JURNAL ILMIAH PENDIDIKAN PROFESI GURU
Volume 7 Nomor 2 2024, pp 371-381
E-ISSN: 2621-5705; P-ISSN: 2621-5713
DOI: https://doi.org/10.23887/jippg.v7i2.84083

Item Quality Analysis of Concept Understanding and Problem Solving in Environmental Change Materials Using ANATES

Intan Hermaya[1], Helendra[2*], Vauzia[3], Fitri Arsih[4]

[1,2,3,4] Biologi, Universitas Negeri Padang, Padang, Indonesia
*Corresponding author: helendras@yahoo.com

Abstrak

Guru kesulitan dalam mengukur kemampuan pemahaman dan problem solving *pada siswa sehingga kemampuan tersebut tidak terukur dengan baik. Tujuan dari penelitian ini adalah menganalisis butir kualitas pemahaman konsep dan* problem solving *dari materi perubahan lingkungan dengan menggunakan alat ANATES. Metode deskriptif kuantitatif digunakan dalam penelitian ini. Alat penelitian terdiri dari soal objektif dan* essay *yang diujikan kepada peserta didik kelas XI SMA setelah diverifikasi oleh para ahli. Metode pengumpulan data menggunakan kuesioner. Instrumen pengumpulan data menggunakan lembar kuesioner. teknik analisis data menggunakan analisis kuantitatif. Temuan analisis menunjukkan bahwa meskipun ada beberapa hal yang perlu dimodifikasi, sebagian besar butir soal memiliki validitas yang baik. Reliabilitas instrumen termasuk dalam kategori tinggi dan sangat tinggi. Indeks kesukaran soal, yang bervariasi dari mudah hingga sulit. Mayoritas soal masuk dalam kategori sangat baik untuk daya pembeda. Berdasarkan hasil penelitian ini, soal-soal secara umum berkualitas tinggi dan dapat digunakan untuk menilai pengetahuan konseptual dan kemampuan problem solving peserta didik terkait materi perubahan lingkungan.*

Kata Kunci: *Kualitas Butir Soal, Pemahaman Konsep,* Problem Solving, *Perubahan Lingkungan, Aplikasi* ANATES

Abstract

Teachers need help to measure students' understanding and problem-solving abilities, so these abilities are not measured well. This research analyses the quality of understanding concepts and problem-solving from environmental change material using the ANATES tool. Quantitative descriptive methods were used in this research. The research tool consists of objective questions and essays tested on class XI high school students after being verified by experts. The data collection method uses a questionnaire. The data collection instrument uses a questionnaire sheet. Data analysis techniques use quantitative analysis. The findings of the analysis show that although several things need to be modified, most of the items have good validity. The reliability of the instrument is in the high and very high categories. Question difficulty index, which varies from easy to complex. The majority of questions fall into the outstanding category for differentiating power. Based on the results of this research, the questions are generally of high quality. They can assess students' conceptual knowledge and problem-solving abilities related to environmental change material.

Keywords: Quality of Test Items, Understanding Concepts, Problem Solving, Environmental Changes, Application of Anates

1. PENDAHULUAN

Success in the teaching and learning includes several components, approaches, and various teaching methods. The aim of holding the learning process is for students' success in education, both in a subject and in education in general (Febliza & Okatariani, 2020; Rusdi et al., 2022; Supriyati & Wijono, 2021). Student factors are critical in addition to other factors. The learning process can occur due to a person's interaction with their environment (Daryanto et al., 2020; Hilliard et al., 2020). One sign that someone has learned is a change in behaviour in that individual that occurs consciously, is continuous and functional (Chairudin

History:
Received : April 17, 2024
Accepted : August 10, 2024
Published : August 25, 2024

Publisher: Undiksha Press

& Dewi, 2021; Ramda et al., 2022). Suitable learning activities expect students to be more critical in analyzing and be more sensitive to global issues and, most importantly, can improve students' understanding of concepts and problem-solving (Lidiawati et al., 2021; Septina et al., 2018; Sukmasari & Rosana, 2017). Therefore, innovative learning approaches can be used as an effort to train students' understanding and problem-solving abilities.

However, the current problem is that many students still have low understanding and problem-solving abilities. Previous research findings also reveal that many students have low comprehension abilities due to ineffective learning activities (Argarini, 2018; Khusniyah & Hakim, 2019). Other research also reveals that low problem-solving abilities in children are caused by students who are only used to listening to teacher explanations, thus not training students' problem-solving abilities (Azizah et al., 2018; Šafhalter et al., 2022; Syaripudin, 2019). Learners' cognitive and metacognitive difficulties can be resolved by receiving either the necessary external support or help finding solutions. The SMA Negeri 6 Kerinci observations found that students were less active in the learning process. This can be seen from the small number of students who ask and respond to questions asked by the teacher. In learning, students do not dare to ask questions if there is material they do not understand. Apart from that, based on the results of interviews with teachers, it was also found that teachers had difficulty measuring the ability to understand concepts and solve problems in environmental change material. This causes children's ability to understand concepts and solve problems less well measured.

Assessing students' conceptual knowledge and problem-solving abilities is an essential step in assessing the effectiveness of the learning process. Creating and analyzing well-formulated question items is an appropriate substitute for conducting assessments (Hutabarat, 2019; Lubis et al., 2022). Incorporating environmental change material into the curriculum is very important for teaching students about real environmental challenges that affect their daily lives, as well as being important from a scientific point of view (Alkhasanah et al., 2022; Kurniasari et al., 2021). To ensure that students understand these concepts and can use them to overcome problems related to environmental change, it is essential to create exciting learning plans (Anugrah et al., 2020; Putri et al., 2020). The topic of environmental change is complex and diverse, covering elements such as pollution, deforestation and climate change, among others.

Understanding these concepts and solving related challenges requires critical thinking and strong analytical skills (Ambaryani, 2017; Anugrah et al., 2020; Putri et al., 2020; Varisa & Fikri, 2022). Questions intended to evaluate students' understanding of ideas and problem-solving abilities should reflect the complexity of the topic and stimulate deep, critical thinking. The effectiveness of the questions in evaluating these two aspects is proven by analyzing the quality of the questions carried out by the Anates application. A thorough study is required to verify that each question is suitable for assessing students' conceptual knowledge and problem-solving abilities. Some indicators of a question item's quality include its validity, difficulty level, and ability to differentiate. The ANATES (Test Analysis) application is helpful in this situation. With the help of ANATES, instructors can improve and hone evaluation tools, provide comprehensive information about the quality of each question item, and evaluate question items statistically (Fietri et al., 2021; Zamzaili & Swita, 2021) .

Previous research states that the ANATES application provides various features that allow educators to analyze test item statistics, such as analysis of item difficulty, item discrimination, and test reliability (Lestari et al., 2024; Zamzaili & Swita, 2021). Teachers can group questions according to their difficulty level, from simple to very challenging, using ANATES. Additionally, the software can assess how healthy questions separate students with high and low ability levels. The findings of the ANATES analysis can be a basis for

improving the quality of questions to create a more reliable and valid evaluation tool (Fiska et al., 2021). Using the ANATES program to assess the quality of question items in environmental change materials is an essential first step in ensuring that student evaluations of educational efforts are accurate and objective. Teachers can increase their confidence that evaluation findings adequately reflect students' conceptual understanding and problem-solving abilities by ensuring high-quality questions. Based on this, this research aims to analyze the quality of conceptual understanding and problem-solving items in environmental change material using ANATES.

2. METODE

This research is a quantitative descriptive research. This indicates that the goal of the study is to use numerical data to accurately depict a phenomenon. To get objective results, the data is subsequently processed using statistical techniques. In descriptive research, the information is given in the form of words that fully and precisely describe the research object based on the available field data, without drawing any broad judgments (Zellatifanny & Mudjiyanto, 2018). The research was conducted at SMA Negeri 6 Kerinci. Students who have studied environmental change-related material are used as research subjects. Sampling is done using a straightforward random sampling technique since every person has the right to be included in the sample. In this case, one class, namely class XI Science 1 consisting of 35 students, was used as a sample. The method used to collect data is a questionnaire. The questionnaire method was used to collect data regarding students' Concept Understanding and Problem Solving abilities. Data collection instrument with questionnaire sheet.

A description question with five questions is used for problem solving, while a questionnaire with twenty objective questions is used as an instrument to grasp the subject. Validity, reliability, degree of difficulty, and question differentiation are the factors that were utilized to gauge the quality of the questions in this study. This study looks at the idea and problem solving components of environmental change materials. Techniques for descriptive statistical data analysis are performed using the Anates program version 4.0 for Windows. The following is how to operate the Anates 4.0 Windows software: 1) Start the Anates Windows application (4.0). 2) Clicking "Create New File" will open a dialog box. 3. After that, fill up the dialog box with the data according to the query parameters you wish to verify. 4) Next, add the data that needs to be assessed, such as the quantity of subjects, queries, and response options. Next, note each student's response in accordance with the assignment number. 6) To confirm that the file has been saved, click "save". If you choose "process all automatically," the information will be handled.

Analyzing comprehension of concepts and addressing issues with the use of the Anates Windows application version 4.0, specifically: Authenticity The test question is administered via empirical validation. Determining the instrument's validity in evaluating learning outcomes is the goal of validation. Using Anates 4.0, an empirical validation analysis was conducted. In data processing, question items are considered legitimate if there is a valid or very valid correlation between the scores on the question and the total scores. Anates version 4.0 was used to test the research instrument's reliability. The researcher's interpretation of the reliability value is as follows (Sundayana, 2016). 1) $0{,}00 \leq r < 0{,}20$: Very Low; 2) $0{,}20 \leq r < 0{,}40$: Low; 3) $0{,}40 \leq r < 0{,}60$: Moderate; 4) $0{,}60 \leq r < 0{,}80$: High; 5) $0{,}80 \leq r \leq 1{,}00$: Very High. The Anates version 4.0 data provided the questions' unique selling point. Academic interpretations of the discriminating power are classified showed in Table 1.

Table 1. Discriminating Power Criteria

Difference Power Index	Classification
0.00 – 0.20	Ugly
0.21 – 0.40	Enough
0.41 – 0.70	Good
0.71 – 1.00	Very Good

(Sudijono., 2018)

The Anates version 4.0 analysis of the problem yields the instrument's difficulty level. The following standards apply for interpreting the questions' difficulty indeks showed in Table 2. comprehension to select the trick response instead of the correct one (Rahayu & Djazari, 2016). The deceptive index values and the distribution of the findings from the examination of the caliber of misleading questions in the science literacy questions that were employed are shown in the Table 3.

Table 2. Criteria for Interpretation of Difficulty Level

Difficulty Index	Classification
0.00 – 0.30	Difficult
0.31 – 0.70	Enough
0.71 – 1.00	Easy

(Sudijono., 2018)

Table 3. Criteria for Interpretation of the Quality of Scammers

Trickster Index	Classification
0	Very Repair
1	Repair
2	Not Good
3	Bad
4	Very Bad

(Arbiatin & Mulabbiyah, 2020)

3. HASIL DAN PEMBAHASAN

Hasil

The Reliability The validity level of an instrument can be used to determine the test of that instrument. Measurement validity encompasses not only measurement accuracy but also the relationship between the target to be measured and the measuring instrument. Validity in the context of exam questions indicates how much the questions truly assess students' conceptual knowledge and problem-solving abilities. The validity value indicated on the instrument table shows this (Retnawati, 2016). An instrument is considered good and can be utilized if it meets one of the three criteria: adequate, high, or very high with a correlation value of $0.40 < rxy \leq 0.60$, $0.60 < rxy \leq 0.80$, or $0.80 < rxy \leq 1.00$.

A correlation value of 0.73 was found based on the findings of the concept comprehension ability test validated utilizing the ANATES 4.09 program. In the meantime, the correlation value was 0.86 in the validation of problem-solving tasks for conceptual understanding using the ANATES 4.09 program. Data analysis revealed that conceptual comprehension questions had a high degree of difficulty and problem solving abilities questions had an extremely high level of complexity. These findings suggest that the research

tool is prepared for usage. The examination of the concept understanding questions revealed that 16 questions met the good criteria and 4 questions met the bad criteria. Any question items deemed unfit for usage will be changed or removed. The message was the same (Putra et al., 2021). To ensure appropriate operation, questions that are no longer legitimate must be changed again, and questions that are no longer useful must be eliminated and replaced with new, better ones. This is done to ensure that the student assessment tool runs well and that the exams students take are actually used to gauge how well they grasp the material. In the meanwhile, the five-question problem-solving set forms part of the legitimate requirements.

In order for the data produced to be reliable, the Reality Test is utilized to see consistent results and be free of measurement mistakes (Darma, 2021). Reliability is the extent to which measurement results hold true when repeated or expanded upon on the same problem using the same instrument. A measuring gadget is deemed dependable if it consistently produces data after multiple tests (Amanda et al., 2019). An accurate measurement device will yield the proper size (Hervi et al., 2023). The reliability test results for concept comprehension ability objective questions were 0.84 and the reliability test results for problem solving ability descriptions were 0.92 with a very high category, according to calculations performed with the ANATES version 4.09 application (Sundayana, 2016). The dependability test's findings indicate a reliable test that is conducted on a regular basis (Loka Son, 2019). A question must be sufficiently challenging to be able to distinguish between pupils who have studied the subject and those who have not in order for it to have discriminating power (Fatimah & Alfath, 2019). Finding the differences in pupils' skills is the aim of the discriminating ability exam. A greater index of question items indicates that the questions are capable of measuring pupils' abilities (Nurhalimah et al., 2022). Distribution of Differences in Objective Questions of Concept Understanding showed in Table 4.

Table 4. Distribution of Differences in Objective Questions of Concept Understanding

Differentiation	Question Item Number	Sum	Percentage
0.00 – 0.20 (Ugly)	-	-	-
0,21 – 0.40 (Enough)	4, 12	2	10%
0.41 – 0.71 (Good)	1, 3, 5, 7, 11, 17, 20	7	35%
0.71 – 1.00 (Very Good)	2, 6, 8, 9, 10, 13, 14, 15, 16, 18, 19	11	55%

Two questions (10%) had sufficient discriminating power, seven questions (35%) had strong discriminating power, and eleven questions (55%) had very good discriminating power, according to the calculation of objective questions for grasping the topic in Table 4. Distribution of Difference in Problem Solving Descriptions showed in Table 5.

Table 5. Distribution of Difference in Problem Solving Descriptions

Differentiation	Question Item Number	Sum	Percentage
0.00 – 0.20 (Ugly)	-	-	-
0.21 – 0.40 (Enough)	-	-	-
0.41 – 0.71 (Good)	2, 4	2	40%
0.71 – 1.00 (Very Good)	1, 3, 5	3	60%

Table 5 indicates that 40% of the questions had high discriminating power and 60% had good distinguishing power based on the findings of the problem-solving calculations. The trial of grasping concepts and solving problems was examined, and the results met sufficient, good, and very good requirements. The findings show that the questions have

educational value. The discriminating power of the question indicates how well it can separate test takers with high and poor talent (Handayani & Iba, 2020).

Over half of the questions have a discriminating power, which is a sign of low quality. The questions with good discriminating power (having sufficient and good criteria) must be added to the question bank after the discriminating power analysis is finished, and the questions with low discriminating power must be changed before being utilized in the following learning session. Question items that are negatively distinctive must be removed and avoided in subsequent learning outcome assessments (Fitriani, 2021). Measuring the question's complexity is the goal of the question difficulty index. A question is deemed good when it strikes a balance between difficulty and validity and reliability standards (Nurhalimah et al., 2022). A query shouldn't be overly easy or difficult. Students won't be motivated to put in more effort to answer questions that are too easy. However, because they believe they are unable to answer the questions, pupils who believe the questions are too challenging will usually quit up fast and get disinterested in trying again (Solichin, 2017a). Distribution of Difficulty Levels of Objective Problems Understanding Concepts showed in Table 6. Distribution of Difficulty Levels for Problem Solving Descriptions showed in Table 7.

Table 6. Distribution of Difficulty Levels of Objective Problems Understanding Concepts

Difficulty Index	Question Item Number	Sum	Percentage
0.00 – 0.30 (Difficult)	-	-	-
0.31 – 0.70 (Enough)	1-20	20	100%
0.71 – 1.00 (Good)	-	-	-

Table 7. Distribution of Difficulty Levels for Problem Solving Descriptions

Difficulty Index	Question Item Number	Sum	Percentage
0.00 – 0.30 (Difficult)	1	1	20%
0.31 – 0.70 (Enough)	1, 2, 3, 5	4	80%
0.71 – 1.00 (Good)	-	-	-

The overall moderate difficulty level (100%) of objective questions requiring conceptual knowledge is based on the results of a difficulty level investigation using Anatès version 4 for Windows software. Conversely, there are four (80%) medium-complex questions and one (20%) challenging question in the descriptive category. Not how many students answer using the answer key, but rather how well the question is received by voters is what determines the quality of the question (Yoshita Cahyaningrum et al., 2023). The way pupils select their responses based on the options provided for every exam question is known as the cheater's functioning (Arbiatin & Mulabbiyah, 2020). The following are the outcomes of the concept understanding fraud: Quality Analysis of Deceivers About Understanding Concepts in Table 8. According to calculations performed with Anates version 4.0, objective questions pertaining to conceptual knowledge have an extremely high deception value of 100%. As a result, the concept comprehension question tool satisfies the requirements for questions that can be applied to learning assessments.

Table 8. Quality Analysis of Deceivers About Understanding Concepts

Trickster Index	Question Item Number	Sum	Percentage
0 (Very Good)	1-20	20	100%
1 (Good)	-	-	0%

Trickster Index	Question Item Number	Sum	Percentage
2 (Not Good)	-	-	0%
3 (Bad)	-	-	0%
4 (Very Bad)	-	-	0%

Pembahasan

The results of the data analysis show that most of the question items have good validity. The reliability of the instrument is in the high and very high categories. Question difficulty index, which varies from easy to complex. The questions were of high quality and could be used to assess student abilities. This is due to the following factors. First, the test items can be used by teachers to assess conceptual knowledge. Well-designed questions can measure how much students understand the basic concepts being taught (Fitriani, 2021; Umi Fatimah & Alfath, 2019). Questions presented asking students to explain the relationship between two concepts or apply theory in new situations can provide an overview of conceptual understanding. Items involving analysis or synthesis allow teachers to assess factual knowledge and students' ability to connect and apply these concepts in a broader context (Lestari et al., 2024; Yoshita Cahyaningrum et al., 2023). This is important for measuring deep conceptual understanding. Well-designed questions will be in line with the learning objectives that have been set. Thus, through test items, teachers can measure the extent to which students achieve these goals, including understanding specific concepts (Lestari et al., 2024; Umi Fatimah & Alfath, 2019).

Second, teachers can use the questions to assess students' problem-solving abilities. Question items can be designed to reflect real-world situations or problems that require applying the concepts studied (Nurhalimah et al., 2022; Yoshita Cahyaningrum et al., 2023). This type of problem allows students to show how they can solve the problem in a relevant and practical way. Problem-solving questions often require students to analyze information, identify critical variables, and understand the relationships between various elements. By assessing students' answers to these questions, teachers can evaluate their ability to analyze problems in depth (Arbiatin & Mulabbiyah, 2020; Fietri et al., 2021). Problem-solving questions often require students to develop and apply problem-solving strategies. This process involves selecting appropriate methods, planning steps, and overcoming obstacles that arise during the problem-solving process. Through assessment questions, teachers can assess how effectively students develop and implement these strategies (Iskandar & Rizal, 2018; Umi Fatimah & Alfath, 2019).

Third, teachers can use the questions to determine students' level of knowledge regarding learning, especially regarding environmental change material. Through assessing questions, students get feedback about how much they understand the material (Solichin, 2017b; Utomo, 2022). This can help students identify areas that need improvement and deepen their understanding of certain concepts. By using test items effectively, teachers can gain deep insight into students' levels of understanding and adapt their teaching strategies to meet students' learning needs better (Nurhalimah et al., 2022; Yoshita Cahyaningrum et al., 2023). Previous findings are essential in assessing students' understanding and knowledge (Masitoh & Aedi, 2020; Suci Mitra & Helendra, 2022). By designing questions that assess various problem-solving aspects, teachers can gain deeper insight into students' abilities to face and solve complex problems. The limitation of this research is that it only tests questions that teachers can use to assess students' understanding and problem-solving abilities, especially regarding environmental change material. This research implies that the analysis results can be used to develop other questions that can measure students' learning abilities.

4. SIMPULAN

The concept of understanding ability test questions using the ANATES 4.09 application has correlation values to research findings. The reliability score of the objective test of the ability to understand ideas and problem-solving ability is at a very high level. Two questions have sufficient discriminating power, seven have discriminating solid power, and eleven have excellent discriminating power based on objective questions of conceptual knowledge. Forty per cent of the questions have good discriminating power, and sixty per cent of the questions have high discriminating power, according to the calculation of the problem-solving description. The descriptive questions include four in the medium category and one in the difficult category. In contrast, the objective questions that assess content understanding have a medium difficulty level. Understanding this meaning objectively has a misleading value, namely a very high value. Conceptual understanding and problem-solvingproblem-solving skills can be used as learning assessment tools in these situations.

5. DAFTAR RUJUKAN

Alkhasanah, N., Wahyuni, S., & Fauziati, E. (2022). Teori Belajar Humanistik dalam Pembelajaran Matematika SD. *Jurnal Penelitian Pendidikan*, *14*(2), 81–89. https://doi.org/10.21137/jpp.2022.14.2.2.

Amanda, L., Yanuar, F., & Devianto, D. (2019). Uji Validitas dan Reliabilitas Tingkat Partisipasi Politik Masyarakat Kota Padang. *Jurnal Matematika UNAND*, *8*(1), 179. https://doi.org/10.25077/jmu.8.1.179-188.2019.

Ambaryani, G. S. (2017). Pengembangan media komik untuk efektifitas dan meningkatkan hasil belajar kognitif materi perubahan lingkungan fisik. *Jurnal Pendidikan Surya Edukasi (JPSE)*, *3*(1). https://doi.org/10.37729/jpse.v3i1.3853.

Anugrah, D., Murwitaningsih, S., Sofyan, D. A., & Susilo, S. (2020). Model pembelajaran kreatif treffinger terhadap kemampuan memecahkan masalah pada materi ekosistem dan perubahan lingkungan. *JPBIO (Jurnal Pendidikan Biologi)*. https://doi.org/10.31932/jpbio.v5i1.601.

Arbiatin, E., & Mulabbiyah, M. (2020). Analisis Kelayakan Butir Soal Tes Penilaian Akhir Semester Mata Pelajaran Matematika Kelas VI di SDN 19 Ampenan Tahun Pelajaran 2019/2020. *El Midad: Jurnal PGMI*, *12*(2), 146–171. https://doi.org/10.20414/elmidad.v12i2.2627.

Argarini, D. F. (2018). Analisis Pemecahan Masalah Berbasis Polya pada Materi Perkalian Vektor Ditinjau dari Gaya Belajar. *Matematika Dan Pembelajaran*, *6*(1), 91. https://doi.org/10.33477/mp.v6i1.448.

Azizah, M., Sulianto, J., & Cintang, N. (2018). Analysis of Critical Thinking Skills of Elementary School Students in Learning Mathematics Curriculum 2013. *Jurnal Penelitian Pendidikan*, *35*(1), 61–70. https://doi.org/10.15294/jpp.v35i1.13529.

Chairudin, M., & Dewi, R. M. (2021). Pengembangan Bahan Ajar Buku Saku Digital Berbasis Problem Based Learning pada Mata Pelajaran Ekonomi. *Edukatif : Jurnal Ilmu Pendidikan*, *3*(3). https://doi.org/10.31004/edukatif.v3i3.491.

Darma, B. (2021). *Statistika Penelitian Menggunakan SPSS (Uji Validitas, Uji Reliabilitas, Regresi Linier Sederhana, Regresi Linier Berganda, Uji t, Uji F, R2*. GUEPEDIA.

Daryanto, D., Purnomo, M. E., & Adib, H. S. (2020). Pengembangan Bahan Ajar PAI Materi Qs. Al-Fil Kelas IV SDN 17 Muara Sugihan Berbasis Multimedia. *Muaddib: Islamic Education Journal*, *3*(1), 1–9. https://doi.org/10.19109/muaddib.v3i1.6294.

Fatimah, L. U., & Alfath, K. (2019). Analisi Kesukaran Soal, Daya Pembeda dan Fungsi Distraktor. *Jurnal Komunikasi dan Pendidikan Islam*, *8*(2), 38–64.

Febliza, A., & Okatariani, O. (2020). The Development of Online Learning Media by Using

Moodle for General Chemistry Subject. *Journal of Educational Science and Technology (EST)*, *6*(1), 40. https://doi.org/10.26858/est.v6i1.12339.

Fietri, W. A., Zulyusri, & Violita. (2021). Analisis Butir Soal Biologi Kelas XI Madrasah Aliyah Sakinah Kerinci Menggunakan Program Komputer Anates 4.0 For Windows. *Jurnal Penelitian Bidang IPA dan Pendidikan IPA*, *7*(1), 28–35.

Fiska, J. M., Hidayati, Y., Qomaria, N., & Hadi, W. P. (2021). Analisis Butir Soal Ulangan Harian Ipa Menggunakan Software Anates Pada Pendekatan Teori Tes Klasik. *Natural Science Education Research*, *4*(1), 65–76. https://doi.org/10.21107/nser.v4i1.8133.

Fitriani, N. (2021). Analisis Tingkat Kesukaran, Daya Pembeda, Dan Efektivitas Pengecoh Soal Pelatihan Kewaspadaan Kegawatdaruratan Maternal Dan Neonatal. *Paedagoria : Jurnal Kajian, Penelitian dan Pengembangan Kependidikan*, *12*(2), 199. https://doi.org/10.31764/paedagoria.v12i2.4956.

Handayani, S. L., & Iba, K. (2020). Karakteristik Tes Keterampilan Proses Sains: Validitas, Reliabilitas, Tingkat Kesukaran dan Daya Pembeda Soal. *Publikasi Pendidikan*, *10*(2), 100. https://doi.org/10.26858/publikan.v10i2.13051.

Hervi, F., Syamsurizal, S., Arsih, F., Advinda, L., & Sari, M. (2023). Analysis of the Quality of Scientific Literacy Question Items on Biodiversity Material Using the ANATES Application. *Jurnal Penelitian Pendidikan IPA*, *9*(SpecialIssue), 1062–1069. https://doi.org/10.29303/jppipa.v9ispecialissue.6304.

Hilliard, J., Kear, K., Donelan, H., & Heaney, C. (2020). Students' experiences of anxiety in an assessed, online, collaborative project. *Computers and Education*, *143*(January 2020), 1–31. https://doi.org/10.1016/j.compedu.2019.103675.

Hutabarat, R. (2019). ESJ (Elementary School Journal) Volume 9 No. 2 Juni 2019. *Deskripsi Implementasi Penilaian Autentik Berbasis High Order Thinking Skills (Hots) Dalam Menjawab Tantangan Abad 21 Di Sekolah Dasar Kota Medan*, *9*(2), 159–169.

Iskandar, A., & Rizal, M. (2018). Analisis kualitas soal di perguruan tinggi berbasis aplikasi TAP. *Jurnal Penelitian dan Evaluasi Pendidikan*, *22*(1), 12–23. https://doi.org/10.21831/pep.v22i1.15609.

Khusniyah, N. L., & Hakim, L. (2019). Efektivitas Pembelajaran Berbasis Daring: Sebuah Bukti Pada Pembelajaran Bahasa Inggris. *Jurnal Tatsqif*, *17*(1), 19–33. https://doi.org/10.20414/jtq.v17i1.667.

Kurniasari, W., Murtono, M., & Setiawan, D. (2021). Meningkatkan Minat Belajar Siswa Menggunakan Model Blended Learning Berbasis Pada Google Classroom. *Jurnal Educatio FKIP UNMA*, *7*(1), 141–148. https://doi.org/10.31949/educatio.v7i1.891.

Lestari, I. D., Yumna K, R. Z., Dewi, S. A. L., M, E. S., Nurhalizah, P., & Rosalina, S. (2024). Analisis Kualitas Butir Soal HOTS pada Mata Pelajaran Biologi yang Diberikan pada Siswa Kelas 12 dan Dievaluasi Menggunakan Aplikasi Anates. *EduInovasi: Journal of Basic Educational Studies*, *4*(2). https://doi.org/10.47467/edu.v4i2.1935.

Lidiawati, K. R., Sinaga, N., & Rebecca, I. (2021). Peranan Self-efficacy dan Intelegensi terhadap Prestasi Belajar pada Mahasiswa. *Jurnal Psikologi Udayana*, *7*(2), 110. https://doi.org/10.24843/jpu.2020.v07.i02.p10.

Loka Son, A. (2019). Instrumentasi Kemampuan Pemecahan Masalah Matematis: Analisis Reliabilitas, Validitas, Tingkat Kesukaran dan Daya Beda Butir Soal. *Gema Wiralodra*, *10*(1), 41–52. https://doi.org/10.31943/gemawiralodra.v10i1.8.

Lubis, S. P. W., Suryadarma, I. G. P., Paidi, & Yanto, B. E. (2022). The Effectiveness of Problem-based learning with Local Wisdom oriented to Socio-Scientific Issues. *International Journal of Instruction*, *15*(2), 455–472. https://doi.org/10.29333/iji.2022.15225a.

Masitoh, L. F., & Aedi, W. G. (2020). Pengembangan Instrumen Asesmen Higher Order Thinking Skills (HOTS) Matematika di SMP Kelas VII. *Jurnal Cendekia : Jurnal Pendidikan Matematika*, *4*(2), 886–897. https://doi.org/10.31004/cendekia.v4i2.328.

Nurhalimah, S., Hidayati, Y., Rosidi, I., & Hadi, W. P. (2022). Hubungan Antara Validitas Item Dengan Daya Pembeda Dan Tingkat Kesukaran Soal Pilihan Ganda Pas. *Natural Science Education Research*, *4*(3), 249–257. https://doi.org/10.21107/nser.v4i3.8682.

Putra, E. N., Syamsurizal, Lufri, & Zulyusri. (2021). The Teacher-Made Multiple-Choice Tests Quality for Natural Science Subject. *Journal of Educational Research and Evaluation*, *5*(3), 422–428.

Putri, C. D., Pursitasari, I. D., & Rubini, B. (2020). Problem Based Learning Terintegrasi STEM Di Era Pandemi Covid-19 Untuk Meningkatkan Keterampilan Berpikir Kritis Siswa. *JIPI (Jurnal IPA & Pembelajaran IPA)*, *4*(2), 193–204. https://doi.org/10.24815/jipi.v4i2.17859.

Rahayu, R., & Djazari, M. (2016). Analisis Kualitas Soal Pra Ujian Nasional Mata Pelajaran Ekonomi Akuntansi. *Jurnal Pendidikan Akuntansi Indonesia*, *14*(1), 85–94. https://doi.org/10.21831/jpai.v14i1.11370.

Ramda, A. H., Delima, M., & Nendi, F. (2022). Dampak Pbl Bermuatan Tutor Sebaya Terhadap Peningkatan Kemampuan Pemecahan Masalah Siswa Smp. *Jurnal Pendidikan Dan Kebudayaan Missio*, *14*(2), 190–203. https://doi.org/10.36928/jpkm.v14i2.1283.

Retnawati, H. (2016). *Analisis Kuantitatif Instrumen Penelitian*. Parama Publishing.

Rusdi, A., Fitriani, A., Hamzah, A., Handayani, T., & Bujuri, D. A. (2022). Development of Diorama Media on Cultural Diversity Materials In my Country Islamic Elementary School. *Al-Mudarris: Journal Of Education*, *5*(2). https://doi.org/10.32478/al-mudarris.v5i2.695.

Šafhalter, A., Glodež, S., Šorgo, A., & Ploj Virtič, M. (2022). Development of spatial thinking abilities in engineering 3D modeling course aimed at lower secondary students. *International Journal of Technology and Design Education*, *32*(1), 167–184. https://doi.org/10.1007/S10798-020-09597-8/METRICS.

Septina, N., Farida, & Komarudin. (2018). Pengembangan Lembar Kerja Siswa Dengan Pendekatan Saintifik Berbasis Kemampuan Pemecahan Masalah. *Jurnal Tatsqif*, *16*(2), 160–171. https://doi.org/10.20414/jtq.v16i2.200.

Solichin, M. (2017a). Analisis Daya Beda Soal, Taraf Kesukaran, Validitas Butir Tes, Interpretasi Hasil Tes dan Validitas Ramalan dalam Evaluasi Pendidikan. *Dirāsāt: Jurnal Manajemen & Pendidikan Islam*, *2*, 192–213.

Solichin, M. (2017b). Analisis Daya Beda Soal Taraf Kesukaran,Butir Tes, Validitas Butir Tes, Interpretasi Hasil Tes Vallidita Ramalan dalam Evaluasi Pendidikan. *Jurnal Manajemen dan Pendidikan Islam*, *2*(2), 192–213. https://doi.org/10.26594/dirasat.v2i2.879.

Suci Mitra, P., & Helendra. (2022). Analisis Kualitas Butir Soal Ujian Akhir Semester Ganjil Tahun Pelajaran 2020 / 2021 Mata Pelajaran Biologi Kelas X SMA Negeri 1 Teluk Sebong Analysis of Quality The Question Final Exam Odd Semester 2020 / 2021 Biology Class X SMA Negeri 1 Teluk Sebong. *Biodidaktika: jurnal Biologi dan Pembelajarannya*, *17*(2). https://doi.org/10.30870/biodidaktika.v17i2.16493.

Sudijono. (2018). *Pengantar Evaluasi Pendidikan*. Rajawali Pers.

Sukmasari, V. P., & Rosana, D. (2017). Pengembangan penilaian proyek pembelajaran IPA berbasis discovery learning untuk mengukur keterampilan pemecahan masalah. *Jurnal Inovasi Pendidikan IPA*, *3*(1), 101–110. https://doi.org/10.21831/jipi.v3i1.10468.

Sundayana, R. (2016). *Statistika Penelitian Pendidikan*. Alfabeta.

Supriyati, & Wijono. (2021). Organizational Culture with Transformational Leadership Style on Production Supervision at PT Diamondfit Garment Indonesia. *JPAI (Journal of Psychology and Instruction)*, *5*(2), 43–49. https://doi.org/10.23887/jpai.v5i2.38858.

Syaripudin, T. (2019). Multiliteration and Higher Order Thinking Skills Implications to Education. *International Journal of Science and Applied Science: Conference Series*, *3*(1), 131. https://doi.org/10.20961/ijsascs.v3i1.32534.

Umi Fatimah, L., & Alfath, K. (2019). Analisis Kesukaran Soal, Daya Pembeda dan Fungsi Distraktor. *Jurnal Komunikasi dan Pendidikan Islam*, *8*(2), 37–64. https://doi.org/10.36668/jal.v8i2.115.

Utomo, B. (2022). Analisis validitas isi butir soal sebagai salah satu upaya peningkatan kualitas pembelajaran di madrasah berbasis nilai-nilai islam. *Jurnal Pendidikan Matematika (Kudus)*, *1*(2). https://doi.org/10.21043/jmtk.v1i2.4868.

Varisa, N., & Fikri, A. A. (2022). Pengembangan media pembelajaran biologi berbasis video blog (vlog) dengan topik perubahan lingkungan. *Penelitian dan Pengembangan Pendidikan*, *2*(1), 33–39. https://doi.org/10.22219/raden.v2i1.22056.

Yoshita Cahyaningrum, I., Fuady, A., & Islam Malang, U. (2023). Analisis Butir Soal Sumatif Akhir Semester Ganjil Mata Pelajaran Matematika Kelas VII dengan Berbantuan Aplikasi Software Anates. *Mathema Journal*, *5*(2), 67–81.

Zamzaili, & Swita, B. (2021). Evaluasi Formatif Berbantuan Komputer Untuk Meningkatkan Kemampuan Analisis Butir Soal Mahasiswa FKIP Universitas Bengkulu. *Jurnal Pendidikan Matematika Raflesia*, *06*(01), 60–69.

Zellatifanny, C. M., & Mudjiyanto, B. (2018). Tipe Penelitian Deskripsi dalam Ilmu Komunikasi. *Jurnal Diakom*, *1*(2), 83–90.