Development of Reasoning Test Instruments Based on TIMSS Framework for Measuring Reasoning Ability of Senior High School Student on the Physics Concept

Muslim*, ASuhandi and M G Nugraha

Physics Education Department, Universitas Pendidikan Indonesia, Jl. Dr. Setiabudhi 229 Bandung, Indonesia, 40154

*Corresponding author’s email: muslim@upi.edu

Abstract. The purposes of this study are to determine the quality of reasoning test instruments that follow the framework of Trends in International Mathematics and Science Study (TIMSS) as a development results and to analyse the profile of reasoning skill of senior high school students on physics materials. This research used research and development method (R&D), furthermore the subject were 104 students at three senior high schools in Bandung selected by random sampling technique. Reasoning test instruments are constructed following the TIMSS framework in multiple choice forms in 30 questions that cover five subject matters i.e. parabolic motion and circular motion, Newton’s law of gravity, work and energy, harmonic oscillation, as well as the momentum and impulse. The quality of reasoning tests were analysed using the Content Validity Ratio (CVR) and classic test analysis include the validity of item, level of difficulty, discriminating power, reliability and Ferguson’s delta. As for the students’ reasoning skills profiles were analysed by the average score of achievements on eight aspects of the reasoning TIMSS framework. The results showed that reasoning test have a good quality as instruments to measure reasoning skills of senior high school students on five matters physics which developed and able to explore the reasoning of students on all aspects of reasoning based on TIMSS framework.

1. Introduction

One of the goals of physics subjects in the curriculum of 2013 is that the students have the competency to develop the ability of reasoning skill within inductive and deductive analysis using the concepts and principles of physics to explain natural events and solve the problems, both qualitatively and quantitatively [1]. This purpose is in line with one of the cognitive domains in TIMSS, namely reasoning dimension. Reasoning can be defined as the ability to think logically to solve the problem using scientific methods [2,3,4], while scientific reasoning skills are the competencies needed to develop critical thinking skills students [5]. Therefore, reasoning skills need to be trained in physics teaching and assessed using an assessment instrument test with good test construction.

Assessment is one important component and it cannot be separated from the learning activities. Assessment was conducted to determine the extent of students’ ability to understand and mastering the concept. In Permendikbud number 23 in 2016, the scope of assessment includes a competence attitudes, knowledge, and skills [1]. Assessment includes all aspects of competencies using appropriate techniques, to monitor the development of the learners ability.

Results of an international test into the cognitive abilities of Indonesian students from Trends in International Mathematics and Social Study (TIMSS) conducted by the IEA (International Association for the Evaluation of Educational Achievement) are in low category. This result due to the
ability of Indonesian students are currently have the knowing ability and not accustomed to solve the application and reasoning problems. This Indonesian student difficulties caused by students’ reading strategy is still very limited so that the level of reasoning is still low, linear, and fragmented (not comprehensive). Learning process should be encourage students to find out from various sources of observation, be able to formulate a problem (queries) not only solve the problem, to train students think analytically (decision-making) and not thinking mechanistic (routine) and have capability to cooperate and collaborate in solving the problem.

Mullis et al (2014) stated that in TIMSS 2015, the cognitive dimension are divided into three domains which describe the thought process that can be used to complete the science developed item [6]. The first domain is knowing, discusses the student's ability to remember, recognize and explain the facts, concepts, and procedures as the foundation in science. The second domain is applying, focusing on the use of knowledge to produce explanations and solve practical problems. The third domain is reasoning, including the use of evidence and scientific understanding to analyse, synthesize, and generalize, often in unfamiliar situations and complex context. Aspects of knowing, application, and reasoning in the cognitive abilities as implemented in TIMSS can be used to show the students’ thinking skills profile. Of these three aspects, knowing and application aspects are basic thinking skills, while reasoning skills is high-level thinking skill. TIMSS 2015 classifies reasoning ability (reasoning) into eight categories: analyse, integrate/synthesize, hypothesize/predict, design, draw conclusions, generalize, evaluate, and justify [7].

Based on analysis result of the Physics national examination (UN), found that most of the problems of physics are tested only measure low-level cognitive abilities that only measures cognitive ability to remember, understanding and applying. The tendency of the assessment construction is more mechanistic, primarily on the use of formulas of physics. While question that can measure high-level thinking skills are rare. Based on aspects of the reasoning, the questions of physics tested yet measure reasoning ability, especially when referring to the TIMSS reasoning skills framework. In line with Wasis (2014) research based on analysis of the Physics national examination in 2013 which showed that the highest proportion of matter is measure applying cognitive processes with the conceptual knowledge dimensions [8].

Considering there are linkages between assessment and learning process and the importance to measure the reasoning ability, as well as national examination in 2016 that based on reasoning which refers to the TIMSS framework, it is necessary to develop the instrument test to measure reasoning skills, so it can serve as a guide for physics teachers in teaching physics at schools to equip students reasoning skills. It is expected that the students will be able to develop the reasoning skill so that they can compete with other students both nationally and internationally. Mardapi (2008) states that the best test instrument can improve the quality of assessment that describe students’ ability profiles [9]. The quality of instrument tests can be determined by analysing the validity, reliability, distinguishing features and the level of test difficulties [10].

2. Research Methodology
This research used research and development (R&D) design. The research focused on the development of instruments test to measure reasoning skills in high school physics material. The development step using 4D stage, which includes: 1) define, 2) design, 3) develop, and 4) disseminate [11], but in this research the development step only from define to develop stage. Research subjects were 104 students at three high schools in Bandung. Sampling was conducted by random sampling technique. Reasoning instrument test is constructed following the TIMSS framework and it has 30 multiple choice questions (MCQs) that cover parabolic and circular motion, Newton’s gravity law, work and energy, harmonic oscilation, momentum and impulse. The quality of reasoning skill tests were analysed using the Content Validity Ratio (CVR) based on assessment results by validator (expert judgment) to obtain the content validity and construct validity items, use the formula:
\[ CVR = \frac{n_c - N}{\frac{N}{2}} \]  

\( n_c \) is number of validator agreeing on item test, \( N \) is total number of validator. The Category of CVR according to Lawshe shown in Table 1 [12].

| Index CVR     | Criteria          |
|---------------|-------------------|
| 0.00 – 0.33   | Not suitable      |
| 0.34 – 0.67   | Suitable          |
| 0.68 – 1.00   | Very suitable     |

Furthermore, analysis conducted to the instrument test, including: the validity of the item, level of difficulty, discriminating power, and reliability of the test. Instrument test that have been constructed and analysed, used to get a profile of students’ reasoning skills, which is based on the achievement of the average score on all reasoning skills aspects.

3. Results and Discussion
Reasoning skills instrument test was developed based on the basic competencies of physics of the 2013 curriculum and referring to the eight categories of reasoning skills by TIMSS 2015 frameworks, which are analyse (A), integrate/synthesize (I), hypothesize/predict (H), design (D), draw conclusions (DC), generalize (G), evaluate (E), and justify (J). Based on the physics material and reasoning skills, matrix for distributing items was created as presented in Table 2.

| No. | Physics subject                | Reasoning skills category and number of item test | Total items |
|-----|--------------------------------|--------------------------------------------------|-------------|
| 1   | Parabolic circular motion      | A : 1    I: 2    H: 3    D: 4    DC: 5    G: -    E: -    J: 5 | 5           |
| 2   | Newton’s gravity law           | A : 6    I: -    H: -    D: 7    DC: 8    G: 9    E: 10   J: -     | 5           |
| 3   | Work and energy                | A : -    I: 11   H: 12   D: -    DC: 13   G: 14   E: 15   J: 16   | 6           |
| 4   | Simple harmonic oscillation    | A : 17   I: 18   H: 19   D: 20   DC: 21   G: -    E: 22   J: 23   | 7           |
| 5   | Momentum and impulse           | A : 24   I: 25   H: 26   D: -    DC: 27   G: 28   E: 29   J: 30   | 7           |
|     | Total item test                | A : 4    I: 4    H: 3    D: 3    DC: 5    G: 4    E: 4    J: 3    | 30          |

CVR test results based on five validators to 30 items of the instrument tests showed that almost all of the questions have a CVR value between of 1 and 0.60 which means that all items meet the suitable criteria. Although the results of the validation have met the criteria, there are some suggestions for improvement given by validator that are associated with changes in answer options and the category of reasoning skills because it does not appropriate, and improvement in writing the questions.

1.1. Characteristics of reasoning skills instrument test
Characteristics of the reasoning instrument tests obtained after limited testing and revision then performed extensive testing with a sample of 104 students in three high schools in Bandung. Based on analysis using the Anates V4 software, the value of validity of the item, level of difficulty and discriminating power presented in Figure 1.
Base on the analysis of the 30 items shown that validity: 11 items low category, 18 items sufficient category, and one item high-category. Thus, for discriminating power: 1 item in enough category and 29 items good category, for level of difficulty: 9 items easy category, 19 items medium category, and 2 items difficult category. Furthermore, analysis result found that the value of reliability test is 0.82, which mean very high category that indicates the reasoning developed test is reliable for use in assessing the students [13]. Then, it can be seen that the reasoning developed instrument test has good quality as a test instrument to measure reasoning skills.

1.2. Profile of student reasoning skills

Students’ reasoning skills profile based on the results of implementation test by using prototype and developed reasoning test are shown in Figure 2.

Based on Figure 2, it can be seen that the achievements of the students’ reasoning skills is good enough with average score of 76.2 and none of the students’ reasoning component that is less than 50. The ability of the students achieved the highest reasoning in the category of formulate hypotheses while analyse ability was lowest components.

4. Conclusion
Characteristics of reasoning skills instruments test referring to TIMSS framework in physics at high school developed based on quality tests results has good quality to measure reasoning skills in physics teaching materials: parabolic and circular motion, Newton’s gravity law, work and energy, harmonic oscillation, momentum and impulse. Reasoning skill instrument test is able to explore the reasoning skills of students on all aspects of reasoning based on TIMSS framework with enough category.

5. References
[1] Kemendikbud 2016 Peraturan Menteri Pendidikan dan Kebudayaan Nomor 23 Tahun 2016 tentang Standar Penilaian Pendidikan
[2] Ammer 2006 Basic Proficiencis in Scientific Reasoning Forsburg (United States : Forsburg State University)
[3] Zimmerman C 2007 The Development of Scientific Reasoning Skill: What Psychologist Contribute to an Understanding of Elementary Science Learning. (United States : Illinious State University)
[4] Koenig K2012 Explicitly Targeting Presevice Teacher Scientific Reasoning Abilities and Understanding of Nature of Science Throught an Introductory Science Course 21(2) 1-9
[5] Fischer F, Kollara I, Uferb S, Sodiana B, et al. 2014 Scientific Reasoning and Argumentation: Advancing an Interdisciplinary Research Agenda in EducationJournal of Frontline Learning Research 5 28-45
[6] Mullis2014 TIMSS Advanced 2015 Assessment Frameworks. USA : Boston College
[7] Mullis I V S & Martin, M O2013 TIMSS 2015 Assessment Frameworks (United States : TIMSS & PIRLS International Study Center)
[8] Wasis 2014 Proceeding of International Conference On Research, Implementation, and Education of Mathematics and Science 8 (1) 147-54
[9] Mardapi, D 2008 Teknik Penyusunan Instrumen Tes dan Nontes(Yogyakarta : Mitra Cendikia)
[10] Matlock, S & Hetzel 1997 Annual Meeting of the Southwest Educational Research Association, Austin
[11] Thiagarajan S, Semmel D S and Semmel M 1974 Instructional Development for Training Teachers of Exeptional ChildrenSource Book (Bloominton: Center for Innovation on teaching the Handicapped)
[12] Lawshe C H 1975 A Conf. Held at Bowling Green State University 28 563-75
[13] Ding L and Beichner R 2009 Physics Education Reseach 5 20103

Acknowledgments
The author would like to thank validators who have provided assessment for the improvement of reasoning instruments test and physics teacher in the city of Bandung who has assisted in the implementation of this study.