Journal of Basic Education Research
Vol. 6, No. 3, September 2025, pp. 441~451
ISSN: 2716-3725, DOI: 10.37251/jber.v6i3.2028

Validity and Reliability Analysis of Bullying Comprehension Test Instrument Development

Ivon Faridatus Syahroch[1], Khusnul Khotimah [2]
[1, 2] Elementary School Teacher Education, Faculty of Education, State University of Malang, Malang, Indonesia

Article Info

Article history:

Received Jul 25, 2025
Revised Aug 17, 2025
Accepted Aug 18, 2025
OnlineFirst Sep 22, 2025

Keywords:

Bullying
Elementary School
Reliability
Validity

ABSTRACT

Purpose of the study: To ensure that the instrument measuring elementary students' bullying comprehension is valid and reliable, so that the data obtained is accurate and trustworthy. This study also aims to validate its content to ensure all measured aspects match the intended construct and are suitable for field trials and further research.

Methodology: A descriptive quantitative method is used in this study, with 28 grade 5 students from Sawojajar 1 Elementary School as research subjects and expert validation. The object of this research is a 50-item bullying understanding test instrument. The data obtained were analysed using SPSS with the Pearson Bivariate Correlation technique and Cronbach's Alpha analysis.

Main Findings: This study demonstrates that the developed instrument meets validity and reliability standards, making it effective for measurement purposes. Content validity was verified through expert validation, resulting in a score of 97.5%, which falls into the category of highly valid. The validity of the 50 question items resulted in 28 valid items. The valid items produced a Cronbach's alpha value of 0,919, confirming the reliability and consistency of the instrument.

Novelty/Originality of this study: Developing valid and reliable instruments to measure elementary school students' understanding of bullying, which were previously unavailable in a standardized and empirically tested form. Through validation and reliability testing, this study ensures that the instrument can be used effectively for measurement purposes, thereby improving the accuracy and consistency of the data obtained.

Corresponding Author:
Khusnul Khotimah
Elementary School Teacher Education, Faculty of Education, State University of Malang
Semarang Street number 5, Sumbersari, Lowokwaru, Malang City, East Java 65145 Indonesia
Email: khusnul.khotimah.fip@um.ac.id

1. INTRODUCTION

Bullying is a social problem that is often found in communities, workplaces, cyberspace, and even schools. Victims of bullying tend to withdraw from social interactions. Most victims of bullying choose not to report the bullying they receive because they consider it a common thing and do not require a serious response [1]. The following factors can cause children to commit acts of bullying, which are: 1)peer influence, 2)personal characteristics, 3)having been a victim of bullying, 4)wrong parenting, 5)victims have a quiet personality, 6)want to seek self-gratification, 7)want to gain popularity, and 8)seek social validation [2], [3]. Based on data from the Indonesian Child Protection Commission (KPAI), the most bullying cases occur in elementary school students, with a percentage of 26%. KPAI has received a complaint report of 2.355 cases throughout 2023. There were 87 cases of violence in the education unit environment, with complaints of children as victims of bullying. However, KPAI also mentioned that bullying cases that occur in elementary schools are often not reported to the authorities

[4]. This is because the deviant behavior of elementary school students, such as mocking, hitting, pinching, and tripping their friends, is often not taken seriously by teachers and parents because it is considered a process of immature student development, so that this behavior occurs repeatedly [5]. Bullying that occurs repeatedly can have a lifelong impact on victims of bullying [6]. One of the real impacts of bullying behavior is the reduced ability of social interaction of bullying victims with the surrounding environment, so that it hampers the learning process at school [7]. In addition, students who are victims of bullying tend to experience anxiety, emotional disturbances, and decreased motivation to learn [8]. Given the many negative impacts of bullying behavior, an instrument is needed to measure students' understanding of bullying behavior so that researchers, teachers, and counselors can develop bullying prevention strategies [9].

Although there have been many socialization and learning activities themed on the importance of understanding bullying, there are still many parties who cannot measure students' understanding of bullying accurately and consistently, because there are no valid and reliable instruments available to measure students' understanding of the dangers of bullying. In fact, knowing how the level of understanding of students related to bullying behavior and the impact caused is the initial diagnostic action to prevent bullying [10]. Data obtained from measuring students' level of understanding of bullying can be used to plan bullying prevention strategies [11]. Therefore, a standardized instrument is needed in terms of validity and reliability that can be used to measure students' level of understanding related to bullying so that the data generated is consistent and accurate.

Valid and reliable instruments are very important in scientific research to produce consistent and accurate data. The research instrument functions as a measuring tool that allows researchers to collect information in accordance with the objectives of the study, so the quality of the instrument directly affects the quality of the research results [12]. An ideal instrument must be valid and reliable. Without validity, the instrument will not be able to produce accurate data, and without reliability, the resulting data will be inconsistent and cannot be used as a strong reference in the research context. Both validity and reliability are needed to ensure that the data collected is not only precise, but also reliable [13], [14]. Validity and reliability are two main aspects that must be considered in developing research instruments. Validity refers to the level of accuracy of a measuring instrument in constructing the intended variable, while reliability indicates the level of stability and consistency of measurement results when the instrument is applied under the same conditions [15]. The quality of a study is determined by research instruments that are systematically arranged and meet the standards of validity and reliability. Through the use of valid and reliable instruments, it is likely that the research results will be valid and reliable [16].

However, many researchers still struggle with designing instruments that actually measure what is meant by high consistency [17]. Common problems include ambiguous questions, inconsistent measurement scales, or indicators that do not adequately represent theoretical constructs. Inappropriate use of research instruments due to these problems can result in inaccurate and irrelevant data[13]. Failure in validity and reliability tests often causes instruments to be unable to measure variables comprehensively. Instruments with a significant value >0,05 and a Cronbach's alpha value <0,60 the variable is said to be invalid and can produce data instability, thereby reducing the generalizability of the findings [18]. To overcome these problems, validity and reliability analysis are important procedures in developing research instruments [19]. Validity, which includes aspects of construct validity, content validity, and criterion validity (empirical validity), ensures that the instrument actually measures the intended variables. Meanwhile, the reliability test ensures the consistency of measurement results over time[20]. Without these two analyses, research instruments risk producing invalid or unstable data, which ultimately harms the credibility of the research as a whole. Therefore, it is necessary to test the validity and reliability of an instrument so that it has the feasibility to be used as a research measurement tool [21].

Research conducted by Rohimah on the relationship between peer group roles and bullying behavior in elementary school children using instruments with a validity range of 0,353-0,799 and reliability of 0,772 showed that there is a relationship between peer group roles and bullying behavior in elementary school children [22]. However, the instrument used in the study did not show students' understanding of bullying, so the results obtained only showed the relationship between peers and behavior or actions carried out at school. Ismiati conducted a study on the effect of verbal bullying on the self-confidence of elementary school students, using an instrument with a reliability value of 0,901, as indicated by Cronbach's alpha in the JAMOVI application [23]. Although a standardized instrument was used, the focus of the study was to measure the difference/influence in self-confidence between students who did not experience verbal bullying and those who did. Suryana conducted a Rasch model analysis to measure the number of bullying victims in a school using an instrument with a person reliability value of 0,49 and a Cronbach's alpha value of 0,59 [24]. This indicates that the consistency of participants in answering questions falls into the weak and inconsistent category, caused by an unbalanced number of items and respondents [25]. Low reliability test values in a research instrument can threaten the consistency of results obtained when used in similar conditions at different times [21]. An instrument analyzed using SPSS can be considered valid if the Sig. (2-tailed) value is <0,05 and has a positive correlation. Meanwhile, an instrument is considered reliable if the Cronbach's alpha value is >0,600, but it is better if the Cronbach's alpha value is at least 0,700 [19]. Hijrianti developed a module focused on providing anti-bullying assertiveness training. This research was based on the high number of bullying cases and the limited availability of bullying prevention modules for students with low

assertiveness skills [26]. Before developing such a module, it is essential to have a standardized test instrument that can measure students' understanding of bullying in terms of validity and reliability, ensuring that the resulting product aligns with students' needs. In the research by Bariyyah and Pratama, the validation of research instruments related to bullying has been conducted, but these instruments are intended for higher education levels than elementary school, namely junior high school and high school [27], [28]. The content of the research instrument needs to be adapted to the educational level of the research subjects so that the instrument can be understood and filled out correctly by the respondents, thereby ensuring that the collected data is valid and reliable.

Conducting research related to bullying behavior and bullying prevention measures will be futile if researchers do not truly understand the level of understanding of the research subjects regarding bullying itself. Based on previous studies, each discusses that validity and reliability play an important role in the development of instruments. The topics discussed also vary, ranging from influence, relationships, to the development of anti-bullying media. However, researchers have not found the results of validity and reliability tests related to standardized instruments that measure students' understanding of bullying. Whereas measuring students' understanding of bullying requires instruments that meet validity and reliability standards so that the data collected can be used as a strong reference in the context of research or development of bullying prevention strategies [29]. Measuring the understanding of bullying to elementary school students is an important thing to do considering the many cases of bullying that occur today [30]. Therefore, it is necessary to develop an instrument that can be used to measure elementary school students' understanding of bullying that is standardized including validity and reliability. The researcher aims to develop an instrument on bullying understanding that is tested based on validity and reliability with the title "Validity and Reliability Analysis of Bullying Comprehension Test Instrument Development".

2. RESEARCH METHOD

A quantitative descriptive method was used in this study to describe the validation of test items on the bullying behavior comprehension test instrument in elementary schools in terms of validity and reliability. This method was chosen because it is appropriate for the characteristics of instrument development research, which aims to present the characteristics of the measuring instrument numerically through descriptive statistical analysis [31]. The subjects of this study were 28 fifth-grade students at Sawojajar 1 Elementary School. The object of the study was 50 items of the bullying understanding test instrument, which were tested for validity and reliability. Data collection was conducted by administering the test instrument to expert validators in the field of elementary school education and learning and to 28 fifth-grade elementary school students. The data obtained from this study consisted of respondents' answers to the items in the research instrument. The expert validation was conducted by one lecturer from Malang State University. The suitability of the content of the questions presented with the established indicators and the use of language were two aspects that needed to be considered. The expert validator assigned scores ranging from 1 to a maximum of 4 for each aspect that needed to be adjusted. The responses from the 28 fifth-grade elementary school students were analyzed using SPSS with the Pearson Bivariate Correlation technique for validity testing and the Cronbach's Alpha method for reliability testing. Data analysis was conducted in several stages:

2.1. Content validity of the test instrument

To determine the content validity of the bullying comprehension test instrument, an expert validation process was conducted by distributing questionnaires to validators who are competent in their fields. Content validity is an important aspect in the development of instruments to measure the extent to which the items represent the entire content domain to be measured. Content validation assessment includes aspects of material/content, construct, and language appropriateness used in the development of instruments [31]. To assess rater consistency, the average percentage of scores given was calculated. The results of this validation were then analyzed to identify items that needed revision based on validator feedback, calculate the content validity index quantitatively, and provide a scientific basis for instrument improvement [32].

2.2. Validity of question items

Pearson Bivariate Correlation in the SPSS program is a parametric statistical technique used to test the validity of test items by analyzing the correlation between the score of each item and the total score of the entire instrument. A test item can be considered valid if the Sig. (2-tailed) value is less than 0,05. The asterisk in the Pearson correlation column also indicates that the test item is valid [33]. This analysis procedure is conducted by entering all student response data for the tested items, calculating the total score for each respondent, and the score for each item relative to the total score using Pearson Bivariate Correlation [34]. In the context of developing a bullying understanding test instrument, the application of Pearson Bivariate Correlation ensures that each retained item truly measures the bullying understanding construct significantly.

2.3. Reliability of test instruments

Each item that has been declared valid is tested for reliability using the SPSS program with the Cronbach's Alpha analysis method. Cronbach's Alpha reliability analysis is the most commonly used technique for measuring the internal consistency of an instrument [16]. A test instrument is said to have high reliability if the Cronbach's alpha value is above 0,700. In instrument development, the minimum acceptable Alpha value is 0,700 for basic research. This testing process is important to ensure that the instrument can produce consistent data when used repeatedly under the same conditions [35].

3. RESULTS AND DISCUSSION

The results and discussion of the validity and reliability of the bullying understanding test instrument can be presented as follows:

3.1. Content validity of the test instrument

Researchers conducted a content validity test by giving an expert validation questionnaire to a lecturer in the field of elementary education and learning and obtained the results in the following table:

Instrument title : Expert Validation Questionnaire
Number of validations conducted : 2 times

Table 1. Expert validation results

No.	Aspects assessed	1st validation score	2nd validation score	Average score	Category
I	Appropriateness of Content Presented				
	1. Questions are presented systematically	4	4	4	Very valid
	2. The questions are in accordance with the indicators on the grid	4	4	4	Very valid
	3. The questions represent all the material presented	4	4	4	Very valid
	4. The question does not contain ambiguous statements	4	4	4	Very valid
	5. Suitability of the answer key with the content of the question	4	4	4	Very valid
	6. The suitability of the cognitive domain with the content of the question	3	4	3,5	Valid
	7. Has a proportional level of difficulty between difficult, medium, easy	4	4	4	Very valid
II	Language Used				
	1. The language used is in accordance with EYD	2	3	2,5	Valid enough
	2. The language used is communicative	3	4	3,5	Valid
	3. The sentences used are clear and easy to understand	3	4	3,5	Valid
	Total average	3,5	3,9	3,7	Very valid

Equation (1) is an example of writing an equation to find the percentage of content validity.

$$percentage\ of\ validity = \frac{Total\ average}{Maximum\ score} x100\% \quad (1)$$

Maximum score for each aspect : 4
Average total score : 3,7

$$percentage\ of\ validity = \frac{3,7}{4} x100\% = 92\%$$

Table 2. Criteria for test question validation
Source: [35]

Percentage value	Category
90% - 100%	Very valid
70% - 89%	Valid

50% - 69%	Valid enough
<50%	Not valid

Source: [35]

Content validity is useful for measuring the extent to which the items in an instrument cover all aspects of the construct being measured, in terms of content, language, and suitability for the measurement purpose. In this study, content validity testing was conducted by an expert who assessed 50 items in a bullying comprehension test instrument for elementary school students. The validation process was carried out in two stages, before and after revisions. The first validation results showed that the instrument obtained a score of 87,5%, which falls into the "valid" category. Although classified as valid, the validator provided several suggestions for improvement, particularly regarding the language used, the cognitive level appropriateness for elementary school students, and the clarity of the question wording. After revisions were made based on these suggestions, the instrument was revalidated in the second stage. In this stage, the validation results improved to 97,5%, which falls into the "highly valid" category. This means that most of the items have met the content validity criteria very well. Based on these results, the validator stated that the instrument is suitable for field testing without further revisions.

Overall, the average content validity rate from both validation stages was 92%, categorized as "highly valid." These results indicate that the instrument has met the criteria for content validity and can be trusted to measure students' understanding of bullying. With high content validity, it can be concluded that this instrument is substantively aligned with the construct being measured [36], [37]. This is important to ensure that the results obtained from the measurement truly reflect students' level of understanding of the topic being studied, namely bullying.

3.2. Validity of question items

The item validation procedure on the instrument was carried out through giving tests to 28 respondents of grade 5 Sawojajar 1 Elementary School Malang as research subjects. The evaluation instrument consists of 50 items with distribution: 32 multiple-choice items, 11 matching items, and 7 true-false items. Analysis of item validity was carried out using the Pearson Bivariate Correlation technique through the SPSS program, with the results of the validation of the bullying understanding instrument presented in Table 3.

Table 3. The results of the validity of the items

		soal_1	soal_2	soal_3	soal_4	...	soal_48	soal_49	soal_50	Total
soal_1	Pearson Correlation	.[a]	.[a]	.[a]	.[a]	...	.[a]	.[a]	.[a]	.[a]
	Sig. (2-tailed)					...				
	N	28	28	28	28	...	28	28	28	28
soal_2	Pearson Correlation	.[a]	.[a]	.[a]	.[a]	...	.[a]	.[a]	.[a]	.[a]
	Sig. (2-tailed)					...				
	N	28	28	28	28	...	28	28	28	28
soal_3	Pearson Correlation	.[a]	.[a]	.[a]	.[a]	...	.[a]	.[a]	.[a]	.[a]
	Sig. (2-tailed)					...				
	N	28	28	28	28	...	28	28	28	28
soal_4 ...	Pearson Correlation	.[a]	.[a]	.[a]	1	...	-0,042	0,042	0,198	0,355
	Sig. (2-tailed)					...	0,831	0,831	0,313	0,064
	N	28	28	28	28	...	28	28	28	28
soal_48	Pearson Correlation	.[a]	.[a]	.[a]	-0,042	...	1	0,167	0,140	.783**

	Sig. (2-tailed)				0,831	..	0,397	0,477	0,000
	N	28	28	28	28	..	28	28	28	28
soal_49	Pearson Correlation	.[a]	.[a]	.[a]	0,042	..	0,167	1	-0,027	0,279
	Sig. (2-tailed)				0,831	..	0,397		0,892	0,150
	N	28	28	28	28	..	28	28	28	28
soal_50	Pearson Correlation	.[a]	.[a]	.[a]	0,198	..	0,140	-0,027	1	0,226
	Sig. (2-tailed)				0,313	..	0,477	0,892		0,248
	N	28	28	28	28	..	28	28	28	28
Total	Pearson Correlation	.[a]	.[a]	.[a]	0,355	..	.783**	0,279	0,226	1
	Sig. (2-tailed)				0,064	..	0,000	0,150	0,248	
	N	28	28	28	28	..	28	28	28	28

*. Correlation is significant at the 0,05 level (2-tailed).
**. Correlation is significant at the 0,01 level (2-tailed).
a. Cannot be computed because at least one of the variables is constant.

To determine the validity of the items tested using the Pearson Bivariate Correlation technique using the SPSS program, it can be seen in the Sig value. (2-tailed) <0,05. The asterisk in the Pearson correlation column also indicates that the item is said to valid. The following is the conclusion of the SPSS validity results in Table 4.

Table 4. Conclusion of the item validity test

No. Soal	Pearson correlation	Nilai Sig.	Conclusion	Interpretation
soal_1	a	0,00	-	-
Soal_2	a	0,00	-	-
soal_3	a	0,00	-	-
soal_4	0,355	0,064	Not valid	-
soal_5	-0,131	0,508	Not valid	-
soal_6	0,067	0,735	Not valid	-
soal_7	0,592**	<0,001	Valid	High enough
soal_8	a	0,00	-	-
soal_9	a	0,00	-	-
soal_10	0,490**	0,008	Valid	High enough
soal_11	0,571**	0,002	Valid	High enough
soal_12	0,571**	0,002	Valid	High enough
soal_13	0,571**	0,002	Valid	High enough
soal_14	a	0,00	-	-
soal_15	0,490**	0,008	Valid	High enough
soal_16	0,197	0,315	Not valid	-
soal_17	0,571**	0,002	Valid	High enough
soal_18	0,260	0,181	Not valid	-
soal_19	0,421*	0,025	Valid	High enough
soal_20	0,332	0,084	Not valid	-
soal_21	0,510**	0,006	Valid	High enough
soal_22	0,523**	0,004	Valid	High enough
soal_23	0,213	0,277	Not valid	-
soal_24	0,471*	0,011	Valid	High enough
soal_25	0,633**	<0,001	Valid	High
soal_26	0,334	0,082	Not valid	-
soal_27	0,528**	0,004	Valid	High enough
soal_28	0,340	0,077	Not valid	-

soal_29	0,734**	<0,001	Valid	High
soal_30	0,370	0,052	Not valid	-
soal_31	0,549**	0,002	Valid	High enough
soal_32	0,555**	0,002	Valid	High enough
soal_33	0,582**	0,001	Valid	High enough
soal_34	0,712**	<0,001	Valid	High
soal_35	0,560**	0,002	Valid	High enough
soal_36	0,230	0,240	Not valid	-
soal_37	0,584**	0,001	Valid	High enough
soal_38	0,712**	<0,001	Valid	High
soal_39	0,424*	0,025	Valid	High enough
soal_40	0,499**	0,007	Valid	High enough
soal_41	0,293	0,130	Not valid	-
soal_42	0,376*	0,049	Valid	Low
soal_43	0,532**	0,004	Valid	High enough
soal_44	0,080	0,687	Not valid	-
soal_45	0,783**	<0,001	Valid	High
soal_46	0,305	0,115	Not valid	-
soal_47	0,474*	0,011	Valid	High enough
soal_48	0,783**	<0,001	Valid	High
soal_49	0,279	0,150	Not valid	-
soal_50	0,226	0,226	Not valid	-

The following is an interpretation of the validity values presented in Table 5.

Table 5. Interpretation of validity values

Pearson Correlation	Interpretasi
0,800 - 1,00	Very high
0,600 - 0,799	High
0,400 - 0,599	High enough
0,200 - 0,399	Low
0,000 - 0,199	Very low

Source: [35]

Validity refers to the degree of accuracy of a measuring instrument in constructing the intended variable. In this study, validity testing was conducted to determine the suitability of the items in the bullying comprehension test instrument for elementary school students. The technique used to test validity was Pearson Bivariate Correlation using the SPSS program [34]. The determination of whether an item was valid or not was based on the significance value (Sig. 2-tailed). If the significance value was < 0,05, the item was considered valid. Additionally, an asterisk (*) in the Pearson Correlation column of the SPSS output also indicates that the correlation is valid [33]. The results of the instrument validation analysis showed that out of 50 items administered to 28 elementary school respondents, 28 items met the validity criteria, while 22 others were deemed invalid. Among the valid items, most were in the “sufficient interpretation” category, while 6 items achieved a high level of validity. The 6 items with high correlations indicate that these questions have a strong relationship with the overall construct being measured, namely understanding of bullying. The remaining 22 items in the “sufficient” category indicate that the questions are still suitable for use, although their validity level is not as high as the previous group.

The invalid items need to be revised or eliminated because they do not contribute significantly to the measurement of the intended variable [15]. Some factors that may cause invalidity are ambiguous question statements, difficulty levels that are not appropriate for the grade level of the students, or question contexts that are not relevant to the experiences of elementary school students [13]. With 28 valid questions obtained, it shows that more than half of the initial questions have met the statistical validity requirements, which are an important basis for developing a quality instrument.

3.3. Reliability of test instruments

The 28 items that have been verified for validity were tested for reliability through the Cronbach's Alpha analysis technique with the help of the SPSS program, the results of which are presented in Table 6.

Table 6. Item reliability results (Item-Total Statistics)

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
soal_7	22,1071	31,951	0,560	0,916

soal_10	22,0357	32,776	0,495	0,917
soal_11	22,0000	32,889	0,649	0,916
soal_12	22,0000	32,889	0,649	0,916
soal_13	22,0000	32,889	0,649	0,916
soal_15	22,0357	32,851	0,470	0,917
soal_17	22,0000	32,889	0,649	0,916
soal_19	22,2143	32,693	0,286	0,921
soal_21	22,2857	31,915	0,408	0,919
soal_22	22,0714	32,439	0,500	0,917
soal_24	22,0357	32,925	0,445	0,918
soal_25	22,1429	31,608	0,586	0,915
soal_27	22,1429	31,831	0,534	0,916
soal_29	22,0714	31,476	0,780	0,913
soal_31	22,2143	31,804	0,469	0,917
soal_32	22,1786	31,634	0,537	0,916
soal_33	22,3214	31,189	0,533	0,917
soal_34	22,1429	30,942	0,746	0,913
soal_35	22,2500	31,602	0,487	0,917
soal_37	22,3929	31,284	0,496	0,917
soal_38	22,1429	31,238	0,675	0,914
soal_39	22,2857	31,767	0,436	0,918
soal_40	22,3571	31,497	0,463	0,918
soal_42	22,1071	32,766	0,353	0,919
soal_43	22,2857	31,471	0,494	0,917
soal_45	22,0714	31,328	0,824	0,912
soal_47	22,0714	32,661	0,436	0,918
soal_48	22,0714	31,328	0,824	0,912

The following is an reability statistic presented in table 7.

Table 7. Reliability statistics

Cronbach's Alpha	N of Items
0,9192	28

The following is an cronbach's alpha reliability classification presented in table 8.

Table 8. Cronbach's alpha reliability classification

Cronbach's alpha value range	Reliability qualification
>0,90	Very high
0,70 – 0,90	High
0,50 – 0,70	High enough
<0,50	Low

Source: [35]

Reliability is useful for indicating the level of stability and consistency of measurement results when the instrument is applied under the same conditions. In this study, reliability testing was conducted to determine the level of stability and consistency of the bullying comprehension test instrument developed. Reliability testing was performed using Cronbach's Alpha analysis, which is one of the most commonly used methods for measuring the reliability of instruments with multiple scales [16]. The reliability test results showed that the Cronbach's Alpha value was 0,919 with a total of 28 items analyzed. An instrument can be said to have high reliability if the Cronbach's Alpha value is above 0,700 [35]. From Table 6, it is known that all valid items have high reliability values of >0,90. Table 7 shows the overall results of the reliability-tested data, which obtained a value of 0,919. Based on the reliability classification table, this indicates a very high level of reliability [38]. This suggests that each item in the instrument is mutually correlated and measures the same construct, namely students' understanding of bullying. This high reliability supports the broader use of the instrument, whether in the context of educational program evaluation, classroom action research, or in developing bullying prevention strategies in elementary schools [21]. Additionally, these results also indicate that the instrument has the potential for use in continuous measurement due to its stability and consistency [39]. Therefore, based on the reliability test results, it can be concluded that the bullying understanding instrument developed is suitable for use as it meets the criteria for a reliable instrument.

This study has several advantages that enhance the validity and reliability of the bullying comprehension test instrument. The use of SPSS in analyzing the test instrument data provides a higher level of confidence in the results [40]. Statistical analysis with SPSS allows the validity and reliability test results to meet good psychometric standards. In addition, the number of questions used is relatively large, namely 50 items, so that the coverage of aspects of bullying understanding becomes more comprehensive. With a large number of questions, this instrument has a better ability to measure various dimensions of students' understanding of bullying, including its forms, impacts, and ways of handling it. The results of this study have practical implications for the development of testing instruments in the field of education, particularly regarding bullying issues in elementary schools. With the availability of a valid and reliable instrument, teachers, counselors, and researchers can measure students' understanding of bullying more accurately. The data obtained can be used to design targeted bullying prevention programs. Additionally, this instrument can serve as a reference for further research exploring other factors related to bullying.

4. CONCLUSION

Valid and reliable instruments are essential in scientific research to produce consistent and accurate data. An ideal instrument must meet validity and reliability standards. Based on previous research, researchers have not found any validity and reliability test results related to standardized instruments that measure students' understanding of bullying in elementary schools. However, measuring students' understanding of bullying requires instruments that meet validity and reliability standards so that the collected data can be used as a strong reference in the context of research or the development of bullying prevention strategies. Therefore, researchers developed an instrument that can be used to measure elementary school students' understanding of bullying, including validity and reliability. A quantitative descriptive method was used in this study to describe the validation of test items on the bullying behavior understanding test instrument in elementary schools in terms of validity and reliability. The subjects of this study were 28 fifth-grade students from SDN Sawojajar 1 and an experienced expert in test item validation. The research object was 50 items of the bullying understanding test instrument with the following distribution: 32 multiple-choice items, 11 matching items, and 7 true-false items. The data obtained were analyzed using SPSS with Pearson Bivariate Correlation and Cronbach's Alpha analysis techniques. After two rounds of content validity testing by the expert validator, the final validity score was 97.5%, categorized as highly valid and suitable for field testing without revisions. Out of the 50 questions tested in the field with 28 respondents, 28 questions were found to be valid, achieving a very high reliability score of 0.019.

However, this study has several limitations. First, content validation involved only one validator, so the level of subjectivity in assessing the quality of the items may still be high. Ideally, validation should be carried out by several experts (expert judgment) to ensure that the instrument truly measures the intended construct. Second, the research subjects were limited to students from one elementary school, so the findings cannot necessarily be generalized to a wider population. School environment factors, anti-bullying policies, and students' social backgrounds in other schools may influence the level of understanding of bullying, so a broader trial is needed in various schools with different characteristics. In order for this instrument to be used more widely, further research needs to be conducted involving more validators and subjects from various schools with different characteristics. Additionally, the development of the instrument could be expanded by adding new aspects, such as students' understanding of cyberbullying or the role of bystanders in bullying cases. Thus, this instrument is not only useful for measuring understanding but can also serve as a tool for evaluating the effectiveness of anti-bullying programs at the school level.

REFERENCES

[1] S. A. Visty, "Dampak bullying terhadap perilaku remaja masa kini," *Jurnal Intervensi Sosial dan Pembangunan (JISP)*, vol. 2, 2021, doi: 10.30596/jisp.v2i1.3976.
[2] A. Jannah, *"Bullying Relasional Pada Siswa Di Sekolah,"* Universitas Negeri Surabaya, Surabaya, 2022.
[3] A. Budiman and F. Asriyadi, *Perilaku Bullying pada Remaja dan Faktor-faktor yang Mempengaruhinya*, Pertama. Purwokerto: CV. Pena Persada, 2021.
[4] Sekolah Relawan, *"Kasus Bullying di Sekolah Meningkat, KPAI Sebut Ada 2.355 Kasus Pelanggaran Perlindungan Anak Selama 2023,"* sekolahrelawan.com.
[5] P. Y. A. Dewi, "Perilaku school bullying pada siswa sekolah dasar," *Edukasi: Jurnal Pendidikan Dasar*, vol. 1, no. 1, pp. 39–48, 2020.
[6] A. Sukawati, D. L. Abdul Muiz, and N. Ganda, "Fenomena bullying berkelompok di sekolah dasar," *PEDADIDAKTIKA: Jurnal Ilmiah Pendidikan Guru Sekolah Dasar*, vol. 8, no. 2, pp. 354–363, 2021, doi: 10.17509/pedadidaktika.v8i2.35344.

[7] R. Emi, Syahrial, and V. A. Hardi, "Hubungan perilaku bullying dengan kemampuan interaksi sosial siswa kelas V SD negeri 37 Pekanbaru," *Indonesian Research Journal on Education: Jurnal Ilmu Pendidikan*, vol. 1, no. 1, 2021, doi: 10.31004/irje.v1i1.1.
[8] R. Candrawati and A. Setyawan, "Analisis perilaku bullying terhadap motivasi belajar siswa sekolah dasar," *PANDU : Jurnal Pendidikan Anak dan Pendidikan Umum*, vol. 1, no. 2, pp. 64–68, May 2023, doi: 10.59966/pandu.v1i2.127.
[9] D. P. Rahmadanti, "Hubungan antara nilai karakter empati dengan perilaku bullying pada siswa kelas XI MIPA 1 SMAN 1 Kembang," Universitas PGRI Semarang, Semarang, 2021.
[10] M. B. Akbar, I. R. Sulistiani, and D. F. Wiyono, "Strategi pencegahan tindakan bullying di lingkungan SMAN 2 Malang," *VICRATINA: jurnal Pendidikan Islam*, vol. 8, no. 5, 2023.
[11] V. Paula, R. Oktarini br Sibuea, K. Lebdawicaksaputri, E. Kasenda, D. Keperawatan, and U. Pelita Harapan, "Edukasi pencegahan tindakan bullying pada anak usia sekolah dasar," *Jurnal Pustaka Mitra*, vol. 2, 2022, doi:10.55382/jurnalpustakamitra.v2i2.204.
[12] A. N. Amalia, Suyono, and R. Arthur, *Penyusunan Instrumen Penelitian: Konsep, Teknik, Uji Validitas, Uji Reliabilitas, dan Contoh Instrumen Penelitian*, Pertama. Jawa Tengah: PT Nasya Expanding Management, 2023.
[13] S. Arikunto, *Prosedur Penelitian: Suatu Pendekatan Praktik*, 15th ed. Jakarta: PT Asdi Mahasatya, 2013.
[14] Y. O. Purba, Fadhilaturrahmi, J. Triani Purba, and K. W. A. Siahaan, *Teknik Uji Instrumen Penelitian Pendidikan*, Pertama. Bandung: Widina Bhakti Persada Bandung, 2021.
[15] D. Budiastuti and A. Bandur, *Validitas Dan Reliabilitas Penelitian*. Jakarta: Mitra Wacana Media, 2018.
[16] sugiyono, *Metode Penelitian: Kuantitatif, Kualitatif, dan R&D*, Cetakan ke-26. Bandung: ALFABETA cv, 2017.
[17] E. Rosyidah and E. Masykuroh, "Memahami strategi dan mengatasi tantangan dalam penelitian metode kuantitatif," *Journal Syntax Idea*, vol. 6, no. 6, p. 1247, Jun. 2024, doi:10.46799/syntax-idea.v6i6.3748.
[18] B. Darma, *Statistika Penelitian menggunakan SPSS*. Jakarta: Guepedia.com, 2021. Accessed: Jul. 02, 2025.
[19] N. M. Janna and Herianto, "Konsep uji validitas dan reliabilitas dengan menggunakan SPSS," *OSFPREPINTS*, 2021, doi:10.31219/osf.io/v9j52.
[20] M. F. Ramadhan, R. A. Siroj, and M. W. Afgani, "Validitas and reliabilitas," *Journal on Education*, vol. 06, no. 02, pp. 10967–10975, 2024, doi:10.31004/joe.v6i2.4885.
[21] P. G. Subhaktiyasa, "Evaluasi validitas dan reliabilitas instrumen penelitian kuantitatif: sebuah studi pustaka," *Journal of Education Research*, vol. 5, no. 4, pp. 5599–5609, 2024, doi:10.37985/jer.v5i4.1747.
[22] A. Rohimah, "*Hubungan Peran Kelompok Teman Sebaya Dengan Perilaku Bullying pada Anak Usia Sekolah di SD Muhammadiyah Mlangi Gamping Sleman Yogyakarta,"* Universitas 'aisyiyah, Yogyakarta, 2016
[23] N. Ismiati, D. A. A. Putri, E. Y. Sari, U. Bariroh, and I. M. Fitriana, "Pengaruh verbal bullying terhadap kepercayaan diri siswa sekolah dasar," *SEMESTA: Jurnal Ilmu Pendidikan dan Pengajaran*, vol. 3, no. 2, pp. 54–65, Jul. 2025, doi: 10.70115/semesta.v3i2.218.
[24] D. Suryana, M. A. Putri, M. Supriatna, and E. S. Yudha, "Analisis rasch model: validitas dan reliabilitas instrumen korban bullying," *Hisbah: Jurnal Bimbingan Konseling dan Dakwah Islam*, vol. 19, no. 2, pp. 199–214, 2022, doi:10.14421/hisbah.2022.192-12.
[25] U. Purwana, D. Rusdiana, and W. Liliawati, "Pengujian kemampuan Menginterpretasikan Grafik Kinematika Calon Guru Fisika: The Polytomous Rasch Analysis," *ORBITA: Jurnal Pendidikan dan Ilmu Fisika*, vol. 6, no. 2, 2020, doi:10.31764/orbita.v6i2.3264.
[26] U. R. Hijrianti, "Validasi modul peduli: pelatihan asertivitas anti-bullying," *Altruis: Journal of Community Services*, vol. 2, no. 3, Sep. 2021, doi:10.22219/altruis.v2i3.18130.
[27] K. Bariyyah *et al.*, "Validasi instrumen illinois bully scale (IBS) untuk mendeteksi tingkat bullying di SMAN 3 Malang," *Education and Social Sciences Review*, vol. 5, no. 2, p. 176, Dec. 2024, doi: 10.29210/07essr491600.
[28] A. M. Pratama *et al.*, "Confirmatory factor analysis (CFA) untuk validasi instrumen sexual bullying questionarry (SEBUQU)," *Jurnal EDUCATIO: Jurnal Pendidikan Indonesia*, vol. 10, no. 2, p. 258, Dec. 2024, doi: 10.29210/1202424887.
[29] F. Sarie, N. T. Sutaguna, P. Suiraoka, and D. Damanik, *Metodologi penelitian*, Manullang, Sardjana. Batam: Yayasan Cendikia Mulia Mandiri, 2022.
[30] R. N. Dewi and L. Maknun, "Urgensi pendidikan karakter bagi anak usia SD untuk mencegah perilaku bullying," *Jurnal Pendidikan Guru Madrasah Ibtidaiyah*, vol. 2, no. 1, 2023.
[31] Sugiyono, *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*, Ke-10. Bandung: ALFABETA, cv, 2010.
[32] T. E. Wibowo and S. Faizah, "Pengembangan soal tes untuk mengukur kemampuan pemecahan masalah siswa pada materi bentuk aljabar," *Alifmatika: Jurnal Pendidikan dan Pembelajaran Matematika*, vol. 3, no. 2, pp. 145–158, Nov. 2021, doi: 10.35316/alifmatika.2021.v3i2.145-158.

[33] B. C. Ngongoloy and M. A. I. Pakereng, "Penerapan metode linear regression dan correlation pearson dalam menganalisis pengaruh kualitas pembelajaran online terhadap prestasi akademik," *Progresif: Jurnal Ilmiah Komputer*, vol. 19, no. 2, pp. 844–857, 2023, doi:10.35889/progresif.v19i2.1529.
[34] F. D. P. Anggraini, A. Aprianti, V. A. V. Setyawati, and A. A. Hartanto, "Pembelajaran statistika menggunakan software spss untuk uji validitas dan reliabilitas," *Jurnal Basicedu*, vol. 6, no. 4, pp. 6491–6504, May 2022, doi: 10.31004/basicedu.v6i4.3206.
[35] S. Arikunto, *Prosedur Penelitian Suatu Pendekatan Praktik.*, vol. 283. Jakarta: Rineka Cipta, 2010.
[36] R. Ramadhani and D. Bagus Pratama Putra, "Pengembangan modul elektronik interaktif berbasis sigil pada materi pitagoras sebagai penunjang pembelajaran jarak jauh," *Jurnal KOULUTUS: Jurnal Pendidikan Kahuripan*, vol. 4, no. 2, pp. 140–152, 2021, doi:10.51158/koulutus.v4i2.621.
[37] E. Hutabri, "Validitas media pembelajaran multimedia pada mata pelajaran simulasi dan komunikasi digital," *Prosiding Seminar Nasional Ilmu Sosial Dan Teknologi (SNISTEK)*, vol. 4, pp. 296–300, Jan. 2022.
[38] A. Maulana, "Analisis validitas, reliabilitas, dan kelayakan instrumen penilaian rasa percaya diri siswa," *Jurnal Kualita Pendidikan*, vol. 3, no. 3, pp. 2774–2156, 2022, doi: 10.51651/jkp.v3i3.331.
[39] M. Iqbal and S. Salsabila, "Validitas dan reliabilitas kuesioner kebiasaan menonton mukbang," *Muhammadiyah Journal of Nutrition and Food Science (MJNF)*, vol. 4, no. 2, p. 102, Nov. 2023, doi: 10.24853/mjnf.4.2.102-109.
[40] I. Isbandiyah and Y. Asmara, "Pelatihan analisis data uji coba instrumen dengan menggunakan program SPSS versi 22," *JURNAL CEMERLANG : Pengabdian pada Masyarakat*, vol. 3, no. 1, pp. 29–37, Dec. 2020, doi: 10.31540/jpm.v3i1.859.