



Developing an Instrument to Detect Science Misconception of an Elementary School Teacher

Anatri Dessty

Ph.D. Scholar, Science Education Program, Yogyakarta State University, Yogyakarta, Indonesia, anatri.dessty2016@student.uny.ac.id

Zuhdan Kun Prasetyo

Prof., Department of Sciences Education, Yogyakarta State University, Yogyakarta, Indonesia, zuhdan@uny.ac.id

Suyanta

Assoc. Prof., Department of Chemistry Education, Yogyakarta State University, Yogyakarta, Indonesia, suyanta@uny.ac.id

Ihwan Susila

Assoc. Prof., Department of Management, Faculty of Economics and Business, Universitas Muhammadiyah Surakarta, Indonesia, ihwan_susila@ums.ac.id

Irwanto

Ph.D. Candidate, Department of Chemistry Education, Yogyakarta State University, Yogyakarta, Indonesia, irwanto.2016@student.uny.ac.id

This study aims to report the development an instrument that is standardized (reviewed by validity, reliability, and difficulty index) to detect science misconception in an elementary school teacher. This study used a 4-D model; defining, designing, developing, and disseminating. First, it was prepared with 47 opened-ended questions, and then it was validated by two experts (Physics and Biology) to find out content validity value and inter-rater reliability. The next stage was arranged an empirical test to 103 elementary school teachers in Surakarta, Indonesia by stratified purposive sampling. The data were analyzed using content validity formula, a measure of Agreement Kappa, Pearson's Correlation, Cronbach's alpha, and difficulty index. The result of this study indicated that there were 22 valid items. It can be concluded that this instrument is worthy to be used to detect the science misconception of an elementary school teacher.

Keywords: diagnostic test, science misconception, validity, reliability, difficulty index

Citation: Dessty, A., Prasetyo, Z. K., Suyanta, Susila, I., & Irwanto. (2019). Developing an Instrument to Detect Science Misconception of an Elementary School Teacher. *International Journal of Instruction*, 12(3), 201-218. <https://doi.org/10.29333/iji.2019.12313a>

INTRODUCTION

Based on Republic of Indonesia Constitution Number 14 Year 2005 concerning Teacher and Lecturer is mentioned that elementary school teacher must have four competencies; personality, social, pedagogy, and professional. These professional competences can be seen by the concept's mastery of five lessons, one of them is Science. Science is a collection of facts, concept, law, and theory that is found by systematic process, with developing supported scientific behavior (Chiappeta & Koballa, 2010). Science is a way of knowing or the values and beliefs inherent to the development of scientific knowledge (Liu & Lederman, 2007).

In the 2013 curriculum which refers to Minister of Education and Culture Regulation Number 37 year 2018 concerning Content Standard, the scope of elementary school science covers: 1) living being and the process of living, they are human, animals, plants and the interaction to the environment, also health; 2) materials, nature and the function that covers: liquid, solid and gas; 3) energy and its changes covers: force, sound, heat, magnetic, electricity, light and simple appliances; and 4) earth and universe covers: soil, earth, solar system, and any other outer space things. Many concepts in this material must be mastered by elementary school teacher. If the teacher not masters those concepts, it will create misconception. Misconception is a condition that shows differentiation on a person conceptual idea with scientific consensus that is accepted in general (Galvin & Simmie, 2015). Misconception definition also refers to incomplete misconception or simple on scientific concept that is disclosed by many students (Vosniadou, 2013). Misconception must be solved, because misconception has negative effect in the learning process (Aydin, 2012). Misconception also a big obstacle in learning Science (Suliyannah, Putri, & Rohmawati, 2018), because the Science concepts also underlies the concept understanding for the next stage (Nakhleh, 1992).

Just like a student, elementary school teacher can also experience misconception (Burgoon, Heddle, & Duran, 2010). The teacher that experience misconception will become the main cause to the student misconception. When this is not corrected, the misconception that is experienced by the elementary students becomes something that is difficult to be straighten and will be brought to the higher level (Kucukozer, 2009). Misconceptions can often pose strong barriers to understanding physics and many of them are detrimental to learning (Clement, Brown & Zietsman, 1989). Misconceptions can be the main source of learning problems at school if they are not detected and clarified immediately. Science misconception is a very significant problem at each level of education (Kumanda, 2015), especially at elementary school level. It is because elementary school is the most basic education. Misconception at the elementary school would bring to the next level. According to Tekkaya (2002), misconceptions will also adversely affect subsequent learning. The need to detect natural science subject misconceptions in elementary schools is essential, especially for elementary school teachers.

An effort to detect a misconception to elementary school teacher is really important to be done, because a teacher is considered as a people who gives a basic concept. This urgency is strengthened because a science misconception will give bad impact

significantly give effect to student's achievement (Ameyaw & Sarpong, 2011). Therefore, it is very important to detect Science misconception of elementary school teacher. This is the same to what is recommended by Ilyas (2018) to conduct this study especially in certain material, for example static electricity (Ersoy & Dilber, 2014).

To detect whether there is misconception, it can used diagnostic test, which is the test that is used to find out limitation of a student so the result can be used as base to give follow up in the form of proper treatment and suitable with the limitation that is lied in a student (Ministry of National Education, 2007). Several experiments to detect misconceptions have been carried out previously with varied types of diagnostic tests. Some of them namely; the detection of elementary school teacher candidates on a concept of the nature of matter and its change of being using a five-tier diagnostic test, with multiple choice tests and reasons open (Shahdan, 2018); identification of misconceptions in elementary school teachers in various concepts on physical and chemical material using multiple choice questions (Lawrenz, 1986); identification of misconceptions in prospective elementary school teachers about the state of the material, using an open ended question (Tatar, 2011); Identification of misconception of elementary school teacher education students in five-tier diagnostic tests (Bayuni, T.C, et all, 2018); identifying the chemical concept to elementary school teacher with open-ended question (Papageorgiou and Sakka, 2000); identify misconceptions in elementary school students using concept maps on Chemical concept material (Sindelkova, M, 2015); identify misconceptions in elementary school teacher candidates on concepts differences among solid, liquid and gaseous states of matter using the open-ended question (Tatar, E, 2011).

The study about misconception of elementary school teacher is still limited. Some of it detects the misconception in force and motion material using instrument of Force Concept Inventory (Gabunilas, 2017). The instrument in the form of open ended questions is used to detect sound material (Kucukozer, 2009); the instrument in the form of objective explanation, is used to detect misconception on photosynthesis material, density concept, pressure and air expansion, motion force concept, electrical stream, and sky object movement (Laksana, 2014). This kind of instrument is also used to detect the concept of pulling and pushing force, gravitation force, friction, magnetic force, simple appliance, the mass of earth thing, and nature of light (Pujayanto, 2012). The instrument in the form of conceptual question formed in two-level multiple-choice is used to detect misconception on force and motion material (Narjaikaew, 2013). It shows that efforts to detect misconceptions in elementary school teachers are still limited, and then need to be implemented in other concepts using appropriate diagnostic tools.

Some kind of diagnostic test that is used has weakness and strength. For the multiple-choice followed by explanation, can minimize guessing answer, and can determine the type error or the problem. The multiple-choice diagnostic test followed by an explanation is not appropriate for detecting misconceptions for elementary school teachers. It is because the type of multiple-choice tests still tends to be in the low order thinking such as remembering, understanding, and implementing the Bloom taxonomy. This is in line with the statement of Wilkie et all (2015) that multiple choice questions

are often regarded as simple concrete knowledge, as in the realm of knowledge and understanding. The availability of these answer choices tends to provide opportunities for guessing, and the reasons for choosing these answers are not explained rationally, but tend to be the origin of answering (Quaigrain & Arhin, 2017). So, understanding the concepts that will be explored from this instrument does not appear. Actually, many diagnostic questions have been developed about the type of reason multiple choice that effectively detects misconceptions for students (Pratiwi, HY, 2016). This present study is not for students but for teachers.

Test in the form of question and brief answer can apprehend complete information, even though it has difficulty in analyzing. But the two-level multiple-choice test is the question that is low-ordered-thinking, there still has possibility to guess (Adodo, 2013). The strength of the open-ended question is that they have practicability as an advantage to them. They are ideal for assessing how well is the students can summarize, hypothesize, finding relations & applying known procedures to new a situation. This kind of test assesses critical thinking skill or higher order thinking (Kurdi, 2015). Open ended question is better for achievement test that has complex relation with the concept application, problem analysis, or decision evaluation, so this explanation test can also be used for diagnostic test that has complex relation concept application, problem analysis, or decision evaluation (Walstad, 2006). This characteristic of the test is suitable for evaluation tool of learning process that is done by elementary school teacher so it becomes the activity that is dynamic and innovative that fit the protégé needs.

Based on the previous study about many Science concepts in other material that is not detected yet, it needs the development of instrument to detect misconception in the form of open-ended question. With the open-ended question instrument to detect science misconceptions in elementary school teachers, elementary teachers can share all concepts and concept application, problem analysis or decision evaluation, so all information can be apprehended also the error source can be found. The aim of this study is to reported the development of instrument to detect science misconception in elementary school teachers in form of open-ended questions. Then it will be obtained better instrument when considered by its validity, reliability, and difficulty index.

CONTEXT AND REVIEW OF LITERATURE

Misconception

Misconception shows the difference between pre-conception and scientific conception. Misconception refers to a concept that incompatible with scientific understanding (Berg, 1991). Science misconceptions occur a lot, because science contains many concepts. Not only in students, teachers can also experience the science of misconceptions and thinking of students as a result of these conditions (Pine et al., 2010). So that it can be said, one of the causes of misconception in students is the teacher. For teachers who do not master the concept, they need to be helped to learn so that they understand the concept correctly. Misconception is difficult to correct. With a good lecture method, misconceptions cannot be eliminated or avoided. Misconceptions are stable and

permanent, and become an obstacle in understanding scientific perspectives. Thus, it is important to consider how to truly uncover misconceptions.

Science Curriculum in Elementary School

Science in elementary school is a subject of learning that is thematically integrated. That is, science is taught along with the content of other lessons in one particular theme. As set out in the Minister of Education and Culture Number 37 year 2018 concerning Standard Content, the minimum competencies that must be mastered by students in elementary school are understanding factual knowledge by observing and asking questions based on their curiosity about themselves, God's creatures and activities, and objects found at home, at school and at the playground. To achieve this minimal competency, it is necessary to further elaborate on some basic competencies that are spread over the scope: 1) living and the process of living, 2) materials, nature and the function, 3) energy and its changes covers, and 4) earth and universe.

Factual knowledge is knowledge of specific, detailed, and elementary facts that can be captured by senses through the process of seeing, hearing, reading, or feeling. Thus, elementary school students need to be facilitated with something concrete in order to help understand the knowledge. In line with Piaget's learning theory, elementary school children (ages 7 to 12 years) have the main characteristics: student knowledge is built through the use of physical objects and real experiences; failed to build knowledge with future prospects without direct experience (Awwad, 2013).

METHOD

Participants

The sample was selected using stratified purposive sampling and was spread in several districts in Surakarta. The participants were 103 fourth-grade elementary school teachers from public (n=52) and private (n=51) schools, with perfect and good quality. The participant in this research has been certified by the Ministry of Education and Culture of Republic Indonesia.

Framework Development

The model of this study is procedural development, it is a descriptive study that shows the steps that should be followed to produce final product. This instrument development uses the 4-D model that is developed by Thiagarajan, Semmel and Semmel (1974), covers the step of defining, designing, developing and disseminating.

Defining

The first step in developing this instrument is defining, that is done in some stages. First, analyzing teacher characteristic that suit the instrument development design to detect misconception, that has relation with the primary competence in Science subjects of 4th grade in elementary school and competence achievement indicators that is arranged. The reason of the material choosing is based on position of integrated thematic Science subject, with minimum competence that must be achieved, so the instrument

arrangement is going to be easier to decide the competence achievement indicators that needs to be achieved. Second, beginning-ending analysis is done to develop misconception detection instrument based on literature that is used by experts in the previous study. Furthermore, it is also done by observing the condition in the field related to instrument forms of misconception detection and the urgency about why misconception on elementary school teacher should be detected. Third, is the process of determining the purpose of the instrument development to detect science misconception in elementary school teachers in form of open-ended questions.

Designing

Second step in developing the instrument is construction designing each question. These steps are divided into three parts, the first one covers curriculum analysis process in Science subject in 4th grade of elementary school. The studiers determine core competence, primary competence, and developing competence achievement indicator. The second is to describe answers criteria that is used to determine misconception or not, designing the beginning question, question script, and assessment rubric. The third is writing about question. Question is written in Indonesian.

Developing

The third step is the instrument developing that involving review by expert judgments. Content validity of the item is set by presenting the expert on physic and biology in elementary school. The expert is chosen based on the experience in education field and based on the study that they pursued. The main purpose of diagnostic instrument developing is explained by the experts before, then they were asked to review the truth of each question with purpose to all complete instrument.

Specifically, experts were asked to review each item based on criteria: a) each item substance to core competence and competence achievement indicators; b) clear presentation of picture or diagram of each item; and c) language aspect in each question. The experts review each item by giving comment to the instrument and relevant or not-relevant explanation in the validation sheet. Suggestions from the reviewer are used to fix the instrument. After it has fixed, this instrument is going to be tested.

Disseminating

This step was the application of the instrument developed wider. For example, for teachers in other schools or another area. The purpose of this dissemination is to find out the effectiveness of instruments in detecting misconceptions.

Instrument

Item Construction

The question is developed based on the content that is occurred in the content standard 2013 curriculum in elementary school of Science subject in 4th grade. There are 8 primary competences that are elaborated into two to three competence achievement indicators. Every indicator is elaborated into two to three questions.

Item Format

First of all, instrument to detect science misconception consist of 47 question of essay test by answer criteria is the result of concepts analysis from the developing question. If the answer is led to misconception condition, it will be continued with the interview based on the answer. To make the process easy in arranging question, it would be better when the question grid is arranged. Table 1 is providing examples on item arrangement of one primary competences.

Table 1
Question Grids in the Instrument to Detect science Misconception

	Competence Achievement Indicators	Question
3.7. Applying the nature of light in the relation of visual sensory	3.7.1. Concluding the nature of light in the relation of visual sensory	. Two same candle that is lit, the first candle is placed outside where the light is so bright. The second candle is placed in a dark room. Which light that will first reach to our eyes as the observer?

Scoring Procedure

To make it easy in giving score by teacher's response, then it is created an instruction in scoring. The giving score instruction is described in Table 2.

Table 2
Giving Score Instruction of Each Item

Score				
5	4	3	2	1
The concept appears and the explanation of the relation between concepts mostly suit the descriptor	The concept appears and explanation of the relation between concepts in parts suit the descriptor	The concepts appear and explanation of the relation between concepts is in fraction with the descriptor	The concept appears and the explanation of the concepts relation is not match with the descriptor	The concept does not appear and the explanation is not match with the descriptor

Data Analysis

This study was done to produce science misconception detection instrument that is valid and reliable. Valid has relation to validity, it is done by seeing the content validity index and validity between items. In this study, content validity can be seen through the agreement that used the expert agreement index that is suggested by Gregory (2004). This calculation is done by considering two categories that is very relevant and less relevant. The amount of expert agreement index is a comparison between the numbers of item that is considered very relevant by the two experts to all of the item number.

Based on the criterion, the amount of Content Validity (CV) is calculated with this formula:

$$CV = \frac{A}{(A+B+C+D)}$$

Where, A = the number of item that is considered very relevant by the two experts; B = the number of item that is considered very relevant by expert I, but Less relevant by

expert II; C = the number of item that is considered very relevant by expert II, but less relevant by expert I; and D = the number of item that is considered less relevant by expert I and II.

The index is around 0 to 1. The criteria of an instrument it must has Content validity that is good if it has the index minimum of .70 (Gregory, 2004).

After fulfilling CV index in the minimum of .70, the 47 item is tested to 8 teachers in empirical way to find out the validity of each items that is calculated by correlating in bivariate between question score with the total constructed score, with help of SPSS 23 program. Item is valid when the correlation score is above .50.

To find out the reliability, used inter-rater approval after the expert testing process was carried out. The inter-rater reliability processing used the SPSS 23. Reliability inter-rater program was analyzed using coefficient Cohen's Kappa in three categories. (Cohen, 1960):

Kappa < 0.4 for poor; 0.4 < Kappa < 0.60 for fair; 0.60 Kappa < 0.75 for good; and Kappa > 0.75 for excellent categories.

The criteria of the instrument must have an inter-rater reliability in a good category, with the Cohen coefficient value no less than 0.4. After the fulfillment of inter-rater reliability, is done by testing the question once to the 103 respondents. Cronbach's alpha Reliability is calculated using statistic software SPSS 23. If the coefficient reliability is higher than $\alpha = .70$, then the instrument will be accepted as high coefficient reliability (Ozdamar, 2013). To find out difficulty index, it is calculated by this formula:

$$\text{Difficulty Index} = \frac{M_{S(i)} + M_{W(i)}}{N_B * m_i}$$

Where, $M_{S(i)}$ = the amount score of higher group; $M_{W(i)}$ = the amount score of lower group; N = the respondents; and m = maximum score of each question.

With the classification: .81 to 1.00 (very easy); .61 to .80 (easy); .41 to .60 (medium); .21 to .40 (difficult); and .00 to .20 (very difficult) (Jandaghi, 2011).

To determined the existence of misconception, it was grouped the level of understanding of the concept from Abraham by Marek (1986). It is divided in 3 categories: do not understanding, misconception and understanding. This is presented in Table 3.

Table 3
Categories of Concept Understanding

Category	Understanding Level	Criteria
Do not understand	<ul style="list-style-type: none"> • Do not respond • Do not understand 	<ul style="list-style-type: none"> • Do not answer • Answer "I do not know" • Repeat the question • Answering but not related to questions • Answering with unclear answer
Misconception	<ul style="list-style-type: none"> • Misconception • Partly understand 	<ul style="list-style-type: none"> • Answering with illogical explanations • The answers indicate the understanding of the concept but also show a misconception.
Understand	<ul style="list-style-type: none"> • Partly understand • Understand the concept 	<ul style="list-style-type: none"> • The answers indicate the partly concept understanding without any misconceptions • The answers show the fully concept understanding with all correct explanations

FINDINGS

Validity

Content Validity

In the beginning, it is developed 47 questions in the form of essay test, which is validated by physic and biology expert. The expert validates the compatibility between competence achievement indicators that is developed by primary competence, also validates compatibility of question to competence achievement indicator. Response by the experts says that in the developing instrument, there are 41 question that is considered relevant by the two experts, 6 questions that is considered relevant by expert II but less relevant by expert I, and there is no question that is very relevant by expert I but less relevant by expert II. The result is presented in the Table 4.

Table 4
Question Number Data with Criteria A, B, C, and D

Question criteria	Question number	Total
Relevant by both experts (A)	1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 31, 32, 33, 34, 35, 37, 38, 39, 41, 42, 43, 44, 45, and 47.	41
Very relevant by expert I, but less relevant by expert II (B)	–	0
Very relevant by expert II, but less relevant by expert I (C)	–	0
Less relevant by both experts (D)	6, 19, 30, 36, 40, and 46	6
Total		47

The content validity (CV) index is obtained 0.87. After this, the instrument is done the empirical test. In this test, the teachers answer the question for 1.5 hours. The highest score is 5 and the lowest is 1 of each item. The score is used to calculate bivariate correlation value between each item score to the total score, by using SPSS program. Correlation value of each question is shown in the Table 5.

Table 5
Validity Score of Each Question

Question	Pearson Correlation	Explanation	Question	Pearson Correlation	Explanation	Question	Pearson Correlation	Explanation
1	.074	Drop	17	.25	Drop	33	.687*	Valid
2	.603	Valid	18	-.264	Drop	34	.688*	Valid
3	.653	Valid	19	.327	Drop	35	.447	Drop
4	.013	Drop	20	.622	Valid	36	.394	drop
5	.648	Valid	21	.571	Valid	37	.422	Drop
6	.421	Drop	22	.318	Drop	38	.662	Valid
7	.841	Valid	23	.620	Valid	39	.685*	Valid
8	.440	Drop	24	.467	drop	40	.227	Drop
9	.653	Valid	25	.077	Drop	41	.545	Valid
10	.175	Drop	26	.536	Valid	42	.644	Valid
11	.764*	Valid	27	.106	Drop	43	.085	Drop
12	.331	Drop	28	.274	Drop	44	.539	Valid
13	.189	Drop	29	.666	Valid	45	.491	Drop
14	.664	Valid	30	.556	Valid	46	.543	Valid
15	.038	Drop	31	-.096	Drop	47	.387	drop
16	.774*	Valid	32	.224	Drop			

According to Table 5, 47 questions that is tested in empirical way, there are 22 questions is valid, which means that the questions is able to measure something need to be measured (Nunnally, 1978), and 25 questions are invalid based on the minimum correlation value that is required ($r > .05$).

Reliability

Inter-rater reliability is obtained at 1.00. Inter-rater reliability from the results of calculations with SPSS 23 is shown in Table 6.

Tabel 6
Inter-Rater Reliability

Symmetric Measures		Value	Asymptotic Standardized Error	Approximate Tb	Approximate Significance
Measure of Agreement	Kappa	1.000	.000	6.856	.000
N of Valid Cases		47			

The Cronbach's alpha coefficient consistency is .902. The reliability test result used SPSS 23 program is shown in Table 7.

Table 7
Statistics Reliability

Cronbach's Alpha	Cronbach's Alpha based on Standardized Items	Valid Item Total
.895	.902	22

Difficulty Index

In this study, from 22 valid questions, there are 8 easy category, 8 medium category, and 5 difficult category, with the difficulty index ranged from .27 to .89. There is 1 question considered as very easy, with the amount of difficulty index .93. So, this question must be fixed and modified to obtain the new better question that is good and acceptable. This distribution is shown in Figure 1.

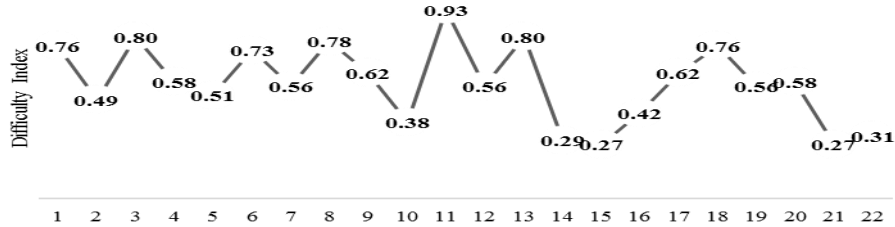


Figure 1
Difficulty Index of 22 Valid Questions

According to 22 valid and reliable items, the following presents the teacher's response to one question that is analysed to determine that if the teacher is in misconception or not.

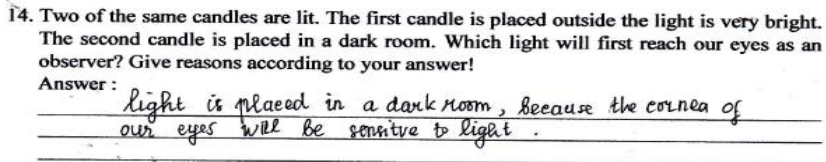


Figure 2
Analysis of Question Number 14 and Teacher's Result

From 22 items valid and reliable, choosed 20 items to detect elementary school teacher's science misconceptions. The results are provided on Table 8.

Table 8
Summary of the Teacher's Misconception in Science

Question	Science Concept	Misconception (%)
1	The eagle's beak	35,72
2	The working process of xylem phloem	85,75
3	Cactus's survive	78,57
4	Mango plants survive	0
5	Life cycles of Butterflies and cockroaches	35,72
6	Concept about force.	71,43
7	Gravitation of force	100
8	Electric static	0
9	Gravitation force in vacuum condition	100
10	Renewable and non-renewable energy sources.	50
11	Changes of the energy	35,71
12	Difference about echo and resonance	35,72
13	Low high and strong weak of the sound	88,23
14	Light velocity	92,86
15	Visibility an object.	71,43
16	Light source	28,57
17	Flat mirror shadow	64,28
18	Natural resource that can be renewed.	0
19	Concept 3R	28,57
20	The effect of pesticide	7,43

Based on Table 8, the highest misconception is a concept of gravitation of force (100%), gravitation of force in vacuum condition (100%), and light velocity (92,86%).

DISCUSSION

The validity of an instrument can be identified through content and empirical analysis from the response data of each item in the instrument (Lissitz & Samuelsen, 2007). Therefore, to find out instrument validity that is developed in this study, is done in two steps, it is the content analysis and empirical test to find out the response of each item.

CV index is around 0 to 1. CV coefficient that is closer to 0 has the meaning that agreement index from validator on the instrument item relevance with their indicators is getting lower. Otherwise, CV coefficient closer to 1 the agreement index from validator on the instrument item relevance with their indicators is getting higher. Content validity is related to the rational analysis of the domain to be measured to determine the representation of the instrument with the ability to be measured (Retnawati, 2016). In this study, it is obtained CV of .87 (closer to 1). It means that agreement index by validator is higher to the item that is in the instrument that suits the indicator that is developed. This thing shows that the instrument has the items that cover all content domains that need to be measured.

This study also reviewed content validity aspect, because it needs to obtain qualified instrument, so when it is used to find out data about the misconception condition on elementary teacher also get the qualified data. It is the same to the statement that content validity is an important topic for the studier that needs quality in good measuring (Polit, Beck, & Owen, 2007). In measuring content validity, two experts on physic and biology education, involved in this study. Next, it needs to be tested in empirical way to receive validated data that is more valid. This is the same as Creswell (2012) that is clearly stated that how much is content validity index can be originated from the evidence during the empirical test and panel expert in the field of study to validate questionnaires.

The inter-rater reliability is obtained at 1.00. It is shows that the reliability of the instrument developed is included in the special category. It was supported by the Cronbach's alpha coefficient consistency is .902. It means that the 90,2% certainty of the consistency of the test items in approximately yielding the same result repeatedly. It shows that the test item is very reliable. This finding is supported by Ceniza and Cereno (2012) that state when reliability coefficient is around .81 to 1.0 it is indicated high reliability, .61 to .80 signified a moderate reliability, .41 to .60 signified a fair reliability, .10 to .40 signified a slight reliability, and less than .10 signified no reliability. This is the same as Ozdamar (2013) and Gelişli, Beisenbayeva, & Gumilyov (2017), if the reliability coefficient is higher than .70, it will be accepted as high reliability coefficient. It refers to Fraenkel and Wallen (2008) that is stated when instrument reliability is higher than .70, the sufficiency of this instrument will be owned (Thaneerananon, 2016). It can be concluded that detector instrument has very high reliability.

Difficulty index is a percentage that gives data to some of the student who can answer correctly (Khoshaim & Rashid, 2016). Difficulty index has the strong effect to the variability test score even on the accuracy where the score test differentiates between

groups (Thorndike, Cunningham, Thorndike, & Hagen, 1991). Teacher's answer response analysis to the question can be used to find out the quality of each item also the item quality in particular. If the difficulty index is around 20% to 90%, the question will be concluded as good and acceptable. Difficulty index is considered to be good when it is 40% to 60%. When the difficulty index is less than 20% (too difficult) and more than 90% (too easy), it is not accepted criteria and need to be modified or fixed (Quaigrain & Arhin, 2017). It is the same to the statement, in this study, from 22 valid questions, there are 8 questions with easy category, 8 questions of medium category, and 5 difficult category question, with the difficulty index ranged from .27 to .89. There is 1 question considered as very easy, with the amount of difficulty index .93. This question that is considered as very easy category must be fixed and modified to obtain the new better question that is good and acceptable (difficulty index around 20% to 90%).

Figure 2 presented about which light will first reach our eyes as an observer when there are two same candles lit at the same time. The first candle is placed outside which the light is so bright, the second candle is placed in the room that is dark/low light.

The answer that makes sense according to scientific purpose is both of the candle lights will reach our eyes at the same time, no one is overtaking first. It is because the speed of light in atmosphere condition is the same. Environment factor (dark and bright) has no effect on the light velocity. It needs explanation to the rational reason that both lights have the same intensity. However, the teacher said that there is a difference light velocity in those two different places. It is the same as the study result of Bintoro et al. (2009) that states teacher is not completely understand that light velocity only determined by the medium density that is going through by the light. Teacher still assumes that the velocity of the candle light is effect by light intensity of the environment. This is the same as the study result that the distance light travels depends on day or night (Sampson, 2013). The lesser the intensity in the environment (low light), then the light travels quicker.

The teacher's answer reveals the concept of light that is relatable to the visual sensory (one of them is cornea), not the concept of light intensity and environment factor. It can be said that the teacher is not good enough in the analyzing the concept, so it can also be said that the teacher is not completely able to analyzing the problem through analysis of the concept relation. So, in making a decision, it is not accurate. Based on this data, it can be concluded that the teacher experience misconception. Based on the answer analysis, it can be seen that teacher has weakness in analyzing the concept in a problem. So, the policy maker such as the headmaster even educational service can determine the follow-up to handle that problem. The headmaster is a party that can directly provide solutions for teachers in their schools who experience misconceptions. However, there are higher authority than the principal, namely the head of the department both at the city and provincial level. They are expected to work together to reduce the misconception, so the quality of the teacher will improve.

CONCLUSION

The finding of this study is proved that the development of science misconception detection instruments for elementary school teachers is valid and reliable in as many as 22 questions. Content validity (CV) coefficient is in the amount of .87, Pearson Correlation is in the ranged of .54 to .84, inter-rater reliability is 1.00, and Cronbach's alpha internal consistency coefficient is in the amount of .902. All items have difficulty index of .27 to .89. Therefore, this instrument can be used to detect science misconception, through the process of concept analyzing, problem analyzing, or decision evaluation.

For the policy maker in the development of elementary teacher professionalism, the valid integrated instrument can be used to find out accurately, where it can reveal the weakness of the teacher in how the teacher understanding in Science material, so it can be decided the more effort on decreasing the weaknesses, for example through the scientific discussion or workshop that can reduce misconception.

REFERENCES

- Abraham, M., R., & Marek E.(1992). Understanding and misunderstanding of eighth graders of five chemistry concepts found in textbooks. *Journal of Research in Science Teaching*, 29(2), 105–20. <https://doi.org/10.1002/tea.3660290203>.
- Adodo, S. O. (2013). Effects of two-tier multiple choice diagnostic assessment items on students' learning outcome in basic science technology (BST). *Academic Journal of Interdisciplinary Studies*, 2(2), 201–210. DOI: 10.5901/ajis.2013.v2n2p201.
- Ameyaw, Y., & Sarpong, L. (2011). The application of some conceptual approaches in rectifying teacher's misconceptions on some sciences topics in the GA South district in the greater-Accra region of Ghana. *Journal of Education*, 1(1), 16–24. <http://www.bioinfo.in/contents.php?id=58>.
- Awwad, A. A. (2013). Piaget's Theory of Learning. *Interdisciplinary Journal of Contemporary Research in Business*, 4(9), 106–129. Retrieved from <https://journal-archievs27.webs.com/106-129.pdf>.
- Aydin, S. (2012). Establishment for misconceptions that sciences teacher candidates have about geometric optics. *The Online Journal of New Horizons in Education*, 2(3), 7–15.
- Bayuni, T.C., Sopandi, W., & Sujana, A. (2018). Identification misconception of primary school teacher education students in changes of matters using a five-tier diagnostic test. *Journal of Physics: Conf Series*, 1013(1), 012086. <https://iopscience.iop.org/article/10.1088/1742-6596/1013/1/012086/pdf>.
- Berg, E. van D. (1991). *Physics's Misconceptions and Remediation*. Salatiga: Universitas Satyawacana.
- Bintoro, R. R., Hambaora, J. J., Radzumaearni, P., Sudarmi, M., & Rondonuwu, F. S. (2009). Student's conception of the propagation and speed of light. *Proceedings of the*

National Seminar on Research, Education and Application of Mathematic and Natural Science, 225–230. <https://eprints.uny.ac.id/12421/>.

Burgoon, J., Heddle, M., & Duran, E. (2010). Re-examining the similarities between teacher and student conceptions about physical science. *Journal of Science Teacher Education*, 21(7), 859–872. <http://www.jstor.org/stable/43156582>.

Ceniza, J. C., & Cereno, D. C. (2012). Development of mathematic diagnostic test for DORSHS. Retrieved from <http://www.doscst.edu.ph/index.php/academics/graduateschool/publication/category/5-volum-1-issue-1-2012>.

Chiappeta, E. L., & Koballa, T. R. (2010). *Science Instruction in the Middle and Secondary Schools: Developing fundamental knowledge and skill*. (7th ed.). Boston, MA: Allyn and Bacon.

Clement, J., Brown, D. E., & Zietsman, A. (1989). Not all preconceptions are misconceptions: finding 'anchoring conceptions' for grounding instruction on students' intuitions. *International Journal of Science Education*, 11, 554-565. <https://doi.org/10.1080/0950069890110507>.

Cohen J. A. (1960). *Coefficient of agreement for nominal scales*. *Educ Psychol Meas.* 20: 37–46. <https://doi.org/10.1177/001316446002000104>.

Creswell, J. W. (2012). *Educational study: Planning, conducting, and evaluating quantitative and qualitative study Boston*. (4th ed.). Boston, USA: Pearson Education.

Ersoy, F. N., & Dilber, R. (2014). Comparison of two different techniques on students' understandings of static electric concepts. *International Journal of Innovation and Learning*, 16(1), 67–80. DOI: 10.1504/IJIL.2014.063374.

Gabunilas, M. L. (2017). Addressing elementary teachers' alternative conceptions in force and motion with an interactive computer simulation. *International Journal of Physics*, 5(5), 147–153. DOI: 10.12691/ijp-5-5-1.

Galvin, E., & Simmie, M. G. (2015). Identification of misconceptions in the teaching of biology: A pedagogical cycle of recognition, reduction and removal. *Higher Education of Social Science*, 8(2), 1–8. DOI: 0.3968/6519.

Gelişli, Y., Beisenbayeva, L., & Gumilyov, L. N. (2017). Scientific inquiry competency perception scale (the case of Kazak post-graduate students) reliability and validity study. *International Journal of Instruction*, 10(1), 273–288. DOI: 10.12973/iji.2017.10117a.

Gregory, R. J. (2004). *Psychological testing: History, principles, and applications*. Needham Heights, MA, US: Allyn & Bacon.

Hola, I. A. (2004). Biological science misconceptions amongst teachers and primary students in Jordan: diagnosis and treatment. *The Internet Society: Advances in Learning, Commerce and Security*, 109–118.

Ilyas, A., & Saeed, M. (2018). Exploring teachers' understanding about misconceptions of secondary grade chemistry students. *International Journal for Cross-Disciplinary*

Subjects in Education, 9(1), 3323-3328. DOI: 10.20533/ijcdse.2042.6364.2018.0444.

Khoshaim, H. B., & Rashid, S. (2016). Assessment of the assessment tool: Analysis of items in a non-mcq mathematics exam. *International Journal of Instruction*, 9(1), 119–132. DOI: 10.12973/iji.2016.9110a.

Kucukozer, A (2009). Investigating prospective science teachers' misconceptions of sound. *Journal of Elementary Education Online*, 8(2), 313–321. <http://ilkogretim-online.org.tr>.

Kumandaş, B. (2015). *Misconceptions in biology education: a Review of Relevant Research*. Unpublished Master's Thesis. Bilkent University, Ankara. <http://hdl.handle.net/11693/28941>.

Kurdi, M. S. (2015). *Essay type questions & their improvements & short answer questions*. Karnataka: Department of Anaesthesiology Karnataka Institute of Medical Hubli. https://www.researchgate.net/publication/281026940_ESSAY_TYPE_QUESTIONS_THEIR_IMPROVEMENTS_SHORT_ANSWER_QUESTIONS/download.

Laksana, D. N. L. (2014). Profile of understanding the science concepts of elementary school teachers in Ngada. *Journal of Scientific Education*, 1(1), 27–38. <http://www.citrabakti.ac.id/ejournal/index.php/jipcb/article/view/37/18>.

Lawrenz, F, (1986). Misconceptions of physical science concepts among elementary school teachers. *Journal of School Science and Mathematics*, 86(8), 654 – 660. <https://doi.org/10.1111/j.1949-8594.1986.tb11669.x>.

Lissitz, R. W., & Samuelsen, K. (2007). Further clarification regarding validity and education. *Educational Researcher*, 36(8), 482–484. <https://www.jstor.org/stable/4621104>.

Liu, S., & Lederman, N. G. (2007). Exploring prospective teachers' worldviews and conceptions of nature of science. *International Journal of Science Education*, 29(10), 1281–1307. <https://doi.org/10.1080/09500690601140019>.

Ministry of National Education. (2007). *Diagnostic Test*. Jakarta: National Education Departement. Minister of Education and Culture Regulation Number 37 Year 2018 about content standard

Nakhleh, M. (1992). Why some students don't learn chemistry: chemical misconceptions. *Journal of Chemical Education*, 69(3), 191–196. DOI: 10.1021/ed069p191.

Narjaikaew, P. (2013). Alternative conceptions of primary school teachers of science about force and motion. *Procedia - Social and Behavioral Sciences*, 88, 250–257. <https://doi.org/10.1016/j.sbspro.2013.08.503>.

Nunnally, J. C. (1978). *Psychometric theory. 2nd Edition*, McGraw-Hill, New York.

Ozdamar, K. (2013). *Statistical data analysis by packet programs (9th ed.)*. Eskisehir: Nisan Press.

- Papageorgiou, G., & Sakka, D. (2000). Primary school teachers' views on fundamental chemical concepts. *Chemistry education: Research and practice in Europe*, 1(2), 237-247. www.chem.uoi.gr/cerp/2000_May/pdf/32-05papageorg.pdf.
- Pine, K. Messer, D., & St. John, K. (2010). Children's Misconceptions in Primary Science: A Survey of teachers' views. *Research in Science & Technological Education*, 19(1). <https://doi.org/10.1080/0263514012004624>
- Polit, D. F., Beck, T., & Owen, S. V. (2007). Focus on study methods is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Study in Nursing and Health*, 30, 459–467. doi:10.1002/nur.20199 PMID : 17654487.
- Pratiwi, H.Y. (2016). The development of multiple choice test instruments to identify the characteristics of the thermodynamic concept of students at the Physical Education at Kanjuruhan University, Malang. *Journal of Educational Inspiration*, 6(2), 842-850.
- President of the Republic of Indonesia. (2005). *The Laws of Republic Indonesia number 14 in 2005 regarding Teacher and Lecturer*. Jakarta: President of the Republic of Indonesia.
- Pujayanto. (2012). Misconception of physics at elementary school teacher. *Journal of Material and Physical Learning*, 1(1), 22–24. jurnal.fkip.uns.ac.id/index.php/fisika/article/view/1856.
- Quaigrain, K., & Arhin, A. W. (2017). Using reliability and item analysis to evaluate a teacher-developed test in educational measurement and evaluation. *Journal of Educational Assessment & Evaluation | Research Article*, 4(1301013), 1-11. <http://dx.doi.org/10.1080/2331186X.2017.1301013>.
- Retnawati, H. (2016). Proving content validity of self-regulated learning scale (the comparison of Aiken index and expanded Gregory index). *Study and Evaluation in Education*, 2(22), 155–164. <http://journal.uny.ac.id/index.php/reid>.
- Sampson, V., & S. Schleigh. (2013). *Scientific argumentation in biology: 30 Classroom activities*. Arlington, Virginia: NSTA Press.
- Sindelkova, M., & Pluckov, I. (2015). The most common primary school pupils' misconceptions based on the concept of protection in the chemical context. *International Letters of Social and Humanistic Sciences*, 65, 130-137. <https://doi.org/10.18052/www.scipress.com/ILSHS.65.130>.
- Suliyannah, Putri, H. N. P. A., & Rohmawati, L. (2018). Identification student's misconception of heat and temperature using three-tier diagnostic test. *Journal of Physics: Conference Series*, 997(1). <https://doi.org/10.1088/1742-6596/997/1/012035>.
- Tatar, E. (2011). Prospective primary school teachers' misconceptions about states of matter. *Journal of Educational Research and Reviews*, 6(2), 197–200. <http://www.academicjournals.org/ERR>.
- Tekkaya, C. (2002). *Misconception as barrier to understanding biology*. Hacettepe

Universitesi Egitim Fakültesi Dergisi, 23, 259–266.
www.efdergi.hacettepe.edu.tr/yonetim/icerik/makaleler/971-published.pdf.

Thaneerananon, T. (2016). Development of a test to evaluate students' analytical thinking based on fact versus opinion differentiation. *International Journal of Instruction*, 9(2), 123–138. http://www.e-iji.net/dosyalar/iji_2016_2_9.pdf. DOI: 10.12973/iji.2016.929a.

Thiagarajan, S., Semmel, D. S., & Semmel, M. I. (1974). *Instructional development for training teachers of exceptional children: A sourcebook*. Minneapolis: Leadership Training Institute/Special Education, University of Minnesota.

Thorndike, R. M., Cunningham, G. K., Thorndike, R. L., & Hagen, E. P. (1991). *Measurement and evaluation in psychology and education* (5th ed.). New York: Mac Millan.

Vosniadou, S. S. I. (2013). Conceptual change from the framework theory side of the fence. *Journal of Science & Education*, 23(7), 1427–1445. <https://link.springer.com/article/10.1007/s11191-013-9640-3>.

Walstad, W. B. (2006). Testing for depth of understanding in economics using essay questions. *Journal of Economic Education*, 37(1), 38–47. <https://www.tandfonline.com/doi/abs/10.3200/JECE.37.1.38-47>.

Wilkie, R. M., Harley, C., & Morrison, C. (2015). High level multiple choice questions in advanced psychology modules. *Psychology Learning and Teaching* 8(2), 30–36. <https://doi.org/10.2304/plat.2009.8.2.30>.