PISA Assessment Framework
DOI 10.37640/ice.1053
In *Proceedings of the 3rd International Conference on Education* (ICE 2024), pages 22-28
ISSN 2964-7479

Development of Student Statistical Thinking Test Instruments Oriented to the PISA Assessment Framework

Siti Nurhasanah[1*], Fiki Alghadari[2], Rusdi Hamdany Nuary[3]
[1,3]Mathematics Education, STKIP Kusuma Negara, Jakarta, Indonesia
[2]Mathematics Education Master's Study Program, Jambi University, Jambi, Indonesia
*sitinurhasanah9@stkipkusumanegara.ac.id

Abstract

This study aims to produce a statistical thinking test instrument that is valid, reliable, and feasible to use for SMK Permata 1 & 2 Bogor students. The instrument was developed through the ADDIE model involving the stages of analysis, design, development, implementation, and evaluation. This instrument is declared valid, reliable, and has adequate variations in difficulty levels and discrimination power. All test items are proven valid, with the r count value greater than r_table indicating that this instrument is able to measure what should be measured. The reliability of the instrument is also high in the first school with a coefficient of 0.774 and the second school with a coefficient value of 0.753, the average of both schools with a coefficient of 0.762, indicating measurement consistency. The level of difficulty in both schools shows that the questions vary, with most questions classified as difficult and one question classified as easy. The discrimination power of the questions in both schools is positive, indicating that each question is able to distinguish students' abilities well. Overall, this instrument is effective and feasible to use to measure students' statistical thinking abilities.

Keywords: Test instrument, thinking, statistics

1 INTRODUCTION

The Abundant Era, or the era of information abundance (Saryadi, W., & Sulisworo, D., 2023), has brought about major changes in the way humans' access and use knowledge. Rapid technological advancements, especially in the fields of information and communication, have accelerated the production and dissemination of information in various forms and mediums. Information that was previously limited and difficult to access can now be easily accessed by anyone, at any time, and anywhere. This change has had a significant impact on the field of education, where individuals no longer rely entirely on conventional learning resources such as printed books or face-to-face teaching in the classroom. Instead, technology has opened up access to abundant digital learning resources, enabling everyone to learn according to their own needs and pace. However, in the midst of this era of abundance, 21st-century skills have become increasingly important to develop in order for students to compete in an increasingly competitive global market. 21st-century skills encompass a variety of abilities needed to face challenges in the digital era, such as technical skills, information management, effective communication, productive collaboration, and creativity in innovation. Additionally, complex thinking skills have also become very important.

Thinking is a dynamic and conscious mental activity in which individuals actively engage in the process of processing information received from various external sources, whether in the form of data, facts, or beliefs. This information is then carefully analyzed,

interpreted, and processed to generate new knowledge that is relevant and can be applied in various everyday life contexts. This thinking process involves a series of complex steps, starting from deeply interpreting data, forming consistent beliefs, to transforming raw information into structured knowledge that can be effectively utilized (Alghadari, F., Sudirman, S., & Kurniasi, E. R., 2022).

In higher-level cognitive learning, students are not only required to passively receive information but also to process it carefully and selectively, connect various concepts they have learned, and develop independent learning strategies. The more connections that are made between various pieces of information and concepts, the more complex the knowledge construction becomes in their minds. The mental processes that occur during thinking activities can vary from simple to highly complex, depending on the number of variables, factors, and components involved in the formation of knowledge. In other words, thinking is not merely the passive processing of information but rather an active activity aimed at generating new understanding, solving problems, and providing critical assessments of various situations and challenges faced. (Wasahua, S., 2021).

In the context of mathematics, particularly statistics, statistical thinking is an essential skill closely related to an individual's ability to collect, analyze, interpret, and present data using various statistical concepts such as mean, median, mode, and standard deviation. This skill has become especially crucial in today's information era, where access to data and information is quick and easy. Statistical thinking enables individuals to make wiser, data-driven decisions, with the ability to critically evaluate various pieces of information and provide appropriate solutions based on relevant evidence. Thus, statistical thinking not only involves mastering basic statistical concepts but also applying that understanding in real-world situations, whether in academic, professional, or everyday contexts. This statistical thinking ability is also a key component of critical thinking skills, which are highly necessary to face complex challenges in the modern world, where making the right decisions often depends on the ability to analyze and interpret data accurately and objectively (Putri, H., 2022).

The importance of statistical thinking skills is widely recognized, with international surveys such as the Programme for International Student Assessment (PISA) showing that students' statistical thinking abilities in Indonesia are still relatively low. Only about 18% of students are able to effectively evaluate problem-solving strategies based on the data they possess (OECD., 2023a). Research conducted in various regions of Indonesia, such as Medan, Palembang, and West Lombok, reinforces these findings, showing that many students still struggle to understand and apply statistical concepts, especially in the context of non-routine problems that require deeper statistical reasoning (Ramadhani, R., & Narpila, S. D, 2018; Nisa, S., Zulkardi, Z., & Susanti, E, 2019; M. Masjudin, et.al, 2020). This reflects an urgent need for more focused and intensive educational interventions to improve students' statistical thinking skills in Indonesia. The goal is for students not only to understand statistical concepts theoretically but also to be able to apply them in real-life situations.

This research aims to develop a statistical thinking test instrument oriented toward the PISA assessment framework, with the hope that this instrument can be used to measure and effectively enhance students' statistical thinking abilities. Through this study, it is expected that a tool will be produced to help educators identify areas where students still face difficulties and design more effective learning strategies to improve statistical thinking skills among students. Ultimately, the enhancement of statistical thinking skills is expected to not only improve students' learning outcomes in mathematics but also their ability to face broader challenges in this increasingly complex information era (Alman, et.al., 2023).

2 RESEARCH METHODS

This research employs a Research and Development (R&D) method using the ADDIE development model, aiming to create a test instrument that can measure the statistical

thinking abilities of 11th-grade vocational high school (SMK) students. The steps in the ADDIE model consist of several phases: (1) Analysis, where the needs for instrument development and the characteristics of the current curriculum are examined; (2) Design, which involves planning the format of the test, determining the time required for completion, developing a test blueprint, and writing the test items; (2) Development, during which the test instrument is revised and reviewed based on feedback from validators; (4) Implementation, where a trial of the test instrument is conducted with students in vocational high schools; (5) Evaluation, which involves analyzing the test items to measure their validity, reliability, difficulty level, and discrimination power.

Data processing techniques in this study include, (1) Validity Test: Using the content validity formula and Aiken's validity index to determine the level of validity of the test items. (2) Reliability Test: Calculating the reliability coefficient using Cronbach Alpha to measure the consistency of the test instrument. (3) Difficulty Level Test: Analyzing the level of difficulty of the questions by calculating the proportion of students who answered correctly. (4) Differentiation Power Test: Measuring the difference in the proportion of correct answers between groups of high and low ability students.

The data obtained from the trial will be analyzed using statistical tools to produce information regarding the quality and effectiveness of the developed test instrument.

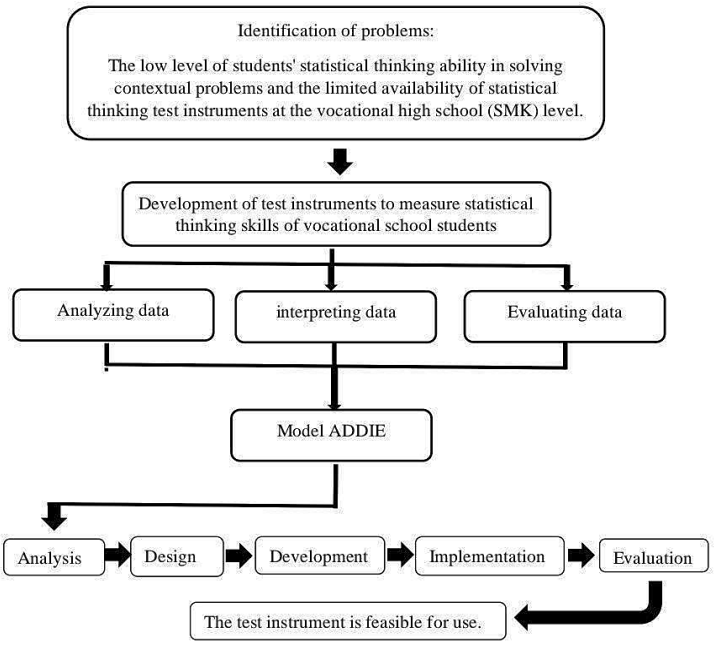

Figure 1. Research Design

3 RESULT AND DISCUSSION

This development research aims to produce a product in the form of a test instrument to measure the statistical thinking abilities of students at SMK Permata 1 & 2 Bogor. The development stages of this test instrument follow the ADDIE development model (*Analysis*, *Design*, *Development*, *Implementation*, And *Evaluate*) (Fatirul & Walujo, 2022).

3.1 Analysis

In this research, there are two main analysis stages carried out to get a comprehensive picture of the learning process that takes place at school. The first stage is curriculum analysis & the second stage is learning resource analysis

3.1.1 Curriculum Analysis

The first stage is curriculum analysis, where researchers examine and evaluate the curriculum used by the school. This step involves identifying the curriculum structure, learning objectives, and competencies that students are expected to achieve. This analysis aims to determine the suitability of the curriculum to national education standards and its relevance to students' needs in daily learning, especially in mathematics subjects.

3.1.2 Analysis of Learning Resources

The second stage in this research is analysis of learning resources, which is an important aspect of the learning process. This stage focuses on various materials and media used by teachers and students during teaching and learning activities. Researchers conducted a comprehensive review of various types of learning resources, such as textbooks used as the main guide, learning modules designed to deepen students' understanding, and student worksheets (LKS) used.

The quality of learning resources is also seen from the aspect of the suitability of their use in achieving the competencies targeted by the curriculum. Researchers reviewed the extent to which these sources were able to facilitate students' understanding of subject matter, especially in mathematics subjects. In this case, it is important to ensure that the learning resources used not only meet academic standards, but also provide a variety of methods that help students with various learning styles. This analysis aims to provide recommendations for improving the quality of learning resources so that they are more in line with student needs and educational developments.

3.2 Design

In designing a statistical thinking test instrument, what the researcher carried out included: (1) determining the form of the test. In this stage the researcher used the form of a description test, where the description test allows students to express students' ideas, understanding and knowledge in written form. (2) determine the length of time to complete the test, generally the test is carried out for 90 to 120 minutes, the duration of the mandatory mathematics lesson hours at the school under study is also an important consideration in determining the time for the test. At the vocational school level, one lesson hour has a duration of 45 minutes. Based on this, the researcher decided that the time given to complete the test was 90 minutes or the equivalent of two class hours. (3) compiling a grid, there are several things that must be considered in compiling the grid, such as the aspects to be tested, indicators for each aspect, question indicators and question items. (4) writing question items, in writing question items based on aspects and indicators of statistical thinking in statistical material.

3.3 Development

At the development stage, the questions that have been prepared will go through a validation process by teachers and lecturers who have expertise in the field of mathematics and mathematics education. The assessors consisted of one mathematics teacher at SMK Permata 2 Bogor and two lecturers in the mathematics education study program at STKIP Kusuma Negara. The results of the expert assessment can be seen in table 1.

Table 1. Expert validation results

Appraiser	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9
I	4	4	4	4	4	4	5	4	4
II	4	4	4	5	5	4	5	4	4
III	4	4	4	4	4	3	4	3	3
V	0,75	0,75	0,75	0,83	0,83	0,67	0,92	0,67	0,67

DOI 10.37640/ice.1053
In *Proceedings of the 3rd International Conference on Education* (ICE 2024), pages 22-28
ISSN 2964-7479

After being analyzed using content validity (aiken), Table 1 shows the assessment results showing that the items analyzed have good validity, with a validity value of 0.76.

3.4 Implementation

After the test instrument was validated, the next step was to test the instrument on 62 students from two different classes at the vocational school level. This trial aims to see how effective and appropriate the test instrument is in measuring students' statistical thinking abilities. Each student is asked to take a statistical thinking test consisting of five descriptive questions, with an allocated time of 90 minutes. Before the test begins, the researcher provides a detailed explanation of the procedures and instructions for working on the questions, and reminds students to read the instructions carefully so they can do the questions well and as expected. After the explanation is complete, students start working on the test with full concentration and focus, following the specified time. Documentation of the test instrument trial implementation can be seen in Figure 2.

Figure 2. Test instrument trial

3.5 Evaluation

The final stage is the evaluation stage, which is carried out by analyzing the results of trials that have been carried out by students. This analysis aims to determine the quality of the test instrument or the level of suitability of the instrument based on four testing criteria, including validity test, reliability test, difficulty level test, and differential power test.

3.5.1 validity test

The test results in class XI BDP show that the first item has a calculated r value of 0.819 and is in the very high category, which shows a very strong level of validity. The second, third and fourth items have calculated r values ranging from 0.744 to 0.749, in the high category, which also shows a very good level of validity. Meanwhile, the fifth item with a calculated r value of 0.597 is categorized as moderate, which shows that the level of validity is sufficient, although not as strong as the other items.

The trial of the test instrument in class third, 0.737 for the fourth item, and 0.652 for the fifth item. These values are then compared with the set value of 0.349. From the results of this comparison, it can be seen that all values are greater than the value.

3.5.2 Reliability Test

The reliability coefficient resulting from this data is 0.762 which indicates a high level of reliability. The value used as a comparison is 0.250. Because the reliability coefficient value is greater than, it can be concluded that this measurement instrument is reliable**.**

3.5.3 Difficulty Level

Comparison of instrument trials in classes XI BDP and For questions number 2, 3, 4, and 5, both are difficult and have low discrimination scores.

3.5.4 Different Power

In the differential power test, a test instrument is said to be good if the item has a differential power value of more than 0. This shows that the item has at least sufficient differential power. The results of calculating the different power of the test instrument show that of the five questions all the different power values are above 0, this means that there are no questions that must be rejected.

4 CONCLUSION

This research produces a statistical thinking test instrument that is valid, reliable and suitable for use for students at SMK Permata 1 & 2 Bogor. The instrument was developed through the ADDIE model which involves the stages of analysis, design, development, implementation and evaluation. This instrument is declared valid, reliable, and has adequate variations in levels of difficulty and differentiation. All question items were proven to be valid, with the calculated r value being greater than the r table which shows that this instrument is able to measure what it is supposed to measure. The reliability of the instrument is also high in the first school with a coefficient of 0.774 and the second school with a coefficient value of 0.753, the average of the two schools with a coefficient of 0.762, indicating consistency of measurement. The level of difficulty in the two schools shows that the questions vary, with most questions classified as difficult and one question relatively easy. The differentiating power of the questions in the two schools shows positive, indicating that each question is able to differentiate students' abilities well. Overall, this instrument is effective and suitable for measuring students' statistical thinking abilities.

5 REFERENCES

Alghadari, F., Sudirman, S., & Kurniasi, E. R. (2022). Student thinking engagement: The status of "being" in learning. *Eureka: Journal of Educational Research*, *1*(1), 29-40.

Alman, A., Herman, T., Prabawanto, S., & Kurino, Y. D. (2023). Statistical Literacy in Elementary Mathematics Learning Through the 2013 Curriculum. *Elementaria Edukasi Journal*, *6*(3), 1454-1466.

Fatirul, A. N., & Walujo, D. A. (2022). *Research Methods for Learning Field Development (special edition for education students and educators)*. Pascal Books.

Masjudin, M., Muzaki, A., Abidin, Z., & Ariyanti, I. A. P. (2020). Analysis of student's statistical thinking ability in understanding the statistical data. *Journal of Physics: Conference Series*, *1521*(3), 032063.

Nisa, S., Zulkardi, Z., & Susanti, E. (2019). Students' statistical reasoning abilities in histogram data presentation material through PMRI learning. *Journal of Mathematics Education*, *13*(1), 21-40.

OECD (2023a), PISA 2022 Results (Volume I): The State of Learning and Equity in Education, PISA, OECD Publishing, Paris, https://doi.org/10.1787/53f23881- en

Putri, H., Susiani, D., Wandani, N. S., & Putri, F. A. (2022). Instrument for assessing cognitive learning outcomes in description tests and objective tests. *Papeda Journal: Journal of Basic Education Publications*, *4*(2), 139-148.

Ramadhani, R., & Narpila, S. D. (2018). Problem-Based Learning with Local Cultural Characteristics as an Alternative to Improve Students' Statistical Thinking Ability. In *Proceedings of the STKIP PGRI West Sumatra National Seminar, Padang, West Sumatra: STKIP PGRI West Sumatra* (pp. 142-50).

Saryadi, W., & Sulisworo, D. (2023). Development of e-module based on the discovery learning to improve the student creative thinking skills. *JTAM (Journal of Mathematics Theory and Applications)*, *7*(1), 11-22.

Wasahua, S. (2021). The concept of developing critical thinking and creative thinking of students in elementary schools. *Educational Horizons*, *16*(2), 72-82.