https://jurnal.unimed.ac.id/2012/index.php/geo/article/view/55504
https://doi.org/10.24114/jg.v16i2.55504 - (Vol 16, No 2 – 2024)
available at http://jurnal.unimed.ac.id/2012/index.php/geo
e-issn (2549-7057) & p-issn (2085-8167)

# Analysis of Multiple Choice HOTS Test Questions on the Final Semester Assessment

**Budi Rahmah Panjaitan*, Epon Ningrum, Bagja Waluya**
Department of Geography Education, Faculty of Social Sciences Education, Universitas Pendidikan Indonesia, Indonesia

**ARTICLE INFO**
Article History:
Received: January 01, 2024
Revision: June 02, 2024
Accepted: June 06, 2024

Keywords:
Item Analysis, Multiple Choice, HOTS

Corresponding Author
E-mail:
budirahmahpanjaitan@upi.edu

**ABSTRACT**
Assessment is essential in the education system to monitor educational development and make appropriate instructional decisions. This study aims to analyze the quality of multiple choice questions based on high-level thinking skills (HOTS) in the Geography subject using Anates Windows Version 4.0.9 software. This evaluative research involves analysis of final semester exam questions for class X Geography at SMA Islam Al Azhar 14 Semarang. Data consisting of questions, answer keys and exam results were collected from 32 students. The analysis was carried out descriptively and quantitatively to assess the level of difficulty, discriminating power, reliability, effectiveness of distractors, and validity of the questions. The analysis results show that the difficulty level of most of the questions is in the medium category (50%). The discriminating power is mostly good, although several questions are wrong or bad. Most of the distractors worked well, but some questions had ineffective distractors. Regarding validity, 17 questions were valid, while 11 questions were invalid. The reliability of the test is in the high category, with a reliability coefficient of 0.71. Anates Windows Version 4.0.9 software effectively analysed the quality of HOTS questions. These findings indicate that although most of the questions are of good quality, some require revision to increase their validity and effectiveness. This research emphasizes the importance of using technology in educational assessment to ensure the quality and fairness of exam questions. It contributes to developing assessment tools to improve students' higher-order thinking skills.

## INTRODUCTION

Assessment has an important role and an integral part of the education system. Assessment is vital to know the progress of educational development from time to time (Setiawan, 2013). Assessment is also an important part of learning, which helps educators make appropriate decisions to determine the next steps related to learning (Am, 2018). The quality of teachers' lessons, the learning experiences of students, and the assessment of curricula, school performance, and system quality are all influenced by educational assessment (Brown, 2022). Through assessment, teachers can determine to what extent learning objectives have been achieved (Santoso, 2007). Assessment can be considered an important part of education because the implementation is integrated with the learning process (Kusainun, 2020). The merits of the assessment instrument determine the proper implementation of the assessment. If the assessment instrument is bad, it will cause the assessment results to be invalid. Meanwhile, the good or bad of the assessment instrument is determined by the scope to be measured, the difficulty level, and the language used (Santyadiputra et al., 2019). Questions are instruments used in the assessment. Quality questions can provide accurate information so that it can be known by students who have mastered the material and those who have not mastered the material (Magdalena et al., 2021).

Assessment requires quality question instruments to ensure the quality of the questions given to students. Therefore, the questions must be analyzed first before being used (Muharromah & Humaisi, 2020). Quality test questions are also composed of quality questions. An analysis of the item's quality is carried out to determine the quality of a test item. Analysis of the quality of the questions is the step that must be passed in order to determine the quality of the questions, both as a whole and per item.

The purpose of item analysis is to obtain information related to the characteristics of each item, either by examining items or empirically (Gusmizain, 2022). Basically, item analysis can be done qualitatively and quantitatively. Qualitative analysis refers to aspects related to material, construction and language. Meanwhile, quantitative analysis includes validity, reliability, difficulty level, discriminating power, and the effectiveness of deception (Wulandari & Pramusinto, 2020). Furthermore, the quality of the questions can be analyzed from various aspects, including validity, reliability, difficulty level, and the distractor's effectiveness (Rahayu & Djazari, 2016).

Efforts to produce quality questions with item analysis are inseparable from the demands of knowledge and technology development in the 21st century. Teachers are expected to be able to produce quality questions (Himmah et al., 2021) that can measure learning progress. One of the most important aspects of the 21st century is the ability to think at a high level, which is known as high-order thinking skills (HOTS). HOTS is a thinking ability that not only remembers but can also develop ideas (Jaenudin et al., 2020; Saraswati & Agustika, 2020). Therefore, teachers are expected not only to be able to provide HOTS-based learning but also to be able to carry out HOTS-based assessments so that learning can be measured properly. The HOTS measuring instrument is important for educators (Arifin & Retnawati, 2015). This HOTS-based assessment can be started by presenting HOTS questions. HOTS questions are instruments that can measure higher-order thinking skills, not only in terms of remembering, restating, or referring without processing (reciting) (Rosyida et al., 2016). This makes it possible to measure higher-order thinking abilities, including critical thinking, creative thinking, and problem-solving (Larasati, 2017).

Of the many types of questions, multiple choice is the one that is widely used (Brown & Abdulnabi, 2017; Tangianu et al., 2018). This is inseparable from the breadth of material that can be covered and the relatively short assessment time (Warju et al., 2020). Multiple-choice questions provide several answer options (Arif, 2014). The form of multiple-choice questions consists of the problems asked and the answer choices (Balitbang-Depdiknas, 2007). Multiple-choice tests can be considered the most common form of assessment in education, and many studies have been conducted to determine best practices for using them to measure learning (Butler, 2018). Multiple choice questions can be used to measure higher-order thinking skills or HOTS if they are prepared correctly (Anderson & Krathwohl, 2021; Javaeed, 2018).

Item analysis can be done by utilizing technology in the form of software. One such software is Anates Windows Version 4.0.9, which Drs Karno developed for M.Pd., a lecturer in Psychology at the Indonesian University of Education. Amidst the problem of teachers' low mastery of media and technology (Wijayanto et al., 2018), it is hoped that even though there is no special training for the use of technology, teachers can develop themselves as autodidacts (Suciani et al., 2021).

Much research has been conducted on item analysis, but most of it still focuses on separate qualitative or quantitative analysis. Technology is also not widely used in test analysis. In fact, this can help simplify test analysis, which is expected to maintain the quality of the questions to be tested. More specifically, there is still little research regarding analyzing geography questions, especially questions oriented toward HOTS (High Order Thinking Skills). Several research studies on the analysis of HOTS Geography questions were found to still use

manual methods using Microsoft Excel (Jati et al., 2023), and other studies used qualitative (theoretical) methods to analyze the distribution of cognitive, construction and language domains (Rachman & Nofrion, 2023).

This research uses Anates Windows Version 4.0.9 software to analyze multiple choice HOTS questions as a form of application of technology to help analyze the quality of the questions. This is expected to make things easier and ensure that the tested questions can measure the abilities you want to measure as well as possible. What is also emphasized in this research is that it specifically analyzes the quality of HOTS question items, not the suitability of the questions according to the HOTS question provisions. This means that the questions analyzed in this study are categorized as HOTS, but their quality needs to be measured from various components of the quality of the questions.

Anates can analyze the items easily and accurately. In addition, Anates can calculate the original score as well as the weighted score, determine the upper and lower groups, calculate discriminating power, calculate test reliability, calculate difficulty level, calculate the correlation of item scores with the total score, and determine the quality of the distractor (Purwati et al., 2021).

This research will analyze multiple choice questions based on high-level thinking skills on Geography test questions using Anates Windows Version 4.0.9 software. This research aims to analyze HOTS-based final semester exam questions in class. It is hoped that this research can contribute to the field of educational evaluation and assist teachers in compiling high-quality questions to improve students' high-level thinking skills, especially in Geography subjects.

## RESEARCH METHODS

This research is included in the type of evaluative research to analyze multiple-choice items from the aspects of difficulty level, discriminating power, reliability, effectiveness of the detractor, and validity. The data collection technique used is a documentation study, which involves collecting data from existing documents. This study obtained data from a Geography teacher in class X SMA Islam Al Azhar 14 Semarang. The data obtained is in the form of questions, answer keys to questions, and results of the Final Semester Assessment of Geography for class X SMA Islam Al Azhar 14 Semarang. This study involved the entire population of 32 tenth-grade students from Al Azhar 14 Islamic High School in Semarang who had completed the Final Semester Assessment for the 2022/2023 academic year. The data analysis technique used in this research was descriptive quantitative analysis, utilizing Windows Anates 4.0.9 software to determine the difficulty level, discrimination power, reliability, distractor effectiveness, and validity of the test items.

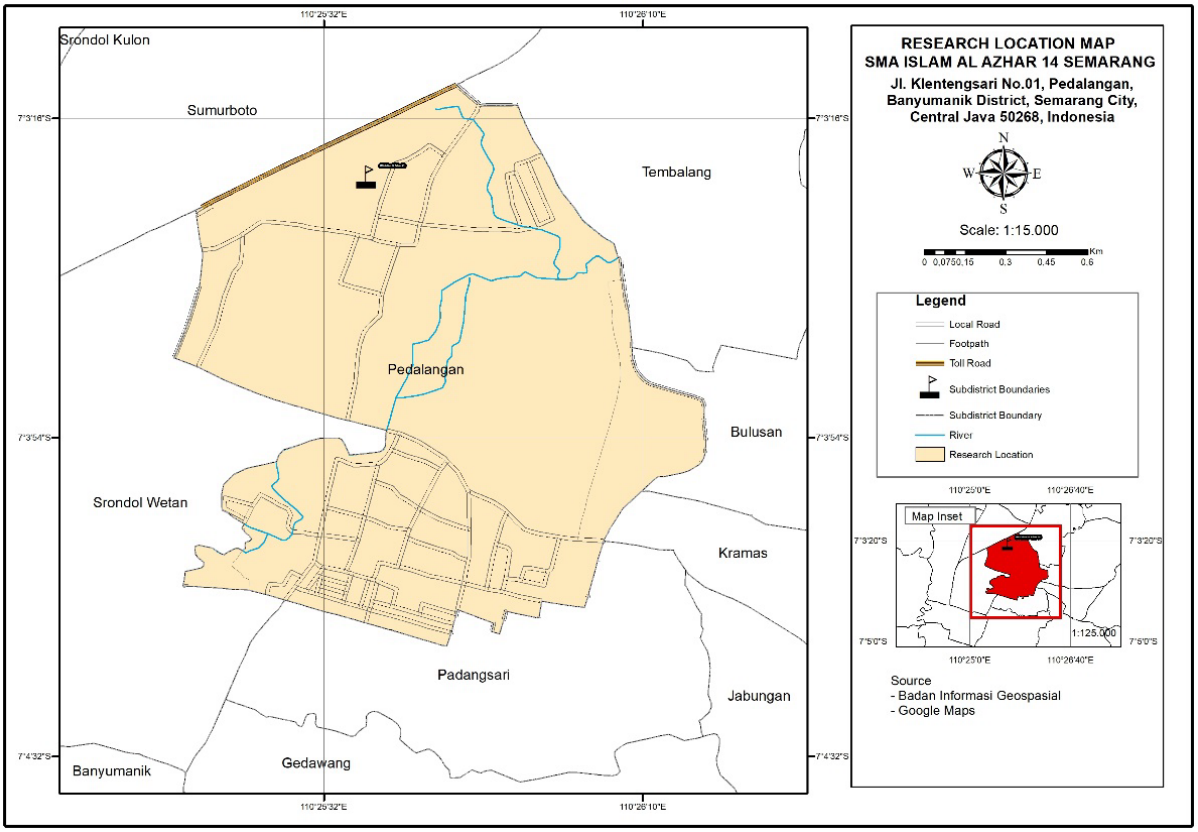


Figure 1. Map of research location (Source: Data Analysis, 2023)

## RESULTS AND DISCUSSION

### Difficulty Level

The difficulty level of the items is calculated by comparing the number of students who answered certain questions correctly with the total number of students who took the exam. The smaller the number of students who answered a question correctly, the more difficult the question. The number that indicates the difficulty level of the question is called the difficulty index. Furthermore, good questions are questions with a proportional or balanced difficulty level. The meaning of balanced here refers to questions with the respective proportions of 25% difficult, 50% medium, and 25% easy questions. The distribution of question difficulty levels can be seen in Table 1.

Table 1. Distribution of Question Difficulty Levels

| No | Difficulty Level | Question Item Number | Amount | Percentage (%) |
|---|---|---|---|---|
| 1 | Difficult | 4,10,13,14,26, | 5 | 17,9 |
| 2 | Medium | 2,5,6,7,9,12,15,16,17,19,23,24,25,28 | 14 | 50 |
| 3 | Easy | 1,3,20,21,22,27 | 6 | 21,4 |
| 4 | Very Easy | 8,11,18 | 3 | 10,7 |

Source: Research Results, 2023.

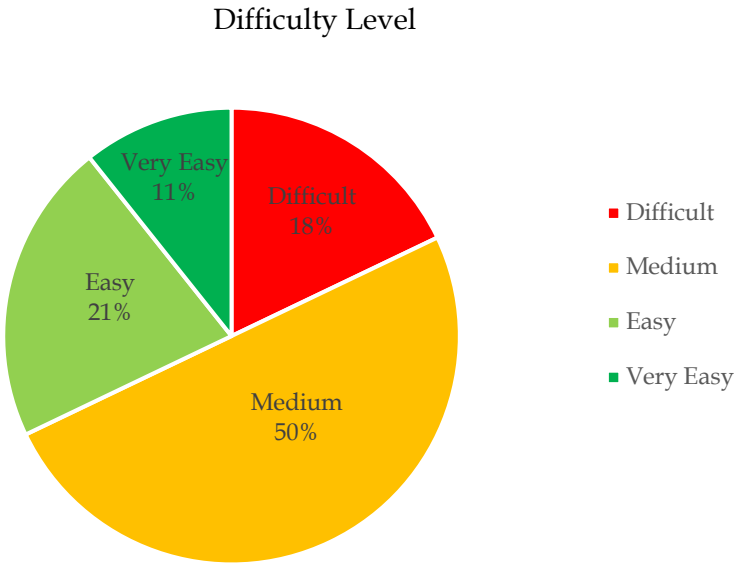


Figure 2. Percentage of question difficulty levels (Source: Research Results, 2023)

The analysis results show that the final semester assessment questions for Geography class X SMA Islam Al Azhar 14 Semarang are dominated by questions with moderate difficulty. While the smallest proportion is in questions with a very easy level of difficulty. Referring to the criteria for good questions with a proportional difficulty level, the final semester assessment questions for Geography for class X SMA Islam Al Azhar 14 Semarang are close to being proportional. However, there is still a slight gap between the percentage of easy and difficult questions.

**Discriminating Power**

Discriminating power in test items is intended to determine the ability of an item to distinguish students' ability to master the material. Karno To (1996), the calculation of discriminatory power can be interpreted in several categories, namely negative - 9% is in the very bad category, 10% -19% is in the bad category, 20% -29% is in the fairly good category, 30% -49% is in the good category, and 50% above is categorized as very good. The distribution of question discriminating power can be seen in Table 2.

Table 2. Distribution of Question Discriminating Power

| No | Discriminating Power | Question Item Number | Amount | Percentage (%) |
|---|---|---|---|---|
| 1 | Very Good | 1,6,7,9,12,15,16 | 7 | 25 |
| 2 | Good | 2,3,5, 13, 19,20,21, 22,24,25, 26,27 | 12 | 42,9 |
| 3 | Fair | 8,11 | 2 | 7,1 |
| 4 | Poor | 4,17,18,23 | 4 | 14,3 |
| 5 | Very Poor | 10,14,28 | 3 | 10,7 |

Source: Research Results, 2023

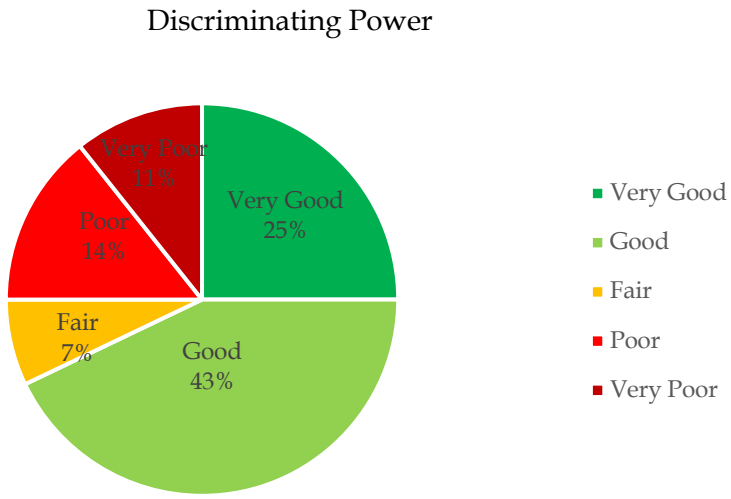


Figure 3. Percentage of question discriminating power (Source: Research Results, 2023)

The analysis results showed that most of the final semester assessment questions for Geography class X SMA Islam Al Azhar 14 Semarang in terms of discrimination were in a good category. However, there are still problems with bad and very bad discrimination, so this needs to be a concern. Follow-up, such as revising the questions or replacing the questions with differentiators that are categorized as good or very good.

**Effectiveness of The Distractor**

The distractors in the multiple-choice questions are said to work well if they are chosen evenly by students who answer incorrectly. The distractor itself is an answer that can trap students because it slightly differs from the actual answer. Each multiple-choice item has three distractor answers. To determine the effectiveness of the distractor on each item, this study adopted the Likert Scale. The effectiveness of the distractor is categorized as very good if it has three distractor answers for the item, the effectiveness of the distractor is categorized as good if it has two answers for the item, the effectiveness of the distractor is categorized as good enough if it has one answer for the item, and the effectiveness of the distractor is categorized as not functioning if there is no distractor's answer to the item. The distribution of the effectiveness of distracting questions can be seen in Table 3.

Table 3. Distribution of the Effectiveness of Distracting Questions

| No | Effectiveness of The Distractor | Question Item Number | Amount | Percentage (%) |
|---|---|---|---|---|
| 1 | Very Good | 17 | 1 | 3,6 |
| 2 | Good | 2,5,7,9,13,14,15,16, ,21,22,25,27 | 12 | 42,8 |
| 3 | Fair | 3,4,6,10,12,23,24,38 | 8 | 28,6 |
| 4 | Not Working | 1,8,11,18,19,20,26 | 7 | 25 |

Source: Research Results, 2023

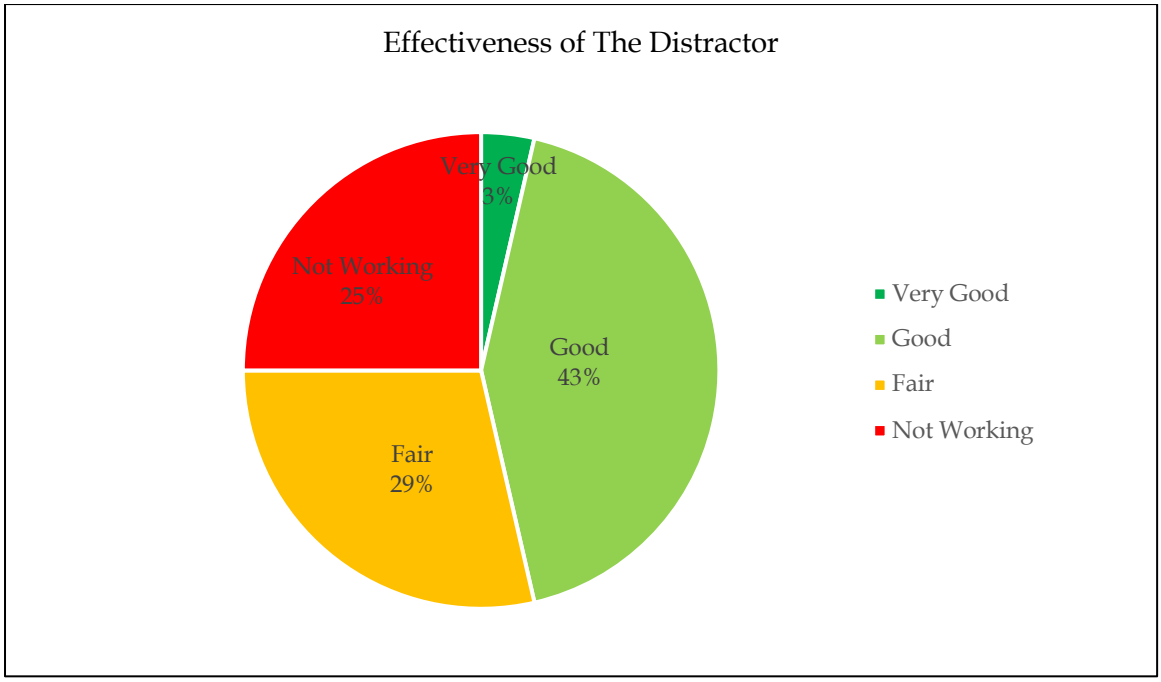


Figure 4. Percentage of distractor effectiveness (Source: Research Results, 2023)

The analysis results show that most of the questions have a good category of distractor effectiveness. However, some items still have the effectiveness of the distractor not working. Thus, items with a distractor effectiveness that does not work can be revised or replaced with other questions with a very good distractor effectiveness category.

**Validity**

Validity is intended to measure the item's contents so that it can be seen whether it is feasible or not to use it (Fink, 2010; Kimberlin & Winterstein, 2008; Richard & Haynes, 2002). The distribution of the question validity can be seen in Table 4

Table 4. Distribution of Question Validity

| No | Validity | Question Item Number | Amount | Percentage (%) |
|---|---|---|---|---|
| 1 | Very Significant | 1,2,3,7,9,12,13,26 | 8 | 28,6 |
| 2 | Significant | 5,6,11,15,16,20,21,24,25 | 9 | 32,1 |
| 3 | Not Significant | 4,8,10,14,17,18,19,22,23,27,28 | 11 | 39,3 |

Source: Research Results, 2023

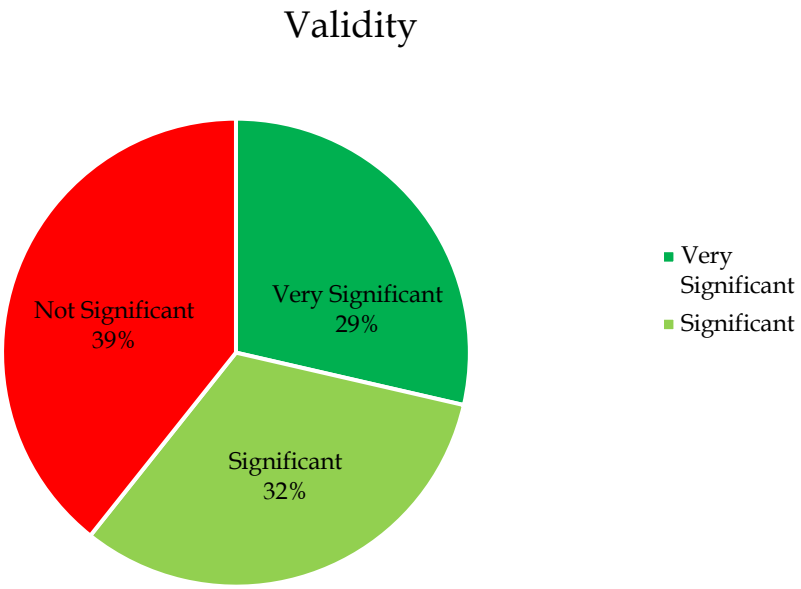


Figure 5. Percentage of validity (Source: Research Results, 2023)

The analysis results showed that there were still invalid items for the end-of-semester assessment of Geography for class X SMA Islam Al Azhar 14 Semarang. This is indicated by item numbers 4, 8, 10, 14, 17, 18, 19, 22, 23, 27, and 28. If we look closely, this occurs because the correlation obtained is smaller than the r table. Based on this, invalid questions can be revised or replaced with valid questions.

**Reliability**

Reliability is intended to determine the consistency of the measuring instrument, in this case in the form of multiple choice questions (Arafah et al., 2021; Sürücü & Maslakçi, 2020). The reliability value in this study refers to (Jihad & Haris, 2008), which categorizes reliability as very high if the r value is 0.90 - 1.00, high if the r value is 0.70 - 0.89, sufficient if the r value is 0.40 - 0.69, low if the value of r is 0.20-0.39, and very low if the value of r is 0.00 - 0.19. The analysis results show that the end-of-semester assessment questions for Geography class X SMA Islam Al Azhar 14 Semarang have a reliability value of 0.71, which is in the high category. This also means that the questions for the final semester assessment of Geography for class X SMA Islam Al Azhar 14 Semarang are in good criteria.

**Recap Item Analysis Questions**

In addition to partial analysis, an overall recap analysis was also carried out. This is done to find out more thoroughly about the quality of the questions that have been analyzed. The overall analysis results showed an average of 16.03, a standard deviation of 4.04, an XY correlation of 0.55, and a reliability of 0.71 of the 28 multiple-choice items analyzed using the Anates 4.0.9 software. 17 items are categorized as valid so that they can be issued again in the learning outcomes test in the future. Meanwhile, 11 items that are categorized as invalid can be discarded or not used anymore. Related to this, these invalid questions can be further investigated to find out the things that cause them to be invalid so that there is a possibility that they can be used in the future if they have been corrected to become valid questions.

**CONCLUSION**

Based on the analysis of the end-of-semester assessment items for Geography class X SMA Islam Al Azhar 14 Semarang

using Anates 4.0.9 software, it can be seen that the level of difficulty of the questions is in the medium category and as a whole, is close to proportional in terms of the distribution of difficulty levels with difficult questions of 14%, moderate 50% and easy and very easy 32%, discriminating power is categorized as good, effectiveness of the distractor is categorized as good, validity shows 17 items are valid, and reliability is categorized as high. The results of the overall item recap show that 17 items are valid, and it is possible to be issued again on the learning outcomes test later. Meanwhile, 11 items that are categorized as invalid can be immediately discarded or further investigated to be repaired so that they can be used again. The use of Anates Windows Version 4.0.9 software has proven effective in analyzing the quality of questions, including level of difficulty, discriminating power, distractor effectiveness, validity and reliability. The use of this technology simplifies the analysis process and helps teachers identify and correct questions that lack quality. Overall, this research shows that analysis of the quality of HOTS questions based on multiple choice in the final Geography semester exam for class X using Anates software is very helpful in ensuring that the questions used can measure students' high-level thinking abilities well. However, several questions require improvement to improve the quality and validity of the assessment. It is hoped that this research can contribute to the field of educational evaluation and assist teachers in compiling high-quality questions to improve students' high-level thinking abilities, especially in Geography subjects.

**REFERENCE LIST**

Am, Z. (2018). Teknik penilaian hasil pembelajaran. Rausyan Fikr: Jurnal Pemikiran Dan Pencerahan, 14(02). https://doi.org/10.31000/rf.v14i02.901

Anderson, L. W., & Krathwohl, D. R. (2021). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. Longman.

Arafah, K., Amin, B. D., Sari, S. S., & Hakim, A. (2021). The development of higher order thinking skills (HOTS) instrument assessment in physics study. Journal of Physics: Conference Series, 1899(1), 012140. https://doi.org/10.1088/1742-6596/1899/1/012140

Arif, M. (2014). Penerapan aplikasi anates bentuk soal pilihan ganda. Jurnal Ilmiah Edutic: Pendidikan Dan Informatika, 1(1). https://doi.org/10.21107/edutic.v1i1.398

Arifin, Z., & Retnawati, H. (2015). Analisis Instrumen Pengukur Higher Order Thinking Skills (HOTS) Matematika Siswa SMA. Seminar Nasional Matematika Dan Pendidikan Matematika Uny, 20, 783–790.

Balitbang-Depdiknas, P. P. P. (2007). Panduan Penulisan Soal Pilihan Ganda. Depdiknas. Jakarta.

Brown, G. T. L. (2022). The past, present and future of educational assessment: A transdisciplinary perspective. Frontiers in Education, 7. https://doi.org/10.3389/feduc.2022.1060633

Brown, G. T. L., & Abdulnabi, H. H. A. (2017). Evaluating the quality of higher education instructor-constructed multiple-choice tests: Impact on student grades. Frontiers in Education, 2, 24. https://doi.org/10.3389/feduc.2017.00024

Butler, A. C. (2018). Multiple-choice testing in education: Are the best practices for assessment also good for learning? Journal of Applied Research in Memory and Cognition, 7(3), 323–331. https://doi.org/10.1016/j.jarmac.2018.07.002

Fink, A. (2010). Survey Research Methods. In International Encyclopedia of

Education (pp. 152–160). Elsevier. https://doi.org/10.1016/B978-0-08-044894-7.00296-7

Gusmizain, A. (2022). Karakteristik Butir Soal Tes Mata Kuliah Matriks & Ruang Vektor Mahasiswa Matematika. Jurnal Evaluasi Pendidikan, 13(2), 127–130. https://doi.org/10.21009/jep.v13i2.28328

Himmah, E. F., Yuliarti, Y., Elmayantie, C., & Rusdiana, L. (2021). Peningkatan Profesionalisme Guru Smk Negeri-1 Palangka Raya Melalui Bimbingan Penelitian Tindakan Kelas. RESWARA: Jurnal Pengabdian Kepada Masyarakat, 2(2), 155–161. https://doi.org/10.46576/rjpkm.v2i2.883

Jaenudin, R., Chotimah, U., Farida, F., & Syarifuddin, S. (2020). Student development zone: higher order thinking skills (HOTS) in critical thinking orientation. International Journal of Multicultural and Multireligious Understanding, 7(9), 11–19. https://doi.org/10.18415/ijmmu.v7i9.1884

Jati, P., Alawiati, A., Jibril, A., Mudin, A., & Prisma, C. (2023). ANALISIS KUALITAS BUTIR SOAL GEOGRAFI. Jurnal Pendidikan Geosfer, VIII. https://doi.org/10.24815/jpg.v%vi%i.32500

Javaeed, A. (2018). Assessment of higher ordered thinking in medical education: multiple choice questions and modified essay questions. MedEdPublish, 7. https://doi.org/10.15694/mep.2018.0000128.1

Jihad, A., & Haris, A. (2008). Evaluasi Pembelajaran. PT Raja Grafindo Pustaka.

Karno To. (1996). Mengenal Analisis Tes. Fakultas Ilmu Pendidikan IKIP Bandung.

Kimberlin, C. L., & Winterstein, A. G. (2008). Validity and reliability of measurement instruments used in research. American Journal of Health-System Pharmacy, 65(23), 2276–2284. https://doi.org/10.2146/ajhp070364

Kusainun, N. (2020). Analisis Standar Penilaian Pendidikan di Indonesia. JP (Jurnal Pendidikan): Teori Dan Praktik, 5(1). https://doi.org/10.26740/jp.v5n1.p%25p

Larasati, D. A. (2017). Pengaruh Model Problem-Based Learning Terhadap Kemampuan Pemecahan Masalah Geografi Sma. Jurnal Geografi. https://doi.org/10.24114/jg.v9i1.6045

Magdalena, I., Fauziah, S. N., Faziah, S. N., & Nupus, F. S. (2021). Analisis Validitas, Reliabilitas, Tingkat Kesulitan dan Daya Beda Butir Soal Ujian Akhir Semester Tema 7 Kelas III SDN Karet 1 Sepatan. ://ejournal.stitpn.ac.id/index.php/bintang/article/view/1291

Muharromah, M. M., & Humaisi, S. (2020). Analisis Butir Soal Penilaian Ujian Semester Gasal Mata Pelajaran IPS di MTs Darul Muna Ponorogo. ASANKA: Journal of Social Science And Education, 1(2), 102–114. https://doi.org/10.21154/asanka.v1i2.2250

Purwati, L. M., Arianty, R., Syakilah, D. M., Ridlo, S., & Susilaningsih, E. (2021). Analisis Soal Tes Pilihan Ganda Berbasis Higher Order Thinking Skill Menggunakan Aplikasi Anates Windows Versi 4.0. 9 For Windows. Jurnal Pendidikan UNIGA, 15(2), 460–473. https://doi.org/10.52434/jp.v15i2.1287

Rachman, F. J., & Nofrion. (2023). Analisis Kualitas Soal UAS Geografi Kelas XI di SMA Negeri 12 Padang. Journal Of Social Science Research, 3(5). https://doi.org/10.31004/innovative.v3i5.5530

Rahayu, R., & Djazari, M. (2016). Analisis kualitas soal pra ujian nasional mata pelajaran ekonomi akuntansi. Jurnal Pendidikan Akuntansi Indonesia, 14(1). https://doi.org/10.21831/jpai.v14i1.11370

Richard, D. C. S., & Haynes, S. N. (2002). Behavioral Assessment. In Encyclopedia of Psychotherapy (pp. 165–183). Elsevier. https://doi.org/10.1016/B0-12-343010-0/00022-2

Rosyida, F., Zubaidah, S., & Mahanal, S. (2016). Memberdayakan Keterampilan Berpikir Kritis dengan Model Pembelajaran Remap TmPS (Reading Concept Map Timed Pair Share). Biologi, Sains, Lingkungan, Dan Pembelajarannya Dalam Upaya Peningkatan Daya Saing Bangsa, 13(1), 209–214.

Santoso, B. (2007). Penilaian Portofolio Dalam Matematika. Jurnal Pendidikan Matematika, 1(2), 31–38.

Santyadiputra, G. S., Pradnyana, I. M. A., & Juniantari, M. (2019). Efektivitas Pengelolaan Bank Soal Berbasis Examview Bagi Guru-Guru Di SMK Negeri 1 Nusa Penida. WIDYA LAKSANA, 8(1), 51–58. https://doi.org/10.23887/jwl.v8i1.15720

Saraswati, P. M. S., & Agustika, G. N. S. (2020). Kemampuan berpikir tingkat tinggi dalam menyelesaikan soal HOTS mata pelajaran matematika. Jurnal Ilmiah Sekolah Dasar, 4(2), 257–269.

Setiawan, F. (2013). Kemampuan Guru Melakukan Penilaian Dalam Pembelajaran Melalui Internalisasi Nilai Kejujuran Pada Pembelajaran Pendidikan Kewarganegaraan. Jupiis: Jurnal Pendidikan Ilmu-Ilmu Sosial, 5(2), 73–81.

Suciani, A., Irafadillah Effendi, D., & Zulida, E. (2021). Geography Learning Media Based On Character Education In The Digital Revolution Era. Jurnal Geografi. https://doi.org/10.24114/jg.v13i2.20305

Sürücü, L., & Maslakçi, A. (2020). Validity and reliability in quantitative research. Business & Management Studies: An International Journal, 8(3), 2694–2726. https://doi.org/10.15295/bmij.v8i3.1540

Tangianu, F., Mazzone, A., Berti, F., Pinna, G., Bortolotti, I., Colombo, F., Nozzoli, C., La Regina, M., Greco, A., & Filannino, C. (2018). Are multiple-choice questions a good tool for the assessment of clinical competence in Internal Medicine? Italian Journal of Medicine, 12(2), 88–96.

Warju, W., Ariyanto, S. R., Soeryanto, S., & Trisna, R. A. (2020). Analisis kualitas butir soal tipe HOTS pada kompetensi sistem REM di sekolah menengah kejuruan. Jurnal Pendidikan Teknologi Dan Kejuruan, 17(1), 95–104. https://doi.org/10.23887/jptk-undiksha.v17i1.22914

Wijayanto, B., Wilis, R., Novio, R., Hamka, J., Tawar Padang, A., & Barat, S. (2018). Analisis Technological Pedagogical and Content Knowledge (TPACK) Guru Geografi Di Kabupaten Solok, Sumatera Barat. Jurnal Geografi. 10(2), 105–116. https://doi.org/10.24114/jg.v10i2.9070

Wulandari, A. R., & Pramusinto, H. (2020). Analisis Kualitas Butir Soal Ujian Akhir Semester Ganjil Mata Pelajaran Otomatisasi Tata Kelola Sarana dan Prasarana Kelas XI Otomatisasi dan Tata Kelola Perkantoran. Economic Education Analysis Journal, 9(2), 366–378.