



# Examining the Validity of an Arabic Test for Islamic Secondary School Students

Mad Ali<sup>1,\*</sup>, Zilva Raihan Nazir Ali Alfathimy<sup>1</sup>, Siti Fatimah Nurhabibah<sup>2</sup>,  
Nadi Azkia Ali Alfathimi<sup>1</sup>, Nida Damia Ramdhani Ali Alfathimy<sup>1</sup>

<sup>1</sup> Arabic Education Department, Faculty of Language and Literature Education, Universitas Pendidikan Indonesia, Bandung, Indonesia

<sup>2</sup> Kementerian Agama Kota Bandung, Indonesia

\*Corresponding author. Email: madali@upi.edu

## ABSTRACT

This study aims to determine the validity of the Islamic Secondary School test instrument in Arabic subject. The method used in this research is descriptive analysis, by analyzing the test instrument and examining three logical sub-validities: surface, content, and construct validities. The results show that surface validity is still low because of many mistakes in pronunciation, letters, syntax, word form and choice, preposition use, writing of *hamzah* letters, matching, and option errors. Content validity shows that the distribution of discussion topics is already represented on the items of the test, but the validity of language goals and actions is still dominated by understanding and grammatical skills, few questions for speech acts and idea analysis skills, and none of them measure the ability to present the results of the analysis of ideas from the text. The logical validity for the sub-construct validity shows that the level of knowledge still dominates the items, followed by the levels of analysis, understanding, application, synthesis, and evaluation. The level of imagination and creation does not get a portion.

**Keywords:** Arabic test, construct validity, Islamic secondary schools, test validation.

## 1. INTRODUCTION

An examination or an exam is an integral part of the activity process in any institution to find out how far the expected goals and targets can be achieved and what feedback for the institution to improve the activity programs in it if the target is not achieved (Arikunto, 2008). Likewise, in educational institutions, an examination, which is identically in the form of a test, is the last phase of the learning process. It can be in the forms of formative tests, midterm exams, final semester exams as summative tests, and a comprehensive summative exam in the form of final exams for student graduation. The results of the school final exam will be used as information by all parties, such as subject teachers, school principals, as well as parents, whether the learning process is running effectively or not, has met the target or not (Munthe, 2015; Ridho, 2018; Martini, Ariyanti, Suyatno, & Prasetyowati, 2022). Therefore, this test as a measurement tool must be able to measure students' abilities, in which the results can be used to make decisions.

Tests can measure students' abilities if they refer to predetermined goals, which include various competency elements that students must master when they have completed a learning activity program. Test instruments that can measure students' abilities are called valid test tools. Therefore, the validity of the test is something that must be fulfilled before the test is used in examination activities. The validity of this test must be tested first by the test creator before it is used so that it is truly effective and efficient in measuring student competency. Test validity consists of logical validity and empirical validity. Logical validity relates to internal tests, while empirical validity relates to external tests. In this context, test makers play a major role in making valid tests so that these tests can measure student behavior (Wang, 2014; Fikriyah, 2021). Logical validity has a significant urgency before empirical validity. This is because logical validity relates to the first step to ascertain whether the test tool is by the curriculum targets

that have been determined, including the level of readability of the test. After all, this logical validity consists of surface validity, content validity, and construct validity.

For tests made by the government, such as tests for final school examinations that are national, for example, the test for obtaining a Pure Evaluation Score (NEM), the level of logical validity has been guaranteed by the national test preparation team. However, for tests made by schools themselves, the level of logical validity has not been guaranteed because they are often not analyzed in advance by the compilers, in this case, the subject teachers. This occurs in the Arabic language subject in madrasas at the end of school exams called the Madrasa Exam (UM) because this subject is not classified as a subject that is tested nationally by the Ministry of Religion, but the questions are left to each school. Because of this, errors often appear in terms of writing (surface validity), not by the curriculum (content validity), and lack of attention to the proportion of the distribution of students' cognitive behavior (construct validity).

The invalidity of the test questions in the Madrasa Exam for Arabic subjects ultimately resulted in the uncertainty of the measuring instrument to determine the quality of students' Arabic language skills. If the exam results are used as information for making decisions by the school, the decisions they make can be wrong, which will be detrimental to all parties involved including students, teachers, principals, parents, and even graduate users.

This study aims to test the logical validity of the Arabic test questions in an exam for Islamic lower secondary school students. This validation is crucial because these test questions are made by Arabic teachers in each school, who sometimes pay little attention to their logical validity. This logical validity is the starting point for making this test a tool that can truly measure students' abilities comprehensively at the end of their schooling, because this validity relates to the clarity of writing and the correctness of the question phrases that can avoid students' misunderstanding of each item, also about language skills by those targeted in the curriculum, the extent to which the test questions can describe the entire content of the curriculum, and is also related to the distribution of cognition which includes knowledge, understanding, application, analysis, evaluation, synthesis, imagination, and creation (Nur'aini, 2011; Muhammad & Ariani, 2021).

On that basis, this research is focused on three problems as follows:

1. How is the surface validity of the Arabic language test questions in the Madrasa Tsanawiyah Exam?
2. How is the validity of the content in the form of ability targets and Arabic learning objectives contained in the Arabic language test questions?
3. How is the construct validity of the test questions?

The results of this study will benefit all parties involved in the Madrasah Tsanawiyah Exam activities specifically for Arabic subjects, in measuring students' abilities in Arabic subjects.

## 2. LITERATURE REVIEW

### 2.1. *Validity of Tests and Their Types*

The validity of the test means that the test is a measuring tool that can measure students' abilities (Soysal & Kou, 2018; Teufl, Miezel, Taetz, Fröhlich, & Bleser, 2018; Alfathimy & Ali, 2021). This is in line with the opinion of Azwar (1986) which states that validity has a sense of the extent to which the accuracy of a measuring instrument performs its measuring function. In the context of learning Arabic, the validity of the Arabic test is a test tool that can accurately measure all elements of language skills, which include listening, speaking, reading, and writing skills (Ilham & Wijati, 2020; Suryanto & Akhadiyah, 2019). In addition to measuring the four skills, for non-Arabic speaking students, there are abilities related to understanding and applicative abilities to the rules of the Arabic language, such as syntax and morphology.

Test validity consists of logical validity and empirical validity. Logical validity relates to the scientific logic being tested, which consists of surface validity, content validity, and construct validity (Arikunto & Jabar, 2014). Empirical validity relates to the validity of tests based on field trial results, by comparing test results to be tested for validity with previous test results, namely standard tests that have been made by other parties and have been tested for validity, as well as test results that are, or compared with predicted results, especially the validity of tests to enter certain educational institutions compared to results achievement test after they attend education for a certain period (Rusli & Mulyana, 2022).

This study only focuses on logical validity which consists of surface validity, content validity, and construct validity.

## **2.2. Logical Validity**

Arikunto (2008) asserts that logical validity for an evaluation instrument is a test tool that meets the valid requirements based on reasoning because the term logical relates to logic or reasoning. This means that a test instrument that has logical validity is an instrument that is by scientific reasoning and test writing rules. Furthermore, logical validity does not need to be tested for conditions but can be known after the test instrument is compiled. Logical validity includes surface validity, content validity, and construct validity.

## **2.3. Surface Validity**

Surface validity according to Arikunto (2009) is a type of validity that has the lowest level and significance, because it is only the result of a cursory glance at the measuring instrument. If the test instrument appears to be by the material to be measured, then the surface validity is fulfilled. This is part of the low level of content validity.

## **2.4. Content Validity**

Syihabuddin (2009) states that content validity refers to whether the test tool is aligned with the objectives and descriptions of the teaching materials. By quoting the opinion of Grondlund (1982), he concluded that to test the validity of the content, two things need to be considered, namely 1) whether the learning outcomes test measures the teaching materials that have been given as a whole; 2) whether the test can measure changes in student behavior as expected. He also emphasized that content validity is carried out logically and rationally. Therefore, content validity is also known as logical validity.

Nazir (1988) citing Egel's opinion states that content validity is validity related to what appears in measuring something, and not to what is to be measured. Here Egel looks more at surface validity only by paying attention to the state of the text on the test, in terms of the shape of the letters, the structure of the presentation of the test, including whether the sentences used on the test are correct or not.

## **2.5. Construct Validity**

Construct validity according to Kusaeri (2012) is validity that pays attention to the extent to which test items can measure what is to be measured by a specific concept or predetermined conceptual definition. Construct validity relates to something abstract, but the symptoms can be seen and can be measured. According to him, the construct validity of a test is also related to the measurement of attitudes, interests, self-concept, leadership style, achievement motivation, and others.

## **3. METHOD**

This research uses a qualitative approach with descriptive analysis methods (Soendari, 2012). The data analyzed is in the form of questions written in the Arabic language subject test document, as an Islamic Secondary School Test to determine student graduation. The total number of questions is 60, taken from Arabic language subjects in class 7 with a portion of 10%, class 8 with a portion of 20%, and class 9 with a portion of 70%. The competencies tested include (1) reading skills which include mastery of vocabulary, phrases, and understanding texts, (2) writing skills which consist of the ability to compose sentences and perfect sentences, and (3) mastery of grammar, which consists of morphology and syntax. This test was written by a team of teachers who are members of the Cimahi City Arabic Language Teachers' Conference organization.

The validity of the test studied in this research is only logical validity, which highlights the extent to which the questions on the test are by scientific reasoning and test writing rules (Arikunto, 2008). This includes three parts, namely surface validity, content validity, and construct validity.

Testing surface validity is done by correcting 1) vowel symbols in each word, 2) *I'rab*/vowel or suffix of each word in the sentence, 3) letter/phoneme writing, 4) word form and choice, 5) appropriate prepositions used, 6) *imla'*/*hamzah* letter writing, 7) matching in masculine-feminine / singular- double/dual- plural and 8) error in answer options for each question.

Construct validity is examined by analyzing each question item in terms of psychological construction, which is related to the domains of learning objectives, by referring to Bloom's taxonomy (Arikunto, 2008), but only focusing on knowledge. As for skills and attitudes, no part can be analyzed, because all the questions are related to knowledge,

as is common in written tests in multiple-choice form. However, the knowledge aspect is divided into (1) memory, (2) understanding, (3) application, (4) analysis, (5) evaluation, (6) synthesis, (7) imagination, and (8) creation.

The implementation of this logical validity examination was carried out by the researcher himself, as a lecturer in Arabic language education evaluation courses in undergraduate and postgraduate degrees at the Department of Arabic Language Education, Faculty of Language and Literature Education, *Universitas Pendidikan Indonesia*/Indonesian University of Education. The references for measuring this logical validity are books on educational evaluation theories in general, and Arabic language evaluation in particular, Arabic language subject curriculum books published by the Ministry of Religion of the Republic of Indonesia, and Arabic language subject books for classes. 7, 8, and 9, as well as Arabic morphosyntactic books.

## 4. FINDINGS AND DISCUSSION

### 4.1. Face Validity

The surface validity of the test questions in the Final Examination of Islamic Secondary School in the city of Cimahi includes 1) vowels in each word, 2) *I'rab*/vowels or suffixes of each word, 3) writing of letters/phonemes, 4) form and word choice, 5) prepositions used, 6) *imla'*/*hamzah* writing, 7) compatibility of masculine-feminine / singular-dual-plural, and 8) error in answer options for each question.

#### 4.1.1. Vowel

From the results of the research on the 60 questions, the researchers found 5 words that were incorrect in terms of the vowel symbols in these words (see Table 1). For example, in the word *المُرَافِقُ* / *al-mura>fiq/ facilities* in question 3, the letter *م* / *mu* should be read *fat-h>ah/vocal A* because the desired meaning in question 3 is *facilities*, whereas if you read *d>ammah/vocal U* then the meaning is accompanying person. Vocal errors in something also result in unclear or non-existent meaning at all, as in questions 52, 58, and 60. Then, sample of words that are wrong in terms of *i'rab* are shown in Table 2.

**Table 1.** Wrong words in terms of vocals

No.	No. Question	Wrong Word Vowel	Correction
1	3	المُرَافِقُ ؟ (م —) (... <i>ccompanying? (mu) al-mura&gt;fiq/</i>	..... المُرَافِقُ.... (م —) (... <i>al-mara&gt;fiq/ facilities? (ma)</i>
2	51	المَزَارِعَ ( ر ) <i>al-maza&gt;ra' / farms (ra)</i>	المَزَارِعَ ( ر ) <i>al-maza&gt;ri' / farms (ri)</i>
3	51	الخَضْرَاءَ (خ —) (... <i>al-khud}&gt;ra&gt; / green (khu)</i>	الخَضْرَاءَ (خ —) (... <i>al-khad}&gt;ra&gt; / green (kha)</i>
4	58	د- واسعٌ ( س — ) <i>wa&gt;sa'un (sa) / wide</i>	د- واسعٌ ( س — ) <i>wa&gt;si'un (si) / wide</i>
5	60	يُنْقَسِمُ ..... (ي —) (... <i>yunqasamu (yu...sa...)/It divides</i>	يُنْقَسِمُ (ي —) (... <i>yanqasimu (ya...si...)/It divides</i>

**Table 2.** Sample of words that are wrong in terms of *i'rab*

No.	No. Of Question (Q)	Words Of Wrong Vocal	Correction
1	3	هل المُرَافِقُ المَدْرَسِيَّةُ فِي مَدْرَسَتِكَ كَثِيرَةٌ ؟ <i>Hal al-mura&gt;fiq al-madrasati&gt;ah fi&gt; madrastika kas&gt;iratun / Are there many school facilities in your school?</i>	.... فِي مَدْرَسَتِكَ.... <i>... fi madrasatika .../ in your school</i>
2	5	عُنْوَانُ مَدْرَسَتِي فِي شَارِعِ "مُحَمَّدٌ نُوحٍ" رَقْمَ 7 كَالِمَائْتَانِ . (ن) <i>'Unwani madrasati&gt; fi&gt; sya&gt;ri' "Muhammad Nu&gt;h" raqmu 7 Kalimantan / The address of my school is at No. 7 Nuh Street, Kalimantan.</i>	عُنْوَانُ ... (ن) <i>'Unwa&gt;nu... ..(nu)/ Adress</i>
...	...	....	....
18	55	وَزَيَّنَهَا بِالْكَوَاكِبِ وَالنُّجُومِ . (م) <i>Wa zayyanaha&gt; bi al-kawa&gt;kibi wa al-nuju&gt;ma / And He decorated it with planets and stars.. (ma)</i>	..... وَالنُّجُومِ (م) <i>wa al-nuju&gt;mi / stars</i>

#### 4.1.2. Writing Letters/Phonemes

In terms of missing or excess letters, errors occurred in 5 words, each of which was in question (Q) item no. 22 (Q22), Q26, Q44, and Q54 (2 words) (see samples in Table 3). For example, in Q22, there appears to be the word المَبَارَاة /*al-muba>rah*, missing the letter *alif* before the letter (ة), it should be المَبَارَاة /*al-muba>ra>h*/ the match, because the origin is المَبَارَاة /*al-muba>rayah*/the match, with the pattern المَفَاعَلَة /*al-mufa>'alah*/reactance. The lack of such letters results in unclear meaning, or even no meaning at all. Likewise, in Q26, it appears that the word فَالَّاحُ /*falla>khun*, should have used the letter ح فَالَّاحُ / *falla>h}un*, because it means a farmer, instead it should have been added the letter (ة) فَالَّاحَةُ /*falla>hatun*/ ; because it indicates a female farmer.

**Table 3.** Sample words containing missing, excess, or mistakes of letters

No.	No. Of Question (Q)	Wrong Words Due To Disadvantages/Excesses/Lettering Errors	Correction
1	22	... لِإِشَاهِدِ الْمَبَارَاةِ . ... <i>li usya&gt;hida al-muba&gt;rah</i> . To watch the match (missing the <i>alif</i> after the letter ج )	.... الْمَبَارَاة . <i>al-muba&gt;rah</i> / the match
2	26	لِأَنَّهَا فَالَّاحُ. (خ) <i>liannaha&gt; falla&gt;khun</i> ./Because she is a farmer.	.... فَالَّاحُ (ح) <i>Falla&gt;hun</i> / a farmer
....	....	.....	.....
5	54	فِيهَا الْجِبَالُ وَالْبَحَارُ وَالْأَنْهَارُ وَالنَّخِيرَاتُ. <i>Fiha&gt; al-jiba&gt;lu wa al-biha&gt;ru wa al-an'aha&gt;ra wa al-buh}aira&gt;ti</i> ./ It has mountains, seas, rivers, and lakes	.... الْأَنْهَارُ ... <i>al-anh}a&gt;ra</i> /rivers ...

#### 4.1.3. Word Form or Word Choice

Regarding errors in word form or choosing the right words, it was found that there were 7 wrong words. Each is at Q13, Q15, Q20, Q 25, Q 26, S28, Q32, Q34, Q38, and Q41 (see sample in Table 4). For example, in Q13, there is the expression ب- يَتَنَاوَلُ الْعِشَاءَ /*b- yata>nawalu al-'isya>'a*/ have dinner). Here there is the word يَتَنَاوَلُ / *yata>nawalu* , it should be written يَتَنَاوَلُ / *yatana>walu* /have, because the correct pattern is يَتَفَاعَلُ /*yatafa>'alu*/he reacted.

**Table 4.** Samples of incorrect words in terms of form and selection

No.	No. Of Question (Q)	Wrong Words in Terms Of Form And Choice	Correction
1	13	....ب- يَتَنَاوَلُ الْعِشَاءَ <i>b-yata&gt;nawalu al-'isya&gt;'a</i> /have dinner	.... يَتَنَاوَلُ ... <i>yatana&gt;walu</i> /have
2	20	.... الرِّيَاضَةُ الْمُخْتَلَفَةُ .../ <i>al-riya&gt;d}ah al-mukhtalifah</i> /... different sports	.... الرِّيَاضَاتُ / <i>al-riya&gt;d}a&gt;t</i> /sports
....	....	.....	.....
5	41	.... بَلْ اسْتَمَرَّ نَزُولُهُ مُفْرَقًا مَدَّةَ 23 سَنَةً. ... <i>bal istamarra nuzu&gt;luhu mufarriqan muddata 23 sanah</i> ./... Rather, its descent continued separately over a period of 23 years.	.... مُفْرَقًا... ... <i>mufarraqan</i> .../ separately

So, here there is a wrong choice of word form. Likewise, in Q20, there is an expression (.../ *al-riya>d}ah al-mukhtalifah* /different sports). Here there is the word الرِّيَاضَةُ / *al-riya>d}ah* /sport, it should be written الرِّيَاضَاتُ / *al-riya>d}a>t* /sports, namely in plural form; because it is characterized by the word المختلفة / *al-mukhtalifah* (which are various). In Arabic texts, it is unacceptable for the word المختلفة / *al-mukhtalifah* to refer to a single word.

#### 4.1.4. Errors or Missing Prepositions

Regarding the absence or misuse of prepositions, there were six errors spread across 6 questions, namely Q20, Q42, Q44, Q47, Q56, and Q60 (see sample in Table 5). For example, in Q20, there is the expression (... / *yaltahiqu>na ila>* / they celebrate). Here there is the preposition إلى /*ila>*, after the verb يلتحقون /*yaltahiqu>na*/ they celebrate. This verb should use the preposition بِـ /*bi* to its object, so the correct one is يلتحقون بِـ /*yaltahiqu>na bi*/ they celebrate.

**Table 5.** Errors or omissions of prepositions

No.	No. Of Question (Q)	Wrong Words In Terms Of Form And Choice	Correction
1	20	... يَلْتَحِفُونَ إِلَى ... yaltahiqu>na <i>ila</i> > ... They celebrate ....	... يَلْتَحِفُونَ بِ... yaltahiqu>na <i>bi</i> .../ They celebrate ....
2	42	لَمْ يَحْتَفِلْ الْمُسْلِمُونَ نُزُولَ الْقُرْآنِ ... Lam yahtafil al-muslimu>na nuzu>la al-Qur'a>ni.../ Muslims did not celebrate the revelation of the Qur'an.	لَمْ يَحْتَفِلْ ... بِنُزُولِ الْقُرْآنِ ... Lam yahtafil al-muslimu>na binuzu>li al-Qur'a>ni.../ ... did not celebrate the revelation of the Qur'an.
...	...	...	...
3	60	مِنْهَا، بِسَبَبِ رَمْيِ النِّفَاقَاتِ ... Minha>, bisababi rami al-nafa>ya>t .../ Including the reason for throwing waste....	مِنْهَا سَبَبُ رَمْيِ النِّفَاقَاتِ ... Minha> sababu rami al-nafa>ya>t .../ Including the reason for throwing waste....

In Arabic, the preposition is what determines the meaning of a verb, which originally had a certain meaning and changed to another meaning. Therefore, the use of the preposition after the verb must be careful; because it can change the meaning. For example, in the verb رَغِبَ /ragiba, if given the preposition في /fi becomes فِي رَغِبَ /ragiba fi, then it means *to like*, whereas if given the preposition عن /an, it becomes رَغِبَ عَنْ /ragiba 'an, it means *to hate*. In Q42, there is an expression (لَمْ يَحْتَفِلْ الْمُسْلِمُونَ نُزُولَ الْقُرْآنِ) / Lam yahtafil al-muslimu>na nuzu>la al-Qur'a>ni/ Muslims did not celebrate the revelation of the Qur'an). Here there is a lack of prepositions in the verb لَمْ يَحْتَفِلْ / lam yahtafil/ *did not celebrate*. There should be the preposition بِ /bi, so it is written بِنُزُولِ الْقُرْآنِ ... / lam yahtafil ... binuzu>li al-Qur'a>ni / *did not celebrate the revelation of the Quran*.

#### 4.1.5. Error Writing of Hamzah Letter

Regarding wrong *imla'* (writing *hamzah*), there is an expression (أَنْظُرْ إِلَى الصُّورَةِ! / unz}ur ila al-s}urati!). The error in writing the *hamzah* letter occurs in the word أَنْظُرْ / unz}ur, which means pay attention. This expression is repeated in 12 question items, namely Q4, Q7, Q11, Q19, Q23, Q27, Q37, Q45, Q49, Q50, Q52, Q53, and Q57 (see sample in Table 6). The editorial is the same, namely the command to pay attention to the picture when answering a question. Supposedly, the *hamzah* letter at the beginning of this command verb is *hamzah washal*, where the *hamzah* letter should not be marked *hamzah* (ء), but written in alif form, so it should be written أَنْظُرْ / unz}ur, not أَنْظُرْ / unz}ur. This violates the rules for writing *hamzah*, which is called *imla'* (Umar, 1993).

**Table 6.** The wrong word in terms of *imla'*/writing the *hamzah* letter

No.	No. Question Items (Q)	The Wrong Words In Terms Of Writing The Hamzah Letter	Correct
1	S4	أَنْظُرْ إِلَى الصُّورَةِ! \ unz}ur ila al-s}urati! Look at the picture!	انظُرْ إِلَى الصُّورَةِ! \ unz}ur ila al-s}urati! Look at the picture!
2	S7	أَنْظُرْ إِلَى الصُّورَةِ! \ unz}ur ila al-s}urati! Look at the picture!	انظُرْ إِلَى الصُّورَةِ! \ unz}ur ila al-s}urati! Look at the picture!
...	...	...	...
3	S57	انظُرْ إِلَى الصُّورَةِ! Look at the picture!	انظُرْ إِلَى الصُّورَةِ! Look at the picture!

#### 4.1.6. Compatibility Errors

Regarding word errors in terms of compatibility between subject and predicate, between adjectives and words characterized in terms of gender (masculine-feminine, singular-double-plural, definite-indefinite), there are three questions which include the words that are not compatible, namely Q15, Q44, and Q46 (see Table 7). For example, in Q15, there is the expression ( ... مِنَ النَّصِّ السَّابِقِ هِيَ ... ) / al-maudu>' al-muna>siba min al-nas}s} al-sa>biq hiya.../The appropriate topic from the previous text is....). The pronoun هي /hiya/she is not matching with its referent, namely the word الموضوع /al-maud}u>', in the masculine form. The word هي /hiya/she should be replaced with the word هو /huwa/he, so that masculine-feminine compatibility occurs. Another example in Q44, there is the expression ( ... تَرْتِيبُ الْجُمْلَةِ الصَّحِيْحَةِ هِيَ ... ) / tarti>b al-jumlah al-s}ahi>hah hiya .../ The correct sentence order is....). The pronoun هي

/hiya /she here is also not compatible with its reference الترتيب /al-tarti>b /the order, which is in the masculine form. The pronoun for compatibility should be هو /huwa/he (Ali, 2022; Ni'mah, 1973). The word الترتيب /al-tarti>b /the order is also wrong in terms of phrase structure because the word الترتيب is *mudhaf*/the first word in the *id}a>fi* phrase/noun phrase, so it cannot begin with the particle *Al* (ال) (Ali, 2019).

**Table 7.** *Muthabaqah/Compatibility errors*

No.	No. Question Items (S)	Wrong Words In Terms Of Compatibility	Correction
1	15	المَوْضُوعُ الْمُنَاسِبُ مِنَ النَّصِّ السَّابِقِ هِيَ... al-maudu>' al-muna>siba min al-nas{s} al-sa>biqi hiya.../The appropriate topic from the previous text is....)	المَوْضُوعُ ... هُوَ... al-maudu>' ... huwa.../The ... topic from ... is....)
2	44	الترتيبُ الجملة الصَّحِيحَةُ هِيَ... /tarti>b al-jumlah al-s}ahi>hah hiya .... / The correct sentence order is....	ترتيب ... هُوَ... /tarti>b ... hiya .... / The .... order is....
3	46	ترتيب هذه الجمل لتكون فكرة مفيدة هِيَ... Tarti>b haz{ih al-jumali litaku>na fikrah mufi>dah hiya .... / Arranging these sentences to make a useful idea is...	ترتيب ... هُوَ... Tarti>b ... huwa .... / Arranging ... is...

The final finding regarding surface validity relates to the error of making options the answer that must be chosen. It turned out that there were four questions in which there were errors in writing options, namely Q24, Q36, Q37, and Q59.

Errors that also disturb test takers (students), or make it easier to guess the answer, are option errors. Of the 60 questions studied, it turned out that there were 4 wrong options, namely Q24, Q36, Q37, and Q59. Types of option errors include the absence of the correct option that can answer the question (Q24), which will be detrimental to test takers; because they cannot answer correctly, because they feel confused about which one to choose; also, errors in the form of options that are not in the same word theme (Q36, Q37, Q59). Mistakes like this will lead students to choose the correct answer. For example, there is a question about a certain city, but the options are not all city names (36), or there is a question about a certain job, but the wrong option is in the form of the place and time (Q37), there are also questions where the answer is easy to guess because it only repeats the words/phrases in the statement being asked (Q59).

Based on the previous explanation of errors covering these eight aspects, this shows that the question items are not valid from a surface perspective, because briefly it appears that there are deviations from the rules for writing words, phrases, and sentences that have been standardized by grammar/ Arabic grammar also violates the rules for writing test questions. This will not happen if the question writer first reviews its surface validity before using it, especially in these eight aspects, either by the writer himself or another party who an Arabic language expert is.

## 4.2 Content Validity

Content validity is the same as face validity, namely the extent to which the question items are in line with curriculum objectives and truly measure basic competencies (Rochman, Abdurrahman, & Ali, 2021). From the results of the analysis of the Arabic language curriculum for Islamic Secondary Schools, an idea was obtained of the extent to which the test question items represented reading themes in the Arabic curriculum, and what basic competencies the question items represented.

Considering that the test questions studied are test questions for Secondary School as a replacement for the National Examination which has been abolished (Rizal, 2022), the topics and basic competencies studied include everything written in the Arabic language curriculum in grades 7, and 8, and 9, both odd and even semesters. Therefore, the content validity analysis includes two aspects, namely 1) the proportion of themes/topics appearing per semester in the question items; and 2) behavioral objectives/measured abilities.

### 4.2.1. Proportion of Appearance of Semester Themes/Topics in Question Items

Table 8 shows the findings regarding the distribution of reading topics/themes in the Madrasah Tsanawiyah Arabic language curriculum. The question items in the Madrasah Examination for Arabic subjects are each examined in terms of the source of the theme they represent. Of the 20 reading topics/themes available in the curriculum, starting from grades 7, 8, to 9, it turns out that all of them are represented in the question items except for two topics/themes.

The distribution of questions for each topic is different. In class 7, there are 5 topics represented by 10 question items (16.7% of the total items). In class 8, there are 6 topics represented by 20 question items (33.3%). In class 9 there are 8 topics/themes represented by 30 question items (50%). The proportional percentage of the number of question items is based on the priority level of topics to be mastered by test takers, which describes the competencies they have (Ruang Kepala Sekolah, 2020). The percentage of 50% of questions in class 9, 20% in class 8, and 10% in class 7, it is assumed that the topics in class 7 have been mastered and have been tested in the first year at the Tsanawiyah madrasa through two summative tests in odd and even semesters. Likewise, topics/themes in eighth grade have also been mastered and tested in the second year through two summative tests in odd and even semesters. Meanwhile, in class 9 the final year of their studies, topics/themes are prioritized to be mastered and have not been fully tested, especially topics in the even semester as a continuation of the topics in the previous odd semester.

However, two topics are not represented by this Islamic Secondary School Examination (SSE), namely one topic in grade 7 and one topic in grade 8. This SSE implies that this is not one hundred percent representative of all topics. To strengthen validation in terms of the content of the materials in the curriculum, there shouldn't be a single topic that is not represented in the question items; so that the comprehensive competencies mastered by students at the end of their studies are known, especially mastery of Arabic language competencies (). However, the test creator has the right to determine the priority of which topics are urgent to be mastered, so he will choose which topics will be tested and which topics will not be tested.

Overall, the content validity of the Islamic Secondary School Examination (SSE) test questions is quite high because the question items are in line with the curriculum content in terms of the topics taught.

**Table 8.** The Proportion of themes/topics appearing per semester in question items

Grade	Topic/Theme	Question Items (Q)	Total	%
7	1) تعارف / <i>ta'a&gt;ruf</i> / introduction	Q1, Q2	2	3.33
	2) المرافق المدرسية / <i>al-mara&gt;fiq al-madrasi&gt;ah</i> / School facilities	Q3, Q4, Q14, Q5, Q17	5	8.34
	3) بيئي / <i>baiti&gt;</i> / my house	Q7, Q8	2	3.33
	4) mixed topics	Q16	1	1.67
	5) Topic not involved: الألوان / <i>al-alwa&gt;n</i> / colores	0	0	0
<b>Total</b>	<b>4 Topics</b>	<b>10</b>	<b>10</b>	<b>6.67</b>
8	1) العنوان / <i>al-'unwa&gt;n</i> / address	Q5, Q6	2	3.33
	2) يوميات الأسرة / <i>Yaumi&gt;a&gt;t al-usrah</i> / Family diary	Q9, Q10, Q13, Q18	4	6.67
	3) الساعة / <i>al-sa&gt;'ah</i> / the hour	Q11, Q12	2	3.33
	4) اللاعبين والرياضيون / <i>al-la&gt;'ibu&gt;n wa al-riyad'iyyu&gt;n</i> / Players and athletes	Q19, Q20, Q21, Q22	4	6.67
	5) المهنة الطبية / <i>al-mihni&gt;ah al-tibbi&gt;ah</i> / Medical professional	Q23, Q24	2	3.33
	6) العيادة / <i>al'iya&gt;d'ah</i> / the clinic	Q25, Q26, Q27, Q28, Q29, Q30	6	10
	7) Topics not involved: الهواية / <i>al-hiwa&gt;yah</i> / hobby	0	0	0
<b>Total</b>	<b>6 topics</b>	<b>20</b>	<b>20</b>	<b>33.33</b>
9	1) السنة الهجرية / <i>ra'su al-sanah al-hijri&gt;ah</i> / Islamic New Year	Q31, Q32, Q33, Q34, Q35, Q36, Q37	7	
	2) mixed topics	Q38	1	1.66
	3) الحفل بذكرى مولد الرسول صلى الله عليه وسلم / <i>al-haftu bi z'ikra&gt; maulid al-rasu&gt;l sallahu 'alihu wa sallamal</i> The ceremony commemorates the birth of the Prophet, may God bless him and grant him peace	Q39, Q40	2	3.33
	4) نزول القرآن والعيدين / <i>nuzu&gt;l al-Qur'a&gt;n wa al'ida&gt;n</i> / The revelation of the Qur'an and the celebration of two holidays	Q41, Q42, Q43, Q44, Q45, Q46, Q47, Q48	8	13.33
	5) جمال الطبيعة / <i>Jama&gt;l al-tabi&gt;'ah</i> / beauty of nature	Q49, Q50, Q51	3	5
	6) من آيات قدرة الله / <i>min a&gt;ya&gt;ti qudratilla&gt;h</i> / One of the signs of God's power	Q52	1	1.66
	7) خالق العالم / <i>Kha&gt;liq al-'a&gt;lam</i> / Creator of the world	Q53, Q54, Q55, Q56	4	6.67
	8) الحفاظ على البيئة / <i>al-hifa&gt;z'u 'ala al-bi&gt;'ah</i> / Preserving the environment	Q57, Q58, Q59, Q60	4	6.67
<b>Total</b>	<b>8 topics</b>	<b>30</b>	<b>30</b>	<b>50</b>



#### 4.2.2. Behavioral Objectives/Measurable Abilities

**Table 9.** Distribution of behavior/abilities measured

No.	Measured Behavior/Abilities	Distribution of Behavior on Question Items			Total
		7th semester	8th semester	9th semester	
1	Understanding the social function, text structure, and linguistic elements (sounds, words, meaning, and grammar) of simple texts.	Q1, Q14, Q15	Q10, Q13, Q23, Q24	Q31, Q32, Q33, Q35, Q36, Q37, Q38, Q39, Q40, Q43, Q44, Q45, Q47, Q49, Q51, Q57, Q60	24 (40%)
2	Analyzing ideas from simple texts	0	Q5, Q20, Q28	Q41, Q46, Q48, Q55, Q56	8 (13.3%)
3	Applying rules about grammatical form, meaning, and function	Q7, Q8, Q16, Q17	Q9, Q18, Q19, Q21, Q22, Q25, Q26, Q27, Q29, Q30, Q34	Q42, Q52, Q53, Q54, Q58, Q59	21 (35%)
4	Demonstrating speech acts	Q2, Q3, Q4	Q6, Q11, Q12	Q50	8 (13.3%)
5	Presenting the results of idea analysis from simple narrative text	0	0	0	0 (0%)
	Total	10 (16,7%)	21 (35%)	29 (48,3)	60 (100%)

Table 9 shows the findings regarding behavioral objectives/Arabic language action skills as measured through tests in the Islamic Secondary School Examination for Arabic language subjects at the Tsanawiyah Madrasah in Cimahi City. Based on the Arabic language curriculum of the Islamic Secondary School, it turns out that there are five Arabic language behaviors/actions taught at this school, namely 1) understanding the social function, text structure, and linguistic elements (sounds, words, meaning, and grammar) of simple texts; 2) analyzing ideas from simple texts; 3) applying rules regarding grammatical form, meaning and function in language; 4) demonstrating speech acts; and 5) presenting the results of idea analysis from simple narrative texts.

The question items measured language actions in the form of the ability to understand the social function, text structure, and linguistic elements (sounds, words, meaning, and grammar) of simple texts. This shows that questions related to understanding text structure and linguistic elements (sounds, words, meaning, and grammar) get the highest proportion. The application of grammar is in second place with a proportion of 35% of question items. Meanwhile, two other language actions in the form of analyzing ideas from a text and demonstrating speech act on each topic of the 20 topics in the curriculum, each receiving a proportion of 13%. This shows that analytical skills and language applications have not received sufficient proportion to be tested. These two language actions are actions that must be empowered at the basic education level at the Islamic Secondary School level because the students have great potential to start practical language habits and be analytical towards available texts, so they need to be trained and tested more through formative tests and summative tests such as the Islamic Secondary School Examination.

### 4.3. Construct Validity

This construct validity includes logical validity; because it is related to the reasoning used to check the validity of the questions. Because, construct validity is related to the construction of the human mind, which psychologists have mapped into several parts, as explained above, which consist of memory, understanding, application, analysis, evaluation, synthesis, imagination, and creation (Utari, Madya & Pusediklat, 2011).

#### 4.3.1. Memory

Of the 60 questions studied, it turned out that there were 18 questions whose answers were related to memory (see sample in Table 10), namely Q1, Q3, Q4, Q6, Q7, Q19, Q29, Q30, Q32, Q33, Q35, Q37, Q49, Q53, Q54, Q57, Q58, Q59. For example, in Q1, the question format is:

*Ha>z|a> akhi>, ismuhu> s|a>lihun. Tarjamatu allati> tahtaha> khat|t|un hiya.../* “This is my brother, his name is Saleh”. The translation of the underlined word is...

The command in this question is only to remember the translation of the words that have been taught. For example, the Q3 question only requires students to know Arabic terms about school infrastructure for which the vocabulary has been taught. Likewise, the other 16 question items which are included in the knowledge category (C1).

**Table 10.** Example of memory questions

No.	No. Question Items (Q)	Question Text	Answer
1	1	هَذَا أَخِي، اسْمُهُ صَالِحٌ. تَرْجِمَةَ الَّتِي تَحْتَهَا خَطَّ هِيَ ... <i>Ha&gt;z a&gt; akhi&gt;, ismuhu&gt; s a&gt;lihun. Tarjamatu allati&gt; tahtaha&gt; khat t un hiya.../</i> “This is my brother, his name is Saleh”. The translation of the underlined word is...	(ج) myboy brother
2	3	المَرَافِقُ الْمَدْرَسِيَّةُ فِي مَدْرَسَتِي كَثِيرَةٌ، مِنْهَا ... <i>Al-mara&gt;fiq al-madrasa&gt;ah fi madrasati&gt; kas i&gt;rah, minha&gt; ...</i> The school facilities in my school are many, including...	معمل (ج) <i>/ma'mal/laboratorium</i>
...	...	...	...
18	59	اسْمُ التَّفْضِيلِ الْمُنَاسِبِ هُوَ ... <i>Ismu al-tafd i&gt;l al-muna&gt;sib huwa.../</i> The appropriate superlative is...	(ج) أَحْسَنُ الْأَعْمَالِ © <i>ahsan al-a'ma&gt;l/</i> The best deeds

4.3.2. Understanding

Regarding the question items belonging to the comprehension category (C2), it was found that there were 10 questions, namely Q5, Q8, Q23, Q34, Q36, Q39, Q42, Q43, Q45, and Q60. This is because, in the ten questions (see sample in Table 11), the form of the command indicates the aspect of understanding, which operationally refers to the activities of exemplifying, classifying, concluding, comparing, and explaining (Budiwan, Hidayatullah, Yusuf, & Asrowi, 2020). For example, the Q5 question requires students to read the previous text, summarize the results of their reading, then they can choose the right answer.

**Table 11.** Sample comprehension questions

No.	No. Question Items (Q)	Question Text	Answer
1	5	، مَا عُنْوَانُ مَدْرَسَةِ عَلِيِّ؟ مِنْ النَّصْنِ السَّابِقِ <i>Min al-nas s  al-sa&gt;biq, ma&gt; 'unwa&gt;nu madrasati 'Ali&gt;?</i> From the previous text, what is the address of Ali’s school?	ب- فِي شَارِعِ مُحَمَّدٍ نُوحٍ رَقْمَ 7 كَالِمَانْتَانَ (b) <i>Fi sya&gt;ri' Muhammad Nu&gt;h Raqm 7 Kalimantan/</i> On Mohamed Nuh Street No. 7, Kalimantan
2	8	عَيَّنْ كَمَا فِي الْمِثَالِ! الْمَائِدَةُ فِي غُرْفَةِ الْأَكْلِ- فِي غُرْفَةِ الْأَكْلِ مَائِدَةٌ. الْكِتَابُ وَالْقَلَمُ فِي غُرْفَةِ الْمَذَاكِرَةِ ... <i>Gayyir kama&gt; fi al-mis a&gt;l!</i> <i>Al-ma&gt;'idah fi gurfah al-akl – Fi gurfah al-akl ma&gt;'idah.</i> <i>Al-akitab wa al-qalam fi gurfah al-muz a&gt;karah- Fi gurfah al-muz a&gt;karah kita&gt;b wa qalam.</i>	
...	...	...	...
10	60	تَلَوْتُ الْبَيْئَةَ الَّتِي قُصِدَتْ مِنْ هَذَا النَّصْنِ هُوَ ... <i>Talawwus  al-bi&gt;'ah allati&gt; qusidat min haz a al-nas s  huwa.../</i> Pollution of the environment that is intended by this text is...	أ- تَلَوْتُ الْمَاءَ A- Water pollution

4.3.3. Application

About the question items in the application category (C3), of the 60 questions studied, it turned out that there were 9 questions (Q), namely Q2, Q10, Q12, Q13, Q17, Q18, Q22, Q25, Q50 (see sample in Table 12). The nine questions are classified as applications because the forms of the questions require students to use teaching materials or apply the materials in new situations, whose action indicators are applying, developing, organizing, restructuring, interpreting, and illustrating (Dewi, Suparwa, & Putra, 2020). For example, in Q2, the question format is a dialogue between Naufal and 'Azam:

Naufal: .... Anta ya 'Azza>m? (... are you from, 'Azzam?)  
'Azam: Ana> min Bandung, Jawa al-Garbi>ah. (I am from Bandung, West Java.)

*Naufal*: .... *anta talibun* , *ya 'Azza>m?* (... you a student, 'Azzam?)

'*Azzam*: *Na'am, ana> t}a>libun fi al-madrasah al-s|a>nawi>ah* "al-Muslimin" Bandung. (I am a student of Senior High School "Muslimin" of Bandung.)

The answer choices are as follows: (a) *hal-aina* /are you...where, (b) *ma> -hal/*what-are you, (c) *hal-min aina/* are you-where... from, (d) *min anina-hal/* where... from- are you.

This question requires students to apply the correct question words in the context of the sentence. With the form of fill-in-the-blank questions that must be answered, test takers are required to apply question words according to sentences in which there are no question words.

**Table 12.** Sample application questions

No.	No. Question Items (Q)	Question Text	Answer
1	2	...	...
2	10	أبي ... القهوة كل صباح في غرفة الجلوس، وأمِّي ... الرُّزُّ في المطبخ . <i>Abi&gt;... al-qawah kulla s}aba&gt;h fi gurfah al-julu&gt;s, wa ummi&gt; ... al-ruzza fi al-mat}bakh.</i> My father... coffee in the living room every morning, and my mother... rice in the kitchen.	أيشرب - تطبخ A- drinks - cooks
...	...	...	...
9	50	يا خالد، ... ! تلك المزارع الجميلة. <i>Ya, Kha&gt;lid ...! Tilka al-maza&gt;ri' al-jami&gt;lah.</i> Khaled, ... ! Those beautiful farms	د- أنظر d- <i>Unz}ur/</i> Look at

#### 4.3.4. Analysis

Regarding the question items in the analysis category (C4), 14 items were found, namely Q9, Q11, Q14, Q15, Q16, Q20, Q21, Q24, Q26, Q27, Q28, Q52, Q55, and Q56 (see sample questions in Table 13). The fourteen question items are categorized as analysis questions; because these questions require the test taker to be able to analyze or break down an object or learning material into smaller or more detailed parts and find the relationship between one part and another (Magdalena, et.al, 2020). For example Q9, the form of the question is: *Nahnu ... fi al-madrasah/*

The answer options are:

(a) *Nagsilu al-mala>bis/* We wash clothes; (b) *Najlisu 'ala al-kursi/* We sit on the chair; (c) *Naktubu al-dars/* We write the lesson, (d) *naqra 'u al-jari>dah/* We read the newspaper.

This question requires the test taker to analyze the meaning of words or phrases, then look at the relationship between these meanings, and choose the correct answer in terms of meaning after the analysis activity.

**Table 13.** Sample analysis questions

No.	No. Question Items (Q)	Question Text	Answer
1	9	نحن ... في المدرسة . <i>Nahnu ... fi al-madrasah./</i> We ... at school.	ج- نكتب الدرس c. <i>Naktubu al-dars.</i> We write the lesson.
2	14	ماذا يعمل هشام في وقت الراحة؟ <i>Ma&gt;z a&gt; ya'malu Hisya&gt;m fi waqt al-ra&gt;hah?</i> What does Hisham do in his time of rest?	أ- يذهب إلى المكتبة a. <i>Yaz}habu ila&gt; al-maktabah./</i> a- He goes to the library.
...	...	...	...
14	56	الجملة التي فيها "الإسم الموصول الصحيح" هي .. <i>Al-jumlah allati&gt; fiha&gt; al-ism al-mawsu&gt;l al-s}ahi&gt;h hiya...</i> The sentence in which there are "the correct relative noun" is...	ج- خلق الله القمر الذي يظهر في الليل c. <i>Khalaqa Alla&gt;hu al-qamara allazi&gt; yaz}haru fi al-lail./</i> God created the moon that appears at night.

### 4.3.5. Evaluation

Next is about question items that are classified as evaluation (C5). Of the 60 questions, it was found that there were four questions in the evaluation category, namely Q41, Q47, Q48, Q51. These four question items are classified as an evaluation because they require the test taker to be able to decide, assess, compare, measure, and criticize (see example sample in Table 14). For example, Q41, the problem is:

*Mahmu>d: Ya Fari>d! Kaifa istamarra nuzu>l al-Qur'a>n?* Mahmud: Oh Farid! How did the revelation of the Qur'an continue?

*Fari>d: Anna al-Qur'a>n lam yanzil daf'atan wa>hidatan, bal istamarra nuzu>luhu> mufarraqa>n muddata 23 sanah.* Farid: The Qur'an was not revealed all at once, but rather it continued to be revealed in separate parts over a period of 23 years.

*Al-ma'na al-sahih min kalimah 23 hiya...*

The correct meaning of the word "23" is...

The answer choices are:

- (a) *13 sanatan fi al-madi>nah wa 10 sanawa>t fi Makkah./* 13 years in Medina and 10 years in Mecca.
- (b) *10 sanawa>t fi Makkah wa 13 sanatan fi al-madi>nah./* 10 years in Mecca and 13 years in Medina.
- (c) *13 sanawa>t fi Makkah wa 10 sanatan fi al-madi>nah./* 13 years in Mecca and 10 years in Medina.
- (d) *13 sanawa>t fi Makkah wa 13 sanatan fi al-madi>nah./* 13 years in Mecca and 13 years in Medina.

Question 41 requires the test takers to read the previous text, then a number appears that is related to the text, and this number must be criticized to find out what it has to do with the information in the text, then the test takers decides to choose the right answer.

**Table 14.** Sample evaluation questions

No.	No. Question Items (Q)	Question Text	Answer
1	41	المعنى الصحيح من كلمة "23" هي ... <i>Al-ma'na al-sahih min kalimah 23 hiya...</i> The correct meaning of the word "23" is...	(ج) (13) سنة في مكة و (10) سنوات في المدينة (c) <i>13 sanatan fi Makkah wa 13 sanawat fi al-madi&gt;nah./</i> 10 years in Mecca and 13 years in Medina.
2	47	الجملة التي تدل على قواعد "لا الناهية + الفعل المضارع" ... <i>Al-jumlah allati&gt; tadullu 'ala&gt; qawa&gt;'id "la&gt; al-na&gt;hiah + al-fi'l al-mud{a&gt;ri' " al-s}ahi&gt;h hiya ...</i> The correct sentence that indicates the rules of "no final + present tense verb" is ...	يا أحمد ، لا تحفل بنزول القرآن إلا في الليلة السابعة عشرة من رمضان <i>Ya, Ahmad la tahtafil binuzul al-Qur'an illa fi al-lailah al-sabi'ah 'asyrata .</i> O Ahmed, do not celebrate the revelation of the Qur'an except on the seventeenth night of Ramadan.
...	...	...	...
4	51	الفكرة الرئيسية من ذلك النص هي ... <i>Al-fikrah al-rai&gt;si&gt;ah min z{a&gt;lik al-nas}s} hiya...</i> The main idea of that text is...	أ- ذهبنا إلى المزارع . <i>A. Z{ahabna ila&gt; al-maza&gt;ri'.</i> We went to the farms.

### 4.3.6. Synthesis

Next are the question items that are classified as synthesis (C6). Of the 60 questions studied, it appears that there are 5 question items in the synthesis question category, namely Q31, Q38, Q40, Q44, and Q46. Classified as synthesis questions, because these five questions require the test takers to be able to put parts together to form a new or coherent unit, organize, develop, build and organize existing elements into one complete idea (Ruwaida, 2019). For example Q31, the form of the question is:

(1) *Ha>jara al-rasu<l sallalla>hu 'alaihi wa sallam wa as}ha>buh min Makkah ila> al-Madinah/* The Messenger, may God bless him and grant him peace, and his companions migrated from Mecca to Medina.

(2) *Taraku> watanahum wa buyu>tahum wa amwa>lahum wa tija>ratahum fi Makkah./* They left their homeland, their homes, their money, and their trade in Mecca.

(3) *Ha>wala al-kuffa>r an yaqtulu> al-nabi>a sallalla>hu 'alaihi wa sallam./* The unbelievers tried to kill the Prophet.

*Tarti>b hazihi al-jumal litaku>na faqrah mufidah huwa...*

Arranging these sentences to form a useful paragraph is... (arrangement of numbers from left ..... )

The answer options are:

(a) 3-2-2 ; (b) 1-2-3; (c) 2-1-3; (d) 1-3-2

Question 31 requires the test takers to understand the previous parts, then look for relationships between one part and another, and then organize them to form a coherent whole of ideas. Likewise with the other four question items (see sample examples in Table 15).

**Table 15.** Sample synthesis questions

No.	No. Question Items (Q)	Question Text	Answer
1	31	<p>تُرْتَبُّ هَذِهِ الْجُمَلُ لِتَكُونَ قَفْرَةً مُفِيدَةً هُوَ .....  <i>Tarti&gt;b ha&gt;zih al-jumal litaku&gt;na faqrah mufi&gt;dah huwa .../</i>            Arranging these sentences to form a useful paragraph is...            (arrangement of numbers from the left)</p>	(c) 2 - 1 -3
2	38	<p>تُرْتَبُّ هَذِهِ الْجُمَلُ لِتَكُونَ قَفْرَةً مُفِيدَةً هُوَ .....            Arranging these sentences to form a useful paragraph is...            (arrangement of numbers from left ..... )</p>	c. 4-3-1-2
...	...	...	...
5	44	<p>يُصَلُّونَ \yus\allu&gt;n pray            - masjid   the mosque al-المسجد - muslims\al-muslimu&gt;na            in\fi - al-Fitr\al-fit\ - Eid\i&gt;d            ترتب الجملة الصحيحة هو ...  <i>Tarti&gt;b al-jumlah al-s\ahi&gt;hah huwa .../</i>            The correct sentence order is...</p>	المسلمون يُصلُّونَ عيدَ الفطر في المسجد Muslims pray Eid al-Fitr in the mosque.

As for the seventh and eighth cognitions, namely imagination and creation, there is not a single form of question item that represents them out of the 60 question items available on the Islamic Secondary School test for Arabic subjects. This is because the question of imagination requires the test taker to be able to combine various concepts of subject matter into a creative imagination. Likewise, the creation category requires test takers to be able to apply the concepts of learning material to a product. These two categories will be realized if the form of the test is an essay or an activity project, it is difficult to make it in the form of a multiple-choice test like this (Novinda, Silitonga & Hamdani, 2019).

To see the proportion of distribution of question items in terms of psychological construction, you can refer to Table 16.

**Table 16.** Psychological constructs/Bloom's taxonomy measured

No.	Bloom's Taxonomy	Questions	Total	%
1	Memory (C1)	Q1, Q3, Q4, Q6, Q7, Q19, Q29, Q30, Q32, Q33, Q35, Q37, Q49, Q53, Q54, Q57, Q58, Q59	18	30
2	Understanding (C2)	Q5, Q8, Q23, Q34, Q36, Q39, Q42, Q43, Q45, Q60	10	16.7
3	Application (C3)	Q2, Q10, Q12, Q13, Q17, Q18, Q22, Q25, Q50	9	15
4	Analysis (C4)	Q9, Q11, Q14, Q15, Q16, Q20, Q21, Q24, Q26, Q27, Q28, Q52, Q55, Q56	14	23.3
5	Evaluation (C5)	Q41, Q47, Q48, Q51	4	6.7
6	Synthesis (C6)	Q31, Q38, Q40, Q44, Q46	5	8.3
7	Imagination (C7)	0	0	0
8	Creation (C8)	0	0	0
<b>Total</b>	<b>8</b>	<b>60</b>	<b>60</b>	<b>100</b>

## 5. CONCLUSION

The logical validity of a test which is a measuring instrument related to the suitability of each test item and the whole with scientific reasoning and theory. Logical validity is internal because it ensures the accuracy of each item and the whole in terms of surface, content, and construct, before being used in exam activities. The clarity of writing and instructions on each item will make it easier for the test takers to give the right answer. This is a valid test from a surface standpoint. Material content as a topic that must be mastered in the final school exam includes all topics that have been studied by the test takers/students from the first year (grade 7) to the ninth year (grade 9) to describe the comprehensive competence for the test takers to graduate from school. This is a valid test in terms of content. The levels of cognition which include knowledge, understanding, application, analysis, evaluation, synthesis, imagination, and creation, must be represented on the test questions so that the level of competence of students/test takers can be ascertained in terms of their cognition in certain subjects. This is what is called construct validity.

Based on the results of the findings and discussion above, the surface validity of the Madrasah Examination test for the MTs Arabic subject was categorized as lacking; because 64 words were incorrect in terms of grammar in the 33 items tested. This should not happen because it will reduce the clarity of the item for the test takers.

Regarding the logical validity of the content sub-validity, the 60 items tested are valid because they contain questions for 90% of the topics taught in grades 7, 8, and 9, which are spread over 60 items, with a proportion of 35% of the topics in grade 9, and 30% of the topics in grade 8, and 25% of the topics in grade 7. Only 10% of the topics were not represented on the Madrasah exam questions.

The logical validity of the sub-validity of the content relating to the goals and behaviors/actions of the Arabic language being tested, including those that are less proportional among the five language actions being taught, so that the validity is quite low. This is because the test questions emphasize aspects of understanding the text and the grammar of the language in it. Meanwhile, speech acts and analytical and critical abilities towards texts do not get adequate proportions in the question items.

As for the logical validity of the construct sub-validity in the above test, it shows that the level of knowledge still dominates the item questions, followed sequentially by the levels of analysis, understanding, application, synthesis, and finally evaluation. The level of imagination and creation doesn't get any portion at all. This means that the items made do not show a high level of thinking.

## REFERENCES

- Alfathimy, Z. R. N. A., & Ali, M. (2021, December). Tanawwu'Žarfı al-Lugah al-‘Arabiyyah fı al-Kitāb al-Madrasi Li al-Lugah al-‘Arabiyyah bi al-Madrasah al-Tsānawıyyah. In *International Conference on Arabic Language and Literature* (pp. 315-356). <http://proceedings.upi.edu/index.php/ical/article/view/1733>
- Ali, M. (2019). *Bahasa Arab Al-Qur'an Jilid II Frasa* [Arabic Language of the Qur'an Volume II: Phrases]. Subang: Royyan Press.
- Ali, M. (2019). *Bahasa Arab Al-Qur'an Jilid III Kalimat* [Arabic Language of the Qur'an Volume III: Sentences]. Subang: Royyan Press.
- Ali, M. (2022). *Morfosintaksis Bahasa Arab Pembelajaran Integratif Sharaf dan Nahwu* [Morphosyntax of Arabic Language in Integrative Learning of Sharaf and Nahwu]. Bandung: UPI Press.
- Arikunto, S. (2008). *Dasar-Dasar Evaluasi Pendidikan Edisi Revisi* [Foundations of Educational Evaluation Revised Edition]. Bandung: Bumi Aksara.
- Arikunto, S. (2009). *Evaluasi Pembelajaran* [Learning Evaluation]. Jakarta: Rineka Cipta.
- Arikunto, S., & Jabar, C. S. A. (2014). *Evaluasi Program Pendidikan: Pedoman Teoritis Praktisi Pendidikan* [Educational Program Evaluation: Theoretical Guidelines for Educational Practitioners]. Bandung: Bumi Aksara.
- Azwar, S. (1986). *Validitas dan Reliabilitas: Interpretasi dan Komputasi* [Validity and Reliability: Interpretation and Computation]. Yogyakarta: Liberty.
- Budiwan, J., Hidayatullah, F., Yusuf, M., & Asrowi, A. (2020). Validity Test of the Family Supported Collaborative Learning (FSCL) Model to Instill Character Values to Junior High School Students. *Universal Journal of Educational Research*, 8, 1-6. <https://doi.org/10.13189/ujer.2020.081701>

- Dewi, N. K. N., Suparwa, I. N., & Putra, A. A. P. (2020). Penerapan Taksonomi Bloom pada Pembelajaran Bahasa Di Kelas 2 SD Bali Kiddy. *Linguistika: Buletin Ilmiah Program Magister Linguistik Universitas Udayana*, 27(2), 184-195. <https://doi.org/10.24843/ling.2020.v27.i02.p10>
- Fikriyah, N. (2021). Analisis Butir Soal Ulangan Tengah Semester Mata Pelajaran Bahasa Arab Kelas VII Semester Ganjil SMP Muhammadiyah 1 Yogyakarta Tahun Ajaran 2019/2020 [Analysis of Items for the Mid-Semester Exam of Arabic Language Subject Grade VII Odd Semester at SMP Muhammadiyah 1 Yogyakarta Academic Year 2019/2020]. *Maharaat: Jurnal Pendidikan Bahasa Arab*, 3(2), 128-140. <https://doi.org/10.18196/mht.v3i2.10501>
- Grondlund, N. E. (1982). *Constructing achievement test. 3rd. ed.*. Englewood Cliffs, NJ.
- Ilham, M., & Wijati, I. A. (2020). *Keterampilan Berbicara: Pengantar Keterampilan Berbahasa* [Speaking Skills: Introduction to Language Speaking Skills]. Lembaga Academic & Research Institute : Pasuruan.
- Kusaeri, K. (2012). Menggunakan model DINA dalam pengembangan tes diagnostik untuk mendeteksi salah konsepsi [Utilizing the DINA Model in the Development of Diagnostic Tests to Detect Misconceptions]. *Jurnal Penelitian dan Evaluasi Pendidikan*, 16(1), 281-306. <https://doi.org/10.21831/pep.v16i1.1118>
- Magdalena, I., Islami, N. F., Rasid, E. A., & Diasty, N. T. (2020). Tiga ranah taksonomi bloom dalam pendidikan [The three domains of Bloom's taxonomy in education]. *EDISI*, 2(1), 132-139.
- Martini, S., Ariyanti, F., Suyatno, S., & Prasetyowati, I. (2022). Competency test to measure the quality assurance in the teaching-learning process of Indonesian public health higher education institutions. *Academia*, (26), 49-60. <https://doi.org/10.26220/aca.3949>
- Muhammad, I., & Ariani, S. (2021). Kognisi Taksonomi Bloom, Kurikulum 2013 dan Penerapannya dalam Pembelajaran Agama Islam di Indonesia [Cognitive Domain of Bloom's Taxonomy, Curriculum 2013, and Its Application in Islamic Education Learning in Indonesia]. *Jurnal MUDARRISUNA: Media Kajian Pendidikan Agama Islam*, 11(3), 425-443. <http://dx.doi.org/10.22373/jm.v11i3.6303>
- Munthe, A. P. (2015). Pentingnya evaluasi program di institusi pendidikan: sebuah pengantar, pengertian, tujuan dan manfaat [The Importance of Program Evaluation in Educational Institutions: An Introduction, Definition, Objectives, and Benefits]. *Scholaria: Jurnal Pendidikan Dan Kebudayaan*, 5(2), 1-14. <http://dx.doi.org/10.22373/jm.v11i3.6303>
- Nazir, M. (1988). *Metode Penelitian* [Research Methods]. Bogor: Ghalia Indonesia.
- Ni'mah, F. (1973). *Mulakhas Qawa'id al-Lughah al-'Arabiyyah*. Dar al-Hikmah : Beirut.
- Novinda, M. R. R., Silitonga, H. T. M., & Hamdani, H. (2019). Pengembangan Tes Pilihan Ganda Menggunakan Model Rasch Materi Gerak Lurus Kelas X Pontianak [Development of Multiple-Choice Tests Using the Rasch Model for Straight Motion Material in Grade X Pontianak]. *Jurnal Pendidikan dan Pembelajaran Khatulistiwa*, 8(6). <http://dx.doi.org/10.26418/jppk.v8i6.33452>
- Nur'aini, E. (2011). Kata Operasional Taksonomi Bloom Versi Baru Untuk Mata Pelajaran Biologi [Operational Words of the New Bloom's Taxonomy for Biology Subjects]. *PEP UNY*, 6.
- Ridho, U. (2018). Evaluasi dalam pembelajaran bahasa Arab [Evaluation in Arabic Language Learning]. *An Nabighoh*, 20(1), 19-26. <https://doi.org/10.32332/an-nabighoh.v20i01.1124>
- Rizal, M. S. (2022). Implementasi Surat Edaran Menteri Pendidikan dan Kebudayaan No. 1 Tahun 2021 (Peniadaan Ujian Nasional) di Madrasah Aliyah Al-Mukhlishin Antibar Kecamatan Mempawah Timur Kabupaten Mempawah [Implementation of Circular of the Minister of Education and Culture No. 1 Year 2021 (Elimination of National Examinations) at Al-Mukhlishin Antibar Islamic High School, East Mempawah Subdistrict, Mempawah Regency]. *Al-Astar*, 1(1), 43-59.
- Rochman, R. N., Abdurrahman, M., & Ali, M. (2021, December). Content Validity Analysis of Arabic School Examination at SMP IT Permata Hati Banjarnegara. In *International Conference on Arabic Language and Literature* (pp. 126-133). Bandung, Indonesia: Universitas Pendidikan Indonesia
- Ruang Kepala Sekolah. (2020). *Cara Menyusun Butir Soal*. Retrieved from <https://www.ruangkepalsekolah.com/2020/02/cara-menyusun-butir-soal-ujian-sekolah.html>

- Rusli, R. K., & Mulyana, M. A. (2022). Komparasi Hasil Belajar Siswa Menggunakan Media Zoom Cloud Meeting Dengan Video Pembelajaran Bahasa Arab [Comparison of Student Learning Outcomes Using Zoom Cloud Meeting Media and Arabic Language Instructional Videos]. *Tatsqifiy: Jurnal Pendidikan Bahasa Arab*, 3(1), 21-36. <https://doi.org/10.30997/tjpba.v3i1.4853>
- Ruwaida, H. (2019). Proses kognitif dalam taksonomi bloom revisi: analisis kemampuan mencipta (C6) pada pembelajaran fikih di mi miftahul anwar desa banua lawas [Cognitive Processes in Bloom's Revised Taxonomy: Analysis of Creating Ability (C6) in Fiqh Learning at MI Miftahul Anwar, Banua Lawa Village]. *Al-Madrasah: Jurnal Pendidikan Madrasah Ibtidaiyah*, 4(1), 51-76. <http://dx.doi.org/10.35931/am.v4i1.168>
- Soendari, T. (2012). *Metode Penelitian Deskriptif [Descriptive Research Method]*. Bandung: UPI Press.
- Soysal, T., & Kou, S. (2018). A simple test for assessing solidification cracking susceptibility and checking validity of susceptibility prediction. *Acta Materialia*, 143, 181-197. <https://doi.org/10.1016/j.actamat.2017.09.065>
- Suryanto, E., & Akhadiyah, S. (2019, December). Developing a competency-based of language skills and literary abilities in the primary school. In *Journal of Physics: Conference Series* (Vol. 1339, No. 1, p. 012066). IOP Publishing.
- Syihabuddin. (2009). *Evaluasi Pengajaran Bahasa Indonesia [Evaluation of Indonesian Language Teaching]*. Bandung: Sekolah Pascasarjana Universitas Pendidikan Indonesia.
- Teufl, W., Miezal, M., Taetz, B., Fröhlich, M., & Bleser, G. (2018). The validity, test-retest reliability, and long-term stability of magnetometer-free inertial sensor-based 3D joint kinematics. *Sensors*, 18(7), 1980. <https://doi.org/10.3390/s18071980>
- Umar, F. A. (1993). *Al-Wasith fi Qowaid al-Imla wa al-Insya*. Riyadh: Maktabah al-Ma'arif.
- Utari, R., Madya, W., & Pusdiklat, K. N. P. K. (2011). Taksonomi Bloom. *Jurnal: Pusdiklat KNPk*, 766(1), 1-7.
- Wang, C. (2014). Communicative validity of the new CET-4 listening comprehension test in China. *Indonesian Journal of Applied Linguistics*, 4(1), 111-122. <https://doi.org/10.17509/ijal.v4i1.608>

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

