**Development of a Holistic Assessment Instrument *Based on Computer Based Test* on Environmental Pollution Learning in Class VII**

**Dini Rosdiane Rusdiyanto1,2, Bibin Rubini,1\*** **Irvan Permana1**

1Department of Science Education, Postgraduate School, Pakuan University, Bogor, Indonesia.

2 SMP Muhammadiyah, Cianjur, Indonesia

**\*Corresponding** Author e-mail: [**bibinrubini@unpak.ac.id**](mailto: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 class VII. The research method used is *Research and Development* (R&D) with ADDIE research design (*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. 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 class VII and provide comprehensive information about student achievement in various dimensions of learning. | **Article History**  Received: 2  Revised:  Published. 2017  **Key Words:**  holistic assessment, *computer-based test*, environmental pollution. |

**How to Cite:** Dini Rosdiane Rusdiyanto, Bibin Rubini, Irvan Permana (2023). Development of a Holistic Assessment Instrument *Based on Computer Based Test* on Environmental Pollution Learning in Class VII. Journal of Education: Journal of Research Results and Literature Review in the Field of Education, Teaching and Learning, vol (no). doi:https://doi.org/10.33394/jk.vxxyyi

|  |  |
| --- | --- |
| C:\Users\IKIP\Pictures\1200px-DOI_logo.svg.png https://doi.org/10.33394/jk.vxxyyi | This is an open-access article under the [CC-BY-SA License.](http://creativecommons.org/licenses/by/4.0/) |

C:\Users\IKIP\Pictures\CC_BY-SA_3.0.png

**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. In response to the dynamics of the global community, UNESCO in 1996 established four pillars of education as the foundation for this global era. 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.

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 academically intelligent, but also able to face and solve the challenges of the global world with a deep and applicable understanding (Rubiyanto, 2010). (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. 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) is one of the strategic steps in improving the efficiency and relevance of assessment. 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.

**Research Method**

The research method chosen is *Research and Development* (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.

**Result and Discussion**

This research is a type of development research. 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 research on the development of holistic assessment instruments based on *Computer Based Test* (CBT) at the needs analysis stage involving literature study, questionnaire distribution, and interviews, include 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). Through this holistic assessment, it is expected that students can be assessed comprehensively, so that teachers can understand not only the level of mastery of the material by students, but also how they apply the knowledge in a real context and how their attitude towards learning. With the implementation of CBT, the assessment process can be conducted more efficiently and objectively, as well as enabling a more in-depth analysis of various aspects of students' learning achievements. This analysis is important to formulate more effective learning strategies that suit students' needs and to improve the overall quality of education.

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 show that teachers still experience 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 conducted in December 2022 opened a window of insight into the constraints faced by science teachers. The focus of the conversation involved three main aspects, namely curriculum, science learning, and assessment instruments, so as to provide a holistic picture of the challenges in implementing learning. The results of the interviews revealed that most teachers are less active in designing assessment instruments and learning tools independently. The process of designing assessment instruments and learning tools is more often done together during the Subject Teacher Conference (MGMP) activities. Obstacles arise when teachers are faced with practical work assessment and attitude assessment. They revealed that learning time during practical work is often used up to provide guidance to students, and finally there is no additional time available to carry out careful assessment. The same thing was also revealed in research conducted by Yunitasari (2019)who conducted research on the development of *authentic* assessment instruments to assess *observing*, *inferring*, and *predicting* skills in science learning.

Not only that, the attitude assessment instrument is also a problem because it still uses a rubric with many items to be assessed, causing inefficiency in the assessment process. Teachers feel limited in providing a thorough assessment due to time constraints and assessment tools that do not support efficiency. As a consequence, science teachers tend to give a general assessment after the learning takes place. Based on the results of these interviews, it can be identified that there is a need for holistic assessment instruments that can help overcome obstacles in assessment. A *Computer Based Test* (CBT) based assessment instrument is a relevant solution, as it can help integrate knowledge, attitude, and skills assessment more efficiently, reduce teacher workload, and provide more accurate assessment results.

Material analysis is also carried out to explore and reveal aspects including facts, concepts, procedural, and metacognitive in learning materials. The goal is to find and use materials that fully support the research and development process carried out. (Branch, 2009). The researcher's decision in selecting materials was based on a deep understanding of the actual situation in the field, especially related to students' awareness of the importance of environmental preservation which is decreasing. This selection was triggered by concerns about the potential for serious impacts if this situation continues to be allowed to develop. Low awareness of the environment can result in increasingly severe damage to the ecosystem (Middleton, 2008). (Middleton, 2008).

Through in-depth analysis of this phenomenon, the researcher selected material focused on Basic Competencies (KD) 3.8 and 4.8 of the second semester VII curriculum, designed to respond to actual problems in the field, especially the problem of environmental pollution and students' lack of awareness of its impact. The selection of this specific material was done with the clear objective of providing a solid foundation for the design and development of relevant comprehensive assessment tools. It is expected that these measures will not only contribute to students' understanding of environmental issues, but also stimulate positive behavioral changes. A solid foundation in assessment is expected to have a positive impact on students' understanding and responsibility in dealing with environmental problems that are increasingly urgent to address. (Wiggins, 1993).

**Design**

The Core Competencies (KI) to be achieved in the research are in line with the foundation stated in Curriculum 2013, which consists of four crucial assessment dimensions: spiritual dimension, social dimension, knowledge dimension, and skill dimension. The focus of the research is on two Basic Competencies (KD) that refer to the knowledge and skills dimensions. KD 3.8, which focuses on the knowledge dimension, requires students to be able to analyze the occurrence of environmental pollution and its impact on the ecosystem . Meanwhile, KD 4.8, which covers the skill dimension, requires students to be able to produce writing that describes solutions to pollution problems in their environment based on observations made.

Competency Achievement Indicators (IPK) are formulated by referring to each Basic Competency (KD) and referring to the level of learning objectives proposed by Anderson, L. W., & Krathwohl (2001). The knowledge dimension, the researcher identified three relevant indicators, namely students' ability to understand (C2), apply (C3), and analyze (C4) content related to environmental pollution. The attitude dimension, there are two indicators selected, namely students' response (A2) to environmental issues and students' ability to appreciate (A3) the importance of environmental maintenance. Communication skills dimension based on Ramadina and Rosdiana (2021) three indicators were selected to measure students' skills in using good language, writing problem solutions precisely and clearly, and organizing concepts well in their writing. The dimension of collaboration skills involves five indicators, which include students' ability to contribute actively, work productively in groups, show flexibility and the ability to compromise, bear responsibility for group tasks, and show respect for the opinions and efforts of group mates. (Greenstein, 2012).

The whole series forms a holistic assessment framework, which not only measures aspects of knowledge, but also involves attitudes, communication skills and collaboration skills. This research seeks to measure students' overall competency achievement and reflects the spirit of inclusive and sustainable curriculum development.

The stage of formulating sub-indicators in the learning process has a very important role in linking the overall indicators with a more specific context, especially in this study, the context of environmental pollution material. Through this process, the indicators that have been determined previously as general guidelines are directed to be more relevant and in accordance with the learning objectives aimed at environmental pollution material.

The process of formulating sub-indicators requires a careful and thorough approach. Educators need to refer to pre-existing indicators, while also adapting and customizing them to suit the main focus of learning, which is understanding the issue of environmental pollution. In this context, careful steps are needed to identify key aspects relevant to the issue of environmental pollution, such as identification of pollutant sources, understanding the ecological impacts caused, and managing solutions to overcome these problems. (Irianto, 2015; Sastrawijaya, 1991).

Through the process of formulating sub indicators, the learning approach can be directed more sharply according to the context at hand, namely the problem of environmental pollution. This provides an opportunity for educators to focus more on directing students' attention to important and relevant aspects related to the material. The process of formulating sub-indicators not only bridges the relationship between the overall indicator and the specific material, but also creates a deeper, more meaningful and empowering learning experience for students. Through the process of formulating appropriate sub-indicators, learning becomes more contextualized, arouses curiosity, and provides students with strong tools to understand and solve real problems related to the environment (Elaine B. Johnson, 2010). (Elaine B. Johnson, 2010).

Referring to the sub-indicators that have been described for knowledge, attitudes, and skills, the next step is to design question indicators. These question indicators have an important role as the main guide in the process of preparing evaluation question items. The question indicator becomes the foundation that directs the creation of questions that are relevant and in accordance with the evaluation objectives that have been previously set. (Mardapi, 2015). The sub-indicators that have been made are converted into more specific and measurable statements, which are used as a reference in making question indicators. Question indicators include key elements to be measured in the evaluation process, both in terms of knowledge understood, attitudes expected, and skills that students must have. Through the design of appropriate question indicators, the process of preparing question items becomes more focused and in accordance with the predetermined learning objectives, so that the learning evaluation process is of higher quality. (Rahman, 2019).

This study, applying the usual form of multiple choice questions with four answer choices (Sudjana, 2020). The total number of questions that have been designed is 30 questions, which are evenly divided between the four dimensions of assessment measured.

The knowledge dimension is represented by 8 questions designed to measure students' understanding of the material on environmental pollution. In the attitude dimension, there are 7 questions that aim to explore students' responses and attitudes towards environmental issues as well as their understanding of the importance of environmental preservation. Meanwhile, the communication skills dimension is measured through 8 questions, which assess students' ability to convey information clearly and precisely regarding solutions to environmental pollution problems. In the collaboration skills dimension, another 7 questions are designed to measure students' ability to work together, contribute actively, and show respect and responsibility in a group or team context.

The selection of the number and distribution of questions is based on the importance of each dimension in achieving holistic evaluation objectives. Therefore, the items designed must be able to cover all aspects of the competencies to be measured, so that the assessment process can provide a comprehensive picture of student achievement in lessons on environmental pollution. (Mustafa, P.S. & Masgumelar, 2022)..

This research adopts the *Computer Based Test* (CBT) approach by using Sublime software as the main platform. In addition, PHP and MySQLi programming languages were used in the development of this CBT. The use of Sublime as the CBT platform provides the flexibility and control needed in designing and organizing online tests. The platform allows for the creation of various types of questions, timing, as well as the ability to securely manage and store test result data.

The choice of PHP and MySQLi programming languages provides a solid framework for the development of CBT systems. PHP is a server-side programming language that can be used to generate dynamic web pages, while MySQLi is an extension to manage connections and operations against the MySQLi database. This combination allows the development of complex features, such as the automation of student answer storage, score calculation, as well as the easy provision of exam result reports. The use of Sublime, PHP, and MySQLi as the main components in the development of CBT guarantees that this system will meet the complex evaluation needs in the research. Through CBT media, the examination and evaluation process can be conducted effectively and accurately, providing strong support in measuring students' comprehensive achievement in knowledge, attitude, and skills related to environmental pollution learning.

This research applies the *Computer Based Test* (CBT) mode by using the *supervised mode* approach. The choice to use *supervised mode* in CBT was made with careful consideration to ensure that computer-based examinations can take place in a monitored and controlled manner. In supervised mode, the examination is monitored by invigilators or proctors, who are responsible for maintaining the integrity of the examination and preventing cheating. This approach allows proctors to monitor student behavior during the exam, including monitoring the use of computer devices, access to outside resources, and communication between students. *Supervised mode* aims to maintain fairness, honesty, and validity of the evaluation results.

The application of *supervised mode* in CBT provides confidence that the data collected is a true reflection of students' abilities. This is important in the context of research that measures multiple dimensions, such as knowledge, attitudes and skills, all of which must be measured accurately and consistently. Therefore, the use of *supervised mode* ensures that the evaluation process takes place with integrity and accuracy, providing results that represent the true achievements in student learning and character building.

This study uses a *Computer Based Test* (CBT) system with the order of questions randomized for each student taking the exam. The concept of randomizing the order of questions aims to prevent cheating or collusion between students, as well as to avoid uniform answer patterns among them. The randomized order of questions provides an opportunity for each student to get a different set of questions from the others, although it still measures the same aspects of knowledge, attitudes, and skills. In addition, randomizing the order of questions can also provide a more varied exam experience for students, minimizing the existence of repetitive answer patterns that might occur if the order of questions is fixed.

The process of randomizing the order of questions is done by the system automatically, so it does not require manual intervention. This can also make it easier for the exam supervisor to monitor the exam, as each student gets a different question sequence. By randomizing the order of questions, the CBT system provides better protection for the integrity of the exam, and maintains the validity and accuracy of the evaluation results.

**Development**

The development stage of a pre-planned holistic assessment instrument is an important step in ensuring the accuracy and quality of the instrument. After designing the instrument, the next step is manufacturing and validation. The validation process is carried out by experts who have competence and expertise in related fields. This validation is carried out both internally and externally to ensure that the holistic assessment instruments that have been developed meet the established standards and have superior quality.

The content validity of the items in this study was carried out by asking for consideration from two expert lecturers in the field of assessment. The validators ensured that the question instrument was well designed and able to measure the desired holistic aspects more realistically and contextually. The results of validation by experts in the field of assessment showed that the items of this instrument were declared suitable for trial use after revision.

A summary of the suggestions given by the validators regarding the content of the items is presented in Table 1. Through this table, areas that require improvement or refinement in the items in the *Computer Based Test* (CBT) can be identified.

**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. |

Suggestions and input from validators were used to revise the questions and then rearranged as initial products. The revised initial product was then tested in a limited trial to evaluate the validity and reliability of the items that had been developed.

**Media Validation**

The media validation process is carried out by validators who have expertise in the field of media as well as lecturers in the Faculty of Computer Science, as well as lecturers in the Faculty of Educational Technology. These validators assessed the validity and provided suggestions for improvements to the features contained in the *Computer Based Test* (CBT) that had been developed. This CBT aims to measure various dimensions in students, including knowledge, attitudes, communication skills, and the ability to collaborate in the context of the theme of environmental pollution. The validation process is carried out through filling out a validation instrument sheet that uses a Likert scale with 5 levels (Sugiyono, 2019). (Sugiyono, 2019). Through this scale, the validators provide assessment and feedback on various aspects of the evaluated media. The results of the media validation process are presented and summarized in Table 2, which shows the results of the validation efforts that have been carried out.

**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 |

1. **Teacher Practitioner Validation of CBT-Based Holistic Assessment Instrument**

In addition to being validated by experts, the assessment development product was also assessed 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 Ease of Use of *Computer Based* Test). The validation results were analyzed using CVR and CVI formulated by Lawshe (1957). (Ayre & Scally, 2014)The results of the validation were analyzed using the CVR and CVI formulated by Lawshe (1957) (Ayre & Scally, 2014), with the total CVR of 0.967 and the total CVI of 0.983. The CVR value of 0.967 indicates that all participants have assessed that the test instrument is fully in accordance with the instrument validation indicators in each aspect. Then, the CVR value is used to calculate the CVI value. The results of the CVI calculation of 0.983 indicate 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**

Researchers continued the field trial on 40 seventh grade students at SMP Muhammadiyah Cianjur even semester of the 2022/2023 academic year. The test was conducted after the learning process was complete, with a processing time of 60 minutes. The test questions consisted of 30 items covering 4 aspects, namely knowledge (8 questions), attitudes (7 questions), communication skills (8 questions), and collaboration skills (7 questions).

These steps aim to detail the extent to which the holistic assessment instrument can reflect students' understanding in various aspects of learning. Questions number 1 to 8 in this exam are specifically designed to measure students' knowledge ability related to the main concepts of environmental pollution, such as sources of pollutants, their impacts, and efforts to control them. The design of these questions takes into account a variety of difficulty levels to cover a range of student abilities, in line with *Bloom*'s taxonomy at the C2 (understanding), C3 (applying) and C4 (analyzing) levels. These questions require students' understanding of the learning materials and connect them to real situations that students can face in their daily lives.

Questions 9 to 15 on this exam were specifically designed to delve deeper into students' attitudes towards the issue of environmental pollution. Through these questions, the researcher sought to explore students' environmental values and ethics, utilize contextual situations and detail their level of involvement in environmental activities or projects. Meanwhile, these questions also attempted to measure the students' level of social awareness of pollution issues facing society. These questions are expected to provide a comprehensive picture of students' attitudes towards environmental pollution, allowing educators to understand their values, ethics and involvement in efforts to maintain environmental sustainability and balance.

Questions 16 to 23 in this exam are specifically designed to measure students' communication skills in the context of the environmental pollution theme. These holistically structured multiple-choice questions integrate various aspects to evaluate students' understanding, analysis, and ability to convey information effectively. The questions are designed to include situational contexts that reflect students' daily lives related to environmental pollution, so that students are tested not only on their theoretical understanding of concepts, but also on their ability to apply that knowledge in real situations. Furthermore, in selecting the types of questions, special attention was paid to the use of specialized vocabulary related to environmental pollution to ensure that students could communicate their ideas clearly and accurately. In addition, the questions also focused on students' ability to communicate the risks and consequences of pollution. This represents an attempt to measure students' analytical skills and their ability to convey information that can build awareness of the risks and consequences associated with human behavior towards the environment. There are questions that ask students to choose the best solution to environmental pollution. This not only tests their understanding of the issue but also measures their ability to effectively organize ideas and recommendations. Not only that, the assessment also covers the dimension of students' personal actions. Through questions that ask students to choose answers that reflect how information about environmental pollution can motivate personal actions, we can measure students' ability to respond to information responsively and make sustainable decisions.

Questions 24 to 30 in this exam are geared towards evaluating students' collaboration skills, particularly in the context of the environmental pollution theme. The questions are carefully designed to create a situation or context for collaboration that raises challenges or problems related to environmental pollution that can be solved together. Through questions number 24 to 30, students are invited to collaborate in solving complex situations related to environmental pollution. Each problem presents a unique scenario that requires active participation and contribution from each team member. The scenarios reflect real challenges that students may face in the real world, such as designing a waste reduction program at school. The cooperative situations or contexts presented in the problems are designed to encourage students to think collaboratively, share ideas, and design solutions that can be implemented together. By including concrete challenges or problems, students are not only tested on their ability to collaborate but also in applying their knowledge related to environmental pollution in a practical context. These situations also provide space for students to determine the role of each team member in solving the problem. This aims to guide students in detailing their respective responsibilities and contributions, creating a balanced and effective cooperation dynamic. Incorporating challenging contextual elements, these questions provide a holistic picture of students' ability to collaborate, communicate and come up with creative and sustainable solutions related to the theme of environmental pollution. Thus, the assessment measures not only the end result but also the collaborative process of students in responding to real challenges in the environmental field.

The results of the test can provide a clearer picture of students' development and progress in terms of knowledge, attitudes, and communication and collaboration skills. Thus, the implementation and testing stages of the holistic assessment instrument are expected to provide a more comprehensive view of the learning achievements of students of SMP Muhammadiyah Cianjur grade VII.

Based on the results of the field test that has been carried out, it can be analyzed:

1. **Validity of the Holistic Assessment Instrument Used.**

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 evaluation technique is said to be valid if it is able to accurately measure the specific abilities expected of learners (Arikunto, 2006). (Arikunto, 2006). 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 student understanding of the material provided, the choice of answers to the items given is less varied so that students easily guess the answer.

1. **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, 2006).

1. **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) |
| 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.

1. **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 a question that can be answered correctly by both groups of students, both high and low ability. Conversely, if both groups of students cannot answer correctly, then the question is considered not to have sufficient differentiating power. Ideally in a test context, questions should only be answered correctly by high-ability students, for example, if high-ability students are able to answer a test correctly, while almost all or all low-ability students answer incorrectly, then the question has a good discrimination index (D).

As a follow-up to the results of this discriminating power analysis, there are several steps that can be taken:

1. Questions that have *satisfactory*, *good*, and *excellent* discriminating power should be included in the question bank and can be reused in subsequent tests because their quality is adequate.
2. Questions that still have low (*poor*) differentiating power need to be further explored and improved. After improvement, these questions can be re-submitted in the learning outcomes test to be analyzed whether their differentiating power has increased or not.

Questions that have a negative discrimination index number should not be used in learning outcome tests, as it indicates that the quality of the question is very poor. This is due to the fact that high-ability students answer incorrectly more often than low-ability students, and only a few students answer correctly.

1. **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.

**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**

In order to gain a more representative understanding of the effectiveness of CBT-based holistic assessment instruments, more thorough trials involving more schools and diverse samples are needed. The use of interactive technologies, such as simulations and animations, can also help students learn the concept of environmental pollution visually and practically, while increasing their participation in the assessment. However, barriers such as limited access to computer devices and internet networks, especially in remote areas or schools with limited funding, can be an obstacle. Lack of digital skills among teachers and students can also reduce the effectiveness of using CBT, so proper training is essential. In addition, differences in access to technology at home can lead to gaps in student readiness, ultimately affecting assessment results.

**Acknowledgment**

We would like to deliver our sincere thanks to the school leaders, teachers and students of SMP Muhammadiyah Cianjur. We also thank the Pakuan University for the source of funds provided in the assignment research. In addition, sincere thanks are also conveyed to experts, practitioners and students who have participated in and supported the completion of this research.

**References**

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). Development of Social Attitude Assessment Instruments on Thematic Learning Subtheme Let's Love the Environment. *Journal of Pedagogy*, *1.* http://jurnal.fkip.unila.ac.id/index.php/pgsd/article/view/17504

Arikunto, S. (2006). *Basics of Educational Evaluation*. Bumi Aksara.

Astuti, S. I., & Darsinah, D. (2018). 2013 Curriculum-Based Authentic Assessment at SD Negeri Mangkubumen Kidul No. 16 Surakarta. *Education Management*, *13*(2), 165-174.

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

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). *Theories of Learning and Learning*. Erlangga.

Elaine B. Johnson. (2010). *Contextual Teaching Learning*. Kaifa.

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

Hairida, H. (2018). Assessment of Student Attitudes in Chemistry Learning Through Self Assessment and Peer Assessment Techniques. *Journal of Mathematics and Science Education*, *9*(2), 37. https://doi.org/10.26418/jpmipa.v9i2.25832

Irianto, I. K. (2015). Teaching Materials for Environmental Pollution. In *Book of Teaching Materials for Environmental Pollution*. Marwadewa University.

Kunandar. (2014). *Authentic Assessment*. Raja Grafindo Persada.

Manurung, A. B., & Rajagukguk, J. (2019). Design of Evaluation of Physics Learning Outcomes Based on Computer Based Test (CBT) on the Subject Matter of Effort and Energy. *INPAFI (Physics Learning Innovation)*, *6*(4). https://doi.org/10.24114/inpafi.v6i4.12479

Mardapi, D. (2015). *Measurement, Assessment, and Evaluation of education* (1st ed.). Nuha Litera.

Maulana, R. (2022). Use of the Computer Based Test (CBT) System to Optimize the Implementation of Odd Semester Final Assessment (PAS) for the 2021-2022 Academic Year. *Widya Borneo Scientific Journal*, *5*(1), 1-14.

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

Mustafa, P.S. & Masgumelar, N. K. (2022). Development of Attitude, Knowledge, and Skills Assessment Instruments in Physical Education. *Biormatics: Scientific Journal of the Faculty of Teacher Training and Education*, *8*(1), 31-49. https://doi.org/10.35569/biormatika.v8i1.1093

Rahman, A. & N. C. E. (2019). *Teaching Evaluation*. Uwais Inspirasi Indonesia.

Ramadina and Rosdiana. (2021). Student's Communication Skills After Applying Active Knowledge Sharing Strategy During Online Learning. *PENSA: E-JOURNAL OF SCIENCE EDUCATION*, *9*, 247-251.

Rubiyanto, N. (2010). *Holistic Learning Strategies in Schools*. Prestasi Pustakaraya.

Samsiadi, S., & Humaidi, M. N. (2022). The Effectiveness of Google Form as a Media for Assessment and Evaluation of Pai Learning at Smk Negeri 1 Berau East Kalimantan. *Research and Development Journal of Education*, *8*(2), 666. https://doi.org/10.30998/rdje.v8i2.13634

Sastrawijaya, A.. (1991). *Environmental Pollution* (1st ed.). Rineka Cipta.

Simarmata, N. N., Wardani, N.S., &, & Prasetyo, T. (2019). Development of Tolerance Attitude Assessment Instrument in Thematic Learning for Grade IV Students. *Basicedu Journal*, *3*(1), 194-199. https://doi.org/10.31004/basicedu.v3i1.101

Sudjana, N. (2020). *Basics of the Teaching and Learning Process* (16th ed.). SBAlgesindo.

Sugiyono. (2019). *Educational Research Methods* (3rd Edition). Alfabeta.

Tristaningrat, M. A.. (2018). Critical Analysis of the Correlation of Authentic Assessment in Learning. *Widyacarya*, *2*(1).

Utami, R. D. (2019). Computer-Based Test Exam Application at SMP Negeri 6 Cimahi. *Journal of Technology and Information*, *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). Development of Authentic Assesment Instruments To Assess *Observing, Inferring, And Predicting Skills In Science Learning.* Yogyakarta State University.