BioEdu
Berkala Ilmiah Pendidikan Biologi
Vol. 12 No. 1 Tahun 2023 Hal: 177-187
https://ejournal.unesa.ac.id/index.php/bioedu

THE DEVELOPMENT OF INTERACTIVE QUIZ BASED ON QUIZIZZ OF RESPIRATORY SYSTEM MATERIALS TO TRAIN HOTS OF GRADE XI OF SENIOR HIGH SCHOOL

Pengembangan Kuis Interaktif Berbasis Quizizz Materi Sistem Pernapasan Untuk Melatihkan HOTS Siswa Kelas XI

Dyah Ayu Ning Tyas
Biology Education, Faculty of Mathematics and Science, State University of Surabaya
e-mail: dyah.18078@mhs.unesa.ac.id

Nur Qomariyah
Biology, Faculty of Mathematics and Science, State University of Surabaya
e-mail: nurqomariyah@unesa.ac.id

Abstract

Higher-order thinking skills (HOTS) are one of the skills that demanded in the 21st century. However, there are still few students who master this skill. The use of technology should be optimized even more in distance learning to support the learning process to achieve maximum results. Therefore, an effort is needed to train students' HOTS. One way to train HOTS is by giving quizzes in the evaluation process by utilizing the Quizizz application. This research aimed to create an interactive quiz based on Quizizz to train HOTS on respiratory system materials that are valid and reliable. This development research referred to the ASSURE model (Analyze Learner, State Objective, Select Media and Materials, Utilize Media and Materials, Require Learner Participation, Evaluate and Revise). The limited trial was conducted on 18 students of class XI MIA 2 SMA Muhammadiyah 10 GKB Gresik. The validity of the interactive quiz was determined based on the results of the validation by material expert and media expert using the validation sheet. The results showed that the validity of the interactive quiz reached 99.73% with a very valid category. Reliability and difficulty level are determined based on the student test results. The reliability of the items scored 0.713 with a reliable category. The proportion of the item difficulty level is 80% with moderate category and 20% with difficult category. Accordingly, the interactive quiz on respiratory system materials for practicing HOTS is valid and reliable.
Keywords: Interactive quiz, HOTS, Quizizz, Respiratory system

Abstrak

Keterampilan berpikir tingkat tinggi (HOTS) adalah salah satu keterampilan yang menjadi tuntutan pada abad ke-21. Namun, pada kenyataannya masih sedikit siswa yang menguasai keterampilan ini. Pemanfaatan teknologi harus lebih dioptimalkan dalam pembelajaran jarak jauh untuk menunjang proses pembelajaran mencapai hasil yang maksimal. Oleh karena itu, diperlukan cara untuk melatihkan HOTS pada siswa. Salah satu cara melatihkan HOTS adalah melalui pemberian kuis dalam proses evaluasi dengan memanfaatkan aplikasi Quizizz. Tujuan dari penelitian ini adalah menciptakan kuis interaktif berbasis Quizizz untuk melatihkan HOTS materi sistem pernapasan yang valid dan reliabel. Penelitian pengembangan ini menggunakan model ASSURE (Analyze Learner, State Objective, Select Media and Materials, Utilize Media and Materials, Require Learner Participation, Evaluate and Revise). Uji coba terbatas dilakukan pada 18 siswa kelas XI MIA 2 SMA Muhammadiyah 10 GKB Gresik. Validitas kuis interaktif didapatkan dari hasil validasi oleh dosen ahli materi dan ahli media menggunakan lembar validasi. Hasil penelitian menunjukkan bahwa validitas kuis interaktif mencapai nilai 99.73% dengan kategori sangat valid. Reliabilitas dan tingkat kesukaran ditentukan berdasarkan hasil tes siswa. Reliabilitas butir soal mendapat nilai 0.713 dengan kategori reliabel. Proporsi tingkat kesulitan butir soal adalah 80% dengan kategori sedang dan 20% dengan kategori sukar. Berdasarkan hasil penelitian, dapat disimpulkan bahwa kuis interaktif materi sistem pernapasan untuk melatihkan HOTS valid dan reliabel.
Kata Kunci: *Kuis interaktif, HOTS, Quizizz, Sistem pernapasan*

BioEdu
Berkala Ilmiah Pendidikan Biologi
Vol. 12 No. 1 Tahun 2023 Hal: 177-187
https://ejournal.unesa.ac.id/index.php/bioedu

INTRODUCTION

Currently, Indonesia and the world are being hit by the Covid-19 pandemic. The Covid-19 affected various fields, especially education in Indonesia. Based on a circular issued by the Ministry of Education and Culture (Kemendikbud) Number 15 of 2020 regarding guidelines for organizing learning from home in the emergency period spreading of Covid-19, all schools in Indonesia carry out learning activities from home (online). The Covid-19 pandemic has urged the implementation of distance learning, which has rarely been implemented simultaneously for all levels of education.

Distance learning is learning that involves the use of the internet and is related to anything that is delivered, activated, or mediated by electronic technology for explicit learning purposes (Waruwu, 2020). The distance learning process is carried out by using pedagogical tools or educational aids such as learning media that allow internet access and good information technology of materials of the learning process and knowledge in the interactions carried out (Pakpahan and Fitriani, 2020). Success in a learning process depends highly on the evaluation process. By evaluating, a teacher can understand the level of students' understanding of the material being taught and can also be used as feedback for teachers to evaluate the teaching strategies applied to improve the quality of learning.

The implementation of the evaluation can be done in various ways but generally using tests. Learning evaluation can be implemented in various forms. One of them is in the form of giving interactive quizzes. An interactive quiz is a form of test that contains questions that students should answer, where each question has a value as an indicator that can be used for student assessment. Giving interactive quizzes is carried out to transfer knowledge from teachers to students and to setting up or conditioning the environment that can stimulate students to improve their learning activities (Indrawati and Hartati, 2021).

The interactive quizzes can be an innovation for teachers to carry out evaluations. Teachers can take advantage of interactive quiz platforms or applications available on the internet that are free of charge, such as Quizizz. Quizizz can use as an evaluation media to make learning more interesting and fun. Because in the Quizizz there are features that can make students not feel bored (Salsabila, et al., 2020).

The development of interactive quizzes as assessment forms is expected as an effort to develop a variety of skills and proficiency that must be possessed in the 21st century. Redhana (2019) revealed that the 21st century as the century of knowledge, the century of knowledge-based economy, the century of information technology, globalization, and the industrial revolution 4.0. One of the skills demanded in the 21st century is higher-order thinking skills (HOTS).

Higher-order thinking skills (HOTS) defined as the widespread use of the mind to discover new challenges. Higher-order thinking skills (HOTS) require a person to apply newly information or prior knowledge and manipulate the information to reach possible answers in new situations (Sucipto, 2017). HOTS is intended for students to learning beyond understanding to encourage application, analysis, synthesis, and evaluation activities in processing information.

The thinking skills of students in Indonesia, especially higher-order thinking skills, are still low. According to Zuriyani research in 2012, level of ability in problem-solving, 41.3% of Indonesian students are at the C-1 level; 27.5% are at the C-2 level; 9.5% are at the C-3 level, and 1.4% is at the C-4 level. There are still few Indonesian students who are at the C5-C6 level. One of the causes of Indonesia's low achievement at the international level is the low cognitive level of students at the HOTS level. Students not used to complete tests or problems related to science process skills (Ramadhan and Wasis, 2013).

Biology learning is one example of learning related to science process skills. This learning can teach students to construct knowledge independently. The subject matter of the respiratory system contained in biology learning is a matter that can teach students' higher-order thinking skills because this matter takes place with real experiences felt by students. However, in reality, there are many students whose higher-order thinking skills are low because they consider that in the respiratory system matter. There are foreign terms that are difficult to understand by students, forcing students to memorize, which causes students to be confused in understanding concepts.

The subject matter of the respiratory system in humans is considered difficult by students. It can be observed from the data analysis of student learning outcomes after the HOTS training test with an average result of 25.07 (Afrita and Darussyamsu, 2020). It means that students' higher-order thinking skills are included in the low category. Higher-order thinking skills can improve if students often trained by being given tests with cognitive levels of C4-C6, so the researchers develop quizzes based on Quizizz as a form of implementation to train students' HOTS in the evaluation process.

This research aimed to create an interactive quiz based on Quizizz on respiratory system materials to train HOTS for class XI students that are valid and reliable and describe

the validity, reliability, and difficulty level of the items developed.

METHODS

This development research referred to the ASSURE model consisting of Analyze Learner, State Objective, Select Media and Materials, Utilize Media and Materials, Require Learner Participation, Evaluate and Revise. This research was conducted from January to April 2022. The product was tested on a limited basis on 18 students of class XI MIA 2 SMA Muhammadiyah 10 GKB Gresik. The end product of this research is an interactive quiz based on Quizizz on respiratory system materials to train students' HOTS.

The Analyze Learner stage (analyzing student characteristics) was carried out to analyze students' learning methods and forms of quizzes used daily by teachers. The State Objective stage (determining learning objectives) was implemented to analyze the curriculum and formulate indicators of questions and learning objectives. The Select Media and Materials stage (selecting media and teaching materials) was carried out to create HOTS items on interactive quizzes. Items are presented in the form of multiple-choice and essays. Items were developed using cognitive levels C-4 and C-5. The aspects of HOTS measured in the research are critical thinking, problem-solving, and decision-making. Preparation of the test items were done by detailing basic competencies into several indicators. The specifications for the HOTS items of respiratory system topic are presented in Table 1.

Basic Competencies	Indicator of Achieving Competencies	Indicators of HOTS	Aspect of HOTS	Cognitive Level	Number	Instrument Form
3.8 Analyzing the relationship between the structure of the tissue in organs of the respiratory system in relation to bioprocess and functional disorders that can occur in human respiratory system.	Analyzing human respiratory mechanisms	Identifying, processing, analyzing, and connecting information to concepts, theories, and opinions.	Critical Thinking	C-4	1, 2	Multiple-Choice
	Associating the relationship of the respiratory frequency with a pulse	Identifying, processing, analyzing, and connecting information to concepts, theories, and opinions.	Critical Thinking	C-4	3	Multiple-Choice
	Analyzing factors that affect the respiratory frequency	Determining decision based on information or problem given.	Decision-making	C-4	4, 8	Multiple-Choice
		Identifying, processing, analyzing, and connecting information to concepts, theories, and opinions.	Critical Thinking		10	
	Associating the influence of factors causing disorders or	Determining decision based on information or problem given.	Decision-making	C-4	5	Multiple-Choice

abnormalities of function in the human respiratory system	Identifying, processing, analyzing, and connecting information to concepts, theories, and opinions.	Critical Thinking		6, 9, 11, 12	
Associating a factor with the volume and capacity of human lungs	Identifying, processing, analyzing, and connecting information to concepts, theories, and opinions.	Critical Thinking	C-4	7	Multiple-Choice
Formulating a solution to prevent disorders of the human respiratory system	Solving problems based on data or information given.	Problem-Solving	C-5	13, 14	Essay
Associating environmental conditions with the respiratory system	Identifying, processing, analyzing, and connecting information to concepts, theories, and opinions.	Critical Thinking	C-4	15	Essay

The Utilize Media and Materials stage (using media and teaching materials) was carried out to validate interactive quizzes based on the assessment of media expert and material expert. The Require Learner Participation stage (developing student participation) was carried out to collect data in the form of students' interactive quizzes result. The Evaluate and Revise stage (evaluating and revising) is carried out to evaluate and revise the developed interactive quiz.

Data collection used validation and test method. Validation was carried out with two validators consisting of media expert and material expert. The test method is carried out by students through interactive quizzes that have been validated by validators.

Validation was carried out to obtain validity values through the results of validation by validators against interactive quizzes developed using the Guttman scale on the research instrument. The validation results were calculated using the following formula:

$$P = \frac{f}{N} \times 100\%$$

Notes:

P : percentage of validity (%)

f : total of test items marked on each aspect (√)

N : total of all test items

The percentage of validity were interpreted based on the criteria as shown in Table 2.

Table 2. Interpretation of Validity

Score (%)	Category
75,1 - 100	Very Valid
50,1 - 75	Valid
25,1 - 50	Less Valid
0 – 25	Not Valid

Interactive quizzes are declared valid if the obtained score were within the average of ≥ 50% (Riduwan, 2012).

Reliability calculations are used to determine whether the items in the developed interactive quizzes are reliable or not. Reliability calculations use the Cronbach Alpha formula using SPSS software. The calculation result was interpreted based on the criteria as shown in Table 3. Then the test items in the interactive quiz are declared reliable if the obtained score were within of $0.60 < r_{11} \leq 0.80$.

BioEdu
Berkala Ilmiah Pendidikan Biologi
Vol. 12 No. 1 Tahun 2023 Hal: 177-187
https://ejournal.unesa.ac.id/index.php/bioedu

Table 3. Interpretation of Reliability

Interval of Reliability	Interpretation of Reliability
$0.80 < r_{11} \leq 1.00$	Very Reliable
$0.60 < r_{11} \leq 0.80$	Reliable
$0.40 < r_{11} \leq 0.60$	Reliable Enough
$0.20 < r_{11} \leq 0.40$	Less Reliable
$0.00 < r_{11} \leq 0.20$	Not Reliable

Source: Sugiyono (2015)

The difficulty level of each test item was determined using the following formula. And the calculation result was interpreted based on criteria as shown in Table 4.

$$P = \frac{Np}{N}$$

Notes:

P : level of difficulty

Np : number of students who answered correctly

N : total number of the students

Table 4. Interpretation of Difficulty Level

Interval of Difficulty Level	Interpretation of Difficulty Level
$0.00 < P \leq 0.30$	Difficult
$0.30 < P \leq 0.70$	Moderate
$0.70 < P \leq 1.00$	Easy

Source: (Arikunto, 2015)

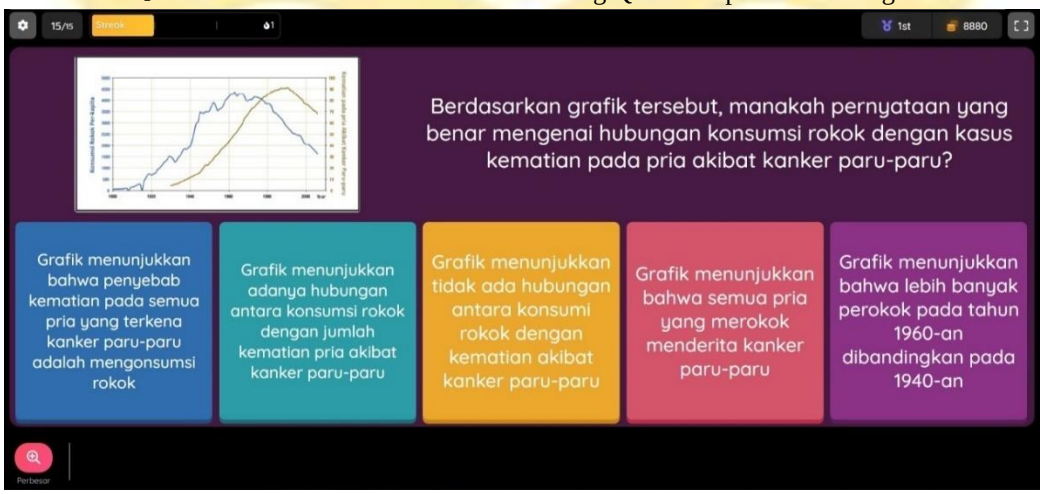
RESULTS AND DISCUSSION

The research aimed to create an interactive quiz based on Quizizz on respiratory system materials to train HOTS for class XI students that are valid and reliable and describe the validity, reliability, and levels of difficulty of the items developed. This development research referred to the ASSURE model (Analyze Learner, State Objective, Select Media and Materials, Utilize Media and Materials, Require Learner Participation, Evaluate and Revise).

Analyzing Learner was the first stage of the development of the HOTS interactive quiz. At this stage, an analysis of students' learning and student habits when doing quizzes was carried out as well as to find out the form of quizzes that are used daily by the teachers.

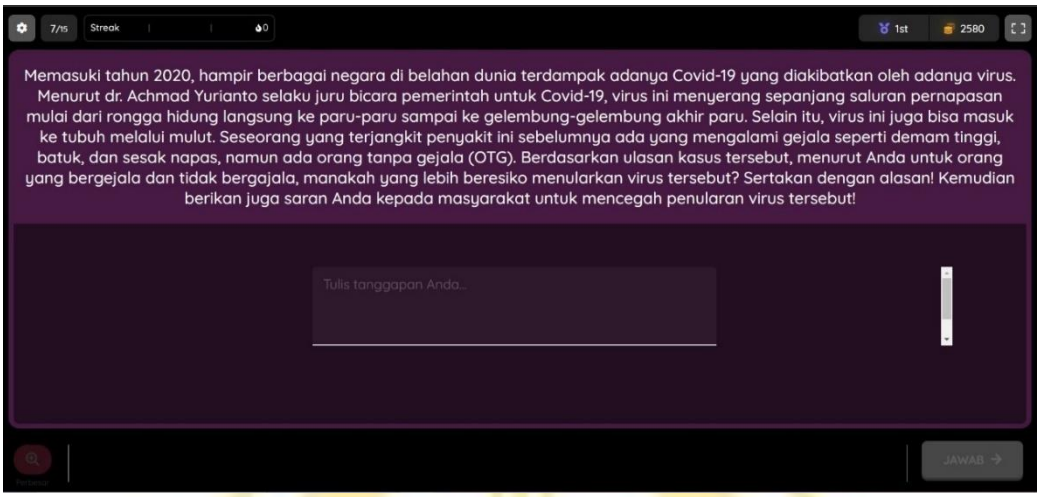
The second stage was to determine the learning objectives. At this stage, curriculum analysis is carried out, formulating indicators of items and learning objectives as specific as possible. There are 7 indicators of competency achievement made in the development of items, namely: 1.) analyzing human respiratory mechanisms; 2.) associating the relationship of the respiratory frequency with a pulse; 3.) analyzing factors that affect the respiratory frequency; 4.) associating the influence of factors causing disorders or abnormalities of function in the human respiratory system; 5.) associating a factor with the volume and capacity of human lungs; 6.) formulating a solution to prevent disorders of the human respiratory system; 7.) associating environmental conditions with the respiratory system.

The third stage was choosing media and developing HOTS items which are presented in the form of multiple-choice and essays. The media used is the Quizizz application which can be accessed via the internet. There are 12 multiple-choice and 3 essay questions. The items developed are by the HOTS indicator, the HOTS aspect, and the cognitive dimensions used. The details of this stage can be seen in Table 1. An overview of interactive quizzes using Quizizz is presented in Figure 1.



(a)

BioEdu
Berkala Ilmiah Pendidikan Biologi
Vol. 12 No. 1 Tahun 2023 Hal: 177-187
https://ejournal.unesa.ac.id/index.php/bioedu



(b)

Figure 1. (a) Interactive Quiz Display on Quizizz with Multiple Choice Question Type; (b) Essay Type

The next stage was to validate the interactive quiz items and display the Quizizz media to the validators consisting of material expert and media expert. Validation was carried out by providing an assessment of the product developed by filling out a validation sheet that was reviewed from the aspects of matter/concept, construction, language, HOTS (Higher-Order Thinking Skills), and media display (Quizizz). The results of the validity of the multiple-choice and essay assessment instruments are presented in Table 5 and Table 6.

Table 5. Validity of Multiple-Choice Test

No	Assessment Aspects	Average (%)		Validity (%)	Interpretation
		V1	V2		
A.	**Matter/Concept**				
1.	Each test item refers to an indicator	100	100	100	Very Valid
2.	Each test item refers to the correct concepts	100	100		
3.	The scope of each test item is clear and there is only one correct answer	100	100		
4.	The content of matter refers to levels of school and class	100	100		
B.	**Construction**				
1.	Each test item does not contain double answers	100	100	100	Very Valid
2.	Each test item does not give any clue of the correct answer option	100	100		
3.	Each test item does not depend on other answers of the other test item	100	100		
4.	The options of each test item are relatively equal	100	100		
5.	The options of each test item are homogeneous and logical	100	100		
6.	The option of each test item does not contain statement that all answers are correct or incorrect	100	100		

No	Assessment Aspects	Average (%)		Validity (%)	Interpretation
		V1	V2		
C.	**Language**				
1.	Grammar and spelling refer to *Ejaan Bahasa Indonesia* (EBI)	100	100	100	Very Valid
2.	The writing is clear and using communicative language	100	100		
3.	Using simple language and being easy to understood	100	100		
4.	Items does not contain multiple interpretations or misunderstandings	100	100		
D.	**Higher Order Thinking Skills (HOTS)**				
1.	Each test item refers to the dimension of cognitive processes of C-4 or C-5	91.7	100	97.23	Very Valid
2.	Each test item refers to one of HOTS aspects, including critical thinking, decision-making, or problem solving	91.7	100		
3.	Each test item uses stimulus that encourages students to think	100	100		
E.	**Media Display (*Quizizz*)**				
1.	The overall media view is interesting	100	100	100	Very Valid
2.	The size and font used are clear	100	100		
3.	How to answer only by pressing the choice of answer that feels right	100	100		
4.	Images, graphs, or tables are clear and functional	100	100		
Average of all aspects (%)				99.45	Very Valid

Based on Table 5., it can be known that the overall validity aspect of the multiple-choice type items obtained a value of 99.45% with a very valid category.

Table 6. Validity of Essay Test

No	Assessment Aspects	Average (%)		Validity (%)	Interpretation
		V1	V2		
A.	**Matter/Concept**				
1.	Each test item refers to an indicator	100	100	100	Very Valid
2.	Each test item refers to the correct concepts	100	100		
3.	The scope of each test item is clear	100	100		
4.	The content of matter refers to levels of school and class	100	100		
B.	**Construction**				
1.	The use of question sentences or questions using question words or commands that demand unraveled answers	100	100	100	Very Valid

No	Assessment Aspects	Average (%)		Validity (%)	Interpretation
		V1	V2		
2.	Test item does not depend on other answers from other questions	100	100	100	Very Valid
C.	**Language**				
1.	Grammar and spelling refer to *Ejaan Bahasa Indonesia* (EBI)	100	100	100	Very Valid
2.	The writing is clear and using communicative language	100	100		
3.	Using simple language and being easy to understood	100	100		
4.	Items does not contain multiple interpretations	100	100		
D.	**Higher Order Thinking Skills (HOTS)**				
1.	Each test item refers to the dimension of cognitive processes of C-4 or C-5	100	100	100	Very Valid
2.	Each test item refers to one of HOTS aspects, including critical thinking, decision-making, or problem solving	100	100		
3.	Each test item uses stimulus that encourages students to think	100	100		
E.	**Media Display (*Quizizz*)**				
1.	The overall media view is interesting	100	100	100	Very Valid
2.	The size and font used are clear	100	100		
3.	How to answer just by typing the correct answer in the field that has been provided	100	100		
Average of all aspects (%)				100	Very Valid

Based on Table 6., it can be known that the overall validity aspect of the essay type items obtained a value of 100% with a very valid category. The results of the overall validity of multiple choice and essay items are presented in Table 7.

Table 7. Validity of Multiple-Choice and Essay Test

Type of Tests	Validity (%)	$\bar{x}$ (%)	Interpretation
Multiple-Choice	99.45	99.73	Very Valid
Essay	100		

Based on Table 7., the validity of multiple-choice and essay items was determined based on aspects of the matter/concept, construction, language, HOTS, and media display (Quizizz). The validity result reached 99.73% with a very valid category.

Validation means "measuring what to measure" (Taherdoost, 2016). Validation refers to the extent to which the evidence obtained supports that the interpretation made is correct and appropriate (Hung and Yang, 2015). With this, validation of the products developed is needed so that the data obtained is fixed, accurate, and trustworthy (Widoyoko, 2012). Validity can be determined based on the accuracy and sensibility of measurement results expertly using validation sheets (Arikunto, 2015). Validation sheets are made adjusted to what you want to measure. This is because the validity of the instrument is related to the accuracy of the measuring instrument used. With a valid instrument, valid data will be obtained as well. The validity of the interactive quiz developed was categorized based on four aspects, namely matter/concept, construction, language, HOTS, and media display. The matter aspect gets validation results with a very valid interpretation, which is 100% for multiple choice test types and essay tests. This is because the aspect of the test matter

has referred to the matter/concept contained in the basic competencies and indicators that have been developed.

The construction aspect was related to the rules in making tests. In this study, the construction aspect got a very valid category of 100% for multiple-choice tests and essay tests. The test is one form of instrument used to measure the purpose of which is to determine the learning achievements or competencies that have been achieved by students for a particular field. Therefore, the test must be arranged by existing rules so that it can be known the extent of students' abilities and skills after completing a certain teaching unit (Khaerudin, 2017).

The language aspect referred to in the development of items is following the Indonesian spelling guidelines (*Pedoman Umum Ejaan Bahasa Indonesia*/PUEBI). The language aspect got a result of 100% for multiple-choice and essay tests. These results indicate that the items developed are following the PUEBI rules. The importance of paying attention to the linguistic aspect is intended so that students can easily understand the meaning of the sentences used and they can avoid misunderstandings or misconceptions. The purpose of using sentences by PUEBI rules is to avoid mistakes in the use of language, the use of language must be written properly and correctly (Kemendikbud, 2016).

The items developed were HOTS (Higher-Order Thinking Skills) questions which were declared very valid with 97.23% results for multiple-choice types and 100% for essay types. This shows that the items developed are following the cognitive dimensions of C-4 and C-5. Both indicators include critical thinking skills, decision-making, and problem-solving. In addition, the items developed have a stimulus that encourages students to think. These HOTS test items can practice skills that are the demands of 21st century proficiency. By frequently practicing these abilities and skills, students are expected to be able to face challenges, problems, and lives in the future (Redhana, 2019).

The interactive quiz developed was packed by utilizing the Quizizz application which can be accessed via the internet for free. The display of the media used can be adjusted as desired. There are many interesting features in the Quizizz application that can be used to create interesting quizzes that can make students not feel bored. Starting from themes, background music, memes, or funny images as a distraction, as well as scoreboards & rankings (Salsabila, et al., 2020).

The interactive quiz was tested on a limited basis to 18 students of class XI to identify reliability and level of difficulty. Reliability identification is used to determine the consistency of the interactive quiz developed, which means that the test instrument will give the same result if it is tested many times. The reliability value was calculated using the Cronbach Alpha formula using SPSS software. The results are presented in Table 8.

Table 8. Reliability of Interactive Quiz to Train HOTS

Reliability Value	Category
0.713	Reliable

Based on Table 8., it can be known that the reliability value obtained is 0.713 with reliable category. Several factors can affect the measurement of reliability, namely: the number of test items (test length), group variability of group heterogeneity, scoring objectivity, reliability estimation method, group level and level of difficulty test, and test homogeneity (Setiyawan, 2014). The interactive quiz developed can be said to be reliable or has a high constant value, meaning that in the future the questions developed can be reused for practice or evaluation with the results obtained will be relatively the same. Items with high reliability mean that they are following the characteristics of good questions.

Identification of item difficulty level was used to determine the distribution of HOTS items. The calculation of each item was presented in Table 9.

Table 9. The Difficulty Level of Test Items

No	Level of Difficulty	Category
Test Items of Multiple-Choice		
1	0.17	Difficult
2	0.56	Moderate
3	0.61	Moderate
4	0.61	Moderate
5	0.67	Moderate
6	0.57	Moderate
7	0.67	Moderate
8	0.39	Moderate
9	0.39	Moderate
10	0.50	Moderate
11	0.33	Moderate
12	0.39	Moderate
Test Items of Essay		
13	0.39	Moderate
14	0.22	Difficult
15	0.28	Difficult

Based on Table 9. It can be known that from the 15 items developed, there are differences in the level of difficulty in each question. There are 12 questions are included in the moderate category, namely at numbers 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, and 13. And 3 questions are included in the difficult category, namely numbers 1, 14,

and 15. The proportion of item difficulty levels were presented in Figure 2.

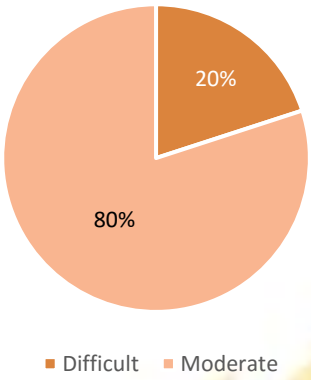


Figure 2. Proportion of Difficulty Levels

Based on Figure 2., it is calculated that 80% of the question items are at the moderate level and 20% at the difficult level. The proportion of the difficulty level of the test items cannot be ascertained, this is because the basis of the determination is seen from the purpose of using the assessment instrument developed (Arikunto, 2015).

In the interactive quiz that was developed, there were no questions with easy categories. According to Yani, et al. (2014), a good question is a question that is neither too easy nor too difficult. Problems that are too easy do not stimulate students to increase their effort to solve them. On the contrary, questions that are too difficult will cause students to become discouraged and not have the enthusiasm to try again because they are beyond their reach. The level of difficulty of the question items cannot only be used to predict the measuring instrument itself, it is also about students' ability to understand the matter taught by the teacher (Khaerudin, 2017). It can be concluded that the items developed in this study are included in the good category.

The items based on the results of the analysis are included in the good category (in the sense that the degree of difficulty of the items is sufficient or moderate), the items should be recorded in the question bank book, then these items can be issued again in tests in the future. For items in the difficult category, several predictions might be the cause: 1.) the sentence of the question is too complex and long; 2.) the matter taught has not been completed; 3.) the question has 2 or more correct answers (Yani, et al., 2014). In addition, there are direct factors that can cause students to be less precise in answering questions, such as environmental conditions, student focus in answering questions, and limited working time.

When viewed from the results of students' answers, three questions numbered 1, 14, and 15 are questions that are categorized as difficult questions. The three questions relate to the mechanism of the respiratory system, provide ideas or solutions to problems of disorders in the respiratory system, and the relationship between environmental conditions and the respiratory system. The items are categorized as difficult, there may be some students who have not yet completed understanding the concept, or maybe less careful in observing the pictures and understanding sentence questions (Yani, et al., 2014). Students assume that the question sentence is too long so it takes a long time to be able to understand the meaning of the question sentence.

CLOSING

Conclusion

Based on the research that has been done, it can be concluded that interactive quiz based on Quizizz of respiratory system materials to train HOTS for class XI students are valid and reliable. The validity results which are include matter/concepts, construction, language, HOTS, and media display reached 99.73% with a very valid category. Reliability test items got a value of 0.713 with a reliable category. The difficulty level of the items obtained two levels of difficulty, 80% included in the moderate category, and 20% included in the difficult category.

Suggestion

Based on the research that has been done, there are suggestions that in developing questions for the evaluating learning outcomes or practice questions, they should be arranged according to the indicators that have been previously made. The indicators made must pay attention to competencies and cognitive domains aimed at achieving learning objectives.

ACKNOWLEDGMENT

Researcher thanks to Nur Qomariyah, S.Pd., M.Sc. as a mentor lecturer, Dr. Sifak Indana, M.Pd. and Erlix Rakhmad Purnama, S.Si., M.Si. as expert validator lecturers, and the participation of students in XI MIA 2 SMA Muhammadiyah 10 GKB Gresik also Ristag Hamida Hanisisa, M.Si. as biology teacher that supporting this research.

REFERENCES

Afrita, M. dan Darussyamsu, R. 2020. Validitas Instrumen Tes Berpikir Tingkat Tinggi (HOTS) Pada Materi Sistem Respirasi Di Kelas XI SMA. *Jurnal Mangifera Edu*. 4(2): 129-142.

Arikunto, S. 2015. *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.

Hung, J. and Yang, D. 2015. The Validation of an Instrument for Evaluating the Effectiveness of Professional Development Program on Teaching Online. *Journal of Educational Technology Development and Exchange*. 8(1): 17-38.

Indrawati, F. dan Hartati, L. 2021. Pelatihan Aplikasi Belajar Berbasis Kuis Dalam Upaya Meningkatkan Aktivitas Belajar Siswa. *E-DIMAS: Jurnal Pengabdian Kepada Masyarakat*. 12 (1): 67-73.

Kemendikbud. 2016. *Pedoman Umum Ejaan Bahasa Indonesia*. Jakarta: Badan Pengembangan dan Pembinaan Bahasa Kementerian Pendidikan dan Kebudayaan.

Kemendikbud. 2020. SE Sesjen: Pedoman Penyelenggaraan Belajar Dari Rumah Dalam Masa Darurat Penyebaran Covid-19.

Khaerudin. 2017. Administrasi, Analisis Butir, dan Kaidah Penulisan Tes. *Jurnal Madaniyah*. 1(12): 97-128.

Pakpahan, R. dan Fitriani, Y. 2020. Analisa Pemanfaatan Teknologi Informasi Dalam Pembelajaran Jarak Jauh di Tengah Pandemi Virus Corona Covid-19. *Journal of Information System, Applied, Management, Accounting, and Research*. 4(2): 30-36.

Ramadhan, D. dan Wasis. 2013. Analisis Perbandingan Level Kognitif dan Keterampilan Proses Sains Dalam Standar Isi (SI), Soal Ujian Nasional (UN), Soal Trends in International Mathematics and Science Study (TIMSS) dan Soal Programme for International Student Assessment (PISA). *Jurnal Inovasi Pendidikan Fisika*. 2(1): 20-25.

Redhana, I.W. 2019. Mengembangkan Keterampilan Abad Ke-21 Dalam Pembelajaran Kimia. *Jurnal Inovasi Pendidikan Kimia*. 13(1): 2239-2253.

Riduwan. 2012. *Skala Pengukuran Variabel-Variabel Penelitian*. Bandung: Alfabeta.

Salsabila, U.H., dkk. 2020. Pemanfaatan Aplikasi Quizizz Sebagai Media Pembelajaran Ditengah Pandemi Pada Siswa SMA. *Jurnal Ilmiah Ilmu Terapan Universitas Jambi*. 4(2): 163-172.

Setiyawan. A. 2014. Faktor-Faktor yang Mempengaruhi Reliabilitas Tes. *Jurnal An Nur*. 6(2): 341-354.

Sucipto. 2017. Pengembangan Keterampilan Berpikir Tingkat Tinggi dengan Menggunakan Strategi Metakognitif Model Pembelajaran Problem Based Learning. *Jurnal Pendidikan*. 2(1): 63-71.

Sugiyono. 2015. *Metode Penelitian Pendidikan (Pendekatan Kuantitatif, Kualitatif, dan R&D)*. Bandung: CV Alfabeta.

Taherdoost, H. 2016. Validity and Reliability of the Research Instrument; How to Test the Validation of a Questionnaire/Survey in a Research. *International Journal of Academy Research in Management (IJARM)*. 5(3): 28-36.

Waruwu, M. 2020. Studi Evaluatif Implementasi Pembelajaran Daring Selama Pandemi Covid-19. *Jurnal Administrasi Pendidikan*. 27(2): 288-295.

Widoyoko, E.P.S. 2012. *Teknik Penyusunan Instrumen Penelitian*. Yogyakarta: Pustaka Pelajar.

Yani, A., Asri, A.F., Burhan, A. 2014. Analisis Tingkat Kesukaran, Daya Pembeda dan Fungsi Distraktor Soal Ujian Semester Ganjil Mata Pelajaran Produktif Di SMK Negeri 1 Indralaya Utara Tahun Pelajaran 2012/2013. *Jurnal Pendidikan Teknik Mesin*. 1(2): 98-115.

Zuriyani, E. 2012. *Literasi Sains dan Pendidikan*. URL: https://sumsel.kemenag.go.id/files/sumsel/file/file/TULISAN/wagj1343099486.pdf. Diakses tanggal 29 Maret 2021.