

## ITEM ANALYSIS OF MULTIPLE-CHOICE QUESTIONS: AN ASSESSMENT OF YOUNG LEARNERS

Maya Marsevani

*English Language Education, Universitas Internasional Batam, Batam, Indonesia*

Email: [maya@uib.ac.id](mailto:maya@uib.ac.id)

APA Citation: Marsevani, M. (2022). Item analysis of multiple-choice questions: An assessment of young learners. *English Review: Journal of English Education*, 10(2), 401-408. <https://doi.org/10.25134/erjee.v10i2.6241>

Received: 26-02-2022

Accepted: 23-04-2022

Published: 30-06-2022

**Abstract:** Although English has been taught for decades in Indonesia, the quality of multiple-choice question given to young learners was hardly examined. To touch on the issue, this study aims at determining the quality of 10 multiple-choice questions (MCQs) in a public elementary school in terms of difficulty level (P), discrimination power (D), and distractor efficiency (DE). This study employed a cross-sectional study to obtain information and evaluate MCQs in the students' tests. It was found the multiple-choice items were efficient in which 80% of them have acceptable difficulty and 20% of them were fairly easy ( $P > 70\%$ ). The discrimination index of 8 (80%) items was excellent and 2 of the MCQs were good. Among 40 distractors, 32 (80%) of them were functional distractors while 8 of them were non-functional distractors (NFDs). It can be concluded that most of the difficulty index was acceptable, the items discrimination power was good. In terms of distractor efficiency, two items were completely inefficient. Furthermore, MCQs that are considered poor should be revised, re-tested, and reviewed before incorporating them into the future test. This study hopefully encourages teachers to identify poorly developed MCQs and to improve their quality.

**Keywords:** *item analysis; English subject; multiple-choice questions.*

### INTRODUCTION

It has been widely acknowledged that assessments are necessary to demonstrate students' level of proficiency and to accurately measure their education progress. Assessment questions aim at helping teachers to succeed in teaching, ensuring their goals in classrooms have been achieved. There are several types of assessment questions, namely true/false, matching, multiple-choice, short answers, and essays.

Among those types, Multiple-Choice Questions (MCQs) are a ubiquitous assessment test type to assess students' knowledge and progress (Bhat & Prasad, 2021; Jayanti, Husna, & Hidayat, 2019; Shin, Guo, & Gierl, 2019) MCQs have potential benefits for teachers and students.

MCQs may provide students with direct feedback about their learning progress. Furthermore, they significantly increase the chances to answer the question correctly by eliminating throw-away options. In addition, teachers can grade their answers easier and quicker without rater bias. MCQs also allow them to include a large amount of material on a single exam. Across the discipline, teachers can use recycled questions that

progressively develop question banks for re-use in different combinations and settings. Several studies also found similar benefits (Brown & Abdulnabi, 2017; Polat, 2020; Patil, Dhole, & Mudiraj, 2016). The studies argue that MCQs are useful to objectively assess students in different educational streams relatively quicker. It is also useful because MCQs can be used to cover a large amount of material.

MCQ is commonly used in other subjects, especially at the language instruction level. Although its application has several benefits, MCQs can be problematic if they are not well designed and developed. To develop the questions properly, teachers need to have the required skills and expertise. However, the act of guessing the answer by students' may decrease the MCQs' validity and reliability (Freatat & Smadi, 2014). Nowadays, students can recognize the correct option without fully understanding the material (Polat, 2020). In this case, teachers are required to develop proper MCQs. The questions need to entail significant, complex, and subjective judgment, which requires time to develop. If MCQs are carefully designed, they may assess higher-order cognition

According to Namdeo and Sahoo (2016), there are three parts of MCQ, which is the stem (question text), key response (the correct answer), and distractors (incorrect options/answers). Regardless of whether they are the most utilized assessment method in classrooms, a clear MCQ format is essential to guide teachers to develop proper MCQ. It was found that teachers barely require much time and effort to develop and score the test (Jannah, Hidayat, Husna, & Khasbani, 2021). In this case, Danuwijaya (2018) states item analysis is a conscious and unconscious process to assess the quality of each item regularly. It is useful to identify difficult and easy options, to check how it discriminates low and high scores, to alternate the function, and to build a good question bank.

Item difficulty is the percentage of students who answer correctly (Karadag, 2016). Item discrimination differentiates between students who understand the test and those who do not. On the other hand, distractor analysis is similar to item difficulty and item discrimination. It is an extension of item analysis to misdirect the test takers from choosing the correct answers. Based on the description, there are three indicators of item analysis, namely item difficulty (P), item discrimination (D), and distractor analysis.

Several studies have examined the quality of MCQs. Test analysis to identify MCQs has been conducted in several countries such as India, Turkey, and Indonesia. These studies analyzed and evaluated discrimination index, difficulty level, and distractor of item analysis of MCQs in various areas (Danuwijaya, 2018; Gajjar, Sharma, Kumar, & Rana, 2014; Shete, Kausar, Lakhkar, & Khan, 2015).

Most of studies focused on analyzing MCQs of medical students. Kowash, Hussein, & Halabi (2019) analyzed item analysis in two postgraduate Pediatric Dentistry (PD) examinations. The result shows that 81% of the questions consisted of information recall which means these questions had a low level of difficulties. Another low difficulty index also found by Rehman, Aslam, & Hassan (2018). They found that the quality of MCQs given to dental undergraduate students in Islamabad was too poor. Most of the MCQs had to revise because of more than 50% of the MCQs had poor discrimination level, high difficulties of the items and less functional distractors.

Another research found by Purwoko and Mundijo (2018) was in line with aforementioned studies above. They carried out the research focused on the medical students in Indonesia. They investigated the quality of MCQs in Medical Faculty of Muhammadiyah University Palembang. The results showed that more than 50% of the questions should be revised because it had low discrimination level, poor difficulty index, and less discrimination efficiency.

While Obon and Rey (2019) investigated item and test quality of pharmacology students using difficulty index and discrimination indices, with distractor efficiency. They found that most of the questions should be revised or discarded. It is because the quality of the items was too poor. However, almost 70% of the distractors were retained.

This situation is supported by Kusumawati and Hadi (2018). They analyzed mathematics assessment in one of senior high school in Yogyakarta. They found that the level of discrimination was low (60%). however, less than 40% of the MCQs had a difficulty level. In line with Obon and Rey (2019), all distractors were well-functioned.

However, positive results showed in a few studies (Gajjar et al., 2014; Harti, Mahapatra, Gupta, & Nesari, 2021; Menon & Kannambra, 2017; Salih, Jibo, Ishaq, Khan, Mohammed, Al-Shahrani, & Abbas, 2020). All of these studies evaluated item analysis of MCQs of medical students. Large number of MCQs they analyzed had acceptable level of P, D and distractor efficiency.

From the previous studies above, it can be highlighted that most of studies investigated item analysis of MCQs in medical areas. Most of the researchers also conducted the study in higher education level. It can be said that there were no researchers investigated and analyzed MCQs quality of English subject especially for young learners. Therefore, this study was conducted to analyze the difficulty level, discriminating power, and distractor efficiency of MCQs of an English subject test for young learners.

## **METHOD**

This study employed a cross-sectional study to obtain information and evaluate MCQs in the students' tests. It was done to determine their

difficulty level, discrimination power, and distractor effectiveness.

In the middle of the term, the fifth-grade students were given a test about the English subject. Ten MCQs were part of the official exam paper that includes matching, short answers, and MCQs. After permission from the English teacher and consent from the participants were given, 40 students were taking the exam in October 2021. The students should select one best answer out of four.

The time given to the students to finish the exam was 60 minutes. They were supervised by the teacher in conducting the exam to avoid cheating. Moreover, a mark was given if they pick the correct answer. No penalties were given for incorrect answers.

The test results were adopted to determine the level of difficulty index, discrimination power, and the efficiency of the distractors. Each item was analyzed by using ANATES Ver. 4.0.9 and SPSS Ver. 20 to reveal their mean and standard deviation of difficulty index (P) and discrimination index (D). Microsoft Excel was also used in the process.

Interpretation:

Difficulty index (P), if:

P < 30% Difficult

P = 30 - 70% Acceptable

P > 70% Easy

Discrimination Index (D), if:

D = Negative, defective item/wrong key

D = 0-0.19 poor discrimination power

D = 0.2 - 0.29 acceptable discrimination power

D = 0.3 - 0.39 good discrimination

D > 0.4 excellent discrimination.

The higher the index value of difficulty items, the lower the item difficulty and the index value of difficulty items, the better the difficulty level of MCQs. On the other hand, the higher the discrimination index, the higher the item discrimination among the students with high- and low-test scores.

## RESULTS AND DISCUSSION

Students' knowledge measurement depends on the assessment that is given to the students. Summative or formative assessments help students to measure their knowledge and skills. It is argued that MCQ is one of the most efficient and effective methods for evaluation (Gajjar et al., 2014). In doing this, it would be essential to analyze and evaluate MCQs based on the difficulty index (P), discrimination

level (D), and distractor efficacy (DE). Bhat and Prasad (2021) explain that to create a good question bank, item analysis of MCQs should be regularly analyzed so it can be used to evaluate students' cognitive skills.

The means and standard deviations of the Difficulty index (P) and Discrimination index (D) were found to be  $60.75\% \pm 17.08$  and  $0.54 \pm 0.19$  respectively (Table 1).

Table 1. *Students' assessment result of 10 items*

Item Analysis Parameters	Mean	Std. Deviation
Difficulty index	60.75%	17.08
Discrimination index	0.54	0.19

The table above shows that difficulty index and the level of the discrimination index were acceptable. This result was in line with most related studies in which the difficulty index was considered in an acceptable level with an excellent discrimination index (Bhat & Prasad, 2021; Harti et al., 2021; Menon & Kannambra, 2017; Salih et al., 2020; Shete et al., 2015).

Item difficulty is related to the proportion of students who correctly respond to a particular item. The level of difficulty can be obtained by analyzing students' responses. It means that question difficulty was not determined by teachers' perceptions, but rather by students' answers (Wijayanti, 2020). In addition, distractors chosen by the students may be useful for the teachers to identify learning difficulties experienced by the students. It can be used to determine weak areas that require improvement. It directly helps teachers' and students' performances (Shete et al., 2015).

A low difficulty index indicates a difficult question. It means that the multiple-choice item was not taught well or it was too difficult for the students to understand. The difficulty level can be used as feedback on the test quality and a modification may be required before reusing the item in another test (Obon & Rey, 2019; Salih et al., 2020).

In this study, 10 MCQs and 40 distractors were analyzed. The first analysis was conducted to determine whether the items were difficult, acceptable, or easy to answer. Each item was categorized based on several criteria, namely Frequency (F) and Percentage (P (%)). The following table shows the final analysis of the difficulty index.

Table 2. Distribution of items difficulty level

Difficulty Level	Criteria	F	P (%)
0.00 - 0.30 (< 30%)	Difficult	0	0%
0.31 - 0.70 (30 - 70%)	Acceptable	8	80%
0.71-1.00 (> 70%)	Easy	2	20%

Out of 10 items, no MCQs were considered to be difficult. Eight items were acceptable (P = 30 - 70%) and two items (20%) were easy (P = > 70%). The table implied that there were no difficult questions given to the students and acceptable questions were dominating. In several previous studies, moderate-level questions were dominating (Elfaki, Alamri, & Salih, 2020; Jannah et al., 2021; Patil et al., 2016; Rao, Kishan Prasad, Sajitha, Permi, & Shetty, 2016). However, Obon and Rey (2019) found that nearly 60% of items should be revised substantially or discarded. It means that the quality of MCQs could be poor.

The discrimination level of the MCQs was also analyzed. The discrimination power can distinguish between higher students (those who have mastered the materials) and lower students (those who have not mastered the materials). To evaluate items discrimination power, the discrimination index was analyzed priorly. Discrimination index analysis results using Anates software are presented in Table 3.

Table 3. Distribution of items discrimination level

Discrimination Level	Criteria	F	P (%)
Negative	Rejected	0	0%
0-0.19	Poor	0	0%
0.2-0.29	Acceptable	1	10%
0.3-0.39	Good	1	10%
0.4-1.00	Excellent	8	80%

It can be seen that items with poor discriminating power could not be found. On the other hand, the table shows an item with acceptable and good discrimination levels each. Interestingly, it was found that 80% of the discrimination level of the items was excellent. This finding indicates that those items did not require modification. Acceptable, good, and excellent items can be reused and stored in the questions bank. However, mis-key questions, poor measurement of material competency, or students' misconceptions may occur

for having more than one answer. Therefore, revisiting and checking the items are recommended before reusing them.

Several studies also found various discrimination levels. Obon and Rey (2019) found most of the items were poor. They obtained poor discrimination power of 19.8% in which the questions were revised or rejected. Other studies found that most of the questions were satisfactory or poor (Namdeo & Sahoo, 2016; Danuwijaya, 2018; Hartati & Yogi, 2019; Jannah et al., 2021). In another study, Manfaat, Nurazizah, & Misri (2021) revealed that 83.33% of the items were good and 16.67% were poor.

Distractor analysis was also conducted in this study. The distractors of each item were analyzed. An item has good distractors if the lower students choose the wrong answers rather than higher students. In this case, distractor analysis is used to measure the efficiency of incorrect options to distract the lower groups (Manfaat et al., 2021). The result of the distractor efficiency analysis is shown in the following table.

Table 4. Distractor analysis

Distractor analysis	
Number of items	10
Number of total distractors	40
Functional distractor	32 (80%)
Non-Functional distractor	8 (20%)

The result shows that 32 (80%) of the distractors were functional while 8 of them were non-functional. Therefore, MCQs with functional distractors could be added to the question bank. On the other hand, the questions or the items with non-functional distractors should be changed.

This result can be compared with several other previous studies (Bhat & Prasad, 2021; Gajjar et al., 2014; Hingorjo & Jaleel, 2012; Kaur, Singla, & Mahajan, 2016; Patil, Palve, Vell, & Boratne, 2016). Earlier studies also revealed many non-functional distractors. The distractors in those studies were completely revised as they failed to misdirect the students (Hartati & Yogi, 2019; Hingorjo & Jaleel, 2012; Mehta & Mokhasi, 2014).

Developing good distractors and reducing non-functional distractors are important to make proper MCQs. Non-functional distractors in MCQs also affect the discrimination level. The more non-functional distractors in an item, the easier it would

be for a student to choose the correct answer and vice versa (Namdeo & Sahoo, 2016).

Table 5. *Distribution frequency of non-functional distractors (NFD)*

Number of NFD	Number of Items	Distractor Efficiency
0 NFD (DE = 100%)	8 (80%)	100%
1-2 NFD (DE = 50-75%)	1 (10%)	50%
3 NFD (DE = < 50%)	1 (10%)	0%

In the table, it can be seen that 8 (80%) items did not have NFDs and only 2 (20%) of them had 1, 2, and 3 NFDs. Similar results were also discovered by the previous studies. From 40 items, Patil et al. (2016) found that only 4 of them should be revised or discarded. Another study by Gajjar et al. (2014) revealed that there were no more than 20 items (out of 150) identified as NFDs. This finding indicates that the designed distractors were useful in each item.

From the discussion, several points can be emphasized including the importance and effectiveness of MCQs, discrimination difficulty index, and distractor analysis. By using the analysis, MCQs that are considered poor should be revised, re-tested, and reviewed before incorporating them into the future test. It would also improve the item quality that would be properly discriminated by the students (Izah, Odubo, Ajumobi, & Torru, 2021; Karkal & Kundapur, 2016; Shaibani, Ali, Deifalla, & Jaradat, 2021).

## CONCLUSION

This study highlighted three main findings, namely difficulty index, discrimination level, and distractor efficiency. The analysis items appear to be efficient. Most of their difficulty was acceptable, which means the items were suitable for the students. It was also revealed that the items discrimination power was good, indicating they are not required to be revised or replaced. In terms of distractor efficiency, two items were completely inefficient. Because several items were problematic, alternative items may be necessary for future use.

The result of this study is expected to encourage teachers to identify poorly developed MCQs and to improve their quality. In addition, teachers should be responsible for the exam and item assessment after it. If teachers analyze the items, the school and the students would be benefited from the feedback.

Developing good MCQs requires knowledge, experience, and practice. Therefore, teachers are suggested to attend seminars about designing proper items. In this case, schools should provide such classes for the teachers to ensure the items are well-developed and properly discriminated with functioning distractors. Teachers can also have a chat with students to know about a test difficulty, discrimination power, distractor efficiency.

For further study, it is suggested to investigate other item analysis factors such as students' ability, the number of questions, the length of the questions, the quality of instructions, and the number of students about the quality of MCQs.

## REFERENCES

- Bhat, S. K., & Prasad, K. H. L. (2021). Item analysis and optimizing multiple-choice questions for a viable question bank in ophthalmology: A cross-sectional study. *Indian Journal of Ophthalmology*, 9(2), 343–346. [https://doi.org/10.4103/ijo.IJO\\_1610\\_20](https://doi.org/10.4103/ijo.IJO_1610_20)
- 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(June), 1–12. <https://doi.org/10.3389/educ.2017.00024>
- Danuwijaya, A. A. (2018). Item analysis of reading comprehension test for post-graduate students. *English Review: Journal of English Education*, 7(1), 29. <https://doi.org/10.25134/erjee.v7i1.1493>
- Elfaki, O. A., Alamri, A., & Salih, K. A. (2020). Assessment of MCQs in MBBS program in Saudi Arabia. *Onkologia i Radioterapia*, 14(5), 12–16.
- Freahat, N. M., & Smadi, O. M. (2014). Lower-order and higher-order reading questions in secondary and university level EFL textbooks in Jordan. *Theory and Practice in Language Studies*, 4(9), 1804–1813. <https://doi.org/10.4304/tpls.4.9.1804-1813>
- Gajjar, S., Sharma, R., Kumar, P., & Rana, M. (2014). Item and test analysis to identify quality multiple choice questions (MCQS) from an assessment of medical students of Ahmedabad, Gujarat. *Indian Journal of Community Medicine*, 39(1), 17–20. <https://doi.org/10.4103/0970-0218.126347>
- Hartati, N., & Yogi, H. P. S. (2019). Item analysis for a better quality test. *English Language in Focus (ELIF)*, 2(1), 59–70. <https://doi.org/https://doi.org/10.24853/elif.2.1.59-70>
- Harti, S., Mahapatra, A. K., Gupta, S. K., & Nesari, T. (2021). All India AYUSH post graduate entrance exam 2019 – AYURVEDA MCQ item analysis. *Journal of Ayurveda and Integrative Medicine*, 12, 356–358.

- <https://doi.org/10.1016/j.jaim.2021.01.013>
- Hingorjo, M. R., & Jaleel, F. (2012). Analysis of one-best MCQs: The difficulty index, discrimination index and distractor efficiency. *Journal of the Pakistan Medical Association*, 62(2), 142–147.
- Izah, S. C., Odubo, T. C., Ajumobi, V. E., & Torru, K. E. (2021). Item analysis of Multiple Choice Questions ( MCQs ) from a formative assessment of first year microbiology major students. *Res Rev Insights*, 5, 1–6. <https://doi.org/10.15761/RRI.1000166>
- Jannah, R., Hidayat, D. N., Husna, N., & Khasbani, I. (2021). An item analysis on multiple-choice questions: A case of a junior high school English try-out test in Indonesia. *Leksika: Jurnal Bahasa, Sastra dan Pengajarannya*, 15(1), 9–17. <https://doi.org/10.30595/lks.v15i1.8768>
- Jayanti, D., Husna, N., & Hidayat, D. N. (2019). The validity and reliability analysis of English national final examination for junior high school. *VELES Voices of English Language Education Society*, 3(2), 127–135. <https://doi.org/10.29408/veles.v3i2.1551>
- Karadag, N. (2016). Analysis of the difficulty and discrimination indices of multiple-choice questions according to cognitive levels in an open and distance learning context. *Turkish Online Journal of Educational Technology*, 15(4), 16–25.
- Karkal, Y. R., & Kundapur, G. S. (2016). Item analysis of multiple choice questions of undergraduate pharmacology examinations in an International Medical School in India. *Journal of Dr. NTR University of Health Sciences*, 5(3), 183–186. <https://doi.org/10.4103/2277-8632.191842>
- Kaur, M., Singla, S., & Mahajan, R. (2016). Item analysis of in use multiple choice questions in pharmacology. *International Journal of Applied and Basic Medical Research*, 6(3), 170–173. <https://doi.org/10.4103/2229-516x.186965>
- Kowash, M., Hussein, I., & Halabi, M. Al. (2019). Evaluating the quality of multiple choice question in paediatric dentistry postgraduate examinations. *19(2)*, 135–141.
- Kusumawati, M., & Hadi, S. (2018). An analysis of multiple choice questions (MCQs): Item and test statistics from mathematics assessments in senior high school. *Research and Evaluation in Education*, 4(1), 70–78. <https://doi.org/10.21831/reid.v4i1.20202>
- Manfaat, B., Nurazizah, A., & Misri, M. A. (2021). Analysis of mathematics test items quality for high school. *Jurnal Penelitian dan Evaluasi Pendidikan*, 25(1), 108–117. <https://doi.org/10.21831/pep.v25i1.39174>
- Mehta, G., & Mokhasi, V. (2014). Item analysis of multiple choice questions- an assessment of the assessment tool. *Historical Aspects of Leech Therapy*, 4(7), 1–3.
- Menon, A. R., & Kannambra, P. N. (2017). Item analysis to identify quality multiple choice questions. *National Journal of Laboratory Medicine*, 7–10. <https://doi.org/10.7860/njlm/2017/25690:2215>
- Namdeo, S., & Sahoo, B. (2016). Item analysis of multiple choice questions from an assessment of medical students in Bhubaneswar, India. *International Journal of Research in Medical Sciences*, 4(5), 1716–1719. <https://doi.org/10.18203/2320-6012.ijrms20161256>
- Obon, A. M., & Rey, K. A. M. (2019). Analysis of multiple-choice questions (MCQs): Item and test statistics from the 2nd year nursing qualifying exam in a university in Cavite, Philippines. *Abstract Proceedings International Scholars Conference*, 7(1), 499–511. <https://doi.org/10.35974/isc.v7i1.1128>
- Patil, R., Palve, S., Vell, K., & Boratne, A. (2016). Evaluation of multiple choice questions by item analysis in a medical college at Pondicherry, India. *International Journal of Community Medicine and Public Health*, July, 1612–1616. <https://doi.org/10.18203/2394-6040.ijcmph20161638>
- Polat, M. (2020). Analysis of multiple-choice versus open-ended questions in language tests according to different cognitive domain levels. *Novitas-ROYAL (Research on Youth and Language)*, 14(2), 76–96.
- Patil, P. S., Dhobale, M. R., Mudiraj, N. (2016). Item analysis of MCQs - Myths and realities when applying them as an assessment tool for medical students. *Int J Cur Res Rev*, 8(13), 12–16.
- Purwoko, M., & Mundijo, T. (2018). Evaluating the use of MCQ as an assessment method in a medical school for assessing medical students in the competence-based curriculum. *Jurnal Pendidikan Kedokteran Indonesia: The Indonesian Journal of Medical Education*, 7(1), 44–48. <https://doi.org/10.22146/jpki.35544>
- Rao, C., Kishan Prasad, H., Sajitha, K., Permi, H., & Shetty, J. (2016). Item analysis of multiple choice questions: Assessing an assessment tool in medical students. *International Journal of Educational and Psychological Researches*, 2(4), 201–204. <https://doi.org/10.4103/2395-2296.189670>
- Rehman, A., Aslam, A., & Hassan, S. H. (2018). Item analysis of multiple choice questions in pharmacology. *Pakistan Oral & Dental Journal*, 38(2), 291–293. <https://doi.org/10.52206/jsmc.2020.10.2.320>
- Salih, K. E. M. A., Jibo, A., Ishaq, M., Khan, S., Mohammed, O. A., Al-Shahrani, A. M., & Abbas, M. (2020). Psychometric analysis of multiple-choice questions in an innovative curriculum in

- Kingdom of Saudi Arabia. *Journal of Family Medicine and Primary Care*, 9(7), 3663–3668. <https://doi.org/10.4103/jfmpe.jfmpe>
- Shaibani, T. A., Ali, F., Deifalla, A. H., & Jaradat, A. (2021). Item analysis of type “A” multiple choice questions from a multidisciplinary units assessment in a problem based curriculum. *Bahrain Medical Bulletin*, 43(2), 471–476.
- Shete, A., Kausar, A., Lakhkar, K., & Khan, S. (2015). Item analysis: An evaluation of multiple choice questions in Physiology examination. *Journal of Contemporary Medical Education*, 3(3), 106–109. <https://doi.org/10.5455/jcme.20151011041414>
- Shin, J., Guo, Q., & Gierl, M. J. (2019). Multiple-choice item distractor development using topic modeling approaches. *Frontiers in Psychology*, 10(April), 1–14. <https://doi.org/10.3389/fpsyg.2019.00825>
- Wijayanti, P. S. (2020). Item quality analysis for measuring mathematical problem-solving skills. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, 9(4), 1223–1234. <https://doi.org/https://doi.org/10.24127/ajpm.v9i4.3036>

**Maya Marsevani**

*Item analysis of multiple-choice questions: An assessment of young learners*