JOURNAL OF INNOVATION AND LEARNING
Volume 1 Nomor 1 2022, pp 43-54
DOI: http://dx.doi.org/10.23887/jil.v1i1.44997

Instrument Evaluation based on STEAM (Science, Technology, Engineering, Art, and Mathematics) for Fifth Grade

Ajeng Wulan Trisna Dewi[1*], I Gusti Agung Ayu Wulandari[2], Gusti Ngurah Sastra Agustika[3]
[1,2] Prodi Pendidikan Dasar, Universitas Pendidikan Ganesha, Singaraja, Indonesia
*Corresponding author: ajeng.wulan@undiksha.ac.id

Abstrak

Masa transisi dari pembelajaran daring ke pembelajaran luring terbatas perlu penyesuaian ulang, contohnya waktu belajar. Tujuan dari penelitian ini adalah membuat rancang bangun dan menganalisis kualitas instrumen evaluasi berdasarkan validitas, reliabilitas, tingkat kesukaran, daya beda, dan efektivitas pengecoh. Penelitian ini merupakan penelitian pengembangan dengan model ADDIE. Subjek penelitian adalah ahli isi mata pelajaran, ahli desain, ahli media, dan 25 siswa kelas V. Pengumpulan data menggunakan metode tes dan nontes. Data dianalisis menggunakan teknik analisis kualitatif dan kuantitatif. Hasil penelitian: (1) rancang bangun pengembangan instrumen evaluasi, (2) Hasil uji kualitas soal yang berkategori cukup baik berdasarkan analisis soal yang mendapatkan 20 soal valid, reliabilitas sangat tinggi, tingkat kesukaran perangkat soal 0,58 yang berkriteria sedang, daya beda perangkat soal 0,20 yang berkategori cukup baik, dan 23% pengecoh berkategori cukup baik, 20% berkategori baik, 17% berkategori kurang baik, 17% berkategori jelek, dan 23% berkategori sangat jelek. Jadi dapat disimpulkan bahwa instrumen evaluasi yang dikembangkan dapat menjadi variasi dalam proses evaluasi.

Kata Kunci: Evaluasi, Instrumen, STEAM.

Abstract

The transition period from online learning to offline learning is limited to need to be readjusted, for example study time. The purpose of this research is to design and analyze the quality of evaluation instruments based on validity, reliability, level of difficulty, discriminating power, and effectiveness of distractors. This research is a development research with ADDIE model. The research subjects were subject content experts, design experts, media experts, and 25 fifth grade students. Data were collected using test and non-test methods. Data were analyzed using qualitative and quantitative analysis techniques. The results of the study: (1) the design and development of evaluation instruments, (2) the results of the quality test of questions that are categorized as good enough based on the analysis of questions that get 20 valid questions, very high reliability, the level of difficulty of the test equipment is 0.58 which has moderate criteria, the different power of the device 0.20 items were categorized as good enough, and 23% of distractors were categorized as good enough, 20% were categorized as good, 17% were categorized as poor, 17% were categorized as bad, and 23% were categorized as very bad. So it can be concluded that the evaluation instrument developed can be a variation in the evaluation process.

Keywords: *Evaluation, Instrument, STEAM.*

History:
Received : November 26, 2021
Revised : December 06, 2021
Accepted : January 08, 2022
Published : February 25, 2022

Publisher: Undiksha Press

1. INTRODUCTION

The evaluation instrument that used since online learning due to the corona pandemic is an online instrument. Online or virtual learning is carried out with the help of communication technology such as smartphone and laptops as well as internet networks(Dewi & Sadjiarto, 2021; Wulandari & Agustika, 2020). Based on an unstructured interview with the teacher of fifth grade SD No. 3 Sibanggede, when learning online the media that are often used are Google Classroom, Zoom Meeting, and Google Form. Google Classroom as a means to provide learning materials and assignments, Zoom Meeting is used as a face-to-face learning tool to explain material that students do not understand, and Google Form is often used as an evaluation instrument. In addition to using Google Forms, teachers also provide evaluations in Google Classroom and ask students to do handwritten questions and then collected in the form of photos in Google Classroom (Chick et al., 2020; Fitriyani &

Astuti, 2021). This of course aims to keep students writing. However, this is a little difficult for the teacher. In which the teacher must correct the students' work one by one. In addition, many students are also lazy to collect assignments so that many of them do not have grades. Many also think that the evaluation process is not important so they do not do the tasks given. Though evaluation is something that must be done in a lesson. Evaluation of the cognitive domain or knowledge competence is useful to determine the level of achievement of learning objectives, the level of success of learning that has been carried out, and also provides an overview of student learning difficulties (Setemen, 2010).

Students who do not work on evaluation questions will hinder the course of learning. The evaluation process is very important to determine the achievement of learning objectives (Haryati et al., 2021). The most important part of learning to determine the quality of learning and measure students' skills is the evaluation process (Hamidah & Wulandari, 2021). Based on previous research, it was found that the evaluation tools used by teachers during online learning have not fully utilized technology-based applications that can increase students' motivation to do it (Kurnia et al., 2022). Teachers are required to make and use a tool to make learning more interesting and students become active and easy to remember (Oktariyanti et al., 2021). Google Form is considered to have been used too often so that students feel boring.

Previous research regarding the development games with Wordwall found that games with Wordwall were considered very practical and suitable for use in the learning process (Oktariyanti et al., 2021). Wordwall is a website for education that has several choices of evaluation games in it, one of which is multiple choice (Çil, 2021). Previous research has shown that Wordwall easier for students to understand the material and makes it easier for teachers to find out the results of student achievement (Wafiqni & Putri, 2021).

To support this, we need an approach that can be used to develop learning plans that will be held. The STEAM approach is one approach that can be applied. The STEAM approach combines Science, Technology, Engineering, Arts, and Mathematics in the learning process with the aim that students easily understand concepts, explore their potential and can apply the concepts learned in everyday life (Dwi Sari & Setiawan, 2020). Learning with the STEAM approach can improve critical thinking skills which is one aspect of the cognitive process needed to overcome problems in the 21st century, Birgili (Conradty & Bogner, 2018; Komala & Rohmalina, 2021). The STEAM approach is also considered suitable for use in elementary school education units. This is because the STEAM approach integrates several disciplines, as well as thematically so that the materials studied are not rigidly restricted (Estriyanto, 2020). Based on research conducted by Bedar and Al-Shboul in 2020, there is a significant positive effect on students' computational abilities after learning using the STEAM approach (Bedar & Al-Shboul, 2020). Computational thinking is a process of solving problems using logic gradually and systematically that is needed by students in various fields (Angeli & Giannakos, 2020; Lee et al., 2014).

Based on those statements, a development of a STEAM-based evaluation instrument with Wordwall was made on the social science knowledge competence of fifth grade elementary school students. 3 Sibanggede to be an innovation or variation in the evaluation process so that students are interested in working on evaluation questions given by the teacher.

2. METHOD

This research is a development research with ADDIE model. The subjects of this study were subject content experts, design experts, media experts, and fifth grade elementary school students. This development research uses the ADDIE model. In this study, the ADDIE

model was used to develop products in the form of interactive evaluation instruments in social studies learning Theme 2 Subtheme 3 Learning 3 for fifth grade elementary school students. The product developed in this research is an interactive evaluation instrument. The ADDIE development research model was used in this study because the steps were easy to understand and could be used to develop interactive evaluation instruments. In addition, this ADDIE model contains evaluations at each stage which can minimize errors in product development of this interactive evaluation instrument. These stages are Analysis (Analysis), Design (Design), Development (Development), Implementation (Implementation), and Evaluation (Evaluation) (Rasvani & Wulandari, 2021).

This study uses 2 methods to collect data for research purposes, namely the test and non-test methods. The test method used in this study is an objective test. Objective tests are given to students to obtain initial data in order to calculate the quality of the questions based on validity, reliability, discriminating power, level of difficulty, and effectiveness of distractors. While the non-test methods used in this study were observation, interviews, questionnaires and questionnaires. Observations were made during the learning process to determine the initial condition of the fifth grade students. Then unstructured interviews were conducted with the homeroom teacher to discuss developments or obstacles that occurred during the transition from online learning to Limited Face-to-face Learning (PTMT). Then a questionnaire or questionnaire is also used.

The instrument used in this study was a questionnaire instrument or a questionnaire sheet and also a test instrument. Questionnaires are used to collect data related to the feasibility of the developed product. The questionnaires will be filled out by experts, namely subject content experts, design experts, and media experts while the test instruments are given to fifth grade students of SD No. 3 Sibanggede. The following is a grid used in the questionnaire instrument and also the test instrument.

Table 1. Grid of Content Expert (Suartama, 2016)

Aspects	Indicator	Number of items
Material	1) The material motivated students	5
	2) Interesting material	6
	3) Accuracy of material	1, 2, 3
	4) Conformity between objectives and learning materials	4
	5) Material is useful for students in real life	7
	6) Easy to understood	8
	7) Help students remember the abilities and knowledge of previous	13
Media	8) Use of pictures that support learning materials	10
	9) Easy to use	9
	10) Instruction are available	12
Evaluation	11) An evaluation is given to measure the ability of students	15
	12) Pemberian umpan balik	11, 14

Table 2. Grid of Design Expert (Suartama, 2016)

Aspects	Indicator	Number of items
Purpose	1) Clarity of general learning objectives	1
	2) Appropriateness of learning indicators	2
	3) Clarity of objectives 3 Strategy	3

Aspects	**Indicator**	**Number of items**
Strategy	4) Clear usage instructions	4, 5
	5) Presentation of interesting material	6
	6) Appropriate presentation of menus	7
Evaluation	7) Evaluation is given to measure the ability of students	9, 10
	8) Feedback	8

Table 3. Grid of Media Expert (Suartama, 2016)

Aspects	Indicator	Number of items
Display	1) Product display quality	1
	2) Design suitability with student characteristics	2
Text	3) Font size suitability	3
	4) Text readability	4
Picture	5) Appropriate image use	5
	6) Image quality	6
	7) Image placement	7
Techniques	8) There are navigation buttons	8
Operational	9) Consistency of button location	9
Color	10) Background color compatibility	10
	11) Image color clarity	11
	12) Font color compatibility	12
	13) A harmonious and balanced color display	13
Audio	14) Appropriate use of sound effects	15

Product trials are carried out with the assessment of experts. Experts who rate this product are subject content experts, design experts and media experts. In addition, the items that will be used in this evaluation instrument are questions that have gone through item analysis to determine the quality of the questions. Some of the tests carried out were validity, reliability, discriminatory, level of difficulty, and distractors. Validity is the accuracy and precision of a measuring instrument, in this case an evaluation instrument (Haryanto, 2020). The validity used is theoretical validity and empirical validity. Theoretical validity is carried out by the supervisor. Meanwhile, empirical validity was measured using the biserial point correlation technique (Koyan, 2007). The criteria used in the test validation is to compare thecalculated with rtable at a significant level of 5%. The test is said to be valid if the rcount is more than rtable at a significant level of 5%.

After calculating the validity of the test, then the reliability of the test is calculated. Test reliability aims to determine the consistency of a test. A test is said to have a high level of confidence if the test can provide constant or steady results. The reliability of the learning outcomes test instrument can be determined by using the Kuder Richadson 20 (KR-20) formula (Koyan, 2007). After calculating the reliability of the items, the item difficulty level is calculated. The calculation of the difficulty level of the question is a measurement of the degree of difficulty of a question. The difficulty level ranges from 0% to 100%, but is more often written as a proportion with a value of 0.00 to 1.00 (Basuki & Hariyanto, 2017). How to calculate the level of difficulty of the test items is as follows.

$$P = \frac{nB}{n}$$

The method of calculating the level of difficulty of the test equipment is as follows.

$$Pp = \frac{\Sigma P}{n}$$

After calculating the level of difficulty of the test items, the discriminatory power of the questions is calculated. The discriminatory power of questions is the ability of the questions to distinguish high-ability students from low-ability students (Arikunto, 2016). The higher the discriminating power coefficient of an item, the more capable the item is to distinguish between students who master competence and students who lack competence (Ratnawulan & Rusdiana, 2014).

$$D_B = \frac{nB_A}{nA} - \frac{nB_B}{n_B}$$

If the value of discriminating power is negative, it means that the question is very bad and must be discarded. A good test according to Fernandes (in Koyan 2007), if it has a discriminatory power between 0.15 – 0.20 or more. To test the quality of the questions, the last thing to do is test the effectiveness of distractors or distractors. Multiple-choice or objective questions have alternative answers (options) that function as distractors. Distractors are not just a complement, but are deliberately given to make students fooled. Therefore, a good distractor is one that is made as similar or as similar as possible to the answer key so that it is chosen by students who are less (Basuki & Hariyanto, 2017; Umi Fatimah & Alfath, 2019).

Good items, distractors will be chosen evenly by students who answer wrongly. On the other hand, items with less good distractors will be chosen unequally. The distractor is considered good if the number of students who choose the distractor is the same or close to the ideal number. The distractor index is calculated by the following formula.

$$\text{IP} = \frac{P}{(N-B)/(n-1)} x100\%$$

If all students answer correctly on an item, then IP = 0 which means the question is not good or can be said to be bad. Thus the distractor doesn't work. Analyze for the validity of the expert test, using a qualitative descriptive analysis. Quantitative descriptive analysis is a method of data processing that is carried out by systematically compiling in the form of numbers and or percentages, regarding an object under study, in order to obtain general conclusions (Agung, 2018). In this study, quantitative descriptive analysis was used to process the data obtained through a questionnaire in the form of Scorees.

Table 4. Criteria for Percentage Scoree (Basuki & Hariyanto, 2017)

Statement	Positif	Negatif
Strongly Agree	4	1
Agree	3	2
Disagree	2	3
Strongly Disagree	1	4

These Scorees will be converted into percentages of each subject. The formula used to calculate the percentage is:

$$P = \frac{\Sigma R}{N} \times 100\%$$

Table 5. Conversion of Achievement Levels with a Scale of 4 (Tegeh & Kirna, 2010)

No.	Achievement Level (%)	Qualification	Description
1	90 - 100 %	Very Good	No need to revise
2	75 - 89 %	Good	Slightly revise
3	65 - 74 %	Enough	Sufficiently revised
4	55- 64 %	Less	Many things were revised
5	0 - 54 %	Very Poor	Repeated product

In particular, content expert validation uses the Gregory formula to find out the results. The mechanism for using the Gregory formula still uses expert judgment, but the results of the assessment have been quantified. The use of the Gregory formula begins with an assessment by experts, then the Scorees are grouped, and the results of the assessment are tabulated in a matrix form such as the following table (Ayuningsih, 2020).

Table 6. Gregory Validity Tabulation Matrix

1st expert		2nd expert		3rd expert	
Less Relevant (Score 1-2)	Very Relevant (Score 3-4)	Less Relevant (Score 1-2)	Very Relevant (Score 3-4)	Less Relevant (Score 1-2)	Very Relevant (Score 3-4)

Table 7. Cross Tabulationof Gregory's Validity with Two Appraisers

		1st Expert	
		Less relevant (Score 1-2)	Very relevant (Score 3-4)
2nd Expert	Less relevant (Score 1-2)	(A) 0	(B) 0

3. RESULT AND DISCUSSION

Result

This study developed an evaluation instrument used in the evaluation process by teachers. The design and development of this evaluation instrument uses the ADDIE development model in which there are 5 stages, namely the analysis stage, design stage, development stage, implementation stage, and evaluation stage. The Analysis stage is the first stage in the ADDIE development model. The analysis is carried out so that the resulting product can be useful for schools in carrying out learning. At this stage, an analysis of the problems experienced by the school, especially in class V, is carried out with respect to evaluation instruments commonly used in the implementation of learning. Three analyzes were carried out, namely an analysis of initial conditions, a needs analysis, as well as an analysis of basic competencies and indicators that will be the focus of the evaluation instrument.

In the analysis of the initial conditions, observations were made during learning and interviews with fifth grade teachers at SD No. 3 Sibanggede to be able to find out the initial conditions of the learning process carried out online. An analysis of the initial conditions is needed so that researchers can adjust the learning evaluation instruments developed to the conditions of the research place and also the students' circumstances. It was found that students carried out online learning and carried out evaluations using Google Forms, but sometimes they still wrote handwritten notes which were then uploaded to Google

Classroom. This makes some students assume that the evaluation carried out by the teacher is not important so they do not work on the evaluation questions given. In the needs analysis, it was found that the fifth grade students of SD No. 3 Sibanggede requires an innovation or variation in undergoing the evaluation process. In addition, adjustments are also made between learning evaluation instruments, needs, and current conditions so that the products used are relevant and supported by facilities owned by students or schools. After that, basic competency analysis and indicator preparation will be carried out which will be the focus in developing STEAM-based evaluation instruments. The following are the basic competencies and indicators that become the reference material for STEAM-based evaluation instruments.

Table 8. Basic Competencies and Indicator

Basic Competency	Indicators
3.3 Analyzing the role of the economy in an effort to improve people's lives in the social and cultural fields to strengthen the unity and integrity of the nation.	3.3.1 Selecting individual economic enterprises. 3.3.2 Analyzing examples of joint or group economic enterprises 3.3.3 Selecting group economic enterprises. 3.3.4 Assessing the attitude of business actors in establishing relations with their competitors in strengthening national unity and integrity 3.3.5 Linking ways of appreciating other people's economic activities in an effort to strengthen national unity and integrity. 3.3.6 Studying humans as social and economic beings 3.3.7 Relating humans as social and economic beings in daily life. 3.3.8 Interpreting economic activity (production) 3.3.9 Interpreting economic activity (distribution) 3.3.10 Interpreting economic activity (consumption)

Then the product will pass through the design stage. The design of the product to be developed is carried out at the design stage. The activities carried out are designing learning evaluation instruments in the form of objective tests in terms of learning materials, learning indicators, learning objectives, basic competencies, and core competencies. At this stage the evaluation instrument display is designed with flowcharts and storyboards. Flowcharts are useful for explaining the course of a program. While the storyboard is used because it contains all the information that will be displayed on the screen. This interactive evaluation instrument is designed based on the STEAM approach.

Stage of development (development). There are several activities carried out at this stage. These activities are making grids, compiling questions, inputting questions into a web page-based application called Wordwall, making an introduction to evaluation instruments using Canva and Power Point.

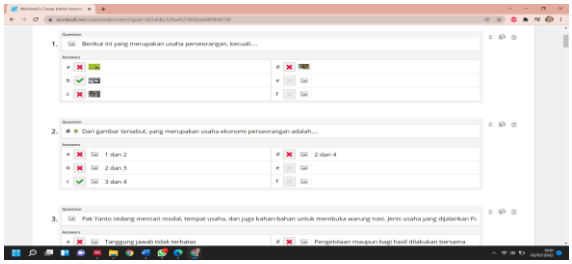

Figure 1. Display of Wordwall

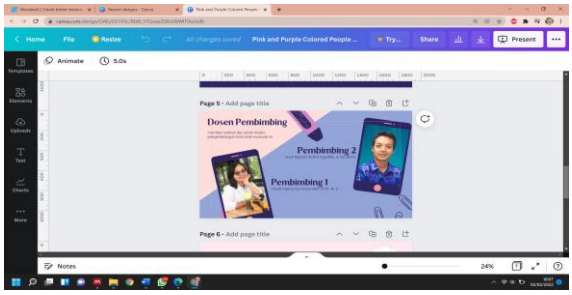

Figure 2. Display of Canva

The completed evaluation instrument, followed by making a product validation questionnaire for experts, then validating the product by the validator. The STEAM-based evaluation instrument got a result of 76.6% which means it is well qualified with few revisions. Subject matter experts (content) suggested replacing some of the questions by adding pictures, stories, or statements so that the questions become HOTS and can train students to think critically. For subject matter experts (questions) the result is 1 which is calculated by the Gregory formula. This means that the questions in the evaluation instrument have very high validity. STEAM-based evaluation instruments were also assessed by design experts. The result of the assessment is 87.5% which is included in the good qualifications with few revisions. Suggestions from design experts are to add learning objectives and ensure there is an ABCD element in the goals created. Another suggestion is that the content of the questions is better made per level to motivate students more.

Design experts also assess the completed STEAM-based evaluation instruments. The result of the design expert test is 83.9% which means that it is included in the good qualifications with a few revisions. The advice obtained is about adding voice over to the evaluation instrument to greet and guide students to use the evaluation instrument. After the expert test was carried out and then corrected or revised several things suggested by the experts, then a trial was carried out on 25 class V students on 40 questions that had been made and the calculation of the quality of the questions which included tests of validity, reliability, level of difficulty, discriminating power., and also the effectiveness of the distractor. The result of the trial is that 20 questions can be used while the rest are discarded. The questions used are questions that have good validity, level of difficulty, discriminating power, reliability, and effectiveness of distractors. After the development stage, the next stage is implementation. The implementation stage is carried out when the product meets the validity according to the specified criteria. At this stage the product is used directly in the learning process of Social Science content Theme 2 Sub-theme 3 Learning 3 in class V. Implementation is carried out in class V when learning takes place using laptop facilities from Badung stakeholder. Implementation is carried out according to school rules, namely 2 sessions according to the distribution of shifts per absent. During the implementation phase, the obstacles or shortcomings in implementation are also seen. If there are obstacles or deficiencies, the product can be repaired or revised at the evaluation stage.

Figure 3. Implementasi instrumen evaluasi

At the evaluation stage, the product is evaluated as a whole to find out the final result of the product whether there are still deficiencies or weaknesses so that improvements or revisions can be carried out. The last stage is evaluation. The evaluation stage is carried out based on the implementation that has been carried out previously. At the implementation stage, the products that have been developed are used directly by students. At that time, it was seen the smoothness in using the product. If there are obstacles and obstacles, they will be refined at this stage.

Discussion

Based on the research and the results obtained, the STEAM-based evaluation instrument on the social science knowledge competence of fifth grade elementary school students no. 3 Sibanggede is suitable for use in learning because it gets a good predicate on the expert test and is also quite good on the question quality test. With the STEAM approach, the learning carried out can train some students' abilities. Among them are social skills, communication, problem solving, concept building, computing skills, design, and so on (Jesionkowska & Wild, 2020). On the evaluation instrument that has been developed, the element of Science is obtained from the way students think systematically working on the questions. Technology is obtained from the use of evaluation instruments made with Wordwall and done online using laptop or cellphone media. Engineering is a concept that is obtained from the way students solve the problems. Art comes from the display of evaluation instruments in which there is music and beautiful graphic displays. Mathematics is obtained from students' ability to solve problems contained in the questions in the evaluation instrument (Rahmawati et al., 2019).

The product development of this research uses the ADDIE development model. The ADDIE development model has the advantage that it is a systematic step, pays attention to a basic stage in the development design, is easily understood by many people, and is also allowed to make periodic improvements in every stage carried out (Hadi & Agustina, 2016; Pranata et al., 2021). The ADDIE development model also begins with a good stage, namely the analysis stage. The analysis stage makes the development carried out to have a definite basis. The results of the trials that have been carried out, for validation using the biserial point formula, 20 valid questions out of 40 questions tested. The number of items used is 13, 15, 16, 18, 19, 20, 23, 24, 25, 26, 27, 30, 32, 33, 34, 35, 36, 38, 39. Based on the results of data processing using KR 20, the result is 0.9, which means the items are included in the very high reliability category (Arikunto, 2016; Koyan, 2007). After knowing the reliability of the item, then tested the differentiating power and the level of difficulty. The level of difficulty test is carried out to find out how much difficulty a question is. A good test is a test that has a difficulty level between 0.25 - 0.75 (Koyan, 2007). The level of difficulty of the test equipment is 0.58 which is in the moderate category (Arikunto, 2016). To find out which questions have a good discriminating power category, a discriminatory power test of the items is carried out. The results of the differentiating power test obtained from 40 items, there are 12 items in the poor category, 6 items in the good category, and 22 items in the fairly good category. There is a difference power of 0.20, which is categorized as good enough. Distractors or distractors on questions are made to deceive students (Arikunto, 2016). Testing the effectiveness of distractors aims to determine whether the distractors that have been made are chosen evenly by students. Distractor effectiveness got 23% very good, 20% good, 17% not too good, 17% bad, dan 23% very bad (Arifin, 2013). Based on the tests that have been carried out by experts, good results were obtained with several revisions. It aims to improve and provide suggestions for the improvement of the STEAM-based evaluation instrument that has been developed. Thus, it was decided that the STEAM-based evaluation instrument

was suitable for use in learning after going through the revision stage because it was suitable for evaluating teaching competencies (Kim & Kim, 2016).

With the STEAM approach, the evaluation (Wahyuni et al., 2020). In addition, by using a STEAM-based evaluation instrument, the evaluation process becomes more efficient because it is carried out online (Nurwulan, 2020). The STEAM approach in the developed evaluation instrument makes the evaluation process more meaningful (Sartono et al., 2020). Because the evaluation aims to measure the achievement of a learning (Damanik & Zuhdi, 2020). The STEAM approach also trains students' abilities to learn to connect one branch of science to another in order to be able to solve problems and make solutions well in all lesson content (Herranen et al., 2021). The implication of this research is that the innovations and variations given in the evaluation process are carried out for social science knowledge competencies. Theme 2 Sub-theme 3 Learning 3. The evaluation instruments that can be used are not only Google Forms or handwriting, but can also use this STEAM-based evaluation instrument.

4. CONCLUSION

STEAM-based evaluation instrument on social science knowledge competence grade V SD No. The 3 Sibanggede developed are of good quality. This was obtained from the results of the quality test of the items (validity, reliability, level of difficulty, discriminating power, and distractor effectiveness) and also the results of expert tests. This evaluation instrument that has been developed can be a new variation and innovation in learning, especially the evaluation process. In-depth studies related to STEAM-based evaluation instruments still need to be carried out so that more comprehensive data will be obtained. Thus, it can be used as a basis for continuously improving the quality of STEAM-based evaluation instruments.

5. REFERENS

Agung, A. A. G. (2018). *Metodologi Penelitian Kuantitatif (Perspektif Manajemen Pendidikan)*. Universitas Pendidikan Ganesha.

Angeli, C., & Giannakos, M. (2020). Computational Thinking Education: Issues And Challenges. *Computers In Human Bihavior*, *105*. Https://Doi.Org/10.1016/J.Chb.2019.106185.

Arifin, Z. (2009). *Evaluasi Pembelajaran (Prinsip, Teknik, Dan Prosedur)*. Pt. Remaja Rosdakarya.

Arifin, Z. (2013). *Evaluasi Pembelajaran*. Bandung: Pt Remaja Rosdakarya.

Arikunto, S. (2016). *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.

Ayuningsih, N. P. M. (2020). Validitas Isi Media Pembelajaran Interaktif Berorientasi Model Problem Based Learning Dan Pendidikan Karakter. *Jurnal Mathematics Paedagogic*, *V*(1). Https://Doi.Org/10.36294/Jmp/Vxix.Xxx.

Basuki, I., & Hariyanto. (2017). *Asesmen Pembelajaran*. Bandung: Pt. Remaja Rosdakarya.

Bedar, R. A. H., & Al-Shboul, M. (2020). The Effect Of Using Steam Approach On Developing Computational Thinking Skills Among High School Students In Jordan. *International Journal Of Interactive Mobile Technologies*, *14*(14). Https://Doi.Org/10.3991/Ijim.V14i14.14719.

Chick, R. C., Clifton, G. T., Peace, K. M., Propper, B. W., Hale, D. F., Alseidi, A. A., & Vreeland, T. J. (2020). Using Technology To Maintain The Education Of Residents During The Covid-19 Pandemic. *Journal Of Surgical Education*, *77*(4), 729–732. Https://Doi.Org/10.1016/J.Jsurg.2020.03.018.

Çil, E. (2021). The Effect Of Using Wordwall . Net In Increasing Vocabulary Knowledge Of 5th Grade Efl Students. *Language Education & Technology (Let Journal)*, *1*(1), 21–28. Http://Www.Langedutech.Com/Letjournal/Index.Php/Let/Article/View/16/3.

Conradty, C., & Bogner, F. X. (2018). From Stem To Steam : How To Monitor Creativity From Stem To Steam : How To Monitor Creativity. *Creativity Research Journal*, *30*(3), 233–240. Https://Doi.Org/10.1080/10400419.2018.1488195.

Damanik, S., & Zuhdi, M. (2020). Model Evaluasi Pembelajaran Aud Berbasis Daring Di Ra Nurun Namirah Medan Marelan (Studi Kasus Selama Masa Pandemi Covid-19). *Al-Fatih: Jurnal Pendidikan Dan Keislaman*, *Iii*(1), 155–172. Http://Jurnal.Stit-Al-Ittihadiyahlabura.Ac.Id/Index.Php/Alfatih/Article/View/84.

Dewi, T. A. P., & Sadjiarto, A. (2021). Pelaksanaan Pembelajaran Daring Pada Masa Pandemi Covid-19. *Jurnal Basicedu*, *5*(4), 1909–1917. Https://Doi.Org/10.31004/Basicedu.V5i4.1094.

Dwi Sari, N., & Setiawan, J. (2020). Papan Gekola Sebagai Media Pembelajaran Matematika Yang Inovatif Dengan Pendekatan Steam. *Jurnal Saintika Unpam : Jurnal Sains Dan Matematika Unpam*, *3*(1), 31. Https://Doi.Org/10.32493/Jsmu.V3i1.4728.

Estriyanto, Y. (2020). Menanamkan Konsep Pembelajaran Berbasis Steam (Science, Technology, Engineering, Art, And Mathemathics) Pada Guru-Guru Sekolah Dasar Di Pacitan. *Jurnal Ilmiah Pendidikan Teknik Kejuruan (Jiptek)*, *13*(2), 68–74. Https://Dx.Doi.Org/10.20961/Jiptek.V13i2.45124.

Fitriyani, H., & Astuti, D. (2021). Pelatihan Pengembangan Instrumen Evaluasi Online Menggunakan Google Form Bagi Guru Sekolah Dasar/Madrasah Ibtidaiyag. *Ethos: Jurnal Penelitian Dan Pengabdian Masyarakat*, *9*(2), 204–215. Https://Doi.Org/10.29313/Ethos.V9i2.7009.

Hadi, H., & Agustina, S. (2016). Pengembangan Buku Ajar Geografi Desa-Kota Menggunakan Model Addie. *Jurnal Educatio*, *11*(1), 90–105. Https://Doi.Org/10.29408/Edc.V11i1.269.

Hamidah, M., & Wulandari, S. S. (2021). Pengembangan Instrumen Penilaian Berbasis Hots Menggunakan Aplikasi Quizizz. *Efisiensi: Kajian Ilmu Adminsitrasi*, *18*(1). Https://Journal.Uny.Ac.Id/Index.Php/Efisiensi/Article/View/36997.

Haryanto. (2020). Evaluasi Pembelajaran; Konsep Dan Manajemen. In *Uny Press*. Http://Staffnew.Uny.Ac.Id/Upload/131656343/Penelitian/Evaluasi Pembelajaran.Pdf.

Haryati, S., Albeta, S. W., Futra, D., & Siregar, A. . (2021). The Development Of Evaluation Instruments In Online Learning Using The Quizizz Application: During Covid-19. *Al-Ishllah: Jurnal Pendidikan*, *4*(2), 1–11. Http://Journal.Staihubbulwathan.Id/Index.Php/Alishlah/Article/View/383.

Herranen, J., Fooladi, E. C., & Milner-Bolotin, M. (2021). Editorial : Special Issue " Promoting Steam In Education ." *Lumat*, *9*(2), 1–8. Https://Doi.Org/10.31129/Lumat.9.2.1559.

Jesionkowska, J., & Wild, F. (2020). Active Learning Augmented Reality For Steam Education — A Case Study. *Education Sciences*, *10*(8), 1–15. Https://Doi.Org/10.3390/Educsci10080198.

Kim, B., & Kim, J. (2016). Development And Validation Of Evaluation Indicators For Teaching Competency In Steam Education In Korea. *Eurasia Journal Of Mathematics, Science & Technology Education*, *12*(7), 1909–1924. Https://Doi.Org/10.12973/Eurasia.2016.1537a.

Komala, & Rohmalina. (2021). Kompetensi Guru Paud Dalam Pjj Melalui Pelatihan Pendekatan Steam Pada Masa Covid-19. *Aulad: Journal On Early Childhood*, *4*(1), 38–52. Https://Doi.Org/10.31004/Aulad.V4i1.96.

Koyan, I. W. (2007). *Assesmen Dalam Pendidikan*. Singaraja: Universitas Pendidikan Ganesha.

Kurnia, L. D., Haryati, S., & Linda, R. (2022). Pengembangan Instrumen Evaluasi Higher Order Thinking Skills Menggunakan Quizizz Pada Materi Termokimia Untuk Meningkatkan Kemampuan Berpikir Tingkat Tinggi Peserta Didik Pendahuluan. *Jurnal Pendidikan Sains Indonesia*, *10*(1), 176–190. Https://Doi.Org/10.24815/Jpsi.V10i1.21727.

Lee, T. Y., Mauriello, M. L., Ahn, J., & Bederson, B. B. (2014). Ctarcade: Computational Thinking With Games In School Age Children. *International Journal Of Child-Computer Interaction*, *2*(1). Https://Doi.Org/10.1016/J.Ijcci.2014.06.003.

Nurwulan, N. R. (2020). Pengenalan Metode Pembelajaran Steam Kepada Para Siswa Tingkat Sekolah Dasar Kelas 1 Sampai 3. *Madaniya*, *1*(3), 140–146.

Oktariyanti, D., Frima, A., & Febriandi, R. (2021). Pengembangan Media Pembelajaran Online Berbasis Game Edukasi Wordwall Tema Indahnya Kebersamaan Pada Siswa Sekolah Dasar. *Jurnal Bas*, *5*(5), 4093–4100.

Pranata, W., Budijanto, & Utomo, D. H. (2021). Buku Suplemen Geografi Berstruktur A-Car Dengan Model Pengembangan Addie. *Jurnal Pendidikan: Teori, Penelitian, Dan Pengembangan*, *6*(2), 185–190. Http://Dx.Doi.Org/10.17977/Jptpp.V6i2.14441.

Rahmawati, Y., Ridwan, A., Hadinugrahaningsih, T., & Soeprijanto. (2019). Developing Critical And Creative Thinking Skills Through Steam Integration In Chemistry Learning Developing Critical And Creative Thinking Skills Through Steam Integration In Chemistry Learning. *International Converence Of Chemistry (Icchem)*. Https://Doi.Org/10.1088/1742-6596/1156/1/012033.

Rasvani, N. L. A., & Wulandari, I. G. A. (2021). Pengembangan Media Pembelajaran Aplikasi Maca (Materi Pecahan) Berorientasi Teori Belajar Ausubel Muatan Matematika. *Mimbar Pgsd Undiksha*, *9*(1), 74–81.

Ratnawulan, E., & Rusdiana. (2014). *Evaluasi Pembelajaran*. Bandung, Pustaka Setia.

Sartono, N., Suryanda, A., Leyli, T. S. A., & Zubaidah. (2020). Implemetasi Steam Dalam Pembelajaran Biologi : Upaya Pemberdayaan Guru Biologi Madrasah Aliyah Dki Jakarta. *Baktimas: Jurnal Pengabdian Pada Masyarakat*, *2*(1), 7–14. Https://Doi.Org/10.32672/Btm.V2i1.2099.

Setemen, K. (2010). Pengembangan Evaluasi Pembelajaran Online. *Jurnal Pendidikan Dan Pengajaran*, *43*(3), 207–214.

Suartama, I. K. (2016). *Evaluasi Dan Kriteria Kualitas Multimedia Pembelajaran*. Singaraja: Teknologi Pendidikan Universitas Pendidikan Ganesha.

Tegeh, I. M., & Kirna, I. M. (2010). *Model Penelitian Pengembangan Pendidikan*. Singaraja: Universitas Pendidikan Ganesha.

Umi Fatimah, L., & Alfath, K. (2019). Analisis Kesukaran Soal, Daya Pembeda Dan Fungsi Distraktor. *Jurnal Komunikasi Dan Pendidikan Islam*, *8*(2), 37–64.

Wafiqni, N., & Putri, F. M. (2021). Fektivitas Penggunaan Aplikasi Wordwall Dalam Pembelajaran Daring (Online) Matematika Pada Materi Bilangan Cacah Kelas 1 Di Min 2 Kota Tangerang Selatan. *Elementar: Jurnal Pendidikan Dasar*, *1*(1). Http://Journal.Uinjkt.Ac.Id/Index.Php/Elementar/Article/View/20375.

Wahyuni, S., Reswita, & Afidah, M. (2020). Pengembangan Model Pembelajaran Sains , Technology , Art , Engineering And Mathemathic Pada Kurikulum Paud. *Jurnal Golden Age, Universitas Hamzanwadi*, *04*(2), 297–309. Https://Doi.Org/Https://Doi.Org/10.29408/Goldenage.V4i02.2441.

Wulandari, I. G. A. A., & Agustika, G. N. S. (2020). Dramatik Pembelajaran Daring Pada Masa Pandemi Covid-19 (Studi Pada Persepsi Mahasiswa Pgsd Undiksha). *Mimbar Pgsd Undiksha*, *8*(3), 515–526. Http://Dx.Doi.Org/10.23887/Jjpgsd.V8i3.29259.