Jurnal Kependidikan:
Jurnal Hasil Penelitian dan Kajian Kepustakaan di Bidang Pendidikan, Pengajaran dan Pembelajaran
https://e-journal.undikma.ac.id/index.php/jurnalkependidikan/index

Vol. 10, No. 4 : December 2024
E-ISSN: 2442-7667
pp. 1496-1506
Email: jklppm@undikma.ac.id

Development of A Holistic Assessment Instrument Based on Computer Based Test on Environmental Pollution Learning

Dini Rosdiane Rusdiyanto[1], Bibin Rubini[2*], Irvan Permana[3]
[1,2*,3]Department of Science Education, Postgraduate School, Universitas Pakuan, Indonesia.
[1]SMP Muhammadiyah, Cianjur, Indonesia.
*Corresponding Author. Email: bibinrubini@unpak.ac.id

Abstract: This study aims to develop a holistic assessment instrument based on Computer Based Test (CBT) that meets the criteria of good assessment in the context of learning Environmental Pollution in Junior High School, especially grade VII. The research method used is Research and Development (R&D) with ADDIE model (Analysis, Design, Development, Implementation, and Evaluation). The assessment instrument consists of 30 multiple-choice test items that cover the dimensions of knowledge, attitudes, communication skills, and collaboration skills. The development of questions through the process of item content validity and media validity validated by two competent experts each, peer science teachers and students. The obtained data is then analyzed using: product moment correlation, the Kuder-Richardson-20 (KR-20) formula, the Du Bois formula, and the item discrimination index. Based on the results of item analysis, 26 questions were tested valid and reliable with a reliability score of 0.862. The results of the difficulty test showed 69.23% of the questions were quite difficult, 11.54% of the questions were too difficult, and 19.23% of the questions were too easy. The differentiating power test showed 61.54% of the questions had good differentiating power and 38.46% of the questions were in the sufficient category. Based on these results, the instrument developed has the potential to increase the effectiveness of learning Environmental Pollution in Junior High School, especially grade VII and provide comprehensive information about student achievement in various dimensions of learning.

Article History
Received: 14-09-2024
Revised: 20-10-2024
Accepted: 13-11-2024
Published: 21-12-2024

Key Words:
Holistic Assessment;
Computer-Based Test;
Environmental Pollution.

How to Cite: Rusdiyanto, D., Rubini, B., & Permana, I. (2024). Development of A Holistic Assessment Instrument Based on Computer Based Test on Environmental Pollution Learning. *Jurnal Kependidikan: Jurnal Hasil Penelitian dan Kajian Kepustakaan di Bidang Pendidikan, Pengajaran dan Pembelajaran, 10*(4), 1496-1506. doi:https://doi.org/10.33394/jk.v10i4.12936

https://doi.org/10.33394/jk.v10i4.12936

Introduction

The emergence of global problems and issues such as human rights violations, drug abuse, and unhealthy competition between students has inspired education experts to formulate new approaches in the learning system (Primantoro, 2016). In response to the dynamics of the global community, UNESCO in 1996 established four pillars of education as the foundation for this global era (Rethinking Education, 2015) (in Tristaningrat, 2018) . The four pillars include *Learning to know*, *Learning to do*, *Learning to be*, and *Learning to live together*. Based on this foundation, there is a follow-up in the form of curriculum needs that are oriented towards the formation of competencies that are relevant to the demands of the real world (Tristaningrat, 2018).

Current education leads to contextual learning, where students are not only given theoretical knowledge, but also invited to relate it to real-world situations (Tristaningrat, 2018). This concept requires teachers to not only be information conveyors, but also facilitators who help students apply knowledge and skills in everyday life. Contextual learning is an important foundation in forming a generation of students who are not only

Vol. 10, No. 4 : December 2024
E-ISSN: 2442-7667
pp. 1496-1506
Email: jklppm@undikma.ac.id

academically intelligent, but also able to face and solve the challenges of the global world with a deep and applicable understanding (Rubiyanto, 2010).

An effective learning process needs to be accompanied by appropriate evaluation to measure student progress and achievement. Therefore, the assessment aspect becomes an integral element of learning activities. In the context of contextual learning, one of the striking features is the application of authentic *assessment*. Authentic *assessment* is an evaluation approach designed to reflect the real-world context and tasks relevant to students' daily lives (Kunandar, 2014). In contrast to traditional evaluation methods that focus more on mastery of theory in isolation, authentic *assessment* emphasizes students' ability to apply knowledge and skills in real-life situations.

The application of authentic *assessment* in contextual learning has a dual purpose. First, it provides a more comprehensive picture of students' abilities in dealing with real-world tasks. Second, it supports the integration of knowledge in practical contexts, ensuring that students not only understand concepts theoretically but can also apply them relevantly in everyday life. Authentic assessment is not only an evaluation tool, but also an integral part of the learning experience that connects education with real-world life. Based on this description, it can be seen that the assessment of students' abilities must be carried out thoroughly (holistically). (Dahar, 2011). However, the phenomenon in the field shows that the assessment of students' abilities is still carried out separately, between the cognitive, affective, and psychomotor domains. (Astuti & Darsinah, 2018).. The cognitive and psychomotor domains are the domains of assessment that are often carried out by teachers, while the affective domain receives less attention. (Apriyana et al., 2019; Hairida, 2018; Simarmata et al., 2019).

The importance of evaluation is not only limited to the national scale, but also extends to the international realm. PISA (*Program for International Student Assessment*) scores are a key indicator to assess the effectiveness of a country's education system. Evaluating students' skills through PISA highlights the importance of critical and creative thinking in processing information to solve real-world problems (Aisyah & Dewi, 2019) . Although PISA provides a comprehensive picture of students' abilities, problems arise in the implementation of holistic assessment at the local level. Teachers still often use conventional paper and pencil-based methods in conducting assessments. Time, cost, and efficiency constraints are challenges faced by teachers in implementing assessments that embrace all dimensions of students' abilities. In facing these challenges, advances in information technology can be an innovative solution. Teachers who are able to integrate information technology into teaching methods can create a learning environment that is dynamic and in line with the times. The utilization of *Computer Based Test* (CBT) as a testing instrument, a computer connected to a network and internet access is used, and scoring is conducted automatically by the computer is one of the strategic steps in improving the efficiency and relevance of assessment (Utami, 2019).

However, the adoption of CBT in the field also faces obstacles, including lack of understanding and logistical challenges, so the development of CBT-based holistic assessment instruments is needed to provide innovative and adaptive solutions to face assessment problems in the modern era. Therefore, this study was designed to develop a CBT-based holistic assessment instrument as a concrete step in overcoming assessment challenges in this global era. The novelty of this research lies in developing an evaluation tool that provides a comprehensive overview of students' development, progress, and achievements in cognitive, affective, and psychomotor domains. Additionally, it links these dimensions to 21st-century skills, such as communication and collaboration with others,

Vol. 10, No. 4 : December 2024
E-ISSN: 2442-7667
pp. 1496-1506
Email: jklppm@undikma.ac.id

while utilizing advanced computer technology to enhance the objectivity and accuracy of assessments.

Research Method

This study used a research and development method (R&D*). Research and Development* is a research method used to produce certain products, and test the effectiveness of these products (Sugiyono, 2019). The type of research and development adopted in this study uses the ADDIE model consists of five stages, namely *Analysis, Design, Development, Implementation,* and *Evaluation* (Branch, 2009). This research was conducted at Muhammadiyah Cianjur Junior High School for Grade VII students in the 2022/2023 academic year. The field trial subjects in this study were 40 students of Class VII even semester of the 2022/2023 academic year at SMP Muhammadiyah Cianjur. Other research subjects were 20 science teachers, and 2 *assessment* and media expert lecturers each.

The research instruments used to obtain the required data were question validation sheets, media validation sheets, holistic assessment items and student questionnaires. The data obtained were then analysed using appropriate analysis techniques. The content validity of the items in the study was carried out by requesting the opinion of two lecturers with expertise in the field of assessment. The assessment development product was assessed using a validation instrument, where the answer 'yes' was scored 1 and 'no' was scored 0. The validation results were calculated using the content validity ratio (CVR) formulated by Lawshe (1975), which measures expert agreement on the essentiality of an item. Items are considered essential if they can represent the purpose of measurement well (Azwar, 2022). CVR test results are interpreted relatively in the range of -1.0 to +1.0. If half of the SMEs (Subject Matter Experts) indicate that an item is essential, the CVR value is 0. If all SMEs indicate that an item is essential, the CVR value is 1. Items that receive negative CVR values are invalid items and must be eliminated.

The feasibility of the computer-based test was evaluated using an assessment instrument developed based on modified media feasibility criteria from the National Education Standards Agency (BSNP). Validators provided ratings and suggestions for improvement using an instrument validation sheet with a 5-point Likert scale. The validation results from the two validators were then calculated using a specific formula, and the percentage of feasibility and suitability of the instrument was determined based on the media validation criteria table (Arikunto, 2013).

Item validity is calculated using the product-moment correlation formula. The correlation coefficient value is interpreted by comparing the calculated r value with the r product moment table. The correlation is considered valid if the r count value is greater than the r table value (ri > rt) (Sugiyono, 2019). Reliability test for multiple choice questions uses the Kuder - Richardson-20 or KR-20 formula (Azwar, 2022). The basis for decision making in reliability testing according to (Sujarweni, 2014) is as follows:

1) If the Cronbach's alpha value is > 0.60, the instrument is considered reliable or consistent.
2) If the Cronbach's alpha value is <0.60, the instrument is considered unreliable or inconsistent.

The calculation of the index number of the difficulty of the question can be obtained using the formula proposed by Du Bois, a question with an index of 0.00 indicates that the question is too difficult, on the other hand, an index of 1.0 indicates that the question is too easy (Anas Sudjiono, 2016). Furthermore, the value of the differentiating power is expressed by the differentiating power index, which is calculated using a specific formula. The higher

Vol. 10, No. 4 : December 2024
E-ISSN: 2442-7667
pp. 1496-1506
Email: jklppm@undikma.ac.id

or greater the index of differentiating power of the question, the more the question can differentiate between high and low groups (Arikunto, 2009).

Results and Discussion

The development of a *Computer Based* Holistic Assessment Instrument on Environmental Pollution Learning in Class VII, has development stages in accordance with the ADDIE development model, which consists of five stages namely *Analysis*, *Design*, *Development*, *Implementation*, and *Evaluation*. Data and discussion results are described as follows:

Analysis

The results of the research at the needs analysis stage involved literature study, questionnaire distribution, and interviews, covering several key aspects. The following is an explanation of the results of each stage:

Literature Study

Literature analysis was conducted to gain an in-depth understanding of holistic assessment and the application of *Computer Based Test* (CBT) in the context of science learning. Assessment applied in learning is not only focused on knowledge, but also students' attitudes and skills in accordance with the views expressed by Dahar (2011). Findings from the literature review highlighted that holistic assessment practices are still fragmented, with separate assessments of knowledge, attitude and skills dimensions. The same thing was also revealed in research conducted by Astuti & Darsinah (2018) who conducted research on 2013 Curriculum-Based Authentic Assessment at SD Negeri Mangkubumen Kidul No. 16 Surakarta. In addition, the conventional method of assessment using pencil and paper creates a number of obstacles, such as time inefficiency, high levels of cheating, and unattractive photocopies. (Maulana 2022; Manurung & Rajagukguk, 2019; Samsiadi & Humaidi, 2022; Utami, 2019).

Questionnaire Distribution

Questionnaires were distributed to science teachers to get their perspectives on the obstacles in implementing holistic assessment. The questionnaire results showed that the teachers experienced difficulties in conducting assessments that cover knowledge, attitudes and skills simultaneously in one assessment session. These difficulties involve aspects such as time management, integration of assessment results, and the need for more efficient instruments. The responses from the science teachers were similar to the research conducted by Manurung & Rajagukguk (2019)who conducted research on the Evaluation Design of *Computer* Based *Test-Based* Physics Learning Outcomes on the Subject Matter of Effort and Energy.

Interview:

Structured interviews in December 2022 with science teachers discussed the curriculum, science learning, and assessment instruments. Most teachers are less active in designing assessment instruments independently, more often done together during MGMP. Obstacles arise in the assessment of practical work and attitudes because learning time is used up for guidance, so the assessment is less in-depth. Yunitasari (2019) also found similar challenges in authentic assessment of observing, inferring, and predicting skills.

Attitude assessment instruments that currently use rubrics with many items are considered inefficient and reduce time for teachers, so a more practical holistic assessment instrument is needed. Computer Based Test (CBT) is seen as a solution that can integrate knowledge, attitude, and skills assessment more efficiently. In addition to the assessment instrument, material analysis also needs to be done to reveal factual, conceptual, procedural,

and metacognitive aspects to support research and development. (Branch, 2009). The selection of materials pays attention to students' low awareness of the environment, which can exacerbate ecosystem damage (Middleton, 2008). (Middleton, 2008). The selected material focuses on KD 3.8 and 4.8 of class VII semester II to improve students' understanding of environmental pollution and develop environmental responsibility (Wiggins, 1993). (Wiggins, 1993).

Design

This research focuses on two Basic Competencies (KD) in the 2013 Curriculum: KD 3.8, which asks students to analyze environmental pollution and its impacts, and KD 4.8, which emphasizes the skill of writing solutions based on observations. (Anderson, L. W., & Krathwohl, 2001).. Knowledge indicators include understanding (C2), application (C3), and analysis (C4), while attitude refers to response and appreciation of the environment. Communication skills include language use, clarity of writing, and organization of concepts. (Ramadina and Rosdiana, 2021)Collaboration is measured by contribution, productivity, flexibility, responsibility, and appreciation (Greenstein, 2012). (Greenstein, 2012). Sub-indicators developed according to the context of environmental pollution include identification of pollutant sources, ecological impacts, and management solutions (Irianto, 2015). (Irianto, 2015; Sastrawijaya, 1991). The instrument is in the form of 30 multiple choice questions, divided into four dimensions: knowledge (8 questions), attitude (7 questions), communication skills (8 questions), and collaboration skills (7 questions). The selection of the number of questions is based on the importance of each dimension in a holistic evaluation, thus providing a comprehensive picture of student achievement in environmental pollution material. (Mustafa, P.S. & Masgumelar, 2022).

This research uses a Sublime-based Computer Based Test (CBT) with PHP and MySQLi for flexibility of test settings and secure data storage. Supervised mode is used in CBT to maintain test integrity, supported by automatic question randomization that prevents cheating, maintaining the accuracy and validity of evaluation results.

Development

The development of holistic assessment instruments is carried out to ensure the accuracy and quality of the instrument through the stages of making and validation by experts in related fields. Validation is carried out internally and externally to ensure the instrument meets quality standards. The content validity of the items was obtained with input from two expert assessment lecturers, who ensured the items were able to measure holistic aspects realistically and contextually. Based on the validation results, the instrument was deemed suitable for testing after revision. Suggestions for improvement from the validators are summarized in Table 1 to identify areas that require revision. The revised questions were then recompiled as an initial product for limited testing, to evaluate its validity and reliability.

Table 1. Summary of Suggestions from Assessment Expert Validators

Input Source	Advice
Validator 1	Revisions related to the sentence structure of questions and answer options for question numbers 4, 6, 7, 8, 9, 10, 11, 12, 14, 16, 17, 19, 21, 22, and 23.
Validator 2	Revisions relate to the sentence structure of questions and answer options for question numbers 1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, and 27.

Media Validation

The media validation process was conducted by validators who are experts in the field of media and lecturers in the faculties of Computer Science and Educational Technology.

They assessed the validity and suggested improvements to the features in the Computer Based Test (CBT) which aims to measure student dimensions, such as knowledge, attitudes, communication skills, and collaboration skills related to the theme of environmental pollution. Validation was conducted through a 5-level Likert scale-based instrument (Sugiyono, 2019). This instrument is used to provide assessment and feedback on aspects of the media. The validation results are summarized in Table 2.

Table 2: Assessment of the suitability of the CBT media used

No.	Aspects	Percentage Score		Average Percentage Score	Criteria
		Expert 1	Expert 2		
1	*Computer Based Test* Display Design.	97%	97%	97%	Very suitable
2	Attractiveness of *Computer Based Test.*	100%	100%	100%	Very suitable
3	Computer Interactivity	100%	95%	97,5%	Very suitable
4	Ease of Use of *Computer Based Test*	90%	100%	95%	Very suitable
5	Computer-Based Test Potential	100%	100%	100%	Very suitable

Teacher-Practitioner Validation of CBT-Based Holistic Assessment Instruments

In addition to expert validation, the assessment development product was also evaluated by 20 science teachers, consisting of 15 statement items covering four aspects (computer-based test display design, computer-based test interactivity, computer-based test potential, and computer-based test usability). The validation results were analyzed using the CVR and CVI formulated by Lawshe (1957) (Ayre & Scally, 2014). The results showed a total CVR of 0.967 and a total CVI of 0.983. The CVR value of 0.967 indicates that all participants judged the test instrument to be fully consistent with the instrument validation indicators in each aspect. The CVR value is then used to calculate the CVI value. The CVI calculation result of 0.983 indicates that the overall validity level of this instrument is in the very suitable category.

Implementation Stage

The implementation stage is carried out by applying holistic assessment instruments that have gone through the validation process. The implementation stage is carried out through two stages, namely initial trials and field tests.

Initial Trial

Initial trials were conducted to determine the level of validity and reliability of the questions, and aimed to obtain empirical tests. The implementation time of the initial trial was May-June 2023. The selection of test subjects was class IX students at SMP Muhammadiyah Cianjur as many as 32 participants who had obtained environmental pollution material. The results of empirical validation show that all items of the instrument meet the valid criteria with a Sig value. (2-tailed) <0.05 with a positive *Pearson Correlation*. The results of reliability testing on the initial trial produced a value of 0.865 and was declared reliable and consistent because the *Cronbach's Alpha* value> 0.60.

Field Trial

In this study, 40 students of Class VII at SMP Muhammadiyah Cianjur in the even

semester of the academic year 2022/2023 were tested using a 30-item, 60-minute test. The questions covered 4 aspects: knowledge (8 questions), attitude (7 questions), communication skills (8 questions), and collaboration skills (7 questions). Questions 1-8 measure understanding of the concept of pollution, questions 9-15 assess attitudes and social awareness of pollution issues, questions 16-23 assess communication skills, and questions 24-30 assess students' collaboration skills through collaborative scenarios. The exam is expected to provide a comprehensive picture of students' knowledge, attitudes, communication skills, and collaboration. Based on the results of the field test conducted, an analysis can be made of:

1) Validity of the applied Holistic Assessment Instrument.

The validity test of the holistic assessment instrument obtained the results of 26 questions in the valid category, and there were 4 (four) questions in the invalid category, namely in questions number 2, 3, 10, and 28 with a calculated r value of 0.163; 0.081; 0.017; and 0.045 smaller than r table with a confidence level of 95% with 40 respondents, namely 0, 304. An assessment technique is said to be valid if it is able to accurately measure the specific skills expected of students (Arikunto, 2009). Invalid questions can be caused by several factors, among others, the language used in the items is less clear and difficult for students to understand, the level of students' understanding of the material provided, the choice of answers to the items given is less varied so that students can easily guess the answer.

2) Reliability Test

Reliability testing is also carried out on the data from the field implementation results to see the extent to which a test or assessment instrument can be relied upon in providing consistent results when used to measure the same aspects of different subjects. Reliability testing was carried out after eliminating invalid question items. The reliability test value shows 0.862 with the number of question items 26, so the instrument is declared reliable or consistent. Based on this, this holistic assessment instrument can be considered an effective tool in measuring students' ability to achieve the expected competencies (Arikunto, 2009).

3) Test Item Level of Difficulty

Based on the results of the calculation of the level of difficulty test of the holistic assessment instrument in this study, it shows that the questions prepared are not too difficult and not too easy. Problems with too easy categories are 5, problems with too difficult categories are 3, and problems with moderate categories are 18. The results of the calculation of the holistic assessment instrument difficulty level test are presented in Table 3:

Table 3. Calculation Results of Holistic Assessment Instrument Level of Difficulty Test

Question No.	Difficulty Level	Category	Question No.	Difficulty Level	Category
1	0,625	Enough (Medium)	17	0,375	Enough (Medium)
4	0,275	Too Difficult	18	0,600	Enough (Medium)
5	0,700	Enough (Medium)	19	0,650	Enough (Medium)
6	0,375	Enough (Medium)	20	0,775	Too Easy
7	0,275	Too Difficult	21	0,500	Enough (Medium)
8	0,450	Enough (Medium)	22	0,450	Enough (Medium)
9	0,725	Too Easy	23	0,550	Enough (Medium)
11	0,850	Too Easy	24	0,425	Enough (Medium)
12	0,650	Enough (Medium)	25	0,850	Too Easy
13	0,725	Too Easy	26	0,525	Enough (Medium)
14	0,700	Enough (Medium)	27	0,500	Enough (Medium)
15	0,475	Enough (Medium)	29	0,650	Enough (Medium)

Question No.	Difficulty Level	Category	Question No.	Difficulty Level	Category
16	0,275	Too Difficult	30	0,425	Enough (Medium)

The test instrument is said to be good if it has a difficulty level between the interval 0.30 - 0.70.

Problems that fall into the difficult category can be caused by some students not understanding the material taught, and some even not preparing themselves optimally. Questions in the medium category indicate that some students were able to answer correctly, but there were still inaccuracies in the answers. This illustrates that students' abilities during the sample test were fairly low. The question items that fell into the too easy category were caused by some students who had a good understanding of the material taught so that they could answer the questions well.

4) Distinguishing Power Test

The distinguishing power test is used to separate students who have an understanding of the material being tested and students who have not mastered the material. The distinguishing power test also plays a role in evaluating the quality of each item. The goal is to determine whether an item has good characteristics, needs revision, or must be rejected. The test instrument is considered effective, if the item has a differentiating power level of at least 0.20 or is stated at least as sufficient. The calculation results of the differentiating power test are presented in Table 4 below:

Table 4. Calculation Results of Item Differentiation Test

Question No.	Distinguishing Power	Category	Question No.	Distinguishing Power	Category
1	0,55	*Good*	17	0,45	*Good*
4	0,35	*Satisfactory*	18	0,60	*Good*
5	0,40	*Good*	19	0,50	*Good*
6	0,45	*Good*	20	0,45	*Good*
7	0,25	*Satisfactory*	21	0,50	*Good*
8	0,40	*Good*	22	0,20	*Satisfactory*
9	0,35	*Satisfactory*	23	0,40	*Good*
11	0,20	*Satisfactory*	24	0,45	*Good*
12	0,40	*Good*	25	0,30	*Satisfactory*
13	0,35	*Satisfactory*	26	0,55	*Good*
14	0,40	*Good*	27	0,30	*Satisfactory*
15	0,35	*Satisfactory*	29	0,50	*Good*
16	0,45	*Good*	30	0,35	*Satisfactory*

A question that does not have good discriminating power is one that can be answered correctly by both groups of students, high and low ability, or by no one. Ideally, questions can only be answered correctly by high ability students, and if this happens, the question has a good discriminating power (D). Based on the discriminating power analysis, the next steps are:

a). Questions with satisfactory, good, and excellent discriminating power should be included in the question bank for reuse.

b). Questions with low (poor) discriminating power must be corrected and retested.

Questions with a negative discriminating power should not be reused because they are of poor quality, since high ability students are more likely to answer incorrectly than low ability students.

5) Distractor Function Analysis

Based on the research results of the holistic assessment test instrument on environmental pollution learning in class VII SMP Muhammadiyah Cianjur, the calculation results show that 25 questions are in the accepted category, and 1 question has a rejected category for one of the answer options.

Evaluation Stage

Evaluate by analyzing the results of the field trial. Then improve the developed product based on suggestions and input from students and science teachers during the implementation stage so that the product can be used. Based on the calculation of the Guttman scale, students' responses to the CBT-based test instrument obtained a result of 89.29 with "very good" criteria. The implications of this research for education are as follows: (1) the use of holistic assessment instruments based on Computer Based Test (CBT) can improve the quality of assessment because it provides a more comprehensive picture of students' understanding and helps teachers to assess critical thinking skills and concept application (Purnawirawan, 2019); (2) technology in assessment, such as CBT, can increase students' motivation and engagement so that students are more interested in the material being taught (Ruswan *et al*., 2024; Sola *et al*., 2022); (3) Computer Based Test (CBT) helps students develop digital skills, including the use of technological tools and digital problem solving, which are increasingly needed in the world of work (Handayani et al., 2024); (4) Computer Based Test (CBT) system allows students to receive feedback immediately after the assessment, so it can more quickly facilitate the improvement of student understanding (Selvia & Prastowo, 2024); (5) Findings from this study can provide a basis for developing better educational policies, especially in terms of technology integration in learning and assessment (Soleh M.I., & Nur Efendi, 2023).

Conclusion

The *Computer Based Test-based* holistic assessment instrument on environmental pollution learning for class VII has passed validation by various related parties. The results of validation by experts in the field of assessment show that the items of this instrument are declared suitable for trial use after revision. Validation by media experts also stated that the *Computer Based Test* used met the criteria very suitable with an average percentage between 97% and 100%. Furthermore, the results of validation by science teacher practitioners resulted in a total CVR value of 0.967 and a total CVI of 0.983. The CVR value which reached 0.967 indicated that all participants assessed this test instrument as fully in accordance with the instrument validation indicators in each aspect, while the CVI value of 0.983 indicated that the overall validity level of this instrument was in the very suitable category. In addition, the instrument was also declared reliable or consistent with a reliability test value of 0.862.

Recommendation

Teachers are encouraged to use holistic assessment that includes cognitive, affective, and psychomotor aspects when evaluating students, especially in the area of environmental pollution. Combine CBT-based assessments with other methods, such as group projects, to get a complete picture of students' abilities. Provide constructive feedback to improve student understanding and regularly monitor the effectiveness of the assessment. Teachers also need technology training to administer CBT effectively.

Policymakers should ensure adequate technology infrastructure in schools to support CBT. Adapt the curriculum to a holistic approach that includes technology in assessment.

Establish national standards to ensure consistency and quality of assessment. Provide funding and support programs for teacher training and encourage research on the effectiveness of technology-enhanced holistic assessment.

References

Aisyah, A., & Dewi, S. (2019). Pengembangan Soal Tipe Pisa Untuk Mengukur Kemampuan Komunikasi Matematis Siswa Tingkat Sekolah Menengah Pertama Di Kota Jambi. *PHI: Jurnal Pendidikan Matematika*, *2*(2), 78. https://doi.org/10.33087/phi.v2i2.34

Anas Sudjiono. (2016). *Pengantar Evaluasi Pendidikan*. Rajawali Pers.

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longmans, Green and Co.

Apriyana, A., Hasyim, A., & Sabdaningtyas, L. (2019). Pengembangan Instrumen Penilaian Sikap Sosial Pada Pembelajaran Tematik Subtema Ayo Cinta Lingkungan. *Jurnal Pedagogi*, *1*. http://jurnal.fkip.unila.ac.id/index.php/pgsd/article/view/17504

Arikunto, S. (2009). *Dasar-Dasar Evaluasi Pendidikan*. Bumi Aksara.

Arikunto, S. (2013). *Prosedur Penelitian: Suatu Pendekatan Praktik*. Rineka Cipta.

Astuti, S. I., & Darsinah, D. (2018). Penilaian Autentik Berbasis Kurikulum 2013 di SD Negeri Mangkubumen Kidul No. 16 Surakarta. *Manajemen Pendidikan*, *13*(2), 165–.

Ayre, C., & Scally, A. J. (2014). Critical values for Lawshe's content validity ratio: Revisiting the original methods of calculation. *Measurement and Evaluation in Counseling and Development*, *47*(1), 79–86. https://doi.org/10.1177/0748175613513808

Azwar, S. (2022). *Reliabilitas dan Validitas* (IV). Pustaka Pelajar.

Branch, R. M. (2009). Approach, Instructional Design: The ADDIE. In *Department of Educational Psychology and Instructional Technology University of Georgia* (Vol. 53, Issue 9).

Dahar, R. W. (2011). *Teori-teori Belajar dan Pembelajaran*. Erlangga.

Greenstein, L. M. (2012). *Assessing 21st century skills: A guide to evaluating mastery and authentic learning.*

Hairida, H. (2018). Penilaian Sikap Siswa Dalam Pembelajaran Kimia Melalui Teknik Self Assessment Dan Peer Assessment. *Jurnal Pendidikan Matematika Dan IPA*, *9*(2), 37. https://doi.org/10.26418/jpmipa.v9i2.25832

Handayani, Nurlaelah, I., Setiawanti, I., Wildani, N., & Rahmah F, D. (2024). Media Digital Dalam Pembelajaran Untuk Mendukung KeterampilanAbad 21 di SMP Plus Miftahul Falah Kuningan. *Pandawa : Pusat Publikasi Hasil Pengabdian Masyarakat*, *2*(3), 17–24.

Irianto, I. K. (2015). Buku Bahan Ajar Pencemaran Lingkungan. In *Buku Bahan Ajar Pencemaran Lingkungan*. Universitas Marwadewa.

Kunandar. (2014). *Penilaian Autentik*. Raja Grafindo Persada.

Manurung, A. B., & Rajagukguk, J. (2019). Desain Evaluasi Hasil Belajar Fisika Berbasis Computer Based Test (CBT) Pada Materi Pokok Usaha Dan Energi. *INPAFI (Inovasi Pembelajaran Fisika)*, *6*(4). https://doi.org/10.24114/inpafi.v6i4.12479

Maulana, R. (2022). Penggunaan Sistem Computer Based Test (CBT) untuk Optimalisasi Pelaksanaan Penilaian Akhir Semester (PAS) Ganjil Tahun Pelajaran 2021-2022. *Jurnal Ilmiah Widya Borneo*, *5*(1), 1–14.

Middleton, N. J. (2008). *Environmental Education and awareness*. EOLSS Publishers.

Mustafa, P.S. & Masgumelar, N. K. (2022). Pengembangan Instrumen Penilaian Sikap,

Pengetahuan, dan Keterampilan dalam Pendidikan Jasmani. *Biormatika : Jurnal Ilmiah Fakultas Keguruan Dan Ilmu Pendidikan*, *8*(1), 31–49. https://doi.org/10.35569/biormatika.v8i1.1093

Primantoro, A. D. (2016). PENDIDIKAN NILAI MORAL DITINJAU DARI PERSPEKTIF GLOBAL Alfian. *Https://Medium.Com/*, *1*(1). https://medium.com/@arifwicaksanaa/pengertian-use-case-a7e576e1b6bf

Purnawirawan, O. (2019). Pengembangan Instrumen Penilaian 4C (Creativity, Critical Thinking, Communication, Dan Collaboration) Sistem Pembelajaran Abad Dua Satu Dalam Pengajaran Bidang Produktif Sekolah Menengah Kejuruan Tesis. In *Program Pascasarjana Universitas Negeri Semarang*.

Ramadina dan Rosdiana. (2021). KETERAMPILAN KOMUNIKASI SISWA SETELAH DITERAPKAN STRATEGI ACTIVE KNOWLEDGE SHARING KETIKA PEMBELAJARAN DARING. *PENSA: E-JURNAL PENDIDIKAN SAINS*, *9*, 247–.

Rubiyanto, N. (2010). *Strategi Pembelajaran Holistik di Sekolah*. Prestasi Pustakaraya.

Ruswan, Acep, et al. (2024). Pengaruh Penggunaan Media Pembelajaran Berbasis Teknologi dalam Meningkatkan Kemampuan Literasi Digital Siswa Sekolah Dasar. *Program Studi Pendidikan Guru Sekolah Dasar, Universitas Pendidikan Indonesia*, *8*(1), 4007–4016. https://doi.org/10.31004/jptam.v8i1.13009

Samsiadi, S., & Humaidi, M. N. (2022). Efektivitas Google Form Sebagai Media Penilaian Dan Evaluasi Pembelajaran Pai Di Smk Negeri 1 Berau Kaltim. *Research and Development Journal of Education*, *8*(2), 666. https://doi.org/10.30998/rdje.v8i2.13634

Sastrawijaya, A. . (1991). *Pencemaran Lingkungan* (1st ed.). Rineka Cipta.

Selvia, N., & Prastowo, A. (2024). Implementasi Guru dalam Pembelajaran Matematika Menggunakan Model Tutorial di Sekolah Dasar Wonosari. *Al Madrasah Jurnal Pendidikan Madrasah Ibtidaiya*, *8*(3), 1024. https://doi.org/10.35931/am.v8i3.3515

Simarmata, N. N., Wardani, N.S., &, & Prasetyo, T. (2019). Pengembangan Instrumen Penilaian Sikap Toleransi Dalam Pembelajaran Tematik Siswa Kelas IV SD. *Jurnal Basicedu*, *3*(1), 194–199. https://doi.org/10.31004/basicedu.v3i1.101

Sola, E., Ismatun Amriyah Bahtiar, Musdalifa, & Azwan Sudarman. (2022). Pengaruh Media Pembelajaran Berbasis Teknologi Terhadap Hasil Belajar Mahasiswa Mpi Kelas B Semester Iv Uin Alauddin Makassar. *Educational Leadership: Jurnal Manajemen Pendidikan*, *2*(01), 48–61. https://doi.org/10.24252/edu.v2i01.30610

Soleh M.I., & N. E. (2023). Integrasi Teknologi Dalam Manajemen Pendidikan Islam: Meningkatkan Kinerja Guru di Era Digital. *Jurnal Tinta*, *5*(2), 104–126.

Sugiyono. (2019). *Metode Penelitian Pendidikan* (Edisi ke-3). Alfabeta.

Sujarweni, W. (2014). *Metodologi Penelitian:Lengkap, praktis, dan mudah dipahami*. Pustaka Baru Press.

Tristaningrat, M. A. . (2018). Analisis Kritis Korelasi Asesmen Autentik dalam Pembelajaran. *Widyacarya*, *2*(1).

Utami, R. D. (2019). Aplikasi Computer-Based Test Ujian pada SMP Negeri 6 Cimahi. *Jurnal Teknologi Dan Informasi*, *9*(2), 151–160. https://doi.org/10.34010/jati.v9i2.2234

Wiggins, G. P. (1993). *Assessing student performance: Exploring the purpose and limits of testing*. Jossey-Bass.

Yunitasari, I. (2019). *PENGEMBANGAN INSTRUMEN PENILAIAN AUTHENTIC UNTUK MENILAI KETERAMPILAN OBSERVING, INFERRING, DAN PREDICTING PADA PEMBELAJARAN IPA*. UNIVERSITAS NEGERI YOGYAKARTA.