Development of Quizizz-Based Interactive Questions to Measure Scientific Reasoning Skills

¹ Adam Malik, ¹Yusup Setiawan, ¹Winda Setya
¹ Prodi Pendidikan Fisika, Universitas Islam Negeri Sunan Gunung Djati Bandung, Bandung 40292, Indonesia
e-mail: adammalik@uinsgd.ac.id; yusupsetiawan1907@gmail.com; suratwindasetya@uinsgd.ac.id

Abstract

This study aims to determine the feasibility level, students' scientific reasoning abilities, and students' responses to the use of quizizz-based interactive questions in measuring students' scientific reasoning skills. This research method uses the 4D model (Define, Design, Development, and Dissemination). The results showed that quizizz-based interactive questions to measure scientific reasoning skills developed were declared very feasible by the validators at a percentage of 82.6%. The results of the percentage of students' scientific reasoning skills on conservation reasoning indicators are 25%, probabilistic reasoning 20.83%, proportional reasoning 8.33%, correlation reasoning 4.17%, variable reasoning 8.33%, and hypothesis-deductive reasoning 12.50 %. The question has positive response criteria as seen from the results of the assessment of students by 88% with very interesting interpretations. Thus, it can be concluded that quizizz-based interactive questions to measure scientific reasoning skills are very feasible to use and get a positive response from students.

Keywords: Interactive questions, Quizizz, Scientific reasoning skills.

Pengembangan Soal Interaktif Berbasis Quizizz untuk Mengukur Keterampilan Penalaran Ilmiah

Abstrak

Penelitian ini bertujuan untuk mengetahui tingkat kelayakan, kemampuan penalaran ilmiah peserta didik, dan respon peserta didik terhadap penggunaan soal interaktif berbasis *quizizz* dalam mengukur keterampilan penalaran ilmiah peserta didik. Metode penelitian ini menggunakan model 4D (*Define*, *Design*, *Development*, dan *Dissemination*). Hasil penelitian menunjukan soal interaktif berbasis *quizizz* untuk mengukur keterampilan penalaran ilmiah yang dikembangkan dinyatakan sangat layak oleh para validator pada persentase 82,6%. Hasil persentase keterampilan penalaran ilmiah peserta didik pada indikator penalaran konservasi 25%, penalaran probabilistik 20,83%, penalaran proporsional 8,33%, penalaran korelasi 4,17%, penalaran variabel 8,33%, dan penalaran hipotesis-deduktif 12,50%. Soal memiliki kriteria respon positif yang dilihat dari hasil penilain peserta didik sebesar 88% dengan interpretasi sangat menarik. Dengan demikian, dapat disimpulkan soal interaktif berbasis *quizizz* untuk mengukur keterampilan penalaran ilmiah sangat layak untuk digunakan dan mendapat respon positif dari peserta didik.

Kata Kunci: Soal interaktif, *Quizizz*, Keterampilan penalaran ilmiah.

INTRODUCTION

The development of technology makes the learning system more innovative. Educators should prepare modern infrastructure facilities and infrastructure to support the quality of education (Malik & Agustina, 2020). This aims to prepare students who are superior, competitive, and able to face global competition. Success to create a quality learning system. The teachers are expected to be able to create a modern learning process by utilizing various current technologies as learning tools or media (Xu et al., 2017).

The use of technology in education is not only a source of learning for students. Technological developments can be used in evaluating student learning activities. Using technology in evaluation can help educators understand and evaluate students in a sustainable manner (Davies, 2011). Technology-based assessments offer a unique opportunity to collect data about learners' cognitive development. Assessment can provide feedback to students and teachers to improve learning (Molnár & Csapó, 2019). Technology can make it easier for teachers to measure individual learners' mastery of 21st-century skills (Xu et al., 2017). Evaluation using technology such as quizizz can increase motivation and students are more careful in working on problems (Göksün & Gürsoy, 2019). Fitriyani, Hertanti, and Nurlaela (2017) Presentation of questions using interactive questions can make questions more varied so that boredom in students can be reduced or even eliminated. Students are more interested in interactive questions than conventional questions or paper-based tests. The use of quizzes in the classroom as a tool for evaluating student learning can increase the interactivity of students and also increase the involvement of students in evaluating learning outcomes (Chaiyo & Nokham, 2017).

Teachers must develop student skills that are relevant to the 21st century, one of which is critical thinking skills. Critical thinking skills consist of scientific reasoning, systems thinking, computational thinking, decision making, and problem-solving (Qian & Clark, 2016). Scientific reasoning is the main focus of science education. Scientific reasoning not only affects students' academic achievement in classroom learning but can also influence daily decision-making and student success(Ding, Wei, & Liu, 2016). Scientific reasoning has a hypothetical-deductive structure and consists of related aspects, such as proportional reasoning, variable control, probability reasoning, and correlation reasoning (Lawson, 1982, 2005). A core component of scientific reasoning is the ability to design controlled experiments. Besides, evaluate the resulting evidence relating to a person's hypothesis (Kant, Scheiter, & Oschatz, 2017). Jensen, Neeley, Hatch, and Piorczynski (2017) argues that reasoning skills are a product of learning. Thus, it is appropriate for teachers to train and develop students' scientific reasoning skills during learning.

In fact, students' reasoning is still difficult to develop even though various ways have been applied. Research of (Xiao, Han, Koenig, Xiong, & Bao, 2018) shows that students find it more difficult to provide the correct explanation than just knowing the answer. Students' scientific reasoning skills are still low in describing problems using a physics approach (Jelicic, Planinic, & Planinsic, 2017). Yuksel and Ates (2018) suggested deficiencies in the development of cognitive skills and scientific reasoning skills of students in the formal and concrete stages of reasoning. Students find it difficult to introduce scientific meaning, solve problems and understand the nature of science. Method of experimenting and discussion obtains a significant increase in scientific reasoning (Marušić & Sliško, 2012).

The results of previous research showed that students had difficulty working on the questions. Students in answering questions sometimes make guesses so that students do not think about the answer correctly (I. G. N. Darmawan, Vosniadou, Lawson, Van Deur, & Wyra, 2020; R. Darmawan, 2015; Moore & Rubbo, 2012; Xiao et al., 2018; S.-N. Zhou, Liu, Koenig, Li, & Bao, 2021). Teachers assess student work really need time. Some students may not provide useful answers and give short, less detailed responses(Fulmer, Chu, Treagust, & Neumann, 2015). Besides, students when working on questions using paper-based tests lack the motivation to solve problems (Nikou & Economides, 2016).

The teacher's job is not an easy one and involves more factors than what can be observed in one study. Teachers must be more creative in filling learning and be more innovative in utilizing technology. The development of interactive questions to measure scientific reasoning skills is a step towards solving the problem. Various types of scientific reasoning instruments and assessments have been developed (Han, 2013; Opitz, Heene, & Fischer, 2017; S.-N. Zhou et al., 2021). Mobilized learning can improve test results, engagement, and reflection from students(Looi, Sun, & Xie, 2014). Test instruments use computers and smartphones can increase students' motivation(Khoshsima & Hashemi Toroujeni, 2017; Nikou & Economides, 2016). Besides, utilizing gamification technology such as quizizz can increase motivation and be more careful in working on problems (Göksün & Gürsoy, 2019). Assessment using quizizz software, teachers do not need to take a lot of time to check student work. Student results are stored in a database in the quizizz software.

Based on previous research related to the development of LCTSR scientific reasoning instruments(Han, 2013; Xiao et al., 2018) there are differences. The novelty in the development of this instrument is found in the questions and answer choices. Problems on each indicator of scientific reasoning are modified. The modifying instrument has an impact on the planning, process, and evaluation of learning. Besides, scientific reasoning instruments are presented using quizizzes. Scientific reasoning problems are presented at a formal operational level that can be used by high school students. Therefore, this study aims to develop quizizz-based interactive questions to measure students' scientific reasoning skills on the concepts of momentum and impulse.

METHODS

The research method used is Design and Development (D&D) by using the 4D model (Define, Design, Development, and Dissemination). The definition stage is consists of a needs analysis, student psychological development, and teaching materials. The design stage is consists of compiling the instrument grid and arranging the question instruments on the quiz. The development stage is carried out by validating the question instruments made to the experts. Instrument validation was related to material, media, and language. After validation, the instrument was tested on students of class XI MIA 1 Cianjur. The dissemination stage was used to measure students' reasoning skills in class XI MIA MAN 2 Cianjur. Besides, questionnaires were distributed to students to find out the response to the use of quizizz.

The research instrument used a questionnaire and test. This research questionnaire sheet consists of validation sheets and student responses. The validation sheet is used to assess the validity or feasibility of quizizz-based interactive questions. The aspects assessed include content, construct, and

language. The validators were experts and teachers. The validation sheet provides a column for comments and suggestions for the validator. Questionnaires are given to students. The aim is to find out student responses related to the use of interactive questions

The reasoning skills test instrument is a multiple-choice reasoned question according to the Lawson Classroom Test of Scientific Reasoning (LCTSR). LCTSR were consists of (1) conservation reasoning, (2) proportional reasoning, (3) co-control of variables, (4) probability reasoning, (5) correlation reasoning (correlation reasoning), (6) hypothetical-deductive reasoning(Lawson, 1982, 2005). The test were consists of 12 questions with indicators to the LCTSR related to the concepts of momentum and impulse.

Processing data of the validity of the scientific reasoning instrument in this study used Anatest software version 4.0.9. The analysis is used to determine the validity, reliability, difficulty level, and distinguishing power of the questions. The analysis of the scientific reasoning skills test used peer assessment techniques. The assessment criteria for the scientific reasoning skills test are shown in table 1 (Han, 2013).

Table 1. Scientific Reasoning Answer Score

Question Type		Score	Category
Statement	Reason		
True	True	2	Good
True	True	1	Enough
False	False	0	Error
False	False	0	Less

The results obtained by students are added up as a whole. Then the value of each student is categorized into three types of scientific reasoning skills. The category of scientific reasoning skills can be seen in table 2.

Table 2. Categories of Scientific Reasoning Skills

Category	Score
Formal	9-12
Transition	5-8
Concrete	1-4

The results of student tests on each indicator of scientific reasoning were processed for analysis. The percentage of students' scientific reasoning skills tests on each indicator were classified according to the criteria as in table 3 (Anas, 2008).

Table 3. Criteria for Levels of Scientific Reasoning Skills

Percentage	Category
81% - 100%	Very good
61% - 80%	Good
41% - 60%	Enough
21% - 40%	Less
0% - 20%	Very less

Questionnaire to measure student responses to the use of quizziz-based interactive tests. The questionnaire uses a Likert scale. The choices in the questionnaire consisted of strongly agree (5), agree (4); quite agree (3), disagree (2), strongly disagree (1). The results of calculating student responses are adjusted to the criteria in table 4.

Table 4. Questionnaire Criteria from Student Responses

Score	Category
80%-100%	Very interesting
60%-80%	Interesting
40%-60%	Less attractive
20%-40%	Not attractive
0%-20%	Very Unattractive

RESULTS AND DISCUSSION

Development of interactive quizizz-based questions to measure students' scientific reasoning skills was using the 4D model (Define, Design, Development, and Dissemination)

A. Define

The definition stage is a series of initial stages of research. This stage was consisting of three steps of analyzing needs, students, and content.

 1. Needs analysis

Needs analysis is a series of early stages of research to find out the problems faced in learning physics. Needs analysis includes literature study by analyzing the problems studied by previous research. The results of the literature study show that students sometimes answer questions by guessing. Students do not think about answers correctly based on understanding (I. G. N. Darmawan et al., 2020; Moore & Rubbo, 2012; Xiao et al., 2018; S.-N. Zhou et al., 2021).

Students have not been trained to work on questions related to scientific reasoning, especially if it is related to contextual in everyday life. This is similar to research conducted by S. Zhou et al. (2016), students are weak in scientific reasoning when faced with scenarios in real life. Learners are easily fooled by provocative answers. Besides, students when working on the paper-based test lacked motivation in working on the questions. Students often provide short written responses that are less detailed(Fulmer et al., 2015). On the other hand, the teacher in assessing student work also takes time. Sometimes students give answers that are less than expected.

Based on the results of previous research, the use of quizizz software in learning evaluation can increase students' learning motivation. The results of the research by Göksün and Gürsoy (2019) using the quizizz application have an impact on students in the exam process. Students are more careful when answering questions using interactive questions. The results of research related to scientific reasoning conducted by Erlina and Susantini (2018) Students' scientific reasoning is in the weak category due to the low level of students' confidence in determining answers. Students belong to the category of wrong scientific reasoning and have a low understanding of physics concepts. This requires innovation to overcome the confidence level of students in terms of answering scientific reasoning questions. Based on these considerations, this study develops an evaluation in the form of quizizz-based interactive questions in measuring scientific reasoning skills.

 2. Student analysis

The results of the analysis related to the characteristics of students by the theory of cognitive by Piaget. Students of senior high school are included in the formal operational category (Bujuri, 2018). Students have the characteristics of

abstract thinking and reasoning logically using possible thinking patterns (Piaget, 2003). High school students have age 15 to 18 years. Students' scientific reasoning abilities are well established. According to research conducted by Ding et al. (2016) the level of scientific reasoning skills in universities is disappointing. Students before entering college need to be trained in their level of scientific reasoning. Having the ability to reason scientifically is not only for entering college but in real life and the world of work. Scientific reasoning as a cognitive ability should be trained and developed from childhood to adolescence. On the other hand, there is a tendency to conceptualize competence as a distinct but specifically coordinated set of skills (Opitz et al., 2017). Based on Piaget's theory, students became research subjects for the development of quizizz-based interactive questions.

3. Content analysis

The content analysis aims to identify concepts used in the reasoning test. The content of the instrument is by the 2013 curriculum for physics lessons. Momentum and impulse content are taught in class X on basic competence 3.10. Students are expected to be able to master the concept and implement it into everyday life. Based on previous research, students find it difficult to apply the concepts of momentum and impulses. students have scientific reasoning skills that are still in the low category (Effendy, Hartono, & Ian, 2018). Mainly related to linking concepts to problem-solving (Bryce & MacMillan, 2009) Students do not yet have a conceptual understanding of solving scientific literacy problems (Al Faizah & Aminah, 2019). The skills of students can be improved by contextual-based practice questions. Interactive quizzes-based questions were developed.

B. Design

The design stage consists of the preparation of the instrument grid and the integration of interactive questions using quizizz software. The preparation of the question instrument grid aims to determine the scope used for instructions for making questions (Mukhlis & Tohir, 2019). This stage is to make it easier to design the right question instrument. The questions develop based on the Lawson Classroom Test Scientific Reasoning (LCTSR) indicator. The scientific reasoning instrument consists of types A and B with 12 questions each. Students answer two questions. In the first question, students must know the concepts of physics that are implemented in real life. The second question, students must provide an explanation or reason with predetermined answer choices. The scientific reasoning question instrument integrates into the quizizz software. Instruments of type A and B reasoning questions judge by experts. Instruments feasible are then tested. A trial of questions determines the best questions (Sholekhah, Maharta, & Suana, 2018).

C. Development

The development stage consists of validating the feasibility of quizizz-based interactive questions by experts and the testing stage. Figure 1 and 2 shows the appearance of the quizizz-based interactive questions that were developed.
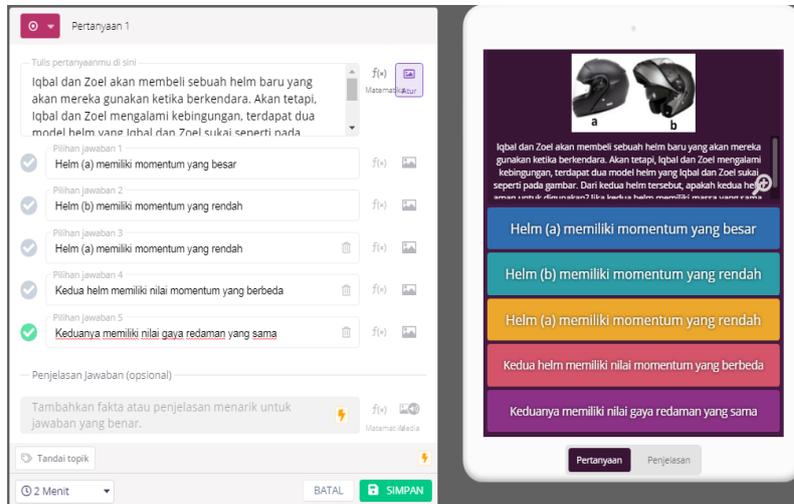
Figure 1. Problems with Scientific Reasoning on Statements
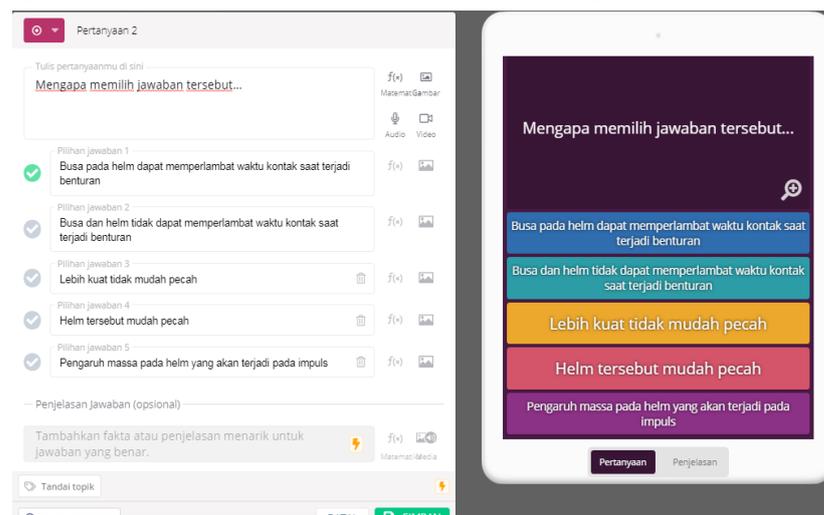
Figure 2. Scientific Reasoning Questions on Reason

Quizizz-based interactive questions are tested for feasibility using a validation sheet that is assessed by content, media, and language experts (Sholekhah et al., 2018). There are several aspects of the assessment of the product feasibility test validated by several experts. The content expert assesses the feasibility of the physics concept contained in the question instrument. The validated material aspects consist of quality content, construct, presentation, and indicator suitability. Media experts assess the feasibility of interactive questions contained in the quizizz software. The validated media aspects consist of a presentation, design, content design, and ease of use. Linguists assess the feasibility of the question instrument in terms of straightforwardness, communicativeness, suitability, and accuracy according to general Indonesian spelling guidelines. The physics teacher assesses the feasibility of the product developed practically in the field. The validator assesses the product that has been developed by filling out the validation

sheet with a Likert scale (Prasetiyo & Perwiraningtyas, 2017). The validator provides suggestions to improve quizizz-based interactive questions so that they are feasible to be tested on students.

The results of the quizizz-based interactive test by experts in terms of various aspects of the assessment are shown in table 5.

Table 5. Aspect Assessment Quizizz-Based Interactive Test

No	Aspect	Eligibility Percentage	Criteia
1	Content Aspect	77,4%	Worthy
2	Media Aspect	84,4%	Very worth
3	Language Aspect	86%	Very worth
	All Aspect	**82,6%**	Very worth

All aspects of the content, media, and language were eligible with very feasible criteria. The content aspect was the lowest feasibility with medium criteria. The language aspect was the highest eligible with very feasible criteria. Expert validation analysis feasible if it meets the interpretation criteria of 61% (R. Darmawan, 2015). Thus, quizizz-based interactive questions to measure scientific reasoning skills feasible for using research instruments.

Before revision	After revision
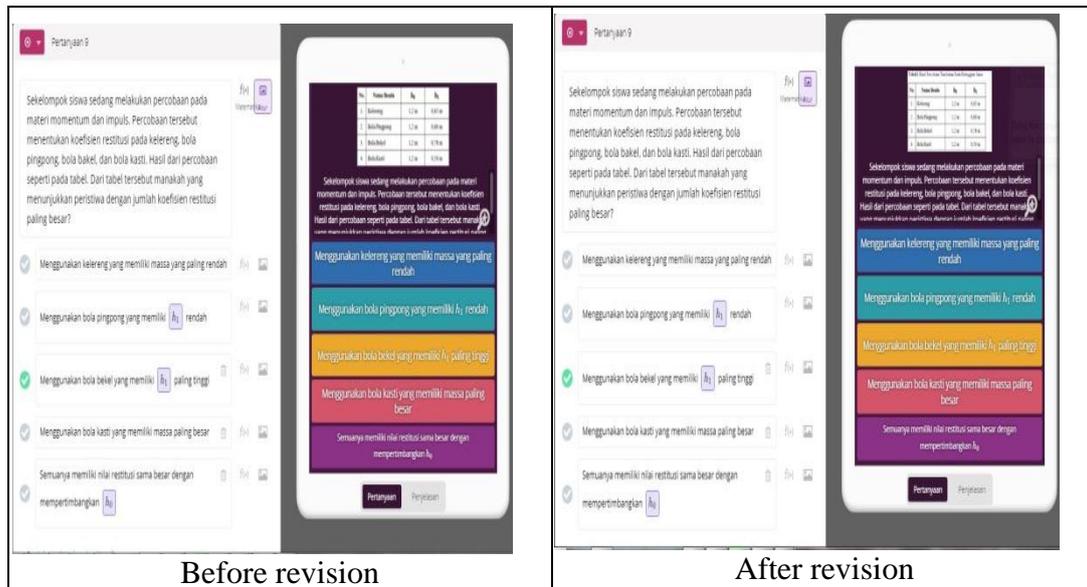
Figure 3. The Revision of Scientific Reasoning Interactive Questions

The improvement for the advice of experts regarding the table information and diction in the questions. The revision has added a description of the table, and the word "amount" in the question was abolished.

Quizizz-based interactive test tests were carried out after being deemed eligible by the validator's assessment. The trial was conducted at MAN 2 Cianjur class XI MIA 1. The interactive type A questions do to 24 students and type B of 29 students.

The results of the validity analysis of quizizz-based interactive questions are said to be significant if the coefficient $r_{xy} > r_{table}$. Type A was five, not significant questions (3, 4, 5, 6, and 8). The five questions were not be used. Type B only number 11 is not significant. This number is not used.

Questions are considered valid if they obtain a criterion of 0.41 (Hamzah, 2014).

The results of the reliability analysis of quiz-based interactive questions in type A are classified as reliable with a value of $r_{xy} = 0.67 > r_{table} = 0.50$. Questions type B $r_{xy} = 0.92 > r_{table} = 0.85$ including reliable. The question instrument can be said to be reliable if the value r_{xy} is greater than r_{table} (Yusup, 2018).

The results of the analysis of the difficulty level of quizizz-based interactive questions were good. Type A questions were ten questions (medium category), one question (difficult category), and one question (easy category). Type B questions have one question (medium category), nine questions (difficult category), and two questions (very difficult category). The question instrument can be said to be good if the question instrument is not too difficult and not too easy (Fatimah & Alfath, 2019).

The results of the analysis of the distinguishing power of the 12 questions were a good category. Type A was 6 questions (enough category), two questions (good category), and four questions (very good category). Type B was two questions (enough category), seven questions (good category), and three questions (very good category). Menurut Fatimah and Alfath (2019) said the discrimination index of 0.20 is the minimum criteria number for a good question instrument.

Based on the results of the quizizz-based interactive test, 12 questions determine to measure students' reasoning skills. The four questions of type A are numbers 1, 2, 11, and 12. The eight questions of type B are numbers 3, 4, 5, 6, 7, 8, 9, and 10.

D. Disseminate

1. Analysis of scientific reasoning skills

Each indicator of scientific reasoning represents two things, namely questions and reasons for an answer to measure students' understanding of scientific reasoning. Students will be given a score of 1 if they can answer correctly. If students answer incorrectly, then give a score of 0. Then, the results obtained by students are added up as a whole. Each student's grades are categorized into three categories of scientific reasoning skills. The results of the assessment of students' scientific reasoning skills are shown in table 6.

Table 6. Results of Scientific Reasoning Skills Category

Category	Total Students
Formal	2
Transition	7
Concrete	15

The category of scientific reasoning skills of the concrete level was 15 students. Seven students are transition level and two students are formal level. Students are scores of 1-4 are classified at the concrete operations. Students are a score of 5-8 are classified at the transition level. Students are scores of 9-12 are classified at the formal operational. High school students are included in the formal operational category. Students have the characteristics of abstract thinking and reasoning logically (Piaget, 2003). The results showed that only 2 out of 24 students were able to reach the formal category. Students are not accustomed to solving problems or questions related to scientific reasoning skills. Students have difficulty in

scientific reasoning or connecting concepts (Setiyoningtyas & Kasmui, 2020).

Student achievement on each indicator of scientific reasoning skills is shown in Figure 4.
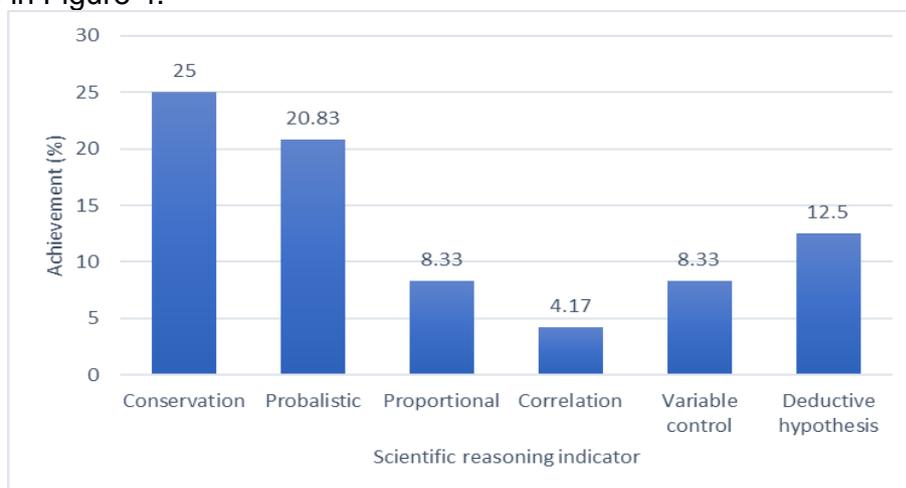
Figure 4. The Results of Scientific Reasoning Skills for Each Student Indicator

Conservation reasoning obtains a percentage of 25%. Students should develop their knowledge in solving problems about what will happen. Probabilistic reasoning obtained a percentage of 20.83%. Students do not understand a result when it is repeated under the same circumstances in a larger context. Students are lacking in maintaining knowledge even though the appearance of the object changes but certain properties of the object will remain the same (Han, 2013). Proportional reasoning obtained a percentage of 8.33%. Students are less able to determine and compare ratios. Correlation reasoning is 4.17%. Students are very less in determining the interrelationships between variables (Subiki & Supriadi, 2018). Pengontrolan variabel memperoleh persentase 8,33%. Controlling variables obtained a percentage of 8.33%. Students' understanding of controlling the dependent and independent variables that affect hypothesis testing needs to be improved. Hypothesis-deductive reasoning obtained a percentage of 12.50%. Students still have difficulty in testing hypotheses and concluding.

The lowest students' scientific reasoning skills occurred in the correlation reasoning indicator. Students have difficulty in connecting two related events. These results strengthen previous research (Subiki & Supriadi, 2018). The highest scientific reasoning skills occur in conservation reasoning indicators. Students can retain knowledge despite the changing context. The low factor of scientific reasoning is produced by students answering guessing questions. The student does not explain the answer correctly (I. G. N. Darmawan et al., 2020; Moore & Rubbo, 2012; Xiao et al., 2018; S.-N. Zhou et al., 2021). Students need to be accustomed to working on scientific reasoning questions as an exercise in reasoning.

2. Student response

The final stage of this research is to find out students' responses to the use of quizizz-based interactive questions to measure scientific reasoning skills. The results of student responses are shown in Table 7.

Table 7. Results of Student Response Recapitulation

No	Indicator	Percentage	Criteria
1	Attractiveness	92,06%	Very interesting
2	Content Quality	87,83%	Very interesting
3	Language	84,38%	Very interesting
4	Convenience	86,94%	Very interesting
	Average	**88%**	Very interesting

The average percentage of the four indicators is 88%, including very interesting criteria. The average results of the four indicators including very interesting criteria strengthen previous research (Solikah, 2020). The largest assessment indicator on the attractiveness aspect is 92.02% including the very attractive criteria. The display on the quizizz software is interesting and can motivate students in working on questions (Purba, 2019). This is also supported by the opinion of Chaiyo and Nokham (2017) that the use of quizizz in the classroom as a learning evaluation tool can increase student interactivity and engagement. The assessment indicator on the linguistic aspect obtained the smallest percentage of 84.38% including very interesting criteria. The questions on the quizizz software have a lot of writing, so the writing on the student's smartphone screen becomes small. When students read the questions, it becomes less clear. According to Nurjanah (2017), making instruments needs to pay attention to linguistic aspects so that students can understand them in answering questions.

The interpretation of the questionnaire results indicated that students stated that quizizz-based interactive questions to measure scientific reasoning skills were very interesting, working on quizizz-based interactive questions was highly motivated, questions were easy to understand, and working on questions using quizizz software was practical and effective for use in evaluating physics learning. Research that has been done has developed interactive questions to measure scientific reasoning skills. Furthermore, there is a need for appropriate learning methods to improve scientific reasoning skills(Moraes, Castellar, Castellar, & Castellar, 2010).

CONCLUSION

We have successfully developed quizizz-based interactive questions to measure scientific reasoning skills on the concept of momentum and impulse. Interactive questions to measure scientific reasoning skills are the form of scientific reasoning questions. Each indicator of scientific reasoning questions consisting of statements and reasons. The results of the validation show an average percentage of 82.6% with an interpretation very suitable. The average student has scientific reasoning skills in the category of less to very less. The interpretation of the questionnaire was an average of 88%. Students stated that quizizz-based interactive questions to measure scientific reasoning skills were very interesting. Thus, quizizz-based interactive questions can be used to measure scientific reasoning skills. The implications of this research are expected to be able to develop quizizz-based interactive questions for other content.

REFERENCES

Al Faizah, W., & Aminah, N. (2019). *Analysis of Student Concepts Understanding in Solving Scientific Literacy on the Topic of Momentum and Impulse.* Paper presented at the Journal of Physics: Conference Series.

Anas, S. (2008). Pengantar statistik pendidikan. *Jakarta: Raja Grafindo Persada.*

Bryce, T., & MacMillan, K. (2009). Momentum and kinetic energy: Confusable concepts in secondary school physics. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching, 46*(7), 739-761.

Bujuri, D. A. (2018). Analisis perkembangan kognitif anak usia dasar dan implikasinya dalam kegiatan belajar mengajar. *LITERASI (Jurnal Ilmu Pendidikan), 9*(1), 37-50.

Chaiyo, Y., & Nokham, R. (2017). *The effect of Kahoot, Quizizz and Google Forms on the student's perception in the classrooms response system.* Paper presented at the 2017 International Conference on Digital Arts, Media and Technology (ICDAMT).

Darmawan, I. G. N., Vosniadou, S., Lawson, M. J., Van Deur, P., & Wyra, M. (2020). The development of an instrument to test pre-service teachers' beliefs consistent and inconsistent with self-regulation theory. *British Journal of Educational Psychology, 90*(4), 1039-1061.

Darmawan, R. (2015). Skala Pengukuran Variabel-Variabel Penelitian. *Bandung: Alfabeta. Hal, 24.*

Davies, R. S. (2011). Understanding technology literacy: A framework for evaluating educational technology integration. *TechTrends, 55*(5), 45-52.

Ding, L., Wei, X., & Liu, X. (2016). Variations in university students' scientific reasoning skills across majors, years, and types of institutions. *Research in Science Education, 46*(5), 613-632.

Effendy, S., Hartono, Y., & Ian, M. (2018). The Ability of Scientific Reasoning and Mastery of Physics Concept of State Senior High School Students in Palembang City. *vol, 247*, 504-509.

Erlina, N., & Susantini, E. (2018). *Common False of Student's Scientific Reasoning in Physics Problems.* Paper presented at the Journal of Physics: Conference Series.

Fatimah, L. U., & Alfath, K. (2019). Analisis Kesukaran Soal, Daya Pembeda dan Fungsi Distraktor. *Al-Manar, 8*(2), 37-64.

Fitriyani, S., Hertanti, E., & Nurlaela, A. (2017). *Pengaruh Metode Drill Berbantuan Soal Animasi terhadap Kemampuan Berpikir Reflektif Siswa Pada Konsep Dinamika Partikel.* Paper presented at the PROSIDING SEMINAR NASIONAL FISIKA (E-JOURNAL).

Fulmer, G. W., Chu, H.-E., Treagust, D. F., & Neumann, K. (2015). Is it harder to know or to reason? Analyzing two-tier science assessment items using the Rasch measurement model. *Asia-Pacific Science Education, 1*(1), 1. doi: 10.1186/s41029-015-0005-x

Göksün, D. O., & Gürsoy, G. (2019). Comparing success and engagement in gamified learning experiences via Kahoot and Quizizz. *Computers & Education, 135*, 15-29.

Hamzah, A. (2014). *Evaluasi Pembelajaran Matematika*: Rajawali Pers.

Han, J. (2013). *Scientific reasoning: Research, development, and assessment.* The Ohio State University.

Jelicic, K., Planinic, M., & Planinsic, G. (2017). Analyzing high school students' reasoning about electromagnetic induction. *Physical Review Physics Education Research, 13*(1), 010112.

Jensen, J. L., Neeley, S., Hatch, J. B., & Piorczynski, T. (2017). Learning scientific reasoning skills may be key to retention in science, technology, engineering, and mathematics. *Journal of College Student Retention: Research, Theory & Practice, 19*(2), 126-144.

Kant, J. M., Scheiter, K., & Oschatz, K. (2017). How to sequence video modeling examples and inquiry tasks to foster scientific reasoning. *Learning and Instruction, 52*, 46-58.

Khoshsima, H., & Hashemi Toroujeni, S. (2017). Comparability of Computer-Based Testing and Paper-Based Testing: Testing mode effect, testing mode order, computer attitudes and testing mode preference. *International Journal of Computer (IJC), 24*(1), 80-99.

Lawson, A. E. (1982). The nature of advanced reasoning and science instruction. *Journal of Research in Science teaching, 19*(9), 743-760.

Lawson, A. E. (2005). What is the role of induction and deduction in reasoning and scientific inquiry? *Journal of Research in Science Teaching, 42*(6), 716-740.

Looi, C.-K., Sun, D., & Xie, W. (2014). Exploring students' progression in an inquiry science curriculum enabled by mobile learning. *IEEE Transactions on Learning Technologies, 8*(1), 43-54.

Malik, A., & Agustina, R. D. (2020). Review Pengaruh Teknologi Gamification Terhadap Psikologis Siswa. *JICTE (Journal of Information and Computer Technology Education), 4*(1).

Marušić, M., & Sliško, J. (2012). Influence of three different methods of teaching physics on the gain in students' development of reasoning. *International Journal of Science Education, 34*(2), 301-326.

Molnár, G., & Csapó, B. (2019). Making the psychological dimension of learning visible: using technology-based assessment to monitor students' cognitive development. Front. Psychol. 10, 1368 (2019).

Moore, J. C., & Rubbo, L. J. (2012). Scientific reasoning abilities of nonscience majors in physics-based courses. *Physical Review Special Topics-Physics Education Research, 8*(1), 010106.

Moraes, J. d., Castellar, S. M. V., Castellar, S., & Castellar, S. (2010). Scientific literacy, problem based learning and citizenship: a suggestion for geography studies teaching. *Problems of Education in the 21st Century, 19*, 119-127.

Mukhlis, M., & Tohir, M. (2019). Instrumen Pengukur Creativity And Innovation Skills Siswa Sekolah Menengah di Era Revolusi Industri 4.0. *Indonesian Journal of Mathematics and Natural Science Education, 1*(1), 65-73.

Nikou, S. A., & Economides, A. A. (2016). The impact of paper-based, computer-based and mobile-based self-assessment on students' science motivation and achievement. *Computers in Human Behavior, 55,* 1241-1248. doi: 10.1016/j.chb.2015.09.025

Nurjanah, N. (2017). Analisis Butir Soal Pilihan Ganda dari Aspek Kebahasaan. *Faktor: Jurnal Ilmiah Kependidikan, 2*(1), 69-78.

Opitz, A., Heene, M., & Fischer, F. (2017). Measuring scientific reasoning – a review of test instruments. *Educational Research and Evaluation, 23*(3-4), 78-101. doi: 10.1080/13803611.2017.1338586

Piaget, J. (2003). *The psychology of intelligence*: Routledge.

Prasetiyo, N. A., & Perwiraningtyas, P. (2017). The development of environment based textbook in biology course at Tribhuwana Tunggadewi University. *Indonesian Journal of Biology Education, 3*(1), 19-27.

Purba, L. S. L. (2019). Peningkatan konsentrasi belajar mahasiswa melalui pemanfaatan evaluasi pembelajaran quizizz pada mata kuliah kimia fisika I. *Jurnal Dinamika Pendidikan, 12*(1), 29-39.

Qian, M., & Clark, K. R. (2016). Game-based Learning and 21st century skills: A review of recent research. *Computers in human behavior, 63*, 50-58.

Setiyoningtyas, R., & Kasmui. (2020). Pengembangan quizizz-assisted test berbasis literasi sains pada materi larutan elektrolit nonelektrolit. *Chemistry in Education, 2*.

Sholekhah, F., Maharta, N., & Suana, W. (2018). Development of Higher Thinking Instrument of Newton 'S Laws of Motion. *Journal of Physics and Science Learning, 2*(1), 17-26.

Solikah, H. (2020). Pengaruh Penggunaan Media Pembelajaran Interaktif Quizizz terhadap Motivasi dan Hasil Belajar Siswa pada Materi Teks Persuasif Kelas VIII di SMPN 5 Sidoarjo Tahun Pelajaran 2019/2020. *BAPALA, 7*(3).

Subiki, S., & Supriadi, B. (2018). Identifikasi Kemampuan Penalaran Ilmiah (Scientific Reasoning) Siswa SMA di Kabupaten Jember Pada Pokok Bahasan Dinamika. *FKIP e-PROCEEDING, 3*(1), 121-126.

Xiao, Y., Han, J., Koenig, K., Xiong, J., & Bao, L. (2018). Multilevel Rasch modeling of two-tier multiple choice test: A case study using Lawson's classroom test of scientific reasoning. *Physical Review Physics Education Research, 14*(2), 020104.

Xu, B., Tu, S., Qiao, S., Qu, X., Chen, Y., Yang, J., . . . Tian, F. (2017). Diagnostic accuracy of angiography-based quantitative flow ratio measurements for online assessment of coronary stenosis. *Journal of the American College of Cardiology, 70*(25), 3077-3087.

Yuksel, I., & Ates, S. (2018). *Examination of the relationship between preservice science teachers' scientific reasoning and problem solving skills on basic mechanics.* Paper presented at the AIP Conference Proceedings.

Yusup, F. (2018). Uji validitas dan reliabilitas instrumen penelitian kuantitatif. *Tarbiyah: Jurnal Ilmiah Kependidikan, 7*(1).

Zhou, S.-N., Liu, Q.-Y., Koenig, K., Li, Q., & Bao, L. (2021). Analysis Of Two-Tier Question Scoring Methods: A Case Study On The Lawson's Classroom Test Of Scientific Reasoning. *Journal of Baltic Science Education, 20*(1), 146.

Zhou, S., Han, J., Koenig, K., Raplinger, A., Pi, Y., Li, D., . . . Bao, L. (2016). Assessment of scientific reasoning: The effects of task context, data, and design on student reasoning in control of variables. *Thinking skills and creativity, 19*, 175-187.