Journal of Turkish Science Education, 2022, *19*(1), 37-51.
DOI no: 10.36681/tused.2022.108

Developing a Problem-Solving Essay Test Instrument (PSETI) in the Instruction of Basic Science Concepts in Ethnoscience Context

W Winarto[1], Sarwi S[2], Edy Cahyono[3], Woro Sumarni[4]

[1]*Universitas Peradaban, Indonesia, Pascasarajana Universitas Negeri Semarang, Indonesia, ORCID ID: 0000-0001-5029-5327*
[2]*Pascasarajana Universitas Negeri Semarang, Indonesia, ORCID ID: 0000-0001-8772-2062*
[3]*Pascasarajana Universitas Negeri Semarang, Indonesia, ORCID ID: 0000-0001-7225-3291*
[4]*Pascasarajana Universitas Negeri Semarang, Indonesia, ORCID ID: 0000-0002-5865-087X*

ABSTRACT
Assessment of problem-solving skills for preservice teachers is important in science instruction. The purpose of this study is to discover; (a) How are the characteristics of the Problem-Solving Essay Test Instrument (PSETI); (b) How is the validity based on experts assessment; (c) how are the validity and reliability based on trials. The research method is research and development. The study covers three stages, those are (1) planning the test, (2) trying out the test, (3) determining the validity, and the reliability. The sample of this study was 80 preservice elementary school teachers of basic science concepts. The PSETI is in the form of an essay test, consisting of 19 items. The essay test contains ethnoscience problems. Content validity based on the Aiken index is 0.84 with a good validity level. The results of the development of the PSETI instrument show 19 items met the validity and reliability of the 21 items/ compiled. The PSETI is an essay test with the scoring which is using a partial credit model based on the three categories in the polytomous data. Instrument reliability was 0.77 and it was concluded that the instruments were good to measure problem-solving skills of preservice teachers in science learning context ethnoscience.

ARTICLE INFORMATION
Received: 23.12.2020
Accepted: 15.08.2021

KEYWORDS:
Test development, PSETI, ethnoscience context, preservice teachers .

Introduction

Individuals should have problem-solving skills to face the 21st century in order to be able to compete globally. Kennedy et al. (2016) stated the basic skills needed by educators related to 21st century skills are critical thinking, problem-solving, collaborative learning, and having the ability to use technology. Dogru (2008), Nurita et al. (2017), and Temel (2015) stated that helping students to develop problem-solving skills is the main target in training preservice science teachers.
Problem-solving skills acquired by preservice teachers become one of the success factors in effective instruction and to improve students' problem-solving skills as well. Teachers with good problem-solving skills are able to manage effective instruction and improve students' learning achievements (Adeyemo et al., 2013), they are able to improve students' problem-solving skills (Mauke & Sadia, 2013), to nurture students constructing their knowledge and to take part in its acquiring process Karatas & Baki (2013), and to encourage them applying their knowledge in creative ways and to develop a deep understanding (Crebert et al., 2011). The teaching having no problem-solving skills would not be able to improve those skills to their own students (Solso et al., 2010). In the end, the students would not have sufficient ability to solve complex problems in their daily lives (Ulger, 2018).

Problem-solving skills should be taught in an instructional process (Mukhopadhyay, 2013). The problems discussed should be closely related to daily lives. It is intended to prepare the students to be able to solve not only structured problems but also unstructured, complex, and various ones (Dixon & Brown, 2012). The students must have the ability to identify and understand the pattern between problems and to select the best way a solution to solve them (Bahtiyar & Can, 2016). However, teaching at the higher education level is not optimum yet to improve problem-solving skills. Studies related to problems solving skills of university students showed that their skills were still in the lower category (Busyairi & Sinaga 2015; Purwandari & Yusro, 2018; Sutarno et al., 2017; Yusuf & Widyaningsih, 2018). Low problem-solving skills were caused by a lack of knowledge, motivational, and emotional aspects (Dostal, 2015) and the use of an instructional model was highly influential to problem-solving abilities (Davis et al., 2019).

Problem-solving is defined as cognitive skills used to solve problems related to real-life (OECD, 2003) in a new, creative, systematic, and analytic way (Bahtiyar & Can, 2016). Experience, knowledge, and intuition applied simultaneously to a problem as a result of creative thinking that can be quickly and effectively used in problem-solving are seen as problem-solving abilities (Ardichvili et al., 2003). Problem-solving skills were connected to critical, analytical thinking, and creating productively which involves quantitative, communicative ability, and the ability to respond critically (Chang, 2010). Polya (1945) as quoted by Selcuk et al. (2008) explained that problem-solving consisted of four solution phases are the following: (1) to understand the problem that students would not be able to solve problems correctly if they do not understand the given problems; (2) to plan the solution, this phase was very dependent to the students' experience in problem-solving. The more experiences they had, the more creative they were to plan the problem-solving; (3) to solve the problem, it was the appropriateness of the planning stage after it was constructed. Then, the execution of problem-solving was done based on the mentioned planning; (4) to conduct a recheck phase. It was done by re-evaluation of what had been done from the phase 1 to phase 3.

Teaching science to improve problem-solving skills by providing realistic problems could arouse challenges in university students and motivate them to get involved in a problem-solving process (Ulger, 2018). Thus, presented problems must be related to daily lives (Heller & Heller, 2010; Jonassen, 2011). Problems introduced to educate students are structured, unstructured, complex, and diverse (Dixon & Brown, 2012). Aksoy (2005) suggested a science instruction based on the discovery scientific process by the method of asking, presenting, applying, and transforming scientific knowledge (Sternberg, 2003). To initiate creativity and problem-solving with discovery learning, teachers could ask students to conduct an independent research or get involved in the training of students' divergent thinking by scientific process skills. The students are encouraged to develop interesting and diverse science subjects, to have scientific observation, classifying, asking scientific research questions, forming hypothesis, planning testing and scientific measurement, utilizing instruments, and drawing conclusion from empirical data (Cheng, 2011). Science learning in the context of ethnoscience is thought to improve problem-solving skills (Sumarni & Kadarwati, 2020).
Ethnoscience is part of knowledge related to things and natural phenomena which is organized in societal science and is produced by certain cultures (Aboyi, 2002) in the fields of agriculture, astronomy, medical practices, mathematic, technique, architect, military science, and ecology (Snively & Corsiglia, 2011). The teaching approach connecting scientific concepts with societal native knowledge produced by certain cultures and students understand natural phenomena they experienced in their lives is defines as ethnoscience approach (Joseph, 2010; Snively & Corsiglia, 2011). The results of science instruction in ethnoscience context showed that it increased university students' positive attitude toward science and it developed their creative thinking in a learning environment (Şener et al., 2015; Kutlu & Gökdere, 2015), it improved creative thinking (Piirto, 2011), chemistry literacy (Sumarni, 2018), character and science literacy (Sarwi & Subali, 2020), problem-solving (Supriyadi et al., 2016; Novia & Kamaluddin, 2015) creative and critical thinking (Sumarni & Kadarwati, 2020).

The role of assessment to measure problem-solving skills is crucial to predict preservice science teachers' accomplishment levels (Kourmousi et al., 2016). There were differences among experts on how to measure problem-solving skills. Istiyono et al. (2019), Suprapto et al. (2020), and Istiyono et al. (2020) developed instruments in the form of multiple choices to measure the skills. Greiff et al., (2015) and Risnita & Bashori (2020) used multiple complex systems essay tests to measure problem-solving skills. Multiple complex systems related independent items to a complex problem. The test of problem-solving skills developed by Butterworth & Thwaites (2013) and Suratno et al. (2020) was grounded on four indicators of problem-solving as the following: (1) to combine skills and to use imagination, (2) to develop a model, (3) to conduct an investigation, and (4) to analyze the data and to draw conclusions. Some of the assessments to problem-solving conducted by Jonassen (2011) and Rokhmat et al. (2020) were *problem schema, analogy, causal, and argumentation*. In this study, the essay test instrument was chosen to measure problem-solving skills are complex and comprehensive cognitive processes in solving problems so that multiple-choice tests do not measure well (Henderson et al., 2001; Kastner & Stang, 2011). The test essay instrument is appropriate for measuring problem-solving because it is able to reveal the thought process in solving problems (Haladyna & Rodriguez, 2013; Aristiawan & Istiyono, 2020). The weakness of the test essay is that the objectivity of assessing the test results decreases with the concentration level of the rater (Reiner et al., 2002), require more time to score (Popham, 2009), and it is especially difficult to score essay test responses reliably (Tuckman, 1993).

Instruments used in the assessment must have certain requisites to be appropriate to be utilized, those are validity and reliability (Mardhapi, 2012). Lissitz & Samuelsen (2007) state that a learning outcome validity test encompasses content and construct validity. Content validity measures the appropriateness of items with intended psychological constructs. Content validity is determined by an agreement among experts in the field of the study (Retnawati, 2016). To find out the level of expert agreement, the index proposed by Aiken (1985) as quoted by Retnawati (2016) can be used. The construct validity in this study used IRT (Item Response Theory). Items analysis used IRT based on one parameter grounded on difficulty level (symbolized as β or b) and it is called as one Parameter Logistic (1-PL) or viewed as RM model (Rasch Model) (Subali, 2012). Composed research instruments were essay tests with varied answers (polytomous) that could be analyzed by a Partial Credit Model (PCM) (Retnawati, 2014). Partial Credit Model (PCM) is a theory of polytomous item responses used to evaluate both test characteristics and test-takers' ability with an assumption that discriminate power in each item is equal and there is no need of sorting the difficulty level in each stage (Wetzel & Carstensen, 2014).

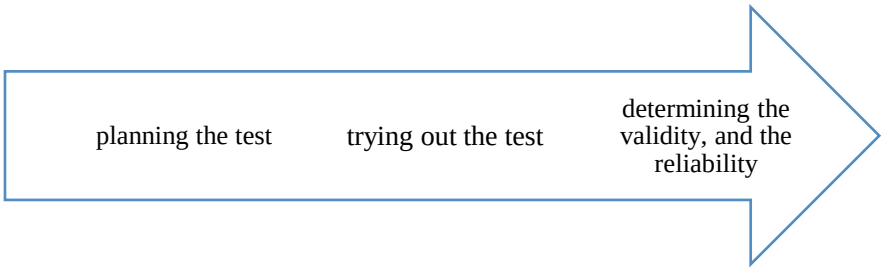
Teachers who have problem-solving skills can develop thinking skills for their students. Conversely, teachers who have low problem-solving skills will fail to develop the thinking skills of their students. Learning science in the context of ethnoscience can develop problem-solving skills. Prospective teachers who choose the measurement of problem-solving skills of prospective teachers need to be carried out in an effort to produce professional teachers. The instrument to measure problem-solving skills was developed by previous researchers in the form of an essay test. The advantages of the essay test are able to reveal higher-order thinking skills in the form of problem-solving skills. The novelty of this research is a problem-solving instrument in the form of an essay test in learning science in the context of ethnoscience. The contribution of this research is to provide an assessment of the problem-solving skills of prospective teachers in science learning in the context of ethnoscience. This study aims get a valid test instrument Problem-solving Essay Test Instrument (PSETI). This research consequently addresses the questions:: (a) How are the characteristics of the Problem-solving Essay Test Instrument (PSETI); (b) How is the validity based on experts assessment; (c) how are the validity and reliability based on trials.

Methods

The research method is research and development. The final product of the research was the instrument to measure the problem-solving skill in science learning context ethnoscience for preservice science teachers. The instruments for assessing the problem-solving skills were implemented from the results of Oriondo & Dallo-Antonio (1998) that covers three stages, those are (1) planning test, (2) trying out the test, (3) determining the validity, and the reliability.

Figure 1

Research Step



Planning the Test

Planning the test means deciding the objectives, the test materials, writing the test grid, writing the item of the test, arranging the scoring guideline, validating the test item, and revising the test item. The objective of the test was to measure the students' problem-solving skills in the context of ethnoscience. The test topics discussed were concepts of science with the subject area (1) quantity and unit, (2) matter, (3) temperature and heat, (4) living things, (5) living things and their environment, and (6) animals that displayed in an essay test. The test was equipped with a grid that presented in tables and contains some information such as materials, indicators of problem-solving skills, and indicators of items. In this research, scoring guidelines were using the analytical rubric scoring.
Content validation was carried out by experts using the Aiken index (Retnawati,2016). Aiken Index V with a value between 0-1 was an agreement index based on experts' assessment of the appropriateness of the items to measured psychological construct and the item was valid if the coefficient validity is more than 0.8 (Retnawati, 2016). The formula to calculate the Aiken index is as the following:

$V = \sum(ri - lo)/[n(c-1)]$ …..(1)

Explanation:
r = value given by assessors
lo = **the** lowest validity value
c = the highest validity value
n = the number of experts conducting the research
i = the number of months from 1
n = the number of assessors

Trying Out the Test

The try-out of the test was done with a sample of this research which consisted of 80 presevice teachers from two universities with the same curriculum. The sampling technique using purposive. Sampling is based on the location of residence where there are many ethnoscience concepts that are used as learning resources. The sample number in the try-out was based on Bond & Fox (2007) that stated the number of participants in analysis using IRT was started from 30 to 300 participants.

Determining the Validity and Reliability

The data of the try-out results were applied to prove the assumptions that underlie the item response theory that was one-dimensional, local independence, and parameter invariance. After the assumptions were completed, then the analysis of the items to find out the quality and parameters of the items could be finished. The item of the test has good quality if the range of mean square is 0.77 to 1.30 and the range of difficulty level was -2 to 2 (Hambleton et al., 1991). The QUEST program presented reliability values from a test item. The internal consistency index value on QUEST output for polytomous scoring is Cronbach alpha (Subali & Suyata, 2012). Internal consistency in the QUEST program as a measure of the reliability of the test instrument. Guilfod (1956) as quoted by (Arikunto, 2016) stated that instrument reliability level of $0.80 < r < 1.00$ was categorized as very high reliability. The value of $0.60 < r < 0.80$ was categorized as high reliability. The value of $0.40 < r < 0.60$ was categorized as medium reliability and the value of $0.20 < r < 0.40$ was categorized as low reliability. The score of $-1.00 < r < 0.20$ was categorized as very low reliability (not reliable

Findings and Discussion

Planning the Test

The test planning was to study the curriculum of an elementary school teacher study program at four universities in Central Java and to conduct an ethnoscience study in the context of basic science concepts. The purpose of the study of learning outcomes is to determine the competencies, to assess the curriculum and integrity between science concepts, ethnoscience, and instructional accomplishment. The results of the study on the relationship between science and ethnoscience were presented in table 1.

Table 1

The Relationship between Science Basic Concepts and Ethnoscience

No.	The Study of Science Basic Concepts	Javanese Society Ethnoscience on the Making of Cirebon Shrimp Paste, the Process of Cilacap Fish Smoking, and the Cultivation of Sidoarjo Milkfish
1	• The magnitude of mass, length, time, and temperature and their measurement units • The measurement of mass, length, time, and temperature	• Composition of brown sugar, acetes, and salt with the ratio of 100 kilograms of acetes, 100 kg of brown sugar, and 20 kilograms of salt to make shrimp paste • The process of a fish pond making with the measurement of 2.5 meters width and 0.5 meters height • The smoking of stingrays in the temperature of 70- 100 ^{o}C • The time of the fish smoking is 60 minutes using heat smoking heat • The time to dry acetes is a half-day (non-standardized measurement) • The time for fermentation is one week/ 7 days • The amount of fertilizer to a new fish pond is one handheld of an adult person (non-standardized measurement) • Milkfish seeds called “nener” are cultivated by pond fish farmers to the size of “glondongan” (non-standardized measurement)

2	• Elements, compounds, mixtures, chemical properties, physical properties, physical, and chemical changes	• The ingredients to make paste shrimp are salt which is a compound (NaCl), brown/ Javanese sugar contains much sucrose ($C_6H_{22}O_{11}$). Those the ingredients for making shrimp paste are salt which is a compound (NaCl), brown sugar / Java contains a lot of sucrose ($C_6H_{22}O_{11}$). The constituent elements of those compounds could be elaborated. • The use of salt before the stingrays smoking • Shrimp paste is a mixture • The physical properties of shrimp paste, salt, and brown sugar • Chemical changes to acetes in the making process • Chemical changes to stingrays
3	• Temperature and measurement • Heat transfer • The effect of temperature and heat changes on an object	• The temperature of fish smoking used by farmers of smoked stingrays using wood fuel or coconut shells • There is conductive heat transfer in the smoking process using live coals. • The drying process of acetes and stingrays using sunlight is a heat transfer process through radiation.
4	• The concept of alkaline and its characteristics • The concept of salt and its characteristics	• The use of salt in the fishes to be processed by smoking and salt to make paste shrimp • The use of fertilizer containing the characteristic of alkaline in the opening of a new fishpond
5	• Abiotic and biotic • Food chain • Food webs • The food pyramid • Predation	• "Klekap" is a natural feeding which is bred by the farmers. Klekap is the mixture of mosses and organism decomposition on the fishpond. • Phytoplankton is natural feeding which growth is attempted in the fishpond location by the farmers. • Moses are natural feeding which growth is attempted in the fishpond location by the farmers. • Milkfishes are living things chosen by fishpond farmers as farm animals to gain economical profit • Brackish water is the blend of sea and underground water used as a medium to live by milkfishes. • The food chain of milkfishes and phytoplankton/ klekap

The relationship between the basic concepts of science and ethnoscience became the basis for the preparation of test instruments in science learning in the context of ethnoscience. The characteristics of the test instrument developed in this study are to present questions in the context of the science concept and problems of local wisdom in the community. In addition, the test instrument measures problem-solving skills based on Polya's (1945) as quoted by Selcuk et al. (2008) theory of problem-solving consisting of recognizing problems, planning and implementing problem solutions, and evaluating. The test instrument construction stage was conducted by creating the items and by creating test results scoring guidance.

Table 2

Problem-Solving Skills Instrument

Problem-solving Skills Dimension	Indicator Question of Problem-solving	Item Number
1. To identify the problem	The students are able to write science concepts observed in the problem of Sidoarjo milkfish fishponds pollution, the making of Cirebon shrimp paste, and the Cilacap stingrays smoking.	1a,2a,3a
	The students are able to write questions on the problems on Sidoarjo milkfish fishponds pollution, the making of Cirebon shrimp paste, and the Cilacap stingrays smoking which have to be solved.	1b,2b,3b
	The students are able to write factors causing Sidoarjo milkfish fishponds pollution, the making of Cirebon shrimp paste, and the Cilacap stingrays smoking which have to be solved.	1c,2c,3c
	The students are able to write important information related to problems on Sidoarjo milkfish fishponds pollution, the making of Cirebon shrimp paste, and the Cilacap stingrays smoking.	1d,2d,3d
2. To plan the problem -solving	The students are able to write the relationship between the correct science concepts related to problems on Sidoarjo milkfish fishponds pollution, the making of Cirebon shrimp paste, and the Cilacap stingrays smoking.	1e,2e,3e
3. To implement the problem -solving	The students are able to write science concepts to solve problems on Sidoarjo milkfish fishponds pollution, the making of Cirebon shrimp paste, and the Cilacap stingrays smoking.	1f,2f,3f
4. To re-evaluate	The students are able to select one solution to problems on on Sidoarjo milkfish fishponds pollution, the making of Cirebon shrimp paste, and the Cilacap stingrays smoking.	1g,2g,3g

Essay Test Instrument

The test presented three problems in making Cirebon shrimp paste, Cilacap stingray smoking and Sidoarjo milkfish cultivation. This test instrument consists of 21 items. The examples of test items are as follows:

The Buntung River is one of the rivers in Sidoarjo that runs through the metal, plastic, car, and electric cable-smelting industrial areas. These industries have great potential as a source of lead (Pb) waste and pollute the pond waters, the source of which is the Buntung river. The presence of lead in pond waters both in water and in sediments causes the accumulation of pollutants in organisms through the food chain. These heavy metals are easily absorbed and buried in phytoplankton/water plants. Milkfish is a type of herbivorous/plant-eating fish. Apart from entering through the food chain, heavy metals can also enter through the gills and diffusion through the surface of the skin. The longer the exposure to heavy metals in the waters, the more damage to the gill organs will be very visible through histology. The research data on the level of river pollution at the Sidoarjo pond locations are presented in Table 3.

Table 3

Concentrations of Lead (Pb) In Water and Sediment Samples.

Sampling point	Location	The concentration of lead (Pb)			
		Water	Sediment	Phytoplankton	Milkfish
Station 1	River	0.018	2.664	0.176	0.177
	Pond	0.005	3.128		
Station 2	River	0.026	3.072	0.188	0.190
	Pond	0.011	2.974		

Station 3	River	0.028	2.849	0.192	0.186
	Pond	0.008	3.220		
Station 4	River	0.021	2.922	0.177	0.175
	Pond	0.000	3.005		

Based on the problem of the milkfish pond pollution in Sidoarjo which is presented with the information above, please answer the following questions:

a. Write down science concepts related to the problem of pollution in the Sidoarjo milkfish pond.
b. Make questions related to the problem of pollution in the Sidoarjo milkfish pond.
c. Write down the factors that cause lead metal pollution to fish ponds from industrial activities in Sidoarjo.
d. Write down the information about the level of lead pollution based on the data presented in Table 3.
e. Based on the problem of pollution in the Sidoarjo milkfish pond, write down the relationship between the science concepts that are in accordance with the problem.
f. Write down at least two science concepts that can solve the problem of lead pollution in milkfish pond operations.
g. Of the two concepts of science / more to solve the problem of embedding pond water. Choose one of the most effective solutions to solve the problem and give the reason.

The testing stage was done through content and construct validity testing. The content validity tested the quality of the instrument to gain ad expert validators' agreement to test items and variables to be measured. The value of expert validators' agreement level was presented in Table 4.

Table 4

Content Validity Data

No	Validator	The Average of on the Assessment
1	I	4.0
2	II	4.3
3	III	4.7
4	IV	4.4
5	V	4.4
Average		4.4
Aiken Index (V)		0.84
Validity		Very Good

Based on Table 4, all items have good content validity. Aiken (V) index was an agreement index between raters toward the appropriateness of the items. Validity based on Aiken index 0.84 with good validity level. Aiken (1985) as quoted by Azwar, (2013) states that an item with a V index of ≤ 0.4 is considered as low validity, it is considered as medium validity for 0,4 ≤ V ≤ 0,8, and the value of ≥ 0,8 is considered as very valid. The instruments in the category of very valid for the theoretical aspect. After revising the suggestions item of the experts, the instrument was stated to be appropriate to the preservice teacher try out. The trials were carried out to prove the assumptions of item response theory that also functioned as empiric validity (Bashooir & Supahar, 2018). Test instruments were given to 80 preservice elementary school teachers.

The Results of Try-Out

The developed instrument is analyzed by the polytomous item response theory. In this theory, some assumptions needed to be completed previously to know the item parameter. The assumption of

unidimensional This assumption is proved through the exploratory factor analysis. The result of exploratory factor analysis can be shown in Table 5.

Tabel 5

The Tesult Of The KMO and Bartlett Tests

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.641
Bartlett's Test of Sphericity	Approx. Chi-Square	246.030
	Df	78
	Sig.	.000

Table 5 shows that the value of KMO > 0.5. It means that the sample size used to fulfill the requirement, hence the analysis factor can be continued (Hair, Black, Babin, & Anderson, 2009). The total variance explained through the instrument of the test can be shown in Table 6.

Table 6

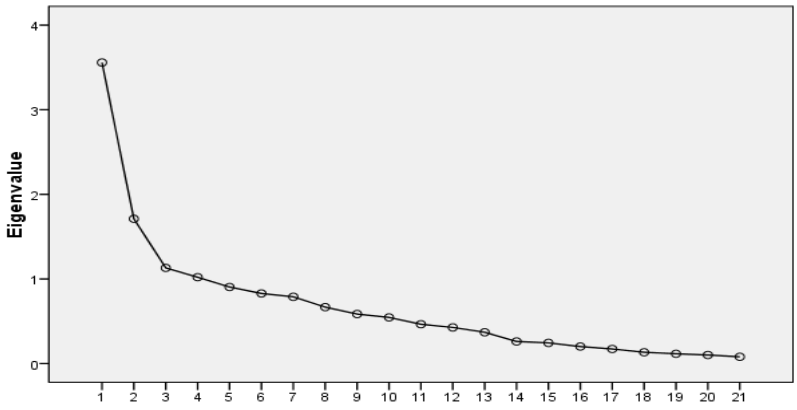
Eigen Value for the Problem-solving Skill Essay Test

	Initial Eigenvalues		
Component	Total	% of Variance	Cumulative %
1	3.373	25.949	25.949
2	1.855	14.270	40.219
3	1.281	9.850	50.069
4	1.167	8.976	59.046

Table 6 shows that the instrument contains four factors. These four factors are in accordance with the theory put forward by Polya (1945) that problem-solving consists of 4 factors/indicators. The result of factor analysis is presented in a scree plot in Figure 2.

Figure 2

Scree Plot Factor Analysis



Based on Figure 2, it can be cleared that the graphic gets the sharp decrease from the first factor to the second factor then it is sloping, so the scree plot formed almost makes the right angle. It

shows that the developed instrument of the test only contains one dominant dimension, so the assumption of one-dimensional is fulfilled (DeMars, 2010).

Item Fit

An item of a test is called suitable with the model when it has INFIT MNSQ value from 0.77 to 1.30 (Adam & Khoo, 1996 as quoted by Subali & Suyata, 2012). Based on the analysis, it gets the item fit in Figure 3.

Figure 3

Item Fit Based On the Mnsq Infit Value

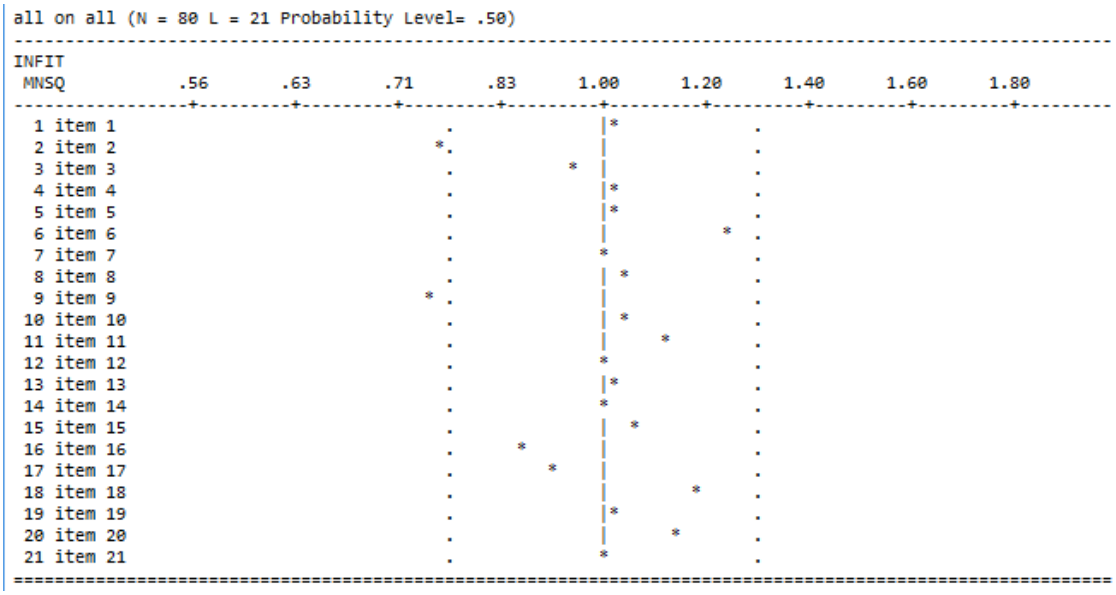


Figure 3 shows that the MNSQ infit value data only items number 2 and 9 are outside the range from 0.77 to 1.30. Item number 2 has an MNSQ Infit value of 0.75 and item number 9 of 0.76 so that the two items are not fit. Except for numbers 2 and 9, all items of the test are suitable with the Partial Credit Model (PCM) so that the items that are fit are 19 items. Based on Table 6, the item is suitable with the Partial Credit Model (PCM). In PCM, it has an assumption about the discriminant value in each item is similar whereas the difficulty index in each step does not have to be in order and similar. The estimation result of item parameters from the try-out can be seen in Table 7.

Table 7

Item Parameter Estimation

Item	Difficulty	Item	Difficulty
1	-0.57	11	-0.3
3	-0.18	13	0.52
4	0.26	14	-0.1
5	0.57	15	0.91
6	0.03	16	-0.75
7	0.32	17	0.61
8	0.25	18	0.24
9	-1.33	19	1.02
10	-1.33	20	1.14
21	1.00		

Based on Table 7, the parameter of difficulty from all items in the instrument of the test is in the range of -0.75 to 1.14. The parameter fulfills good criteria based on Hambleton & Swaminathan (1985) because the difficulty index of the item is still in the range $-2 < b < 2$. The internal consistency in the QUEST program is a measure of the reliability of the test instrument. The internal consistency value (reliability) from test items was 0.77. The test instrument has a reliability value of 0.77, high reliability.

The characteristics of the PSETI instrument are a description test that presents the ethnoscience problems of the Cirebon shrimp paste-making process, the smoking of Cilacap stingrays, and Sidoarjo milkfish ponds. PSETI is an essay test with the scoring which is using a partial credit model based on the three categories in the polytomous data. PSETI instrument has 19 items. Developing a problem-solving instrument in the form of an essay test is researched by Aristiawan & Istiyono (2020); Warsono et al. (2020); Kurniawan & Taqwa (2018); Lestari, Purwanto, & Sakti (2019). The essay test is considered to measure the complex cognitive level (Kubiszyn & Borich, 2013) and to push them to activate high order thinking skills in answering the question (Baig et al. 2014). The essay test is suitable to measure the students' problem-solving skills. PSETI presenting ethnoscience problems. The problems discussed should be closely related to everyday life (Heller & Heller, 2010; Jonassen, 2011). It is intended to prepare students to be able to solve problems that are not only structured but also unstructured, complex, and diverse (Dixon & Brown, 2012). Science instruction in ethnoscience increased university students' positive attitude toward science and it developed their problem-solving & creative thinking in a learning environment (Şener et al., 2015). PSETI assessment by experts was analyzed based on the AIken index of 0.84. The item was valid if the coefficient validity is more than 0.8 (Retnawati, 2016). PSETI assessment, based on experts, falls in the very valid category.

PSETI validity, based on field testing in terms of exploratory unidimensional testing of factor analysis based on scree plots can be cleared that the graphic gets the sharp decreasing from the first factor to the second factor then it is sloping, so scree plot formed almost makes the right angle. It shows that the developed instrument of the test only contains one dominant dimension, so the assumption of one-dimensional is fulfilled (DeMars, 2010). PSETI validity in terms of PCM testing. An item of the test is called suitable with the model when it has an INFIT MNSQ value from 0.77 to 1.30 (Adam & Khoo, 1996). Items number 2 and 9 are outside the range from 0.77 to 1.30. Item number 2 has an MNSQ Infit value of 0.75 and item number 9 of 0.76 so that the two items are not fit. Except for numbers 2 and 9, all items of the test are suitable with the Partial Credit Model (PCM) so that the items that are fit are 19 items. The parameter fulfills good criteria based on Hambleton & Swaminathan (1985) because of the difficulty index. The difficulty level of items that meet the criteria is between -2 to +2. The difficulty level of the PSETI instrument is between -0.75 to 1.14 so it is concluded that it is a good criterion. PSETI instrument reliability is based on internal consistency value. The internal consistency value from test items was 0.77. The test instrument has a reliability value of 0.77 has high reliability. The results of this study are in line with the research Hidayat et al. (2017) concluded that essay test instruments to measure problem-solving skills had the reliability value of 0.8. Thus, it was good for measurement. Nadapdap & Istiyono (2017) developed an essay instrument to measure problem-solving using an item difficulty index ranges from -1.47 to 0.88, meaning that all items are good, and information function analysis and SEM show that the test fits the ability between -1.3 and 2.7. Haeruddin et al. (2020) concluded that essay test instruments to measure problem-solving skills had a reliability value of 0.8. Thus, it was good for measurement.

The advantage of the PSETI instrument is that it has a high-reliability value and meets the theoretical validity and empirical testing. The weakness of this instrument is like an essay test instrument so that the objectivity of assessing, scoring guidelines, and the time it takes to evaluate it is a long time. The results of this study can be used to measure the problem-solving skills of prospective teachers so that they can be used as an evaluation of learning that equips these skills.

Conclusion and Implications

Characteristics of PSETI in the form of an essay test and contains ethnoscience problems. Content validity based on the Aiken index was 0.84, with a good validity level. The results of the development of the PSETI instrument consisting of 19 items met the validity and reliability of the 21 items compiled.

References

Aboyi, O. S. (2002). Effects of ethnoscience-based instructional package on students' interest in science. *Journal of the Science Teachers Association of Nigeria*, *37*(1 & 2), 60–68.

Adams, R. J., & Khoo, S. T. (1996). Quest (program komputer). The interactive test analysis system. Victoria: ACER

Adeyemo, S. A., Babajide, F. T., & Amusa, J. O. (2013). The relationship among teachers' problem solving abilities students' learning styles and students' achievement in Physics Department of Science and Technology Education, University of Lagos. *Lagos Nigeria Australian Journal of Basic and Applied Sciences*, *7*(4), 654–660.

Aksoy, G. (2005). *Create a button science education scientific method processbased learning products, the effect of UNMEE* [Unpublished Master's Thesis]. Zonguldak Karaelmas University.

Ardichvili, A., Cardozo, R., & Ray, S. (2003). A theory of enterpreneurial opportunity identification and development. *Journal of Business Venturing*, *18*(1), 105–123.

Arikunto, S. (2016). *Research procedure: A practical approach*. Rineka Cipta.

Aristiawan, & Istiyono, E. (2020). Developing Instrument of Essay Test to Measure the Problem-Solving Skill in Physics. *Jurnal Pendidikan Fisika Indonesia*, *16*(2), 72–82.

Azwar, S. (2013). *Research method*. Pustaka Pelajar.

Bahtiyar, A., & Can, B. (2016). An investigation of problem-solving skills of preservice science teachers. *Educational Research and Reviews Academic Journal*, *11*(23), 2108–2115.

Baig, M., Ali, S. K., Ali, S., & Huda, N. (2014). Evaluation of Multiple Choice and Short Essay Question Items in Basic Medical Sciences. *Pakistan Journal Medical Science*, *30*(1), 3–6. https://doi.org/10.12669/pjms.301.4458

Bashooir, K., & Supahar. (2018). Validitas dan reliabilitas instrumen asesmen kinerja literasi sains pelajaran fisika berbasis STEM. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *22*(2), 168–181. https://doi.org/https://doi.org/10.21831/pep.v22i2.20270

Bond, T. G., & Fox, C. M. (2007). *Applying The Rasch Model: Fundamental Measurement in The Human Sciences* (2nd ed.). Lawrence Erlbaum.

Busyairi, A., & Sinaga, P. (2015). The profile of problem solving skills creatively of senior high school students in the main discussion of dynamic electricity. In *The Proceeding of Physics National Seminar (E-Journal)* (pp. 23–28).

Butterworth, J., & Thwaites, G. (2013). *Thinking skills: Critical thinking and problem solving* (2nd ed.). Cambrige University Press.

Chang, C.-Y. (2010). Does problem solving = prior knowledge + reasoning skills in earth science? an exploratory study. *Res. Sci. Educ*, *40*(2), 103–116.

Cheng, V. M. Y. (2011). Infusing creativity into Eastern classroom: Evaluations froms students perspectives. *Journal of Thinking Skills and Creativity*, *6*(1), 67–68.

Crebert, G., Patrick, C.-J., Cragnolini, V., Smith, C., Worsfold, K., & Webb, F. (2011). *Ethical behaviour and social responsibility toolkit.* http://www.griffith.edu.au/data/assets/pdf_file/0009/290691/Ethicalbehaviour.pdf

Davis, D. D., McDuffi., A. R., Drake., C., & Seiwell., A. . (2019). Teachers'perceptions of the official curriculum: Problem solving and rigor. *International Journal of Educational Research*, *93*, 91–100.

DeMars, C. (2010). *Item Response Theory*. Oxford University Press.

Dixon, R. A., & Brown, R. A. (2012). Transfer of learning: Connecting concepts during problem solving. *Jounal of Technology Education*, *24*(1), 2–17.

Dogru, M. (2008). The application of problem solving method on science teacher trainees. *Journal of Environmental & Science Education*, *3*(1), 9–18.

Dostal, J. (2015). Theory of problem solving. *Procedia: Social and Behavioral Sciences*, *174*, 2798–2805.

Greiff, S., Stadler, M., Sonnleitner, P., Wolff, C., & Martin, R. (2015). Sometimes less is more: Comparing the validity ofcomplex problem solving measures. *Intelligence*, *50*, 100–113.

Haeruddin, H., Prasetyo, Z. K., & Supahar, P. (2020). The development of a metacognition instrument for college students to solve physics problems. *International Journal of Instruction*, *13*(1), 767–782. https://doi.org/https://doi.org/10.29333/iji.2020.13149a

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). *Multivariate Data Analysis* (7th ed.). Upper Saddle River: Prentice Hall.

Haladyna, T. M., & Rodriguez, M. C. (2013). *Developing and Validating Test Items*. Routledge.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of Item Response Theory Library*. Sage.

Heller, P., & Heller, K. (2010). *Problem solving labs, in cooperative group problem solving in physics. Research Report*. University of Minnesota.

Henderson, C., Heller, K., Heller, P., Kuo, V., & Yerushalmi, E. (2001). Instructors' ideas about problem solving setting goals. In *Physics Echration Research Conference* (pp. 2–5).

Hidayat, S. ., Setyadin, H. A., Hermawan, Kaniawati, I., Suhendi, E., Siahaan, P., & Samsuih, A. (2017). Pengembangan Instrumen Tes Keterampilan Pemecahan Masalah Pada Materi Getaran, Gelombang, dan Bunyi. *Jurnal Penelitian Dan Pengembangan Pendidikan Fisika*, *3*(2), 2461–1433.

Istiyono, E., Mustakim, S. ., Suranto, Widihastuti, & Mukti, T. . (2019). Measurement of physics problem-solving skills in female and male students by phystepross. *JPII*, *8*(2), 170–176.

Istiyono, E., Widihastuti, Supahar, Hamdi, S. (2020). Measuring creative thinking skills of senior high school male and female students in physics (CTSP) using the IRT-based PhysTCreTS. *Journal of Turkish Science Education, 17*(4), 578-590.

Jonassen, D. (2011). Supporting problem solving in PBL. *Interdisciplinary Journal of Problem-Based Learning*, *5*(2), 95–119.

Joseph, M. . (2010). Etnhoscience and problem of method in the social scientific study of religion. *Oxford Journal*, *39*(3), 241–249.

Karatas, I., & Baki, A. (2013). The effect of learning environments based on problem solving on students' achievements of problem solving. *Journal of Elementary Education*, *5*(3), 249–268.

Kastner, M., & Stang, B. (2011). Multiple choice and constructed response tests: Do test format and scoring matter? *Procedia Social and Behavioral Sciences*, *12*, 263–273.

Kennedy, I. G., Gloria, L., & Hélia, J. (2016). *Education skills for 21st century teachers: Voices froma global online educators' forum*. Springer Briefs in Education.

Kourmousi, N., Xythali, V., Theologitou, M., & Koutras, V. (2016). Validity and reliability of the problem solving inventory (PSI) in a nationwide sample of Greek educators. *Journal Soc. Sci.*, *5*(25), 1–11.

Kubiszyn, T., & Borich, G. D. (2013). *Educational Testing and Measurement: Classroom Application and Practise* (10th ed.). Willey

Kurniawan, B., & Taqwa, A. R. (2018). Pengembangan instrumen tes kemampuan pemecahan masalah fisika pada materi listrik dinamis. *Jurnal Pendidikan*, *3*(11), 1451–1457.

Kutlu, N., & Gökdere, M. (2015). The effect of Purdue model based science teaching on creative thinking. *International Journal of Education and Research*, *3*(3), 589–599.

Lestari, P. E., Purwanto, A., & Sakti, I. (2019). Pengembangan instrumen tes keterampilan pemecahan masalah pada konsep usaha dan energi. *Jurnal Kumparan Fisika*, *2*(3), 161–168.

Lissitz, W., & Samuelsen, K. (2007). Further clarification regarding validity and education. *Educational Researcher*, *36*(8), 282–484.

Mardhapi, D. (2012). *Measurement of Educational Assessment and Evaluation*. Nuha Medika.

Mauke, M., & Sadia, I. W. (2013). The effect of contextual teaching model and learning towards the understanding of concepts and problem solving skills in a physics-science instruction at MTs Negeri Negara. *Jurnal Pendidikan IPA*, *3*(1281–293).

Mukhopadhyay, R. (2013). *Problem solving in science learning-some important considerations of a teacher.* http://www.iosrjournals.org/iosr-jhss/papers/Vol8-issue6/C0862125.pdf

Nadapdap, A., & Istiyono, E. (2017). Developing physics problem-solving skill test for grade X students of senior high school. *REiD (Research and Evaluation in Education), 3*(2), 114-123. doi:http://dx.doi.org/ 10.21831/reid.v3i2.14982

Novia, N., & Kamaluddin. (2015). Penalaran kausal dan analogi berbasis etnosains dalam memecahkan masalah fisika. In *Prosiding Simposium Nasional Inovasi dan Pembelajaran Sains* (pp. 445–448).

Nurita, T., Hastuti., P. W., & Sari., D. A. (2017). Problem solving ability of science students in optical wave course. *Jurnal Pendidikan IPA Indonesia*, *6*(2), 341–345.

OECD. (2003). *The PISA 2003 assessment framework – mathematics, reading, science and problem solving, knowledge and skills*. OECD.

Oriondo, L. L., & Dallo-Antonio, E. M. (1998). *Evaluating Educational Outcome*. Rex.

Piirto, J. (2011). *Creativity for 21st century skills: How to embed creativity into the curriculum*. Sense.

Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory Into Practice*, *48*(1), 4–11.

Purwandari, P., & Yusro, A. (2018). Pembelajaran fisika menggunakan inkuiri terbimbing dengan metode eksperimen dan proyek ditinjau dari kreativitas dan kemampuan berpikir kritis siswa. *Jurnal Momentum: Physics Education Journal*, *2*(1), 39–46.

Reiner, C. M., Bothell, T. W., Sudweeks, R. R., & Wood, B. (2002). *Preparing Efective Essay Questions A Self-directed Workbook for Educators*. New Forums Press.

Retnawati, H. (2014). *Teori Respon Butir dan Penerapannya*. Nuha Medika.

Retnawati, H. (2016). *Validitas Reliabilitas dan Karakteristik Butir*. Nuha Medika.

Risnita, R.& Bashori, B. (2020). The effects of essay tests and learning methods on students' chemistry learning outcomes. *Journal of Turkish Science Education, 17*(3), 332–341.

Rokhmat, J., Marzuki, Wahyudi, & Putrie, S. D. (2019). A strategy of scaffolding development to increase students' problem-solvingabilities: The case of physics learning with causalitic-thinking approach. *Journal of Turkish Science Education, 16*(4), 569-579.

Sarwi, A., S, F., & Subali, B. (2020). The analysis of ethnoscience-based science literacy and character development using guided inquiry model. *Journal of Physics: Conference Series*, *1567*, 1–6.

Selcuk, G. S., Çalışkan, S., & Erol, M. (2008). The effect of problem solving instruction on physics achievement, problem solving performance and strategy use. *Latin America Journal Physics Education*, *2*(3), 151–166.

Şener, N., Türk, C., & Taş, E. (2015). Improving science attitude and creative thinking through science education project: A design, implementation and assessment. *Journal of Education and Training Studies*, *3*(4), 57–67.

Snively, G., & Corsiglia. (2011). Discovering indigenous science: Implications for science education. *Science Education*, *85*(1), 7–34.

Solso, R. L., MacLin, M. K., & MacLin, O. H. (2010). *Cognitive Psychology* (8th ed.). Pearson Education.

Sternberg, R. J. (2003). A broad view of intelligence: The theory of successful intelligence. *Consulting Psychology Journal: Practice and Research*, *55*(3), 139–154.

Subali, B. (2012). *Prinsip Asesmen & Evaluasi Pembelajaran*. UNY.

Subali, B., & Suyata, P. (2012). *Pengembangan Item Tes Konvergen dan Divergen: Penyelidikan Reabilitasnya Secara Empiris*. Diandara Pustaka Indonesia.

Sumarni, W. (2018). The influence of ethnoscience-based learning on chemistry to the chemistry's literacy rate of the prospective teachers. *Unnes Science Education Journal*, *7*(2), 198–205.

Sumarni, W., & Kadarwati, S. (2020). Ethno-Stem project-based learning: Its impact to critical and creative thinking skills. *Jurnal Pendidikan IPA Indonesia*, *9*(1), 1–11.

Suprapto,E., Saryanto, Sumiharsono, R.& Ramadhan, S. (2020). The analysis of instrument quality to measure the students' higher order thinking skill in physics learning. *Journal of Turkish Science Education, 17*(4), 520-527.

Supriyadi, S., Haeruddin, H., & Nurjannah, N. (2016). Peningkatan kemampuan memecahkan masalah antara model penalaran kausal berbasis etnosains dan sains modern. *JRKPF UAD*, *3*(2), 164–172.

Suratno, S., Wahono, B., Chang, C-Y., Retnowati, A., & Yushardi, Y. (2020). Exploring a Direct relationship between students' problem-solving abilities and academic achievement: a stem education at a coffee plantation area. *Journal of Turkish Science Education, 17*(2), 211-224

Sutarno, S., Setiawan, A., Suhandi, A., Karniawati, I., & Putri, D. H. (2017). Keterampilan pemecahan masalah mahasiswa dalam pembelajaran bandul fisis menggunakan model problem solving virtual laboratory. *Jurnal Pendidikan Fisika Dan Teknologi*, *3*(2), 164–172.

Temel, S. (2015). The problem-solving skills of the teachers ın various branches. *Educational Research and Reviews*, *10*(5), 641–647.

Ulger, K. (2018). The effect of problem-based learning on the creative thinking and critical thinking disposition of students in visual arts education. *Interdisciplinary Journal of Problem-Based Learning*, *12*(1), 3–6.

Warsono, D., Nursuhud, P. I., Darma, R. S., & Supahar, P. (2020). Empirical analysis of diagramatic representation test instruments using partial credit model in realizing learning outcomes. *International Journal of Instruction*, *13*(3), 617–632.

Wetzel, E., & Carstensen, C. H. (2014). Reversed thresholds in partial credit models: A reason for collapsing categories. *Assessment*, *21*(6), 765–774.

Yusuf, I. & Widyaningsih, S. W. (2018). Implementasi pembelajaran fisika berbasis laboratorium virtual terhadap keterampilan proses sains dan persepsi mahasiswa. *Berkala Ilmiah Pendidikan Fisika*, *6*(1), 18–28.