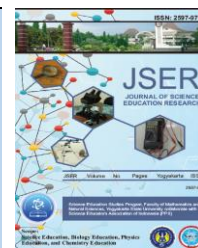




Journal of Science Education Research

Journal homepage: www.journal.uny.ac.id/jserr



Development of HOTS Science Test: Ethno-Science Technology Engineering and Mathematics based on Javanese Gamelan

N. A. Awaliyah¹, Y. Widiyawati², I. Nurwahidah³

^{1,2,3}Study Program of Science Education, Faculty of Sciences and Technology, Ivet University

Corresponding Author. E-mail: nadaafida14@gmail.com

Keywords:

Ethno-STEM, High Order Thinking Skills, Javanese Gamelan.

Abstract

Ethno-STEM is appropriate to accommodate High Order Thinking Skills (HOTS). The study aimed to develop an ethno-STEM-based HOTS science test instrument for Javanese gamelan musical instruments for junior high school students. The development research referred to the Thiagarajan 4-D (Four-D) development model with define, design, develop, and disseminate. The sample in this study was 24 students of VIII grades. The learning expert validation sheets, practitioner validation sheets, student questionnaires, and student ability test results were used in this particular research. The data analysis of the science test instrument feasibility showed that: (1) both experts' and practitioners' judgment with overall score classified into a very good category and students' response classified into good categories; and (2) the reliability test showed a correlation coefficient of 0.8 so that it was declared as reliable. In short, the instrument based on Javanese gamelan Ethno-STEM instruments on Vibration and Sound Waves was appropriate to improve and optimize students' HOTS.

INTRODUCTION

Due to the obvious swift advancement of science and technology, educators must be flexible to meet new challenges. Aiming the students to apply their cognitive skills to solving problems in daily life, educators today must also be able to help students develop higher-order thinking skills (Priyani and Nawawi, 2020). Higher Order Thinking Skills (HOTS) is an ability that focuses on the ability to analyze and make the right decisions to solve a problem, so it is closely related to reasoning (Yuliandini, Hamdu and Respati, 2019).

Higher-order thinking trains students to think at a higher level (Sari *et al.*, 2019). In line with this opinion, Haryanti and Suwarma (2018) also stated that these 21st century skills consist of several specific skills that will support an individual to be able to face the challenges of the 21st century. There are four types of skills studied by the US-based Partnership for 21st Century Skills (P21), which are part of the known individual competencies with "the 4Cs"- communication,

collaboration, critical thinking, and creativity (Partnership for 21st Century learning, 2015). Ethno-STEM is an approach that can accommodate 4C (Sudarmin *et al.*, 2019) to improve students' HOTS in preparation for facing global competition in this century.

STEM allows students to learn academic concepts appropriately by applying four disciplines, i.e.: science, technology, engineering, and mathematics (Azalia, 2020). Ethno-STEM is STEM based on culture or local wisdom. Sudarmin *et al.* (2019) stated that the use of the term ethnoscience comes from the Greek "Ethnos", which means nation and Latin "Scientia", which means knowledge. So the notion of ethnoscience is related to community customs or local wisdom of the community. According to Kamid, Saputri, and Hariyadi, (2021) education and culture are something that cannot be avoided in everyday life because culture is a unified whole and comprehensive and applicable in society. Meanwhile, education is a basic need for every individual in the society.

Based on the previous opinion, ethno-STEM is a bridge between natural science and culture. Applying ethno-STEM as a learning approach would be very possible opportunity to connect the material studied and culture, so that student's understanding of the material becomes easier. This certainly helps educators as facilitators in learning to make it easier for students to understand the material.

The latest results of PISA achievement in 2018 showed that Indonesia was ranked 73rd out of 78 countries that took the PISA tests (Masfufah and Afriansyah, 2021). The test results showed that the performance of Indonesian students is still relatively low. Based on the results of an interview with a class VIII science teacher in Kendal regency, it was obtained that the application of the HOTS test instrument was rarely applied in SMP NU 05 Awwalul Hidayah Gemuh so the students still had difficulty understanding the content of the questions. It is easier for students to solve problems with low cognitive levels. However, when the HOTS questions were tested, their HOTS had not been trained because the assessment instruments used by the teacher were usually taken from various books or exam questions. According to Kusuma *et al.* (2017) HOTS ability of students is still relatively low because students are not familiar with the types of HOTS questions. Not many students can reason on complex problems (Nurwahidah, 2018). Therefore, students need to practice HOTS questions to develop thinking skills to a high level (Nurwahidah *et al.*, 2020).

Students may be given practice questions that invite them to think at a higher level (Devi, 2012). In line with Dewi and Riandi (2016); and Husnawati, Hartono, and Masturi (2019) who state that the lack of training of Indonesian children in completing tests or questions that require analysis, evaluation, and creativity is one of the causes of their low thinking ability. Training HOTS can be done by giving questions that connect the material with the real life of students to achieve meaningful learning (Widana, 2017). According to Krathworl & Anderson in Bloom's revised taxonomy, HOTS involving analysis (C4), evaluating (C5), and creating or creativity (C6) is considered HOTS (Nafiati, 2021).

Several studies on test instruments containing

ethno-STEM, such as on the design of test instruments containing ethnosience to measure the critical thinking ability of high school students stated that the ethnosience-charged test instrument developed was effective in improving critical thinking skills (Agustin, Susilogati and Addiani, 2018), whereas Kamid, Saputri and Hariyadi (2021) examined the development of the Jambi culture-based HOTS questions, which concluded that the questions were effective for increasing learning motivation in students because they linked the real world and local culture that had become a habit. In this study, the material chosen to be developed in the HOTS science test instrument was Vibration and Sound Waves based on Javanese gamelan instruments.

The introduction of gamelan as a cultural identity of the Indonesian nation is carried out through karawitan extracurricular activities (Nursulistiyo, 2019). This is, surely, very potential as a real learning tool to understand science learning. One of the popular traditional music in Indonesia even abroad is the Javanese gamelan, which has a unique set of musical instruments (Yuda and Azis, 2019). The majority of Javanese gamelan musical instruments are used by hitting, for example drum, *bonang*, *saron*, *kenong*, *kethuk*, *kempyang*, and *gong*. Through the activity of playing and observing gamelan instruments, students can examine the Vibration and Sound Wave material that has been taught.

The lack of teachers' information and knowledge on science content of Javanese gamelan musical instruments drives the need for the development of an ethno-STEM-based HOTS instrument. This study aims to develop a HOTS science test instrument based on ethno-STEM Javanese gamelan musical instruments for junior high school students, which is feasible to improve and optimize students' HOTS.

RESEARCH METHODS

The research developed referred to the Four-D (4-D) model of instrument development by Thiagarajan (1974). The 4-D development model consisted of 4 stages, i.e.: (1) define, (2) design, (3) develop, and (4) disseminate. The research procedure is presented in Figure 1.

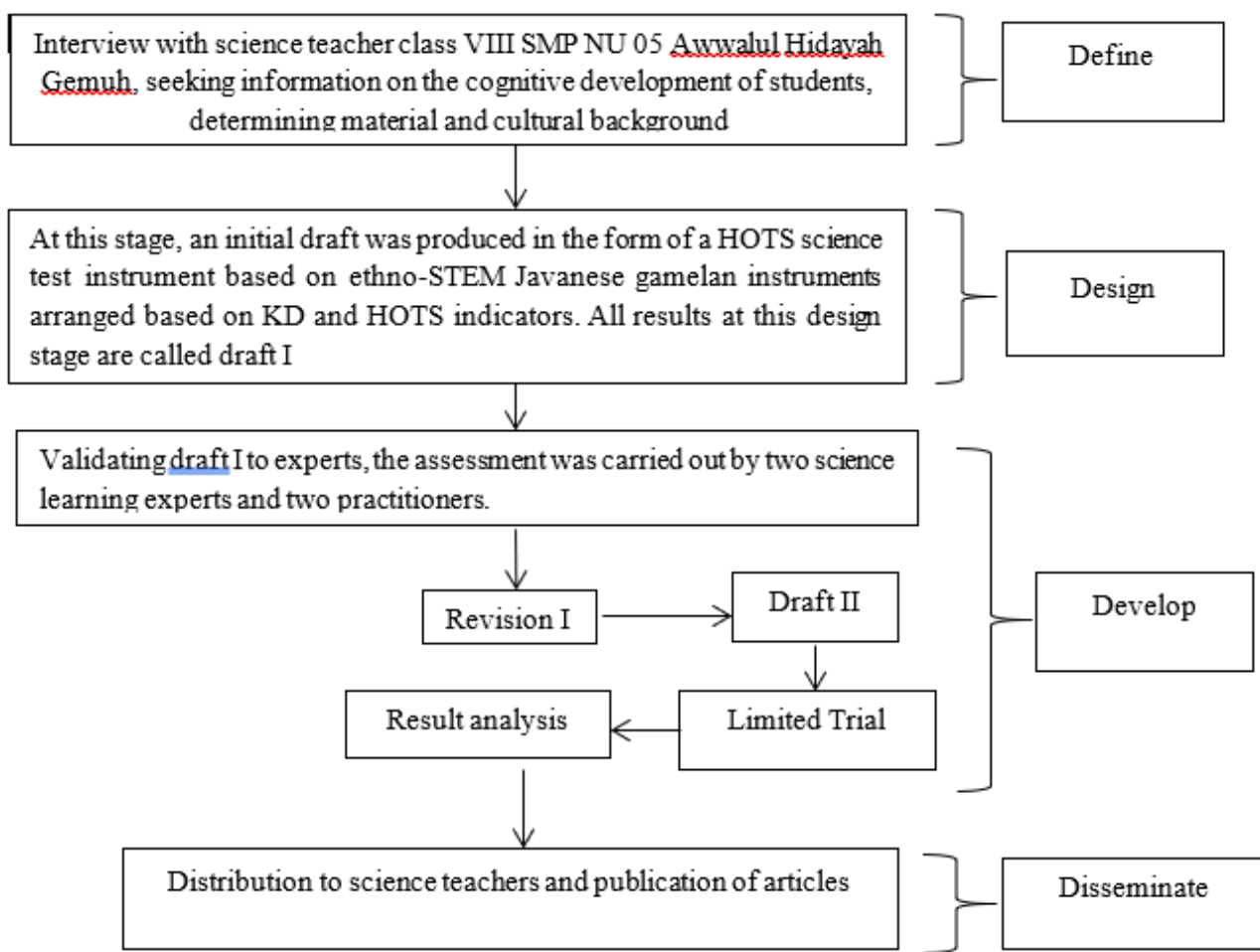


Figure 1. Product design of the HOTS science test instrument based on Javanese gamelan for junior high school students

This research was conducted at SMP NU 05 Awwalul Hidayah Gemuh starting from February to June 2022 with a population of Class VIII students. The sample was 24 students, consisting of 22 female students and 2 male students. The data was collected through data analysis of the validation of science learning experts, science learning practitioners, and data analysis of student responses, in the form of a questionnaire with a Likert scale that had a maximum score of 4.

Then, the data analysis was performed on the results of the student’s ability to do the HOTS science test based on ethno-STEM Javanese gamelan. The validation score data was tabulated for each aspect of the assessment. Then, the average score was calculated. It was then converted into interval data on a scale of four. The quantitative data was then interpreted with qualitative sentences.

The formula used to calculate the average total score of each aspect was as follows:

$$\text{average score} = \frac{\text{Total Points}}{\text{Total Score}} \quad (1)$$

The reference for converting scores to a four-quality scale of the HOTS science test instrument

from learning experts and practitioners is presented in Table 1.

Table 1. The conversion of the HOTS science test instrument quality assessment scores from learning experts and practitioners

No	Score Interval	Criteria
1	$X \geq \bar{X} + SBx$ $X \geq 3,00$	Very good
2	$\bar{X} + 1.SBx > X \geq \bar{X}$ $3.00 > X \geq 2.50$	Good
3	$\bar{X} > X \geq \bar{X} - 1.SBx$ $2.50 > X \geq 2.00$	Fair
4	$X < \bar{X} - 1.SBx$ $X < 2.00$	Poor

(Mardapi, 2008)

Where:

\bar{X} : (Maximal score + minimal score) \times $1/2$,

SBx : standard deviation of overall score.

In this study, the feasibility was determined with a minimum value with fairly good criteria. If the results of the assessment provided by science learning experts and practitioners showed more than the same final results with fairly good criteria, then the instrument was feasible for use in the evaluation of science learning. Then, the data on

student's response score on the quality of the HOTS science test instrument was analyzed using a descriptive percentage technique (Akbar and Hartono, 2017). The formula was as follows:

$$\% = \frac{\sum \text{average score}}{\sum \text{maximum score}} \times 100\% \quad (2)$$

The reference for the conversion of the percentage of student's response scores is presented in Table 2.

Table 2. The conversion of the quality assessment scores of the HOTS science test instruments from students

No	Score Interval (%)	Criteria
1	81 - 100	Very good
2	61 - 80	Good
3	41 - 60	Fair
4	21 - 40	Poor
5	0 - 20	Very Poor

(Riduwan, 2014)

Based on these criteria, the HOTS science test instrument based on ethno-STEM Javanese gamelan musical instruments for junior high school students was considered feasible if the percentage of students' response scores reached an average percentage of more than fairly good. Analysis results of the student's ability to work on questions was tested using the test of item validity, reliability, level of difficulty, and distinguishing power.

Pearson moment correlation test was used to test the validity of the items, meanwhile the reliability test was tested using the Alpha Cronbach formula, i.e. (Arikunto, 2018):

$$r_{11} = \left(\frac{n}{n-1} \right) \left(1 - \frac{\text{Total varians}}{\text{Total varians}} \right) \quad (3)$$

The calculation of the difficulty level for the descriptive questions used the following formula (Arikunto, 2018):

$$TK = \frac{\text{Every points avarage}}{\text{Maximum ideal score}}, \quad (4)$$

while the following formula was used to calculate the discriminatory power coefficient of the question (Arikunto, 2018):

$$DB = \frac{\text{upper mean} - \text{lower mean}}{\text{every points total ideal score}} \quad (5)$$

RESULTS AND DISCUSSION

The development of HOTS science test instruments based on ethno-STEM Javanese gamelan musical instruments for junior high school students has reached the final stage and meet very feasible criteria. The ethno-STEM-based science HOTS test instrument is a subjective

test in the form of a description of 10 items. This is divided into three higher-order thinking indicators proposed by Krathworl and Andreson in Bloom's revised Taxonomy (Krathwohl, 2001) i.e.: HOTS involving analysis (C4), evaluation (C5), and creativity (C6) consisting of 7, 2, and 1 questions, adapted to the subject of Vibration and Sound Waves. The results of the research refer to the development of the 4-D model developed by Thiagarajan (1974). The 4-D development model is a learning instruments consisting of 4 stages, i.e.: (1) define, (2) design, (3) develop, and (4) disseminate, which are explained as follows.

1) Define

In the defining stage, a preliminary study was conducted in the form of interviews with science teachers. It was obtained that the methods on science learning often used in the classroom were lectures, discussions, and memorization. Practicums are carried out several times. Only printed books were used as the source of learning. This is very unfortunate because the school facilities are quite adequate.

The assessment instrument used in the evaluation does not yet contain the HOTS indicator. So students often understand science concepts by remembering or memorizing formulas. The problem is that the HOTS test instrument has not been implemented, which involves a local wisdom approach in the form of Javanese gamelan. Karawitan is an extracurricular at SMP NU 05 Awwalul Hidayah Gemuh, which can be used as a means to understand a science concept besides books and practicums.

2) Design

In the designing stage, mapping is conducted on the basic competencies according to the subject of Vibration and Sound Waves, and HOTS competencies that must be mastered by students. Next, guidance is arranged for the questions of the HOTS science test instrument. In this study, higher order thinking dimensions used the revised Bloom's taxonomy, i.e.: analysis (C4), evaluating (C5), and creativity (C6) are considered HOTS (Nafiati, 2021).

Then, the HOTS science test instrument was arranged based on Javanese gamelan according to the question guidelines. The instrument in the form of a description consisted of 10 items and equipped with working instructions, consisting of general instructions and special instructions, and designing scoring guidelines. The product produced in the planning stage is called draft I. An example of the HOTS science test instrument guideline is presented in Table 3.

Table 3. Example of an ethno-STEM-based HOTS science test instrument grid

No Question	Basic competencies	Question indicator	Cognitive Level	Question
3	3..11 Analysing the concepts of vibration, waves, and sounds in everyday life, including the human hearing system and the sonar system in animals.	A simple harmonic vibration statement is presented with a vibration amplitude of 12 cm and a vibration period of 36 seconds, students can make a graph of the time deviation of the vibration.	C6 (Creating)	A <i>nayaga</i> beats with a <i>gong</i> and produces simple harmonic vibrations with amplitude of vibration of 12 cm and a period of vibration of 36 seconds. Make a graph of the deviation vs. time of the vibration!

3) Develop

This development stage aims to produce a revised HOTS science test instrument based on the suggestions of learning experts, practitioners, and students. The assessments from learning experts and practitioners become the basis for making revisions so that the products developed are feasible to continue to be tested. Furthermore, draft I was validated by two learning experts and two science learning practitioners.

The ethno-STEM-based HOTS test instrument in terms of content components consists of 20 assessment indicators consisting of 5 material components, 4 ethno-STEM components, 5 construction components, 5 grammatical components, and 1 time component. Then, the score obtained is converted into a value. The total scores of the ethno-STEM-based HOTS science instrument validation scores from learning experts are presented in Table 4.

Table 4. Science learning expert scores conversion result

Aspect	N	Total score	Total score	Average score	Criteria
Theory (Subject)	5	20	17.5	3.5	Very good
Ethno-STEM	4	16	14	3.5	Very good
Construction	5	20	17.5	3.5	Very good
Grammar	5	20	16.5	3.3	Very good
Time	1	4	3.5	3.5	Very good

Based on Table 4, in assessing the quality of the ethno-STEM-based HOTS science test instrument for the subject aspect with a total score of 20, the total score was 17.5 and the average score for each indicator was 3.5, which is in the very good criteria. In the ethno-STEM aspect with a total score of 16, the total score is 14 and the average score of each indicator was 3.5 in the very good criteria. In the construction aspect with a total score of 20, an average score of 17.5 is obtained and the average score for each indicator is 3.3 in the very good criteria. Then on the grammatical aspect with a total score of 20, the total score obtained is 16.5 and the average score on each indicator is 3.5 in the very good criteria. In the aspect of time with a total score of 4, the total

and average scores are 3.5, which is in the very good criteria. The quality of the HOTS science test instrument is declared feasible for use in learning because the average score of each indicator reached the very good criteria or passed the minimum reliability limit score.

The validation of the science test instrument of ethno-STEM-based HOTS is also carried out by two expert practitioners as users, a teacher of class VIII, and a science teacher of class IX at SMP NU 05 Awwalul Hidayah Gemuh. The results of the expert and practitioner assessment scores regarding the quality of the ethno-STEM-based HOTS science test instrument are presented in Table 5.

Table 5. Science learning practitioners scores conversion results

Aspect	N	Total Score	Total score	Average score	Criteria
Theory	5	20	20	4	Very good
Ethno-STEM	4	16	16	4	Very good
Construction	5	20	19	3.8	Very good
Grammar	5	20	19	3.8	Very good
Time	1	4	4	4	Very good

Based on Table 5, the assessment of the ethno-STEM-based HOTS science test instrument for the theory aspect with a total score of 20 gives a total score of 20 and an average score of 4 for each indicator in very good criteria. In the ethno-STEM aspect with a total score of 16, the total score obtained is 16 and the average score for each indicator is 4 in very good criteria. In the construction aspect with a total score of 20, a total score of 19 is obtained and the average score for each indicator is 3.8 in very good criteria. In the grammatical aspect with a total score of 20, a total score of 19 is obtained and the average score for each indicator was 3.8 in very good criteria. Then, the aspect of time with a total score of 4 produces a total score of 4 in the very good criteria. Thus, the ethno-STEM-based HOTS science test instrument based on the practitioner's assessment is declared feasible because it passes the minimum reliability limit score.

The assessment of the ethno-STEM-based HOTS Science test instrument is also carried out by distributing a student's response questionnaire of 10 people containing 20 statements with the following answer choices, i.e.: strongly agree (4), agree (3), disagree (2), and strongly disagree (1). The results of the student's response scores are then converted into percentages. The conversion of student response percentage scores is presented in Table 6.



Table 6. Conversion results of the percentage score of student responses

n	SM	R	PS = $\frac{\sum R}{SM} \times 100\%$	Criteria
20	80	60	75%	Good

Based on Table 6, the results of the distribution of student's responses with a maximum score of 80, produces an average score of 60 of 10 students. So the percentage score obtained is 75% in the good criteria and is declared feasible because it passes the minimum reliability limit score. Thus, the validation results from science learning experts, practitioners, and student's responses indicate that the ethno-STEM-based HOTS science test instrument of Javanese gamelan musical instruments is feasible for use by teachers in evaluating the learning process to measure students' HOTS. The results of the revision according to the suggestions provided by the validators of learning experts, practitioners, and students are called as draft II. An example of the revision of draft I is presented in Table 7.

After the improvements according to revisions and suggestions from experts, draft II is tested in a limited way with a sample of 24 people consisting of 22 female students and 2 male students of class VIII SMP NU Awwalul Hidayah Gemuh who are randomly selected.

Table 7. Examples of revisions by science learning experts

Number of question	Comments	Revised Results
1	1) Adding sentences that support the article. 2) The description on the picture is not correct	1) Correction of sentences that support the article. 2) Improvements of the image captions.
4	 <p>Gambar 3. Seseorang meniup seruling</p> <p>Sumber : repository.kemendikbud.go.id</p> <p>Seseorang memainkan seruling yang memiliki kolom udara terbuka. Pada kedua ujung seruling tersebut memiliki nada atas kedua dengan frekuensi 1700 Hz. Jika kecepatan suara di udara adalah 340 m/s, maka bagaimana menentukan panjang serulingnya ?</p>	 <p>Gambar 3. Sekelompok orang dalam grup musik sedang meniup seruling</p> <p>Sumber : repository.kemendikbud.go.id</p> <p>Beberapa orang memainkan seruling yang memiliki kolom udara terbuka Pada kedua ujung seruling tersebut memiliki nada atas kedua dengan frekuensi 1700 Hz. Jika kecepatan suara di udara adalah 340 m/s, maka bagaimana menentukan panjang serulingnya ?</p>
5.9	3) Incorrect punctuation.	3) Fixed punctuation.

The results of the scoring on each student's answer are then tested for validity, reliability, level of difficulty, and differentiating power. The results of the validity test of the HOTS science test instrument and its criteria are presented in Table 8.

Table 8. The results of the item validity test

Criteria	Question Number
Invalid	1 - 4,7 - 10
Valid	5,6

In the validity test with the Pearson moment correlation test, the criteria for valid or invalid of each item in the instrument may be obtained from the comparison between the Pearson correlation coefficient (r) and the Pearson moment value table (r_{table}). If the value of $r_{count} > r_{table}$, then, the test item is declared as valid (Arikunto, 2018). With a sample of 24 people and a significance level of 0.05, the r_{table} is 2.07. Based on Table 8, the items numbered 1 to 4 and 7 to 10 were declared invalid. The valid items are 5 and 6.

Table 9. Reliability test results

n	Total variance $\sum \sigma_1^2$	Total variance σ_1^2	Reliability coefficient r11	Criteria
10	1960.11	6257.8	0.8	Very high

Furthermore, the reliability test is carried out. A measuring instrument is declared to be reliable if the measuring instrument is consistent. The instrument can be used several times to measure

the same object, and produces the same data. The results of the reliability test are presented in Table 9.

Based on Table 9, the reliability coefficient of 0.8 is obtained in very high criteria. Then, the HOTS science test instrument based on ethno-STEM is reliable. A description test item is reliable if the instrument reliability coefficient is higher than 0.5 (Hidayat *et al.*, 2017).

After carrying out the reliability test, it calculates the level of difficulty. The comparison between the number of students who answered correctly a test number and the number of students who answered the test number is known as the difficulty index or level of difficulty (Fitriarosah, 2016). The results of the difficulty level test are presented in Table 10.

Table 10. Difficulty level test results

Criteria	Question Number
Moderate	1,2,4,5,6,8
Hard	3,7,9,10

Based on the data of the difficulty level test in Table 10, these ten questions include questions that are classified as medium and difficult. The moderate questions are 1,2,4,5,6, and 8. The questions that are difficult consist of number 3,7,9, and 10.

The steps for the discriminatory power test begin by sorting the total sample score from the highest to the lowest scores, then taken from 27% of the upper group students and 27% of the lower group students, with a total sample of 24. Then, 10 students have an average score of the upper group,

while 10 students have the average score of the lower group. The discriminatory power test is presented in Table 11.

Table 11. Discriminatory power test results

Criteria	Question number
Fair	1,5,6
Bad	2,3,4,7,8,9
Very Bad	10

Based on the results of the discriminatory test of the ethno-STEM-based HOTS science test instrument, three questions have distinguishing power with fair criteria for question numbers 1,5, and 6. Six questions with bad criteria are question numbers 2,3,4,7,8, and 9. Finally, question number 10 is in very bad criteria.

4) Disseminate

At the dissemination stage, the HOTS science test instrument based on the Javanese gamelan ethno-STEM is distributed to science teachers and also published in the form of articles.

Based on the analysis, the indicator of the ability to think at a higher level is very low. The low level of HOTS in the research results may also be caused by the difficulty level of the questions and the number of sample, which only consists of one group. The questions tested have moderate and difficult levels. In addition, this can also be caused by several factors that may be very influential, for example, students are not used to working on questions that train HOTS in collaboration with local wisdom.

So exercises in working on the ethno-STEM-based HOTS questions need to be performed so that efforts to improve HOTS in students can be consider optimal. In accordance with the research reviewed by Lestari, Ashari, and Nurhidayati (2021) students gave a positive response to the STEM-based physics learning outcome test instrument, which is practical to use because it could be carried out in limited trials. Marvia and Rahmawati (2020) stated that the analysis of the results of HOTS showed that the ability of students to answer the questions of HOTS is still low. This is because students had not been trained in answering questions with cognitive levels of C4 - C6.

Research carried out by Kamid, Saputri, and Hariyadi (2021) regarding the development of test instruments of Jambi culture of the flat-sided building subject in class VIII produces the following results. The development of HOTS questions increases students' learning motivation because the questions connect the subject matter with the real-world context, especially Jambi

culture so that learning becomes meaningful. Jambi culture, which is included in the questions, is used as a stimulus for the questions to attract the students in answering the questions.

The questions in the ethno-STEM-based science test instrument of Javanese gamelan musical instruments present information in the form of texts, pictures, and graphics related to Vibration and Sound Waves. This can help students connect the concepts of the material that has been taught. So they no longer rely on memorizing formulas but use HOTS in answering questions. The ethno-STEM approach has not been widely developed, while the STEM approach has been developed in many countries, even in Indonesia (Sudarmin *et al.*, 2019). The product of the HOTS science test instrument based on the ethno-STEM Javanese gamelan as a result of this development can be used as an alternative to measure HOTS in students and reintroduce local wisdom that live around them, especially Javanese gamelan, which is very potential as a learning tool.

CONCLUSION

Based on the results of research and development regarding the ethno-STEM-based HOTS science test instrument for Javanese gamelan musical instruments, the following conclusions can be drawn. The stages of development carried out are define, design, develop, and disseminate. The assessments from learning expert and practitioner validators in terms of material (theory), ethno-STEM, construction, grammar, and time aspects produce higher average scores than the minimum average score. The results of the students' response scores are in a good category. Moreover, the reliability test shows that the test instrument is declared reliable. Thus the HOTS science test instrument based on ethno-STEM Javanese gamelan musical instruments is feasible to improve and optimize students' HOTS.

ACKNOWLEDGEMENT

We would like to thank the Ministry of Education, Culture, Research and Technology and the Education Fund Management Institute (LPDP). This study is a part of the authors' independent scientific research for 2021-2022, which is funded by the Ministry of Education and Culture and LPDP.

REFERENCES

- Agustin, N., Susilogati, S. & Addiani, K. (2018). Desain Instrumen Tes Bermuatan Etnosains untuk Mengukur Kemampuan Berpikir Kritis Siswa SMA. *Jurnal Inovasi Pendidikan Kimia*, 12 (2), 2159–2169.

- Akbar, F. I. & Hartono, R. (2017). Pengembangan Lembar Kegiatan Peserta Didik dengan Model Pengembangan 4-D pada Materi Mitigasi Bencana dan Adaptasi Bencana Kelas X SMA. *Jurnal Pendidikan Geografi*, 22 (2), 135–147. doi: 10.17977/um017v22i22017p135.
- Arikunto, S. (2018). *Dasar-dasar Evaluasi Pendidikan. Edisi ketiga Cetakan Pertama*. Jakarta: PT Bumi Aksara.
- Azalia, I. (2020). Pengaruh Penerapan E-Book Bermuatan STEM Terintegrasi Etnosains Terhadap Keterampilan Generik Sains Peserta Didik Pada Materi Kesetimbangan Kimia. *Skripsi*, Tidak dipublikasikan. Universitas Negeri Semarang.
- Devi, P. K. (2012). Pengembangan Soal ‘Higher Order Thinking Skill’ Dalam Pembelajaran IPA SMP/MTs. *Jurnal Pendidikan IPA*, 2 (2), 32–40.
- Dewi, N. & Riandi, R. (2016). Analisis Kemampuan Berpikir Kompleks Siswa Melalui Pembelajaran Berbasis Masalah Berbantuan Mind Mapping. *Jurnal Edusains*, 8 (1), 98–107. doi: 10.15408/es.v8i1.1805.
- Fitriarosah, N. (2016), Pengembangan Instrumen Berfikir Kreatif Matematis Untuk Siswa SMP. *Prosiding Seminar Nasional Pendidikan Matematika*, Universitas Kanjuruhan Malang 1(1997), 243–250.
- Haryanti, A. & Suwama, I. R. (2018) Profil Keterampilan Komunikasi Siswa Smp Dalam Pembelajaran IPA Berbasis Stem. *WaPFI (Wahana Pendidikan Fisika)*, 3 (1), 49. doi: 10.17509/wapfi.v3i1.10940.
- Hidayat, S. R. *et al.* (2017). Pengembangan Instrumen Tes Keterampilan Pemecahan Masalah pada Materi Getaran, Gelombang, dan Bunyi. *Jurnal Penelitian & Pengembangan Pendidikan Fisika*, 3 (2), 157–166. doi: 10.21009/1.03206.
- Husnawati, A., Hartono, H. & Masturi, M. (2019). Pengembangan Soal Higher Order Thinking Skill (HOTS) Fisika Kelas VIII SMP Materi Gerak Pada Benda. *UPEJ Unnes Physics Education Journal*, 8 (2), 133–140.
- Kamid, K., Saputri, R. & Hariyadi, B. (2021). Pengembangan Soal Higher Order Thinking Skills Berbasis Budaya Jambi. *Jurnal Cendekia : Jurnal Pendidikan Matematika*, 5 (2), 1793–1806. doi: 10.31004/cendekia.v5i2.678.
- Krathwohl, L. W. (2001). *A Taxonomy for Learning, Teaching, and Assesing: A Revision’s Taxonomy of Educational Objectives*. New York: David McKay Company.
- Kusuma, M. D. *et al.* (2017). The development of higher order thinking skill (HOTS) instrument assessment in physics study. *IOSR Journal of Research & Method in Education (IOSRJRM)*, 7 (1), 26–32. doi: 10.9790/7388-0701052632.
- Lestari, R. P., Ashari & Nurhidayati. (2021). Pengembangan Instrumen Tes Hasil Belajar Fisika Berbasis STEM untuk Mengukur Keterampilan Proses Sains Peserta Didik. *Jurnal Inovasi Pendidikan Sains (JIPS)*, 2 (1), 18–24.
- Mardapi, D. (2008). *Teknik penyusunan instrumen tes dan non tes*. Yogyakarta: Mitra Cendikia.
- Marvia, A., & Rahmawati, D. (2020). Validitas Instrumen Tes Berpikir Tingkat Tinggi (HOTS) pada Materi Sistem Respirasi di Kelas XI SMA. *Mangifera Edu*, 4 (2), 129–142. doi: 10.31943/mangiferaedu.v4i2.83.
- Masfufah, R., & Afriansyah, E. A. (2021) Analisis Kemampuan Literasi Matematis Siswa melalui Soal PISA. *Mosharafa: Jurnal Pendidikan Matematika*, 10 (2), 291–300. doi: 10.31980/mosharafa.v10i2.825.
- Nafiati, D. A. (2021). Revisi taksonomi Bloom: Kognitif, afektif, dan psikomotorik. *Humanika*, 21 (2), 151–172. doi: 10.21831/hum.v21i2.29252.
- Nursulistiyo, E. (2019). Pemanfaatan Siter, Kendang, Saron, Kenong, dan Gender sebagai media pembelajaran fisika. *Jurnal Riset dan Kajian Pendidikan Fisika*, 6 (1), 5. doi: 10.12928/jrpkpf.v6i1.13393.
- Nurwahidah, I. (2018). Pengembangan soal penalaran model TIMSS untuk mengukur high order thinking (HOT). *Thabiea : Journal of Natural Science Teaching*, 1 (1), 20–29. doi: 10.21043/thabiea.v1i1.3874.
- Nurwahidah, I. *et al.* (2020) Development of science test to measure HOTS and digital literacy of junior high school students on the topic of city noise. *Edusains*, 12 (2), 203–213.
- Partnership for 21st Century learning. (2015). 21st CENTURY STUDENT OUTCOMES, 1–9. available on: <http://www.p21.org/our-work/p21-framework>.
- Priyani, N. E. & Nawawi, N. (2020). Pembelajaran Ipa Berbasis Ethno-Stem Berbantu Mikroskop Digital Untuk Meningkatkan Keterampilan Proses Sains Di Sekolah Perbatasan. *WASIS: Jurnal Ilmiah Pendidikan*, 1 (2), 99–104. doi: 10.24176/wasis.v1i2.5435.
- Riduwan. (2014). *Metode dan teknik menyusun proposal penelitian*. Bandung: Alfabeta.
- Sari, Y. *et al.* (2019). Meningkatkan kemampuan menyusun soal IPA berorientasi HOTS bagi

- guru Sekolah Dasar Gugus Pandanaran Dabin IV UPTD Semarang Tengah. *Indonesian Journal of Community Services*, 1 (2), 175. doi: 10.30659/ijocs.1.2.175-183.
- Sudarmin, S. *et al.* (2019). Implementing the model of project-based learning: integrated with ETHNO-STEM to develop students' entrepreneurial characters. *Journal of Physics: Conference Series*, 1317 (1). doi: 10.1088/1742-6596/1317/1/012145.
- Thiagarajan, S. *et al.* (1974). *Instructional Development for Training Teachers of Exceptional Children A Sourcebook*. Idiana Uni.
- Widana, I. W. (2017). Higher order thinking skills assessment (HOTS). *JISAE*, 3 (1), 32–44. doi: 10.21009/JISAE.031.04.
- Yuda, Y. P. & Azis, M. N. L. (2019). 3D modeling the gamelan of saron as a documentation of cultural heritage preservation efforts. *Journal of Physics: Conference Series*, 1375 (1). doi: 10.1088/1742-6596/1375/1/012036.
- Yuliandini, N., Hamdu, G. & Respati, R. (2019) Pengembangan Soal Tes Berbasis Higher Order Thinking Skill (HOTS) Taksonomi Bloom Revisi di Sekolah Dasar. *Pedadiktika: Jurnal Ilmiah Pendidikan Guru Sekolah Dasar*, 6 (1), 37–46.