



Development of Multiple Choice Learning Evaluation based on Automatic Assessment of Hydrosphere Material

Ridha Sulistyو Budi^{a,1*}, Aldina Eka Andriani^{a,2}

^a Universitas Negeri Semarang, Indonesia

¹ ridhasulistyo3@students.unnes.ac.id*; ² aldinaekaandriani@students.unnes.ac.id

* corresponding author

Article Information

Article History:

Received May 2025.

Accepted June 2025

Published July 2025

Keywords:

Evaluation;

Multiple Choice;

Automatic Assessment;

Hydrosphere.

How to Cite:

Budi, S.R., & Andriani, A. E. (2025). Development of Multiple Choice Learning Evaluation based on Automatic Assessment of Hydrosphere Material. *Jurnal Dimensi Pendidikan dan Pembelajaran Universitas Muhammadiyah Ponorogo*, 13(2), pp 189-204.

Abstract

The preparation of learning evaluations that refer to various forms of questions, stimuli, and cognitive levels, as well as the utilization of digital technology, is still not optimally carried out by teachers, especially in hydrosphere learning. This causes the assessment to be monotonous, less challenging, and unable to describe students' thinking abilities comprehensively. This phenomenon becomes the urgency of the need to develop innovative learning evaluations. Therefore, this study aims to describe the design, test the feasibility, and effectiveness of developing a multiple-choice learning evaluation based on automatic assessment of hydrosphere material. This Research and Development (R&D) study follows the Borg and Gall model, with 27 students from class VA at Purwoyoso 04 State Elementary School as subjects. Data collection used non-test methods (observation, interviews, questionnaires, documentation) and tests (pre-test and post-test), analyzed using normality, t-test, and N-Gain tests. The developed evaluation includes various question types (standard and complex multiple choice), stimuli (text, images, video, audio, tables), and cognitive levels (MOTS and HOTS) using the Quizizz platform. Expert validation showed very feasible criteria: 92.18% (material) and 91.66% (media), with 100% positive responses from teachers and students. Effectiveness was proven through an average score increase of 33.3, a significance value < 0.001 , and an N-Gain score of 0.73 (high category). This study concludes that the multiple-choice learning evaluation tool based on automatic assessment was successfully developed through systematic research stages. The device was declared very feasible based on expert assessment and user response, and proved effective in improving student learning outcomes on hydrosphere material.

INTRODUCTION

Learning evaluation in elementary schools is often done conventionally and has not been utilized optimally. This can be seen in the learning practices in the VA class of Purwoyoso 04 State Elementary School on Hydrosphere material, where the teacher still uses ordinary multiple-choice questions without any stimulus innovations such as video, audio, or tables. This evaluation approach has not fully accommodated students' development of critical thinking skills. In fact, according to (Sutrisno et al., 2022) evaluation activities are often only carried out as a formality and have not been optimally utilized to identify students' needs in the learning process. Therefore, there is a need for transformation in learning evaluation design that is more interactive and adaptive with technological support. According to (Nahdiyah et al., 2024) Improving the quality of learning must align with developments in communication and information technology that are developing in society at large.

Learning approaches that focus too much on low-level cognitive abilities are one of the causes of low student reasoning skills. Many teachers still apply learning approaches that are dominant in Lower Order Thinking Skills (LOTS), such as remembering (C1), understanding (C2), and applying (C3), according to Bloom's taxonomy. As a result, learning activities do not encourage the development of students' critical, analytical, and creative thinking skills. According to (Ichsan et al., 2020) students' reasoning ability is closely related to the non-optimal application of Higher Order Thinking Skills (HOTS) based learning. HOTS includes the ability to analyze, evaluate, and create, which requires students to think logically, reflectively, and initially. In this context, HOTS-type questions are important because they not only test memory, but also challenge students to think critically and creatively in solving problems according to (Arum et al., 2023).

Teachers have not been optimal in preparing technology-based learning evaluations and automatic assessments during the evaluation process. Teachers conduct evaluations conventionally, namely dictating questions to students, while the techniques used in correcting are done manually, which takes longer to analyze each question and student. Digital-based learning evaluation offers various conveniences in its implementation, one of which is the ability to provide feedback or corrections to student answers objectively and quickly, according to (Mahardika et al., 2023). Digital-based evaluation comes as a more innovative, efficient, and adaptive alternative to the needs of the times. Educational technology now allows teachers to utilize various digital platforms in designing, managing, and evaluating learning automatically and in real-time. One of them is the use of Quizizz, which provides an automatic assessment system so that teachers no longer need to correct student answers one by one manually. According to (Zhao, 2019) Ranking is based on the level of question correction alone or the level of question correction and the time taken to complete it. Thus, Quizizz can facilitate the system automatically to check answers, provide scores, and present immediate feedback.

The learning outcomes of students on hydrosphere material are still low. Learning outcomes are a form of evaluation that aims to measure how students understand the material delivered in the learning process. According to (Afifah & Putranto, 2023), Evaluation results can be presented as scores as an indicator to assess students' success level after following the learning process. From some research findings related to existing problems regarding the preparation of the evaluation of learning Science learning VA class of Purwoyoso 04 State Elementary School. According to (Damasanti & Nuroh, 2023), to realize success in the education system, educators must continue to improve the quality of the learning process by applying various methods, strategies, and innovations that focus on the use of technology. The support of information

technology and multimedia media allows the creation of a creative and interactive learning environment to encourage active student involvement in learning activities. With the help of information technology and multimedia, an innovative learning environment can be implemented to create students who are active in learning (Kusnadi & Azzahra, 2024). Efforts that can be made are the need to develop interactive and fun learning evaluations by utilizing Automatic Assessment-based technology, one of which is through the Quizizz application.

LITERATURE REVIEW

The following are some of the literature reviews that exist in this study:

Quizizz

Definition of Quizizz

In recent years, especially in education, the development of technology-based evaluation has become the focus of attention, along with demands for more interactive, efficient, and relevant assessments to 21st-century learning needs. One platform that is starting to be widely utilized is Quizizz. The use of Quizizz in the context of learning evaluation is considered because it can present a learning atmosphere that is fun, motivating, and supports active student involvement. Quizizz according to (Suhartatik, 2022) It is an interactive quiz used in the classroom learning process, which can be applied as a daily evaluation tool, a midterm assessment, and an end-of-semester assessment.

According to (Jong & Tacoh, 2024). Quizizz is a game-based learning application that allows the simultaneous involvement of many students in the classroom, creating a more interactive and fun learning atmosphere. Using Quizizz, students can work on practice questions directly in class using their mobile devices. The characteristic distinguishing Quizizz from other applications is the avatar feature, interesting themes, and background music that excite the learning process. In this research, Quizizz is used as a platform to implement multiple-choice-based automatic assessment, due to its ability to provide instant feedback and customize student needs. With an automated assessment system, teachers no longer need to check answers manually, making the evaluation process more efficient and objective. In addition, Quizizz allows real-time analysis of results, so teachers can immediately identify students' weaknesses and strengths for further improvement.

Quizizz features

Quizizz provides a variety of interactive question types, such as multiple choice, essay, short form, and answers in the form of images. In addition, some premium features include question types such as match pairs, drag and drop, reorder, drop-down options, and video and audio responses. This application is also equipped with entertainment elements such as memes and background music that add excitement when used in learning. Meanwhile, according to (Kurniawan & Huda, 2020) Some features that can attract students' interest include adding background music when working on questions, a visually appealing display, and adding background music when working on questions. When working on questions, an attractive visual display with attractive font designs and colors, immediate feedback when answering right or wrong, and a leaderboard that shows students' scores and positions in real-time.

The advantages of quizzes in research

- 1) In question development, the multiple choice formats include ordinary and complex multiple choice. According to (Jannah et al., 2022) Ordinary multiple-choice questions are suitable for mastery of factual knowledge, while complex multiple-choice questions can test conceptual understanding and critical thinking skills.
- 2) These questions contain varied and interesting stimuli, such as text, images, audio, video, and tables. According to (Zulham, 2020) evaluation questions emphasizing cognitive aspects can be enriched with visual stimuli, increasing concept understanding and reflective thinking.
- 3) The assessment system used is equipped with automatic assessment reports, according to (Wahyuni, 2025) through automatic reports, teachers can analyze student learning outcomes more accurately and quickly, including speed of answering and error trends. This analysis includes evaluating concept understanding, speed of answering, and patterns of errors that appear.

Student Learning Outcomes

In carrying out the teaching and learning process, many things must be considered to achieve the learning and learning objectives. Likewise, with learning, there are changes in a person in the form of knowledge, attitudes, and skills. The following is a description of the theory regarding learning outcomes. According to (Sudjana, 2016) states that "student learning outcomes are essentially changes in behavior due to a broader sense covering cognitive, affective, and psychomotor fields. Learning outcomes are the abilities that students have after receiving their learning experience. Students are said to be successful in learning if they can develop abilities, knowledge, and attitudes. Bloom in (Arifudin, 2020) divides learning outcomes into three domains: 1) Cognitive domain, regarding intellectual learning outcomes, 2) Affective domain, regarding attitudes, and 3) Psychomotor domain, regarding skills and abilities to act.

The domain used in this study is the cognitive domain, which refers to students' intellectual abilities in understanding, processing, and using information. According to Bloom in (Gunawan & Palupi, 2016) The cognitive domain is the ability to absorb the meaning of the material or material learned. This domain includes various levels of thinking, ranging from remembering, understanding, applying, analyzing, evaluating, to creating. Student learning outcomes can be known through evaluation activities. This evaluation aims to obtain valid and objective data to show the extent to which students have achieved the learning objectives that have been formulated. Evaluation not only functions as a measuring tool, but also as a basis for reflecting on the learning process that has taken place. The data obtained from the evaluation becomes an important reference in designing follow-up learning, such as providing reinforcement, remediation, or enrichment.

Natural Science

Science learning has been found in every school, leading to student development. It can be seen as science, which has four dimensions: products, processes, attitudes, and technology. Science as a product is a series of knowledge; for example, we enter the classroom and then deliver science material, which is the nature of science as a product. This knowledge can be classified as factual, conceptual, procedural, and metacognitive (Muria & Budianti, 2021). Natural Science (IPA) learning in elementary schools should be done outside the classroom by utilizing the surrounding environment as a learning resource. This approach allows students to

interact directly with nature to explore real and contextual phenomena. Through outdoor activities, students gain conceptual knowledge and develop observation skills, critical thinking, and curiosity about the environment.

METHODS

This type of research is Research and Development using the Borg and Gall development model. (Sugiyono, 2021). The research and development step consists of 10 steps, namely: 1) potential and problems, 2) data collection, 3) product design, 4) design validation, 5) design revision, 6) product trial, 7) product revision, 8) usage trial, 9) product revision, 10) mass production. The R&D approach was chosen because the main objective of this research is to produce a product in the form of a learning evaluation tool that is valid, feasible, and effective for use in the teaching and learning process. This approach allows researchers to design and systematically test and revise products based on empirical data from the field.

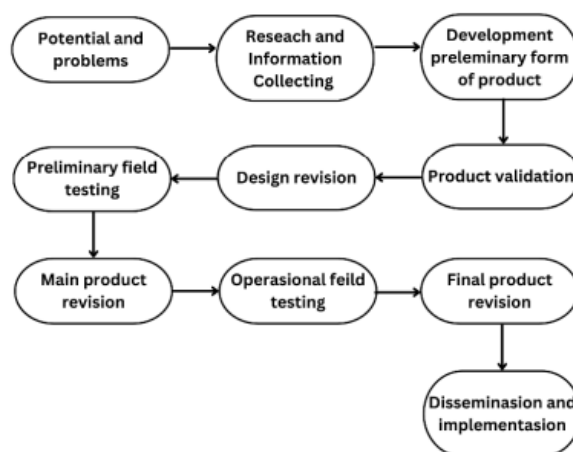


Figure 1. Borg and Gall's development model (Sugiyono, 2021)

The urgency of the content approach arises from conditions in the field, which show that learning evaluations in elementary schools are still dominated by questions that are less varied and have not touched on higher cognitive aspects, especially on hydrosphere material. Factors contributing to the findings in this study include teacher involvement in the trial process, students' enthusiasm for technology-based evaluation, and the quality of evaluation tools that have gone through the expert validation and revision process.

The subjects of this study were VA class students of Purwoyoso 04 State Elementary School in Semarang City, totaling 27 students, with 15 male students and 12 female students. Data collection techniques consisted of tests in ordinary multiple choice and complex multiple choice (pre-test and post-test), as well as non-test (observation, interview, questionnaire, and documentation). Data analysis techniques used normality test, t-test, and N-Gain test.

RESULTS AND DISCUSSION

This research focuses on developing the design, feasibility, and effectiveness of *multiple-choice* learning evaluation based on the *Automatic Assessment of* hydrosphere material. Multiple-choice questions were chosen because they can measure cognitive outcomes widely and

efficiently, and can be processed automatically through digital systems. According to Astuti, W., & Huda (2022), *Automatic assessment* is a technology-based assessment approach that can provide instant feedback, encourage reflection on student learning, and facilitate teachers in analyzing evaluation results.

Development of Multiple Choice Learning Evaluation Design based on Automatic Assessment

This research is a research and development using a model from Borg and Gall with 10 stages, namely:

This development research was conducted using the Borg and Gall model. The first stage is potential and problem analysis. At this stage, identifying problems in the field and the potential that can be developed (Sugiyono, 2021) This stage involves a series of observation processes of learning activities, interviews with class teachers, documentation, and data on the learning outcomes of VA class students at Purwoyoso 04 State Elementary School, Semarang City. Based on the potential and problem analysis stage, it shows that: teachers have not utilized technology in the process of evaluating Natural Science learning; Learning evaluation in class VA is still carried out conventionally, namely correcting in manua; the length of time to correct and score students; the types of questions given have not facilitated students to think critically, this is seen from the cognitive level that adjusts to the learning outcomes without the development of higherlevel thinking skills, which results in low student learning outcomes.

The potential owned by SD Negeri Purwoyoso 04 includes various learning support facilities that are pretty complete. Each class has a projector or LCD that allows teachers to present learning materials visually and interactively. Based on the potential and existing problems, researchers analyze student needs and develop products that aim to improve their learning outcomes. The product developed is the development of a multiple-choice learning evaluation based on Automatic Assessment of the hydrosphere.

The second stage carried out by researchers is research and information collection. Based on the results of observations and interviews with VA class teachers at Purwoyoso 04 State Elementary School, it is known that student learning outcomes on hydrosphere material are low. The results of the analysis of teacher and student needs questionnaires stated that an Automatic Assessment-based Multiple Choice learning evaluation was needed, considering the limitations of existing facilities and infrastructure. Students need an Automatic Assessment-based multiple-choice learning evaluation with two multiple-choice formats, namely ordinary multiple choice and complex multiple choice, accompanied by question stimuli in the form of text, images, video, audio, and tables, which are attractively packaged in the content of Natural Science lessons about the hydrosphere. In line with the opinion of (Maulidia & Pahlevi, 2020) the stimulus presented in this question item should be able to attract the attention of students and enable students to think critically, think creatively, and find the correct answer to the question item.

The third stage is the development of a preliminary form of the product. Researchers design products tailored to the questionnaire results on the needs of teachers and students. At this stage, the research produced a product prototype to be developed according to (Qoriah, 2021). Multiple Choice learning evaluation based on Automatic Assessment is made based on the results of the questionnaire analysis of the needs of teachers and students. Researchers' development of a preliminary form of product using Quizizz, which contains a display of multiple-choice learning evaluation questions based on Automatic Assessment on Quizizz, and a display of multiple-choice

learning evaluation values based on Automatic Assessment on Quizizz, which can be used to measure the level of understanding and learning outcomes of students.



Figure 2. Display of the regular multiple-choice question form

Based on Figure 2, according to (Masidjo, 2005) In ordinary multiple-choice questions, students are faced with statements or questions that contain problems and several alternative answers; students must choose one of the most correct or most appropriate answers.



Figure 3. Display of a complex multiple-choice question form

Based on Figure 3, according to (Santosa, 2023) suggests that complex multiple-choice questions are designed to measure students' higher-order thinking skills, especially in analyzing various possible answers to a problem. This complex multiple-choice is almost the same as the usual multiple-choice form, but the complex multiple-choice form has a combination of statements; from the combination of statements, there is only one correct answer.

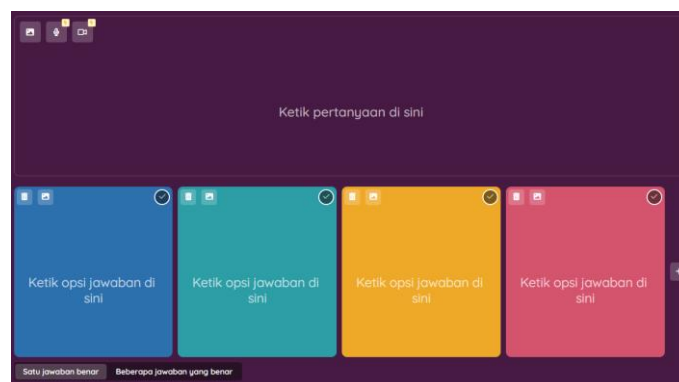


Figure 4. Stimulus questions on Quizizz

Based on Figure 4. Variations of question stimulus on Quizizz: The first icon is a feature that adds images as part of the question. The second icon refers to adding audio or sound to the question. This feature allows quiz makers to insert voice recordings, conversations, or specific audio effects. The third icon is a feature that adds a video as a stimulus to the question. According to (Zulham, 2020) evaluation questions emphasizing cognitive aspects can be enriched with visual stimuli, increasing concept understanding and reflective thinking.

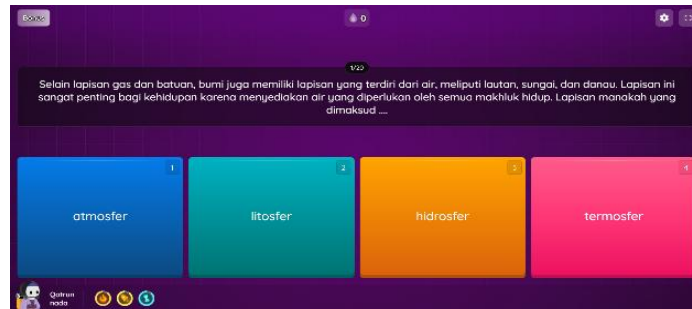


Figure 5. Stimulus display of text questions

Based on Figure 5. shows a question stimulus in the form of text on the Multiple Choice learning evaluation based on Automatic Assessment, according to (Susanti, 2021) text provides an initial context for developing literacy and logic skills and increasing understanding of concepts through narrative or description.



Figure 6. Stimulus display of questions in the form of images

Figure 6 shows a question stimulus in the form of an image on the Multiple Choice learning evaluation based on Automatic Assessment, according to the use of images, it is instrumental in clarifying the context of the question, especially to activate students' visual and imaginative power, according to (Nurfadhillah et al., 2021).



Figure 7. Display of the question stimulus in the form of a video

Based on Figure 7, the question stimulus in a video on the Multiple Choice learning evaluation based on Automatic Assessment can be uploaded directly or taken from sources such as YouTube. With videos, students can understand complex concepts more concretely.

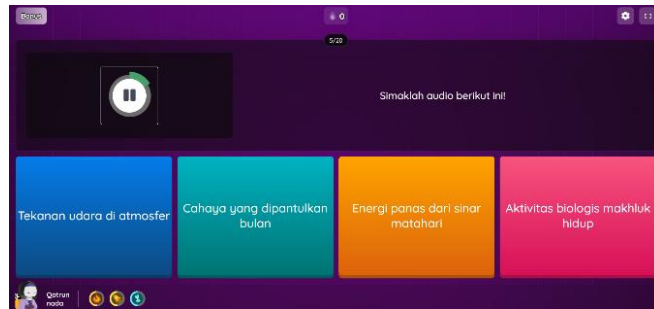


Figure 8. Display of audio question stimulus

Figure 8 shows the question stimulus in the form of audio on the Multiple Choice learning evaluation based on Automatic Assessment, according to (Fitri & Ardipal, 2021) Supports auditory learning and improves capture through hearing, students can develop listening skills more actively and purposefully.



Figure 9. Stimulus display of questions in the form of a table

Figure 9 shows the stimulus question as a table on the Multiple Choice learning evaluation based on Automatic Assessment. The table can present data, comparisons, or statistical information, which students then analyze.

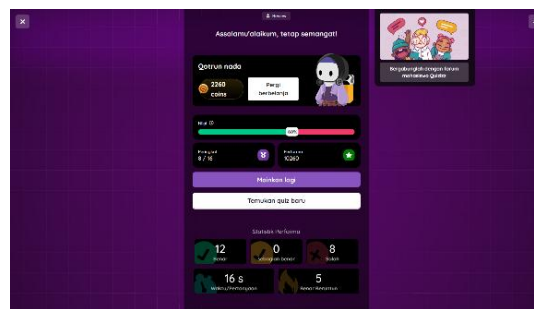


Figure 10. Value display

Figure 10 shows a recapitulation of student performance, including the number of correct answers, incorrect answers, fastest time, and medals earned.

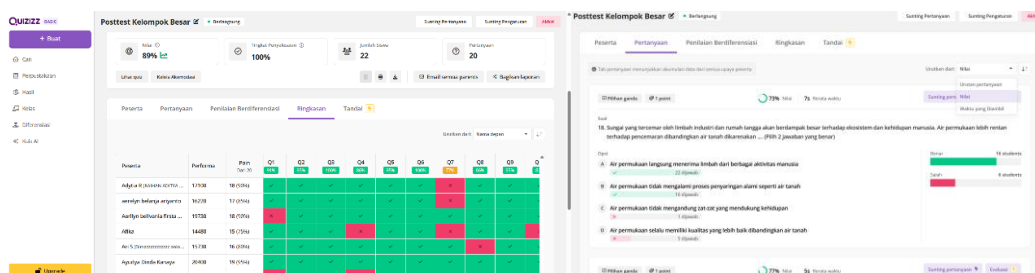


Figure 11. Automatic Assessment view

Based on Figure 11. shows the individual student performance in detail, including: percentage of average score, correct questions, wrong questions, and student performance. Thus, teachers can easily track which questions are the most challenging and which students need further assistance.

The fourth stage is product validation. The product to be tested must be validated (Wulandari, 2017). Design validation of multiple-choice learning evaluation based on Automatic Assessment was carried out by material experts and media experts. Researchers made a validation questionnaire for material experts and media experts. Researchers gave validation questionnaires to material and media experts. Then the results of the material and media expert questionnaires were analyzed to determine whether or not the product was feasible to be applied in learning activities. After going through design validation, we obtained expert validators' assessments, suggestions, and input. The results of material and media expert validation obtained very feasible criteria.

In the fifth stage is design revision, according to (Firmansyah, 2025) Input from material experts ensures that the content presented in the media is aligned with curriculum standards and learning objectives. Researchers must revise the design according to the material and media experts' assessment, suggestions, and input until the development product is suitable for testing.

The sixth stage is preliminary field testing. Product trials were conducted to determine the feasibility level of the product developed in terms of content, design, and operational functions according to (Nafi'ah et al., 2025). In this study, product testing activities were carried out on small groups of students to obtain suggestions for developing multiple-choice learning evaluations based on Automatic Assessment. The small group trial was conducted at SD Negeri Purwoyoso 04, Semarang City, with six students in class VA. After the trial was conducted, researchers distributed questionnaires of teacher and student responses to find out the responses of teachers and students to the practicality of Multiple Choice learning evaluation based on Automatic Assessment on hydrosphere material.

Table 1. Pre-test and Post-test Results of Product Trial

<i>Action</i>	<i>Average</i>	<i>Improved</i>
<i>Pre-test</i>	40	43,3
<i>Post-test</i>	83,3	

Based on Table 1, the average pre-test of product trials is 40, and the average post-test of product trials is 83.3, so there is an increase of 43.3. So there is an increase in student learning outcomes after using the Automatic Assessment-based Multiple Choice learning evaluation.

In the seventh stage, the main product revision is made. Based on the results of the questionnaire analysis of teacher and student responses in the small group trial, there was no product revision because it was appropriate and met the practical criteria. The eighth stage is operational field testing. The use trial was conducted in a large group, which was carried out in the VA class of Purwoyoso 04 State Elementary School, Semarang City, with 21 students.

Table 2. Pre-test and Post-test Results of the User Trial

<i>Action</i>	<i>Average</i>	<i>Improved</i>
<i>Pre-test</i>	54,5	33,3
<i>Post-test</i>	87,8	

Table 2 shows that the average pre-test of the usage trial was 54.5, and the average post-test of the usage trial was 87.8, so there was an increase of 33.3. So there is an increase in student learning outcomes after using the Automatic Assessment-based Multiple Choice learning evaluation.

The ninth stage is the final product revision stage. The data obtained from the usage test is used as a basis for making final stage revisions to the effectiveness of the product developed (Gumantan & Mahfud, 2020). The analysis of the results of the usage trial shows a very high level of effectiveness in helping students learn.

The tenth or final stage is the dissemination and implementation production stage. The purpose of this stage is to introduce the product so that it can be accepted by users, both individually and in groups (Santi et al., 2022). The Multiple Choice Learning Evaluation dissemination was limited to the elementary school where the research was conducted, namely SD Negeri Purwoyoso 04, Semarang City.

Feasibility of Multiple Choice Learning Evaluation based on Automatic Assessment

Table 3. Expert Validation Results

<i>Expert</i>	<i>Percentage</i>	<i>Criteria</i>
Material Expert	92,18%	Very Feasible
Media Expert	91,66%	Very Feasible

Table 3 shows that the results of the material expert validation scored 92.18% with very feasible criteria, and the results of the media expert validation scored 91.66% by obtaining very feasible criteria. Based on the results of the material expert validation questionnaire which contains 16 statements consisting of (1) aspects of presentation (suitability of material with learning outcomes; conciseness of concepts; suitability of material; practical and efficient evaluation instruments) (2) aspects of content quality (completeness of questions; accuracy; communicative), (3) aspects of construction (suitability of questions; provide motivation), (4) language aspects (accuracy of language use) which is a systematic and continuous determination of the quality (value and importance) of something based on specific considerations and criteria for decision making (Noly Shofiyah, 2018). Through these aspects, the Multiple Choice learning evaluation based on Automatic Assessment obtained a score of 59 out of a maximum score of 64, thus getting a percentage of 93.75% and obtaining very feasible criteria.

Based on the results of the media validation questionnaire which contains 15 statements consisting of (1) aspects of the quality of content and objectives (accuracy; importance; interest of students; suitability for the situation of students), (2) instructional aspects (clarity of instructions for using the media; motivational quality of students; impact on students; impact on teachers), (3) technical / appearance aspects (readability; ease of use of the media; effectiveness and efficiency of the media; display quality; quality of program management; quality of images presented) in line with the opinion of (Fitri Lutfia Zahroh et al., 2024) in making evaluation instruments not only involves technical aspects, but also considers pedagogical and psychological principles, (2024) in making evaluation instruments not only involves technical aspects, but also considers pedagogical and psychological principles. This is important to ensure the instrument can effectively support a better learning process. Through these aspects, the Multiple Choice learning evaluation based on Automatic Assessment obtained a score of 45 out of a maximum score of 60, thus getting a percentage of 92.18% and obtaining very feasible criteria.

So, the Multiple Choice learning evaluation based on Automatic Assessment is feasible for the hydrosphere material in the VA class of Purwoyoso 04 State Elementary School, Semarang City. In addition, the feasibility of multiple-choice learning evaluation based on Automatic Assessment is also supported by the responses of teachers and students. The response results from the VA class teacher amounted to 100% with very feasible criteria. In addition, researchers also received a response from VA class students in the usage trial of 100% with very feasible criteria, so the Multiple Choice learning evaluation based on Automatic Assessment is very feasible to use in learning.

This is supported by previous research discussing the development of Quiz-based evaluations in Islamic religious education, concluding that interactive multimedia is feasible and well applied in elementary schools. In line with previous research, research that discusses the feasibility of developing evaluations using Quizizz concluded that the evaluation was declared feasible for use in learning based on the results of validation by material experts, media experts, teacher responses, and student responses.

Effectiveness of Multiple Choice Learning Evaluation based on Automatic Assessment

The effectiveness of Multiple Choice learning evaluation based on Automatic Assessment is tested through normality test, t-test, and N-gain test using SPSS 30

Table 4. Normality Test Results

Action	Sig	Criteria
Pre-test	0.333	Normal
Post-test	0.095	Normal

Based on Table 4, which refers to the Shapiro-Wilk normality test, the test results on the trial show a pre-test significance value of $0.333 > 0.05$, so the data distribution meets the normality assumption. In the post-test, normality test results showed a significance value of $0.095 > 0.05$, so the data distribution met the normality assumption. After the pre-test and post-test results on the usage trial were declared, the data distribution met the assumptions of normality, and the researchers conducted a t-test (Paired sample t-test).

Table 5. Paired sample t-test result

Action	Mean	Improved	Sig. (2-tailed)
Pre-test	54,5	33,3	<0,001
Post-test	87,8		

Based on Table 5. shows that the average value of the pre-test of the usage trial was 54.5, and the average value of the post-test of the usage trial was 87.8, so there was an increase of 33.3. Based on the Paired sample t-test results, the significance value (2-tailed) is H_0 rejected and H_a is accepted. So there is an increase in student learning outcomes after using the Automatic Assessment-based Multiple Choice learning evaluation.

Table 6. N-Gain result

Action	Mean	Improved	N-gain	Criteria
Pre-test	54,5	33,3	0,7342	High
Post-test	87,8			

Based on Table 6, the results of the N-Gain test in the usage trial (large group) obtained a mean of 0.7342. The N-Gain test in the usage trial showed a result of 0.7342, including high

criteria. This shows that the large group pre-test and post-test results meet high criteria. The results of the moderate N-Gain test indicate that the multiple-choice learning evaluation based on Automatic Assessment is efficacious in improving the learning outcomes of the hydrosphere material. This shows that the multiple-choice learning evaluation based on Automatic Assessment can improve student learning outcomes.

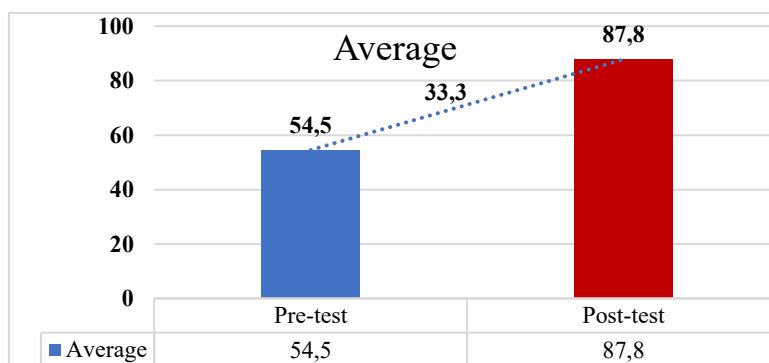


Figure 12. Average pre-test and post-test results

The results of the N-Gain test show high criteria, the data is said to be normally distributed if the data values are symmetrically distributed around the average value and form a bell curve (Tanzimah, 2024). The data distribution is symmetrical to the middle value or mean, so that most of the data values are close to the mean, and the amount of data will be less if it moves away from the mean to the right or left. This is influenced by several factors, including variations in the questions used between ordinary and complex multiple-choice. Complex multiple-choice questions require students to analyze several statements at once, which contributes to developing higher-order thinking skills. With this variation, students recall information and are challenged to understand, evaluate, and relate knowledge in a broader context. Furthermore, various question stimuli, such as text, images, audio, video, and tables, can broadly stimulate students' understanding and cognitive engagement. As well as student learning readiness, especially in the aspect of understanding and skills in using technology, is also a determining factor in the success of learning. According to (Irsan et al., 2024) Using quizzes that include images, video, and audio as a question stimulus affects student understanding and participation, supported by students' readiness to use technology.

Learners' learning outcomes have improved after using the Automatic Assessment-based Multiple Choice learning evaluation. Students' learning outcomes can increase, as this is due to the application of evaluations that have a good impact on improving students' understanding of the material and learning outcomes. Learning outcomes can be improved by applying a multiple-choice learning evaluation based on Automatic Assessment.

CONCLUSION

Based on the results of the analysis and discussion, it can be concluded that the description of the development of Multiple Choice learning evaluation based on Automatic Assessment of hydrosphere material using quizzes that pay attention to variations in the form of questions, variations in question stimulus, cognitive levels, and automatic assessment. This research has limitations on the scope of material, which is still limited to the topic of the hydrosphere, as well as the use of one type of digital evaluation application, namely Quizizz. Based on these findings, it is recommended that future research develop and examine various other digital evaluation

applications on different learning material topics. This aims to expand the application of similar evaluation models and assess their impact on various learning contexts and more diverse learner characteristics.

The main contribution of this research lies in the development of an automatic assessment-based multiple-choice learning evaluation that takes into account the variety of forms of multiple-choice questions, namely ordinary multiple-choice and complex multiple-choice, as well as the variety of question stimuli, including text, images, video, audio, and tables. This evaluation is also designed based on two cognitive levels, namely MOTS (Medium Order Thinking Skills) for level C3 and HOTS (Higher Order Thinking Skills) for levels C4 and C5, which allows a more thorough assessment of learning outcomes.

REFERENCES

- Afifah, R. F. N., & Putranto, A. (2023). Penerapan Metode Outdoor Study Terhadap Minat Dan Hasil Belajar Pada Mata Pelajaran IPS Kelas VIII Di MTsN 4 Tulungagung. *Jurnal Dimensi Pendidikan Dan Pembelajaran*, 11(2), 208–218. <https://doi.org/10.24269/dpp.v11i2.7144>
- Arifudin, O. (2020). *Psikologi Pendidikan (Tinjauan Teori Dan Praktis)*. Widina Bhakti Persada.
- Arum, F. P., Ngazizah, N., & Khaq, M. (2023). Analisis Kemampuan HOTS Ditinjau dari Gaya Belajar Peserta Didik Bermuatan Ips Kelas V Semester 1 di SD Muhammadiyah Purwodadi. *JIMPS: Jurnal Ilmiah Mahasiswa Pendidikan Sejarah*, 8(4), 5279–5287. <https://doi.org/https://doi.org/10.24815/jimps.v8i4.27243>
- Astuti, W., & Huda, M. (2022). Pengembangan Asesmen Otomatis Berbasis Digital dalam Pembelajaran Sekolah Dasar. No Title. *Jurnal Inovasi Teknologi Pendidikan*, 9(1), 45–54.
- Damasanti, L., & Nuroh, E. Z. (2023). Penggunaan Media Powerpoint Pada Pembelajaran Tematik di SD Muhammadiyah 1 Krian. *Jurnal Dimensi Pendidikan Dan Pembelajaran*, 11(1), 19–25. <https://doi.org/10.24269/dpp.v11i1.5902>
- Firmansyah, M. (2025). Pengembangan Buku Digital Berbasis Visual Mnemonik Untuk Meningkatkan Keterampilan Membaca Pemahaman Literal Anak Disleksia. *Jurnal Pendidikan Khusus, Universitas Negeri Surabaya*.
- Fitri, F., & Ardipal, A. (2021). Pengembangan Video Pembelajaran Menggunakan Aplikasi Kinemaster pada Pembelajaran Tematik di Sekolah Dasar. *Journal Basicedu*, 5(6), 6330–6338. <https://doi.org/10.31004/basicedu.v5i6.1387>
- Fitri Lutfia Zahroh, Muhammad Muflih, & Fitri Hilmiyati. (2024). Pembuatan Instrumen Evaluasi Pembelajaran; Analisis Literatur Review. *Jurnal Paris Langkis*, 5(1), 319–328. <https://doi.org/10.37304/paris.v5i1.17507>
- Gumantan, A., & Mahfud, I. (2020). Pengembangan Alat Tes Pengukuran Kelincahan Menggunakan Sensor Infrared. *Jendela Olahraga*, 5(2), 52–61. <https://doi.org/10.26877/jo.v5i2.6165>
- Gunawan, I., & Palupi, A. R. (2016). Taksonomi Bloom – Revisi Ranah Kognitif: Kerangka Landasan Untuk Pembelajaran, Pengajaran, Dan Penilaian. *Premiere Educandum : Jurnal Pendidikan Dasar Dan Pembelajaran*, 2(02). <https://doi.org/10.25273/pe.v2i02.50>
- Ichsan, I. Z., Sigit, D. V., Ristanto, R. H., Luthfi, I. A., Muharomah, D. R., Efendi, M., Panjaitan, R. G. P., Marhento, G., Widiyawati, Y., & Susilo, S. (2020). LOTS dan HOTS tentang tanaman obat: pembelajaran sains dan lingkungan saat new normal COVID-19. *Edubiotik : Jurnal Pendidikan, Biologi Dan Terapan*, 5(02), 91–102. <https://doi.org/10.33503/ebio.v5i02.888>

- Irsan, Syamsurijal, & Suarti. (2024). Meningkatkan Interaksi dan Evaluasi Pembelajaran dengan Quizizz: Kuis Interaktif untuk Guru Masa Depan. *Jurnal Abdidas*, 5(6), 845–860. <https://doi.org/10.31004/abdidas.v5i6.1058>
- Jannah, Karnalim, F., P., O., Murad, A., Aditya, D. F., Rima, B. A., & Nurhas, I. (2022). Pelatihan Desain Kuis Hots Interaktif Dengan Aplikasi Kahoot! Dan Quizizz Di Masa Pandemi: Studi Kasus Guru Sekolah Dasar Gugus Pangeran Antasari Kota Banjarbaru. *JCES (Journal of Character Education Society)*, 5(1), 243–251. <https://doi.org/10.31764/jces.v3i1.6026>
- Jong, A., & Tacoh, Y. T. B. (2024). Pemanfaatan Aplikasi Quizizz untuk Meningkatkan Motivasi Belajar Siswa. *Jurnal Dimensi Pendidikan Dan Pembelajaran*, 12(1), 131–147. <https://doi.org/10.24269/dpp.v12i1.7344>
- Kurniawan, C. D., & Huda, M. M. (2020). Pengaruh Penggunaan Quizizz Sebagai Latihan Soal Terhadap Hasil Belajar Siswa Kelas V Sd. *Jurnal Pena Karakter*, 3(1), 23–41.
- Kusnadi, E., & Azzahra, S. A. (2024). Penggunaan Media Pembelajaran Interaktif Berbasis Wordwall dalam Meningkatkan Motivasi Belajar Peserta Didik pada Mata Pelajaran PPKn di MA Al Ikhlah Padakembang Tasikmalaya. *Jurnal Dimensi Pendidikan Dan Pembelajaran*, 12(2).
- Mahardika, A. I., Saputra, N. A. B., Muda, A. A. A., Riduan, A., Luzuardi, N. S., & Nurmalinda, N. (2023). Pelatihan Pengembangan Evaluasi Pembelajaran Digital Menggunakan Quizizz bagi Guru di Kota Banjarmasin. *Jurnal Abdimas Prakasa Dakara*, 3(1), 1–9. <https://doi.org/10.37640/japd.v3i1.1540>
- Masidjo, I. (2005). *Penilaian Hasil Belajar Siswa di Sekolah*. Kanisius.
- Maulidia, F., & Pahlevi, T. (2020). Pengembangan Instrumen Penilaian Tes Soal Pilihan Ganda Berbasis HOTS Pada Mata Pelajaran Administrasi Umum Jurusan OTKP SMK Negeri 1 Lamongan. *Jurnal Pendidikan Administrasi Perkantoran (JPAP)*, 8(1), 136–145. <https://doi.org/10.26740/jpap.v8n1.p136-145>
- Muria, A. L., & Budianti, Y. (2021). Model pembelajaran core sebagai solusi untuk meningkatkan hasil belajar IPA siswa sekolah dasar. *Jurnal Pedagogik Pendidikan Dasar*, 8(1), 1–6.
- Nafi'ah, A., Sa'diyah, L., Hermawan, A., & Hadi, S. (2025). Pengembangan Media Pembelajaran Audio Visual Berbasis Aplikasi Canva Untuk Mengidentifikasi Isi, Sistematis, Dan Kebahasaan Teks Proposal Pada Siswa Kelas Xi Ma Syekh Subakir Nglegok. *J-Symbol: Jurnal Magister Pendidikan Bahasa Dan Sastra Indonesia*, 13(1), 86–97.
- Nahdiyah, D., Hadi, M. N., & Sholekhudin, A. (2024). Inovasi E-Modul Berbasis Moodle Pada Mata Pelajaran Al-Qur'an Hadis Dalam Meningkatkan Motivasi Belajar Siswa Kelas X MA Ma'arif Durensewu Pandaan. *Jurnal Dimensi Pendidikan Dan Pembelajaran*, 12(2), 169–184. <https://doi.org/10.24269/dpp.v12i2.9585>
- Noly Shofiyah, S. B. (2018). *Buku Ajar Mata Kuliah Assessment pembelajaran. In Paper Knowledge. Toward a Media History of Documents*.
- Nurfadhillah, S., Cahyani, A. P., Haya, A. F., Ananda, P. S., Widyastuti, T., & Tangerang, U. M. (2021). Penerapan Media Audio Visual Berbasis Video Pembelajaran Pada Siswa Kelas IV Di SDN Cengklong 3. *Jurnal Pendidikan Dan Dakwah*, 3(2), 396–418. <https://doi.org/10.36088/pandawa.v3i2.1272>
- Qorih, H. (2021). *engembangan media pembelajaran scrapbook pada materi permasalahan ketenagakerjaan di Indonesia SMA N 1 Labuhan Ratu (Skripsi, Universitas Muhammadiyah Metro)*. Universitas Muhammadiyah Metro Repository.
- Santi, L. N., Nuriman, N., & Mahmudi, K. (2022). Pengembangan Buku Ajar Berbasis

- Augmented Reality (Ar) Menggunakan Assemblr Apps Pada Tema 9 “Menjelajah Angkasa Luar” Kelas Vi Sekolah Dasar. *Sekolah Dasar: Kajian Teori Dan Praktik Pendidikan*, 31(2), 78.
- Santosa, H. (2023). Analisis penggunaan model soal HOTS dalam penilaian pembelajaran. *JUPIIS (Jurnal Pendidikan Ilmu-Ilmu Sosial)*.
- Sudjana, N. (2016). *Penilaian Hasil Proses Belajar Mengajar*. PT. Remaja Rosdakarya.
- Sugiyono. (2021). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Alfabeta.
- Suhartatik, T. (2022). *Best Practice Implikasi Media Quizizz Berbasis Android Terhadap Kualitas Pembelajaran Dalam Mencetak Siswa Berprestasi di Tingkat Nasional*. Ahlimedia Book.
- Susanti, A. I. (2021). *Media Pembelajaran Berbasis Teknologi Informasi dan Komunikasi*. Deepublish.
- Sutrisno, S., Yulia, N. M., & Fithriyah, D. N. (2022). Mengembangkan Kompetensi Guru Dalam Melaksanakan Evaluasi Pembelajaran Di Era Merdeka Belajar. *ZAHRA: Research and Thought Elementary School of Islam Journal*, 3(1), 52–60. <https://doi.org/10.37812/zahra.v3i1.409>
- Tanzimah. (2024). Pengaruh Penggunaan Media Flash Card terhadap Hasil Belajar Matematika Sekolah Dasar Kelas 1. *Indonesian Research Journal on Education*, 4(4), 3430–3435. <https://doi.org/https://doi.org/10.31004/irje.v4i4.1476>
- Wahyuni, N. (2025). Penggunaan Aplikasi Quizizz untuk Meningkatkan Antusiasme dan Hasil Belajar Siswa pada Mata Pelajaran IPAS di Kelas V MIN 10 Aceh Besar. *JURNAL Studi Tindakan Edukatif (JSTE)*, 1(1), 47–51.
- Wulandari, Y. (2017). Kelayakan Aspek Materi Dan Media Dalam Pengembangan Buku Ajar Sastra Lama. *Gramatika STKIP PGRI Sumatera Barat*, 3(2), 162–172. <https://doi.org/10.22202/jg.2017.v3i2.2049>
- Zhao, F. (2019). Using Quizizz to Integrate Fun Multiplayer Activity in the Accounting Classroom. *International Journal of Higher Education*, 8(1), 37–43.
- Zulham, M. (2020). Keefektifan Multimedia Interaktif Berbasis Mobile dengan Pendekatan Kontekstual untuk Meningkatkan Social Skill dan Kemampuan Berfikir Kritis Pada Materi Gerak dan Gaya. *Jurnal Penelitian Pembelajaran Fisika*, 11(2), 209–214. <https://doi.org/10.26877/jp2f.v11i2.6138>