The Development of Validation, Practicality, and Effectiveness Instruments in Research and Development: Identifying and Mapping Common Errors

Budi Halomoan Siregar[1*], Kairuddin[2], Roslah Arsad[3], Asmin[4], Alpina Rosalia[5]

{budihalomoan@unimed.ac.id[1], kairuddin@unimed.ac.id[2], rosla280@uitm.edu.my[3], asmin@unimed.ac.id [4], alpina@unimed.ac.id[5]}

[1,2,3]Department of Mathematics Education – Universitas Negeri Medan, Medan, Indonesia
[4]Mathematical Sciences Studies, College of Computing, Informatic and Mathematics – Universiti Teknologi Mara (UiTM), Malaysia

Abstract. High-quality research outcomes depend on the use of credible research instruments. However, this study reveals a disparity between the demand for quality educational products and the accuracy of the instruments used in their development. The research aims to identify prevalent errors in the construction of validation, practicality, and effectiveness instruments in undergraduate and master's theses focus on the development of educational products. This study uses a descriptive and qualitative method to analyze documents to identify and map the common mistakes found in these works. The study reveals four primary categories of errors: (1) inappropriate language in statement or questions items; (2) misalignment between items and the instrument type; (3) insufficient verification of the number of questions or statements; and (4) the improper choice of measurement scales and evaluations. These findings emphasize the necessity of providing specific assistance to students in order to assist them in creating reliable research instruments, which will eventually enhance the quality of educational research outcomes.

Keywords: Validation Instruments, Practicality Instruments, Effectiveness Instruments, Research and Development, Qualitative Method.

1 Introduction

In the context of development research, instruments designed to measure the validity, practicality and effectiveness of products are crucial components that determine the quality and reliability of research results [1,2]. However, in practice, students often experience difficulties in developing instruments that meet these standards, which has a negative impact on the quality of the products produced [3]. This problem prompts the need for an in-depth analysis of the errors that often occur in the development of validity, practicality and effectiveness instruments in development research, so as to provide appropriate solutions to improve the quality of research [4].

The definition of product quality standards in this study refers to the opinion of Nieveen [1, 5]. Instrument validity refers to the extent to which the instrument actually measures the

ICONSEIR 2024, November 28, Medan, Indonesia

DOI 10.4108/eai.28-11-2024.2355344

intended concept or variable [5, 6]. To ensure the high quality of the material, the content itself (the planned curriculum) must be well-designed. Nieveen states that the material should be based on current knowledge (content validity), and all its components should be coherently related (construct validity) [1, 2]. If these criteria are met, the product is considered valid. Instrument validity includes content validity, which is the extent to which the content on the instrument covers all relevant aspects of the material or variable being measured [5, 7, 8]. The second characteristic of high-quality materials is that teachers (and other experts) find the materials easy to use, so that teachers and students can use them in ways that are in line with the developers' objectives [9]. That is, there must be consistency between the curriculum as designed and understood and the curriculum as implemented and operated [1, 5]. In the case of multimedia development, practicality includes factors such as device compatibility and resource affordability [10, 11, 12]. The third characteristic of high-quality materials is that students appreciate the learning programme and that the expected learning actually occurs [5, 13, 14]. With effective materials, alignment is achieved between the expected and piloted curriculum, as well as between the expected and successfully achieved curriculum [1,5]. In this study, the effectiveness instrument is only limited to the student response instrument to the developed product.

Based on the results of the analysis of errors in developing validation instruments, practicality and student responses, the main problems in developing instruments by researchers include: (1) errors in formulating questions or statements on validation, practicality, and effectiveness instruments. In many cases, students could not distinguish clearly between the elements that should be included in each instrument, such as including aspects of ease of use that should be in the practicality instrument into the validation instrument. (2) Question items or statements that are not relevant to the type of instrument (measurement purpose). (3) Insufficient number and coverage of questions or statements in the instrument, which resulted in the instrument not being able to accurately measure the validity, practicality, and effectiveness aspects of the developed product. (4) Errors in determining the measurement scale. This error occurs when the measurement scale used in the instrument is inappropriate or inadequate to measure the variables under investigation.

Addressing this issue, possible solutions include intensive training for students on the principles of effective instrument design, with a focus on an in-depth understanding of validity, practicality, and effectiveness. This training should include practical guidance on how to appropriately formulate questions or statements according to different instrument purposes. In addition, the development of comprehensive guides and resources that provide examples of tested instruments is also an important step. These guides should provide templates and real-life examples of instruments that have been successfully implemented in research, so that students have a clear reference point and can adapt them for their own context.

While several studies on instrument development have been conducted, there is still a need to formulate practical guidelines and hands-on applications that can be applied by university students [16, 16]. This study aims to provide an in-depth analysis of frequent errors in instrument development for validation, practicality, and effectiveness, and offer solutions that can improve university students' understanding and skills in designing effective instruments. The urgency of this study lies in its efforts to improve the quality of academic research by identifying and addressing fundamental problems in instrument development, which will ultimately contribute to improving the quality of research outcomes and practical relevance in the field of mathematics education and development research in general.

2 Research Method

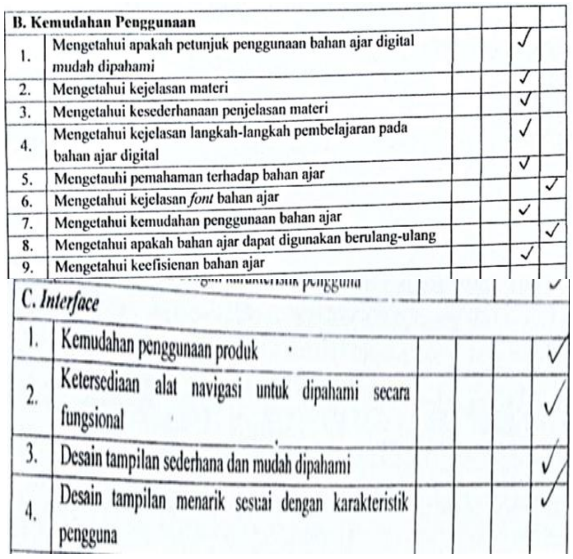
The object studied was the validity, practicality and effectiveness instruments contained in the theses of undergraduate and postgraduate students. This study used a qualitative descriptive method to analyse errors in the development of validation, practicality, and effectiveness instruments in research and development [17]. This method is carried out through the following stages: (1) Data collection: the first stage involved collecting thesis documents from undergraduate and postgraduate students related to the development of educational products. These documents were drawn from various sources, including university libraries, academic repositories and research archives. This collection aimed to obtain a representative sample of relevant academic products [18]. (2) Document classification: once the documents were collected, the next step was to classify the theses relevant to the type of educational product development research. The classification process involves assessing the relevance of the documents based on certain criteria, such as the type of educational product developed and the research focus of the validation, practicality and effectiveness instruments [19]. (3) Error analysis: at this stage, the researcher analysed the types of errors found in the validation, practicality and effectiveness instruments in the theses. This analysis involved identifying common errors in the formulation of questions or statements, insufficient number and scope of questions, and mismatch between the instrument and the research context. (4) In-depth Interviews: to obtain more in-depth information regarding the errors found, the researcher conducted interviews with several thesis writers. These interviews aimed to further explore the causes of the errors, the challenges faced in developing the instruments, and their views on possible solutions. (5) Data generalisation and interpretation: the data obtained from document analysis and interviews were then used to generalise the findings and interpret the meaning of the data found. The findings were processed to understand existing patterns of errors and provide insights into best practices in instrument development.

The research data was analysed through four stages as follows: (1) Reducing data: data from document analysis and interviews were reduced by filtering out relevant and significant information. This process involved grouping the data based on the categories of errors identified, as well as eliminating data that were not relevant to the focus of the study. (2) Displaying data: the data that has been reduced is then displayed in the form of narratives, tables, pictures, or charts. This data presentation aims to facilitate understanding and visualisation of the findings, as well as to show the relationship between the types of errors and the context of instrument development. (3) Interpreting the data: this stage involves interpreting the displayed data to identify patterns, trends and meanings of the errors found. This interpretation is based on in-depth analyses of the findings and interviews, as well as related literature. (4) Summarising the data: conclusions are drawn based on the analysis and interpretation of the data. These conclusions include a summary of the errors found, recommendations for improvements in instrument development, and implications for future research practices. Using this method, the study aims to provide a comprehensive understanding of errors in instrument development and offer practical solutions to improve the quality of research instruments at the academic level.

3 Results and Discussion

To obtain data for this study, the researcher implemented the following steps: (1) collecting undergraduate and postgraduate student research instruments, (2) selecting instruments related to the development of educational products, (3) selecting instruments developed based on the concept of Nieveen's quality standards. At this stage, the remaining instruments that met the criteria consisted of 10 instruments derived from undergraduate student researchers and 10 undergraduate students, (4) reading and analysing the instruments, (5) interpreting student errors in developing instruments, (6) interviewing students to obtain in-depth information, and (7) reducing data in the form of tables and diagrams.

After collecting, sorting and classifying the type of thesis on the development of an educational product, the researcher analysed the validity, practicality and response instruments in detail. A snapshot of the questionnaire is outlined in figure 1.



B. Kemudahan Penggunaan					
1.	Mengetahui apakah petunjuk penggunaan bahan ajar digital mudah dipahami			✓	
2.	Mengetahui kejelasan materi			✓	
3.	Mengetahui kesederhanaan penjelasan materi			✓	
4.	Mengetahui kejelasan langkah-langkah pembelajaran pada bahan ajar digital			✓	
5.	Mengetauhi pemahaman terhadap bahan ajar			✓	
6.	Mengetahui kejelasan *font* bahan ajar				✓
7.	Mengetahui kemudahan penggunaan bahan ajar			✓	
8.	Mengetahui apakah bahan ajar dapat digunakan berulang-ulang				✓
9.	Mengetahui keefisienan bahan ajar			✓	
	[illegible]				✓

C. *Interface*					
1.	Kemudahan penggunaan produk				✓
2.	Ketersediaan alat navigasi untuk dipahami secara fungsional				✓
3.	Desain tampilan sederhana dan mudah dipahami				✓
4.	Desain tampilan menarik sesuai dengan karakteristik pengguna				✓

Fig. 1. Snapshot of Product Validation Instrument

After conducting an in-depth analysis of the two validation instrument documents, there were several errors in developing the instrument: (1) the instrument included the easy to use indicator in the validity instrument. In contrast, based on Nieveen's definition, the usefulness and ease of use of the developed product should be included in the measurement indicators of product practicality [1, 5]. (2) On the validity instrument, there is a statement ‘ease of use of the product’. The statement should be in the practicality instrument. In addition, the statement is considered to be general, so that the purpose of measuring practicality is not specifically measured, which can reduce the accuracy of evaluating the practicality of the product [20, 21].

Based on the results of the interviews, this error was caused by the students' incomprehension in distinguishing between the concepts of validity and practicality in the context of development research. Instrument validity should assess the extent to which the instrument can measure the variables to be evaluated, while practicality measures operational aspects, such as ease and usefulness in using the product. In this case, the indicator of ‘ease of

use' in validity indicates a confusion of concepts that can affect the final results of the study. If students do not construct the instrument appropriately, the results of the analysis will be invalid because they do not reflect the aspects to be measured [22, 23].

Secara keseluruhan konsep tampilan modul digital berbasis *microlearning* mampu meningkatkan minat belajar	Podcast, Poster, Video, dan ppt dalam modul pembelajaran mobile mempermudah guru dalam menyampaikan konsep materi pda siswa
Penataan materi dalam modul digital ini menjadikan materi tidak menarik.	Modul pembelajaran mobile membuat proses pembelajaran menjadi menyenangkan.

Fig. 2. Snapshot of the m4s1 and m6s2 product practicality instruments

After analysing the practicality instruments m4s1 and m6s2, it was found that some statements on the instruments were more likely to measure attractiveness or the level of user pleasure with the developed product. For example, in the statement 'the arrangement of material in the digital module makes the material uninteresting' or 'the mobile learning module makes the learning process fun'. The statement shows the measurement of attractiveness or the level of user pleasure towards the digital module. In fact, based on Nieveen's theory, the practicality instrument must evaluate two aspects, namely 'usable' and 'easy to use' [1, 5]. Thus, measuring users' attractiveness or level of interest in the digital module should not be included in the practicability indicator. Instead, the indicator is more appropriate if it is included in the product user response instrument.

The aspect of practicality that is associated with attractiveness can obscure the measurement process. This can lead to bias in the evaluation results, because practicality should measure the usability and easy to use aspects of a product, not how attractive the product is. Instruments that assess users' "attractiveness" or " excitement" towards the developed product should stand alone in the category of user response instruments, so as not to interfere with the objectivity of practicality measurement emphasised by Nieveen [1, 5].

No.	Pernyataan	Alternatif Pilihan				
		SS	S	R	TS	STS
1.	Bahasa yang digunakan dalam LKPD memudahkan saya memahami materi Pola Bilangan.					
2.	Istilah-istilah yang terdapat dalam LKPD sulit untuk dimengerti.					
3.	Bahasa yang digunakan dalam LKPD mudah dipahami.					
1.	Bahan ajar digital berbasis pendekatan saintifik memberikan kesempatan untuk belajar sesuai dengan kecepatan berpikir saya.					
2.	Saya dapat mengulang-ulang membaca materi yang belum saya pahami dalam modul ini sendiri.					
3.	Bahan ajar digital berbasis pendekatan saintifik mampu menjelaskan konsep materi kubus dengan baik.					
4.	Penjelasan materi kubus dalam bahan ajar digital dijabarkan secara baik dan jelas.					

Fig. 3. Snapshot of student response questionnaire

After analysing the user response instrument, it was found that some statements in the instrument tended to measure the content validity or practicality of the product. Examples of statements such as 'the language used in the LKPD makes it easier for me to understand the Number Pattern material' and 'the language used in the LKPD is easy to understand' show that

the focus is more on the content aspect or ease of use and accessibility, which are indicators of validity or practicality in a product.

A statement such as 'the language used makes it easy for me to understand the material' is actually more in line with the practicality indicator, as it measures the extent to which the product helps users understand the learning content easily. In contrast, the user response instrument should focus on users' perceptions and emotional experiences, such as interest or attraction to the LKPD, rather than on the functional aspects of the product. This misplacement of practicality indicators on the user response instrument may result in a biased evaluation, where functional aspects and user perceptions are mixed, resulting in inaccurate measurement results. Consequently, the concept of evaluation proposed by Nieveen becomes unclear, because practicality and user response should be evaluated through different instruments to achieve a more focused and accurate assessment [1, 5].

After analysing the instrument thoroughly, the researcher found several types of errors experienced by students. Broadly speaking, these types of errors are categorised into 4, namely: (1) Language errors in compiling statement or question items, (2) Question or statement items that are irrelevant to the purpose of measurement (type of instrument), (3) The number of questions or statements does not meet the adequacy requirements of an instrument validity, practicality or effectiveness, and (4) determination of measurement and assessment scales in research instruments. The four types of errors and error codes are outlined in Table 1.

Table 1. Types of errors in developing validity, practicality and effectiveness instruments

No	Type	Description	Code
1	Language errors in drafting statement or question items	Statements or questions on the instrument contain more than one item of measurement purpose. Questions are not designed to be specific questions with a single purpose. For example, 'How easily did you find the material and how quickly were you able to understand it?' This question combines two different measurement objectives, namely: (1) ease of finding the material and (2) quickness of understanding. This may cause respondents to have difficulty in providing consistent and accurate answers, as they may have different views on the two aspects. This error reduces the validity of the data collected, as the answers given do not clearly reflect the respondents' views on each of the items measured.	1A
		Use of imprecise or unclear language. For example, 'Product quality is adequate'. Without a clear definition of what is meant by 'adequate,' respondents may interpret the term differently. This vagueness in language can result in variations in respondents' interpretations. This can cause the accuracy of the data collected to be less accurate.	1B
		Leading questions. Questions designed to lead respondents to a particular answer can result in biased data. Example: 'How useful is this product in improving concept understanding?'. This question has the potential to encourage respondents to give positive answers.	1C
2	Question or statement items that are not relevant to the purpose of	Questions or statements are not relevant to the purpose of measurement (type of instrument). For example, in the validation instrument there are questions related to the ease of use of the prototype to achieve the research objectives, such as 'How useful is the electronic student worksheet for improving concept	2

	measurement (type of instrument)	understanding?', this question item should tend to be contained in the practicality instrument. Conversely, if the practicality instrument contains a statement 'The suitability of multimedia with relevant learning theories', this statement item should be the focus of the validation instrument. This mismatch can cause the data collected to not reflect the aspects actually measured by each type of instrument.	
3	The number of questions or statements does not fulfil the adequacy requirements of an instrument for validity, practicality or effectiveness.	Lack of Questions or Statements: This error occurs when the instrument does not include enough questions or statements to measure all dimensions of the variable under study. For example, in the instrument for the practicality of multimedia products, there are only a few questions aimed at evaluating the usable aspect and no questions measuring the easy to use aspect. The statements on the instrument do not cover all important dimensions of the practicality aspect of the multimedia. The lack of questions can reduce the data and not reflect all aspects needed for a comprehensive assessment.	3A
		This error occurs when the instrument gives an unbalanced proportion to the various dimensions of the variable being measured. For example, in a multimedia validity instrument, if the instrument focuses more on technical aspects, such as software features, but does not consider pedagogical aspects. This causes the instrument to measure product validity to be incomplete. An imbalance in assessment can lead to biased and inaccurate results.	3B
4	Determination of measurement and rating scales in research instruments	Response scale inaccuracy: this error arises when the response scale is not properly designed to capture the required range of answers. For example, in a practicality instrument, if the response scale only includes 'Easy,' or 'Difficult,' without providing options for more detailed judgements such as 'Fairly Easy,' Neutral, or 'Very Difficult,' then the scale may not be sufficient to accurately capture the level of ease of use. This imprecision may lead to less detailed and uninformative data.	4

Based on the identification of errors committed by students in instrument development, four common types of errors were classified. These errors are summarised in Table 2.

Table 2. Types of errors in validity, practicality and user response instruments.

Student's Initials	**Error Type**			**Total**						
	Validity Instrument	Practicality Instrument	Student Response Instruments	1A	1B	1C	2	3A	3B	4
M1S1	1A, 2	2, 3A	1A, 2	2	0	0	3	1	0	0
M2S1	1A, 1C	2, 3A	2, 3A	1	0	1	2	2	0	0
M3S1	2	2, 4	2	0	0	0	3	0	0	1
M4S1	1C, 2	2, 3A	2	0	0	1	3	1	0	0
M5S1	1A, 2	2, 3B		1	0	0	2	0	1	0
M6S1	2			0	0	0	1	0	0	0
M7S1				0	0	0	0	0	0	0
M8S1	1A, 2	2, 3A	1B	1	1	0	2	1	0	0
M9S1	1C, 2	2, 3A	2	0	0	1	3	1	0	0
M10S1				0	0	0	0	0	0	0
M1S2				0	0	0	0	0	0	0

M2S2	1A, 1B, 2	3A, 3B, 2, 4	2	1	1	0	3	1	1	1
M3S2	1C, 2	2, 3A, 4	2	0	0	1	3	1	0	1
M4S2	1A, 2, 3B	2, 3A	1A, 2	2	0	0	3	1	1	0
M5S2				0	0	0	0	0	0	0
M6S2	2	2	2	0	0	0	3	0	0	0
M7S2				0	0	0	0	0	0	0
M8S2		2	2	0	0	0	2	0	0	0
M9S2	1C, 2	3A	2	0	0	1	2	1	0	0
M10S2	1A, 2, 3B	1A, 2, 3A	1A, 2	3	0	0	3	1	1	0
Total	28	29	17	11	2	5	38	11	4	3

The information in Table 2 shows that there were 28 errors in the validity instrument, 29 errors in the practicality instrument, 17 errors in the student response instrument and 4 instruments with no errors identified. Furthermore, the number of each type of error is shown in Figure 1. There are 18 errors in composing language, 38 aspects of question incompatibility with measurement objectives, 15 errors in the proportion of statement items, and 4 errors in determining the rating scale.

Based on the results of in-depth interviews, it was revealed that the main cause of students to commit language errors in compiling statement items or questions is the lack of in-depth understanding of the concept of instrument quality standards according to Nieveen. Students generally do not clearly understand how the concepts of validity, practicality, and effectiveness should be translated into appropriate language in instrument items. Students often used less measurable sentences to measure certain aspects. Some students also admitted that they found it difficult to interpret the quality indicators according to Nieveen, thus making inaccurate statements in their instruments. They also admitted that guidance and concrete examples related to the concept of instrument development were still quite minimal, so they developed with their limited understanding.

Language errors made by students in preparing statement or question items are generally caused by a lack of understanding of the basic principles of instrument preparation based on clear and measurable quality standards. Students have not been able to identify the need to use language that is objective, precise, and free from ambiguity, which is essential to ensure that items can accurately measure the intended variable. This lack of understanding is due to their limited literacy regarding the concept of quality instrument development. Siregar emphasises the importance of using clear and specific language in item development to illustrate indicators of validity, practicality and effectiveness [3]. In addition, there is a lack of understanding on how to formulate questions that can generate valid and reliable data. This causes students to tend to use sentences that are general or undirected, thus reducing the credibility and accuracy of the instrument. Limited academic guidance is also a determining factor, where students do not get adequate direction to avoid language errors in the process of preparing the research instrument.

Student errors in compiling question items or statements that are irrelevant to the purpose of measurement are generally caused by a lack of deep understanding of the relationship between the concept to be measured and the items that must be compiled in the research instrument. Students do not understand that each item in the instrument must reflect indicators that are in accordance with the variable being measured, such as validity, practicality, or product effectiveness, in accordance with the principles outlined by Nieveen [1, 5]. This ignorance causes students to tend to include questions or statements that are not directly related to the purpose of measurement, so that the resulting instrument becomes biased and cannot provide valid data. In addition, students do not understand the relationship between theory and practice

in instrument development, this is due to the lack of guidance to ensure that the items compiled are truly relevant and on target. These errors reflect students' inability to connect theoretical concepts with practical applications in measurement, which should be the main focus in the process of developing research instruments.

Student errors in determining the number of questions or statements that do not meet the adequacy requirements of an instrument for validity, practicality, or effectiveness are often caused by a lack of understanding of the basic principles in the preparation of instruments that can measure variables comprehensively. Students tend not to realise that to achieve an adequate level of validity, the number of items used in the instrument must be sufficient to cover all dimensions or aspects to be measured, in accordance with the quality standards proposed by Nieveen [1, 5]. This misconception can result in the preparation of instruments with a limited number of questions or no depth, so that the instrument is insufficient to capture the variability or complexity required to properly assess the validity, practicality or effectiveness of the product. In addition, university students often do not consider the importance of the balance between depth and coverage of items in the instrument, leading to instruments that are too narrow or unrepresentative.

Students' errors in determining measurement and assessment scales in research instruments are caused by a lack of understanding of the basic concepts of appropriate measurement scales and the relevance of their use. Students do not fully understand that the selection of an appropriate scale-whether nominal, ordinal, interval, or ratio-should be adjusted to the type of data to be measured and the measurement objectives to be achieved, such as validity, practicality, or effectiveness, in accordance with the theory proposed by Nieveen. This misconception can cause students to choose an inappropriate measurement scale, which can reduce the validity of the instrument and the accuracy of the measurement results. In addition, students often do not realise the importance of consistency in determining the rating scale for each item in the instrument, which serves to provide clear and measurable information about the aspects being tested. The lack of understanding of the aspects of the right scale, coupled with the lack of adequate guidance, are the main factors why students make mistakes in determining measurement and assessment scales.

4 Conclusion

Student errors in developing research instruments are caused by a lack of in-depth understanding of the concept of instrument quality, such as validity, practicality, and effectiveness in accordance with Nieveen's concept. These errors include aspects of language, the suitability of questions to the measurement objectives, the accuracy of the number of items, and the determination of a consistent measurement scale. These errors are also caused by limited literacy related to the concept of instrument quality and the lack of adequate academic guidance. As a result, the instruments that are prepared become biased, irrelevant, and cannot accurately reflect the variables being measured. This can reduce the credibility and accuracy of the research results. The results of this study contribute to strengthening the literature that focuses on instrument quality standards in the context of educational product development, which has not been widely described in practice in higher education. These results emphasize the importance of the role of lecturers in guiding students to better understand the concept of instrument development and how to apply it to produce accurate and reliable instruments. Practically, the findings of this study provide recommendations that the curriculum in the field of higher

education, especially in mathematics education, can include special modules that explore instrument development material. This can be done by implementing intensive training and a more structured guidance system, which will improve student literacy in terms of clarity, accuracy, and relevance of items designed to measure the intended variables validly and reliably. These implications are also relevant to higher education policies that aim to improve the quality of graduates with better research skills. The limitations of this study include the limited scope of the sample of students at one institution, which may affect the generalization of the results to a wider context. In addition, the lack of longitudinal data makes the results of this study unable to measure the development of student understanding in the long term. Therefore, further research that includes a wider context and sample, and uses a longitudinal approach, is needed to enrich and strengthen these findings. Overall, this study emphasizes the importance of a deep understanding of the concept of product quality standards and instrument quality literacy. With a significant contribution to understanding the root causes of effective instrument development, this study is expected to have a positive impact on improving the quality of educational research, thus potentially enriching the development of science at the global level.

Acknowledgement. This research was fully funded by LPPM Unimed through contract number 0016/UN33.8/PPKM/PPT/2024. We would like to thank all parties who have supported this research.

References

[1] Nieveen, N. & Folmer, E. (2013). Formatif Evaluation in Educational Desain Research. In Plomp, T. & Nieveen, N. (Eds) Educational Design Research, Part A: An Introduction. New York: Routledge Nieveen, N. & Folmer, E. (2013). Formatif Evaluation in Educational Desain Research. In Plomp, T. & Nieveen, N. (Eds) Educational Design Research, Part A: An Introduction. New York: Routledge.

[2] Sugiyono (2022). Metode Penelitian & Pengembangan (Research and Development). Bandung: PT. Alfabeta.

[3] Siregar, B. H., Mansyur, A., Rangkuti, S. L D., Rahmadhani, F. (2022). Multimedia Pembelajaran Interaktif. Medan: UMSU Press.

[4] Crotty, M. (1998). The Foundations of Social Research: Meaning and Perspective in The Research Process. Crows-Nest, NSW: Allen & Unwin.

[5] Nieveen, N., (1999). Prototyping to Reach Product Quality. In J. van den Akker, R.M. Branch, K. Gustafson, N. Nieveen, & T. Plomp (Eds), Design Approaches and Tools in Education and Training. Boston: Kluwer Academic, 125-136.

[6] Nieveen, N., McKenney, S. & Van den Akker, J. (2006). Educational Design Research: The Value Of Variety. In: Van den Akker, J., Gravemeijer, K, McKenney, S. & Nieveen, N. (Eds). (2006). Educational Design Research. London: Routledge,151-158.

[7] Plomp, T. (2010). Educational Design Research: An Introduction. In T. Plomp & N. Nieveen (Eds), Educational Design Research-Part A: An Introduction (pp.10-51). Enschede, Netherlands: Netherlands Institute for Curriculum Development (SLO).

[8] Gravemeijer, K., & Cobb, P. (2013). Design Research from the Learning Design Perspective. In Plomp T., & N. Nieveen (Eds.), Educational Design research. Part A: An introduction (pp. 72-113). Enschede: Netherlands: SLO.

[9] M., D., Kruizinga., Frederik, E., Stuurman., V., Exadaktylos., Robert, Doll., D., T., Stephenson., Geert, Jan, Groeneveld., G., J., A., Driessen., Adam, F., Cohen. (2020). Development of Novel, Value-Based, Digital Endpoints for Clinical Trials: A Structured Approach Toward Fit-for-Purpose Validation. Pharmacological Reviews, 72(4):899-909. doi: 10.1124/PR.120.000028.

[10] Chien-Hung, Liu. (2019). 1. A compatibility testing platform for android multimedia applications. Multimedia Tools and Applications, doi: 10.1007/S11042-018-6268-Y.

[11] Ana, Dominguez., Julian, Florez., Alberto, Lafuente., Stefano, Masneri., Inigo, Tamayo., Mikel, Zorrilla. (2020). 3. A Methodology for User Interface Adaptation of Multi-Device Broadcast-Broadband Services. IEEE Access, doi: 10.1109/ACCESS.2020.3039616

[12] Zihe, Song., Yingfeng, Chen., Lei, Ma., Shangjie, Lu., Honglei, Lin., Changjie, Fan., Wei, Yang. (2022). 5. An Empirical Analysis of Compatibility Issues for Industrial Mobile Games (Practical Experience Report). doi: 10.1109/ISSRE55969.2022.00030

[13] Sebastian, Thiele., Robert, Kordts. (2024). 1. You know it, you like it – student appreciation of active learning methods. Scandinavian Journal of Educational Research, doi: 10.1080/00313831.2024.2318461

[14] Xiaoming, Xu., Nancy, Bos., Hongbin, Wu. (2022). 2. The relationship between medical student engagement in the provision of the school's education programme and learning outcomes. Medical Teacher, doi: 10.1080/0142159X.2022.2047168

[15] Cecilia, K., Y., Chan., Lillian, Y., Y., Luk. (2020). 1. Development and validation of an instrument measuring undergraduate students' perceived holistic competencies. Assessment & Evaluation in Higher Education, doi: 10.1080/02602938.2020.1784392

[16] Lisa, Schurer, Lambert., Daniel, A., Newman. (2022). 3. Construct Development and Validation in Three Practical Steps: Recommendations for Reviewers, Editors, and Authors*. Organizational Research Methods, doi: 10.1177/10944281221115374

[17] Creswell, J. W., & Creswell, J. D. (2018). Research Design: Qualitative, Quantitative, And Mixed Methods Approaches. Los Angeles, California: Sage Publications

[18] Johnson, R.B. and Christensen, L.B. (2008). Educational Research: Quantitative, Qualitative, and Mixed Approaches. 3rd Edition, Lose Angeles: Sage Publications, Inc

[19] Denzin, N. K., & Lincoln, Y. S. (1994). Handbook of Qualitative Research. Thousand Oaks, CA: SAGE.

[20] Akker, J., Gravemeijer, K., McKenney S., & Nieveen. (2006). Educational Design Research. London: Routledge Taylor and Francis Group.

[21] Bakker, A. (2018). Design Research in Education: A practical Guide for Early Career Researchers. New York, NY: Routledge

[22] Van den Akker, J. (1999). Principles and Methods of Development Research. In J. van den Akker, R.M. Branch, K. Gustafson, N. Nieveen, & T. Plomp (Eds), Design approaches and tools in education and training. Boston: Kluwer Academic, 1-14.

[23] Van den Akker, J., Gravemeijer, K, McKenney, S. & Nieveen, N. (Eds). (2006). Educational Design Research. London: Routledge. ISBN10: 0-415-39635-2

[24] Sriadhi, S., Hamid, A., Restu, R. (2022). Web-Based Virtual Laboratory Development for Basic Practicums in Science and Technology. TEM JOURNAL - Technology, Education, Management, Informatics, Vol.11, No.1, pp. 396-402.

[25] Sriadhi, S., Restu, R., Sitompul, H. (2021). Multimedia Simulation Model for Electrical Laboratory Learning, In: IoP Conf. Series: Materials Science and Engineering, Vol.1098, doi:10.1088/1757-899X/1098/3/032020.