Psychometric and Structural Evaluation of the Physics Metacognition Inventory Instrument

Haeruddin*
Yogyakarta State University, INDONESIA

Zuhdan Kun Prasetyo
Yogyakarta State University, INDONESIA

Supahar
Yogyakarta State University, INDONESIA

Elisa Sesa
Tadulako University, INDONESIA

Gazali Lembah
Tadulako University, INDONESIA

Abstract: The purpose of this study is to evaluate the psychometric and structural instruments of the Physics Metacognition Inventory (PMI) developed by Taasoobshirazi, Bailey, and Farley. The PMI consists of 26 items in six factors. The English and Indonesian versions were tested on an English course (N = 37) in the Geophysics study program at Tadulako University. The trials were conducted separately within a two-week interval. The data collected from 364 students of the Physics Education Department, University of Tadulako were analyzed using the Exploratory Factor Analysis (EFA). Later, data were collected from 351 students of some Indonesian universities which have physics education study programs, and the data were analyzed using the Confirmatory Factor Analysis (CFA). The EFA result reveals six factors based on the rotation result with the maximum loading factor. The CFA result shows the RMSEA values of .018, χ^2 (284) = 316.32 (χ^2 / df = 1,11), GFI = .93, CFI = .99, AGFI = .92 and NFI = .93 which meet the cut-off statistic value, and therefore, the model is considered fit, with the Construct Reliability Estimation (CR) of .93, Composite Reliability of ω = .95, and maximum reliability of Ω = .96. The results obtained reveal that the PMI scale has good, valid and reliable psychometric properties. Therefore, PMI can be used to measure the level of metacognition of students when solving physics problems. Future studies using PMI are also discussed.

Keywords: *Psychometric evaluation, physics metacognition inventory, problem solving.*

Introduction

Metacognition is an important component in the teaching of science, and thus it must be introduced to students. In the teaching of science, there is a metacognitive process that makes learning meaningful and enables students to learn science independently (Kipnis & Hofstein, 2008). The metacognitive strategy in teaching affects students' learning achievement at the cognitive- affective, and psychomotor levels (Srinivasan & Pushpam, 2016), improves conceptual understanding of science (Colthorpe, Sharifirad, Ainscough, Anderson, & Zimbardi, 2018), develops students' higher-order thinking skills (Ghanizadeh, 2017), and improves students' attitude towards science (Jahangard, Soltani, & Alinejad, 2016). For this reason, schools must become the place for metacognition to develop because there is a lot of self-awareness teaching going on. At school, students have a lot of opportunities to monitor and manage their cognition, and to get metacognitive knowledge about themselves, tasks, and strategies (Flavell, 1979).

For the last few decades, studies on metacognition have emphasized problem-solving processes. Physics problem solving involves a complex cognitive process and metacognition is the biggest factor to success in physics problem solving (Balta, Mason, & Singh, 2016). The students who use metacognition in a problem-solving process have a higher probability to get a correct solution (Akben, 2018; Koch, 2001). The students with improved metacognition tend to have better reasoning which can function to set the interaction between intuitive reasoning and analytic reasoning

---

* **Corresponding author:**
Haeruddin, Graduate School, Yogyakarta State University, Jl. Kolombo Karangmalang Yogyakarta, PO. 55281, Indonesia.
✉ haeruddin.spd2015@student.uny.ac.id

(Kryjevskaia, Stetzer, & Grosz, 2014). The extensive contribution of metacognition becomes a challenge for researchers and is considered as one of the most important problems for students to be successful in problem-solving.

One of the most important aspects of metacognition is the process of contemplating and focusing on one's own thinking. Metacognition is very important for effective thinking and problem solution and it is one of the superiorities of expertise in a certain field of science and skills. Experts use metacognitive strategies to monitor understanding during problem solutions and for self-correction. Therefore, the assessment must try hard to determine whether or not someone has good metacognitive skill. The first effort to include metacognitive thinking in a problem-solving process is to discover students' metacognitive awareness. For this reason, it is important to determine the students' metacognitive level before their metacognition is developed (Ozturk, 2017). At this point, one important element is the adaptation of a standard instrument for the students.

Since Flavell (1979) introduced the concept of metacognition, many researchers have reported the challenge in metacognition evaluation. For example, metacognition cannot be observed directly in students because the process occurs internally (Sperling, Howard, Miller, & Murphy, 2002). To measure students' awareness and metacognitive regulation, researchers have used varieties of the instrument, such as questionnaires of self-report, think-aloud protocols, observation, performance evaluation, and interviews (Dinsmore, Alexander, & Loughlin, 2008; Winne & Perry, 2000). Practitioners with limited resources prefer to use a self-report questionnaire.

When definitions related to metacognition are examined, it is observed that they generally focus on two components. Metacognition is defined as knowledge of metacognition and regulation of metacognition (Brown, 1978), in which the former involves our awareness, our thinking process, particularly declarative knowledge of our memory (Flavell, 1979), and the latter involves our planning and control of these processes (Jacobs & Paris, 1987). Metacognition contains the information which is used in the problem-solving process, which is cognitive knowledge related to one's cognitive competence and cognitive regulation. Cognitive knowledge consists of declarative, procedural, and conditional knowledge. Cognitive regulation consists of planning, monitoring, evaluating, debugging, and managing information (Taasoobshirazi et al., 2015; Taasoobshirazi & Farley, 2013).

Scrutinizing the psychometry of an instrument is important because it is related to its future reliability. Instrument reliability and validity are the main characteristics of measurement. Reliability is the ability to reproduce results consistently in time and space. Validity refers to the instrument's properties to measure exactly what it proposes. Psychometric assessment is useful for establishing a valid and reliable instrument, to ensure the quality of study results (Souza, Alexandre, & Guirardello, 2017). Metacognitive Awareness Inventory (MAI) has been used extensively, but its scoring method from the respondents is not consistent and the empirical data are inadequate to support the theory (Harrison & Vallin, 2017). There are pros and cons in the metacognitive instrument that needs big scale evaluation (Schellings & van Hout-Wolters, 2011) and there needs to be research investigating the characteristics of a self-report instrument to develop a new procedure. One of the problems they showed is the correlation assumption explained by cultural background factors. In other words, out of the measured metacognitive construct, the latent interfering variable contributes to metacognition score variabilities.

In line with the research mentioned above, it appears that some gaps must be filled regarding the validity and reliability of the PMI structure. Therefore, given the problem of structural reliability, it was decided to use CFA and EFA on PMI data. Thus, the first objective of this study was to examine the psychometrics and structure of the PMI developed by Taasoobshirazi et al. (2015). One potential reason is the reliability and validity of the instrument is the main characteristic of a measurement that can ensure the quality of research results. Unlike the MAI which has been evaluated by many researchers, the new PMI needs to be further evaluated. Also, the development of PMI is based on the MAI pattern which consists of 6 (six) dimensions and focuses on the use of diagrams in solving physics problems (Taasoobshirazi & Farley, 2013). Most research examines metacognition in solving physical problems using oral interviews by developing several items (Balta et al., 2016; Kryjevskaia et al., 2014; Mansyur, Lestari, Werdhiana, & Rizal, 2018; Mundilarto, 2003; Pathuddin, Budayasa, & Lukito, 2019).

The lack of research on the role of metacognition in physics problem solving is a problem, considering the importance of problem-solving in improving physics learning achievement (Chi, 2006; Yuberti et al., 2019). Practitioner-oriented studies have investigated the correlation between awareness and metacognitive achievement (Sart, 2014), with some studies showing evidence that the students who managed their cognition tend to work better in problem-based learning (Hmelo-Silver, 2004; Uyar, Yilmaz Genc, & Yasar, 2018), and in improving academic achievement (Winston, Van Der Vleuten, & Scherpbier, 2010).

However, there has been a doubt that the positive correlation between metacognitive skill and academic achievement is not as strong as it has been expected (Jacobse & Harskamp, 2012; Veenman, 2011). A weak correlation is due to individuals with low achievement seldom using metacognitive thinking, and those with high achievement using metacognitive thinking automatically (Veenman, Kok, & Blöte, 2005). Weak correlations are caused by a lack of quality instruments used to measure metacognition. In the current study, the PMI developed by Taasoobshirazi et al. (2015) retesting with adaptation to Indonesia. In this context, the questions whose answers are sought in this study are (1) Do

the data support the factor structure proposed by Taasoobshirazi et al. (2015)?, (2) Is there a more appropriate factor structure supported by the data collected?

Methodology

Research Goal

This study aims to evaluate the psychometry and structure of PMI (Appendix). Previously, the instrument was adapted which refers to processes developed in other languages that have proven to be reliable and valid (Beaton, Bombardier, Guillemin, & Ferraz, 2000). The first phase of this research is to adapt the scale which was originally written in English and then translated into Indonesian. The translated PMI is then read by two experts to evaluate the results of the translation. After making the suggested changes, an evaluation by an Indonesian linguist is then carried out to determine the suitability of the sentence in Indonesian. Furthermore, the English and Indonesian versions were given to 37 students of the Geophysics study program at Tadulako University for 2 weeks. The second phase is to evaluate psychometric and factor structure by using EFA and CFA.

Sample and Data Collection

The study group is chosen by the criteria of the sampling technique, which is the sampling based on the consideration of the population or certain characteristics (Patton, 2014). The criteria in this study are groups of students registered at the physics department of a university. In the data collection, voluntary participation becomes the basis for determining that the students give a true answer. As many as 715 students participated in this study. The scale which has been translated into Indonesian is first given to 364 students, 63% (N = 228) females, and 37% (N = 136) males, registered at the Department of Physics Education and Pure Physics, University of Tadulako. The data collected were used to scrutinizing the structure of EFA. Furthermore, the data source was expanded to those collected from several universities in Indonesia, which run physics education study programs. The data collected were used to scrutinizing the structure of CFA. The number of students who filled in the scale is 351 – 74% (N = 259) females and 26% (N = 92) males.

It is mentioned in the article that there is no potential conflict of interests reported by the author as our basis for adapting the scale into Indonesian (Taasoobshirazi et al. 2015). The PMI instrument which is based on the information processing theory consists of 26 items. The main two-dimension factor consists of knowledge of cognition and regulation of cognition. Knowledge of cognition consists of declarative, procedural, and conditional knowledge, and regulation of cognition is planning, monitoring, evaluation, debugging, and information management. Table 1 shows the PMI factor items.

Table 1. PMI Factor Items.

Factor	Items
Knowledge of cognition: declarative, procedural, and conditional	11,13, 12, 6, 7, 5
Regulation of cognition: information management	23, 10, 4, 18
Monitoring	15, 16, 21, 2
evaluation	9, 8, 17, 25
debugging	3, 22, 26
planning	1, 20, 24, 14, 19

The scale was translated by the researcher, and then it was read by two English language experts to evaluate the result of the translation. After the suggested modification was made, the instrument was scrutinized by an Indonesian language expert to determine the accuracy and appropriateness of the Indonesian sentences. Afterward, the scale was given to an instrument expert and three experts in physics education. The language expert suggested that the sentence "I draw free-body diagrams" be changed into "I draw" because the meaning can be confusing. The students responded to each of the 26 items on the five-point Likert scale, from 1 (never) to 5 (always) from the perspective of when they were solving physics problems without any negative statements. Having been revised following experts' suggestions, the English and Indonesian version scales were given to 37 students (23 females and 14 males) of the Geophysics Department, University of Tadulako. Trials were conducted separately within the two-week interval. The high correlation (r = .841, p = .00) was found in the participants' responses to the Indonesian and English version scales. Afterward, the scale was given to 364 students to measure the EFA structure and 351 students to examine the CFA structure.

Analyzing of Data

The data analysis began with the exploratory factor analysis (EFA) to know the number of factors that are formed. The EFA is aimed at revealing the factors which construct PMI when solving physics problems. This analysis used SPSS version 25.00. In EFA, the Chi-Squared at Bartlett's test shows the sample is sufficient. When the significance value is

lower than .01, the sample is said to be sufficient. These data were supported by the value of Kaiser Meyer Olkin measure of sampling adequacy (KMO MSA) which is higher than .50 (Hair Jr, Black, Babin, & Anderson, 2014).

CFA is a statistical method used for examining model compatibility which is suggested to use LISREL 8.80. The statistic value of χ^2 for the model fit is not significant (*p-value* > .05) because the result indicates that there is no difference between the model and the data (Joreskog & Sorbom, 1993). Also, the index used as the reference of model fit evaluation is as follows. (1) The Goodness Fit Index (GFI) which is the model fit index in describing the developed model. In determining the model fit based on GFI, the expected GFI value is ≥ .90. The GFI value ranges from .00 (poor fit) to 1.00 (perfect fit) (Joreskog & Sorbom, 1993). (2) The Root Mean Square Error of Approximation (RMSEA) shows the residue found in the model. The expected value of RMSEA is ≤ .05. The RMSEA value of ≤ .05 indicates close fit, and when the value is in the range of .05 < RMSEA ≤ .08 the model is still acceptable as a fit model (good fit) (Bowen & Guo, 2013). (3) The Comparative Fit Index (CFI) is the index of a comparative model that is derived from an ideal model. The expected CFI value is above .90 (Hoyle, 2012). (4) The Adjusted Goodness of Fit (AGFI) is the criteria of a fit index developed from the GFI adjusted to the degree of freedom ratio for the proposed model with the degree of freedom for the null model. The recommended AGFI value to indicate a model fit is ≥ .90 (Schumacker & Lomax, 2016). (5) The Normed Fit Index (NFI) is the comparison between the proposed model and the null model. The expected NFI value is ≥ .90 (Schumacker & Lomax, 2016). The reliability estimation is conducted in three ways: reliability estimation with construct reliability (CR), composite reliability (ω), and maximum reliability (Ω) (Geldhof, Preacher, & Zyphur, 2014).

Findings / Results

Exploratory Factor Analysis (EFA)

Before the EFA analysis was the calculation of Cronbach's Alpha, which shows a value of .91. The EFA result of the correlation among the 26 items in the factor analysis determined by using Bartlett's test shows χ^2 = 3286.33, df = 325, p < .001, and Kaiser Meyer Olkin measure of sampling adequacy (KMO) of .911, which is higher than .50. The analysis based on anti-image correlation (AIC) shows that none of the scores is below .50. The analysis using the Kaiser-Guttman formula concerning the eigenvalue of higher than or equal to 1 shows that there are six factors formed. All of the six factors together contribute 58.22% of the total variants. Parallel analysis is used to compare the size of eigen values which is obtained from randomly generated data sets by using a Monte Carlo PCA simulation (Pallant, 2011). Acceptance criteria is if the calculated value is greater than the simulation value. otherwise will be rejected. By using the number of variables as many as 26, the number of samples is 364 and the number of replications is 100, the results obtained that meet the criteria are 4 factors. The eigenvalue, percentage of variants, and cumulative percentage of each factor are shown in Table 2. Furthermore, the researchers analyzed by considering the value of the loading factor of each item by using the varimax method.

Based on the result of the rotation of the maximum loading value, there are six factors formed as shown in Table 3. Based on the number of factors loaded, the factors were labeled. The labeling of the factors is based on the factor load after the rotation, taking into account the factor load above .40 (Retnawati, 2016).

Table 2. Eigenvalue, variant percentage, and cumulative percentage of each factor.

Factor	Initial Eigenvalues			Parallel Analysis	
	Eigenvalues	% of Variance	Cumulative %	Random Eigenvalue	Standard Dev.
1	8.01	30.82	30.82	1.52	.043
2	1.95	7.50	38.32	1.44	.033
3	1.47	5.65	43.97	1.38	.025
4	1.39	5.35	49.32	1.33	.027
5	1.17	4.49	53.81	1.24	.024
6	1.15	4.41	58.22	1.20	.021

Table 3. Results of rotation with maximum loading value.

Item	Component					
	1	2	3	4	5	6
it05	.666					
it06	.658					
it07	.694					
it11	.675					
it12	.567					
it13	.599					

Table 3. Continued

Item	Component					
	1	2	3	4	5	6
it01			.533			
it14			.606			
it19			.672			
it20			.692			
it24			.676			
it04				.757		
it10				.736		
it18				.787		
it23				.551		
it08		.709				
it09		.790				
it17		.689				
it25		.667				
it16					.530	
it02					.760	
it15					.690	
it21					.557	
it03						.808
it22						.747
it26						.495

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 6 iterations.

Confirmatory Factor Analysis (CFA)

The result of the data analysis using LISREL 8.80 shows that the construct that forms the PMI model in the CFA process has met the designated goodness of fit criteria. The probability value of the tested goodness of fit shows a score of .09 (< .05), RMSEA 0.018 (< .08) and the value of χ^2 (284) = 316.32 (χ^2/df = 1.11). The results of other tests of the model goodness of fits such as GFI (.93), CFI (.99), AGFI (.92) and NFI (.93) show the value of > .90, and thus the model is considered fit. The complete result of the calculation of the loading factor can be seen in Table 4.

Table 4. Comparison of item loadings on the PMI between Taasoobshirazi et al. (2015) and this study.

Factor	Item	Item loading	
		Indonesian	English
Factor 1. Knowledge of cognition: declarative, procedural, and conditional	11	.44	.74
	12	.61	.77
	13	.60	.73
	5	.49	.58
	6	.62	.63
	7	.54	.62
Factor 2. Regulation of cognition: information management (free-body diagrams)	10	.62	.91
	18	.47	.45
	23	.55	.90
	4	.64	.84
Factor 3. Regulation of cognition: monitoring	15	.84	.91
	16	.58	.91
	2	.56	.71
	21	.59	.81
Factor 4. Regulation of cognition: evaluation	17	.71	.61
	25	.62	.58
	8	.78	.91
	9	.86	.95
Factor 5. Regulation of cognition: debugging	22	.66	1,00
	26	.61	.13
	3	.77	.96
Factor 6. Regulation of cognition: planning	1	.77	.47
	14	.72	.35
	19	.54	.52
	20	.68	.31
	24	.61	.46

Based on the result of each indicator in each item, the items have met the requirement, i.e. their loading factors are above .40 (Bowen & Guo, 2013), and therefore they are acceptable. The result of the validity measure by taking into account the relevant loading factor with the t-test shows the *t-count* of > *t-critical* ($\alpha = .05$, $t > 1.96$) (Hoyle, 2012).
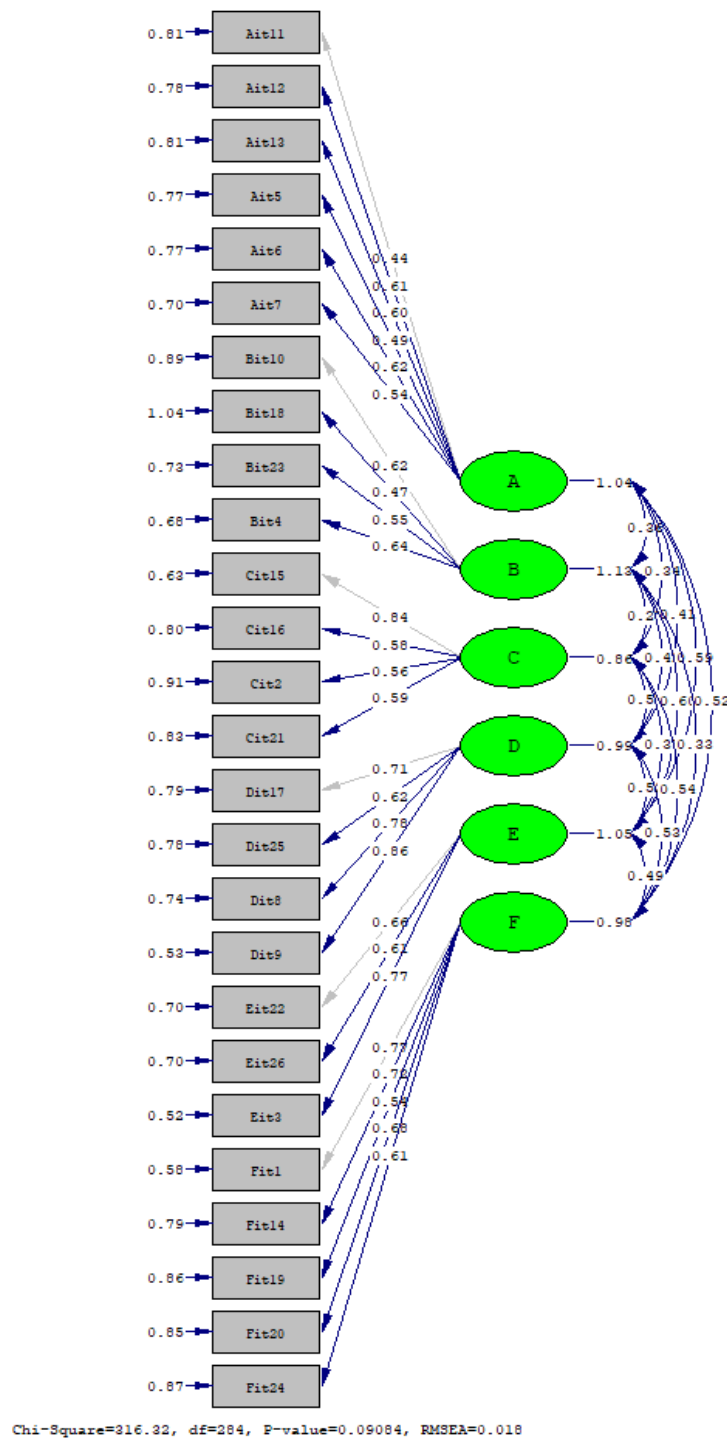

Figure 1. Confirmatory factor analysis model.

Reliability

Cronbach's alpha as a reliability measure has been widely used but based on the unidimensional assumption to measure multidimensions, it may disregard the actual reliability (Kamata, Turhan, & Darandari, 2013; Widhiarso & Ravand, 2014). For this reason, this research uses the coefficient which is in line with the characteristics of measurement. The estimation used is the reliability estimation with construct reliability (*CR*), composite reliability (ω), and maximum reliability (Ω).

Using the values shown in Figure 1 to measure the reliability estimation results in: (1) the estimation of construct reliability (*CR* = .93) using the loading factor of each indicator that forms the instrument (λ) and the index of errors of

each indicator (δ); (2) composite reliability (ω = .95) using the heterogeneous item-construct relationship and estimating the actual variant scores as the function of the loading factor of each item; (3) the maximum reliability (Ω = .96) which compares the actual score variants with unit-weighted scale variants which represent the optimal composite reliability-weighted scale (Geldhof et al., 2014).

Discussion and Conclusion

This research aims to scrutinize the PMI psychometry related to the construct validity in university students. The adapted scale shows a good correlation (r = .841) between the Indonesian version scale and the English version scale. The adaptation is made so that Indonesian practitioners can easily measure students' metacognition while they are solving physics problems. This can be done by physics teachers and physics education researchers in understanding and supporting students' success in solving problems, and, in turn, in achieving the physics learning outcome.

Although the PMI has not been used much in the studies on metacognition, this research is limited to the measurement of the factor structure. The result of our research supports the conclusion that 26 items function better to form six factors (Schraw & Dennison, 1994; Taasoobshirazi et al., 2015). We also found that this theoretical structure is highly in line with the EFA and CFA of the previous studies. Some psychometry that we analyzed shows that all of them meet the psychometry cutoff value. We also scrutinized the result of the PMI scale adaptation to Turkish (Unlu & Dokme, 2019) using 24 items (Taasoobshirazi & Farley, 2013).

Table 5 shows that all psychometric values meet the measurement criteria. One of the striking differences between this study and the previous ones is in the calculation of the construct reliability. This study looks at three aspects of reliability: construct reliability (*CR*), composite reliability (ω), and maximum reliability (Ω). Kamata et al. (2003) state that this method assumes the actual reliability is much better than coefficient-alpha in all conditions. For this reason, researchers have to use the coefficient which is suitable for the measurement characteristic to estimate the reliability of multidimensional procedures (Widhiarso & Ravand, 2014).

Table 5. Comparison of PMI Psychometric Scales.

Psychometric Aspects	Indonesian (26 items)	Turkish (24 items)	English (26 items)
Sample size	715	458	285
Cronbach's Alpha	.91	.93	.87
$\chi 2/df$	1.11	3.55	2.39
GFI	.93	.86	-
AGFI	.92	.82	-
CFI	.99	.89	.93
RMSEA	.018	.07	.07

The result of this study shows that the structure of the PMI proposed by Taasoobshirazi et al. (2015) and in Turkish scale (Unlu & Dokme, 2019) is in line with the model adapted to Indonesian. The scale which has good psychometric characteristics can be used to determine students' metacognition as a valid and reliable instrument. The finding of this study is limited to the data collected from the students of the physics education department. Further studies can be conducted with the sample of the students of chemistry, mathematics, and biology education departments, to generalize the finding.

Suggestions

Future studies are suggested to explore the differences between genders (Pathuddin et al., 2019) and their impact on the quality of the use of problem-solving strategies and the achievement in problem-solving. The six factors identified by EFA will interpret the dimension of students' understanding of metacognition when they are solving physics problems. Some items are needed to be added, which are by the dimension, such as when the students understand a problem starting with reading activities (Haeruddin, Prasetyo, & Supahar, 2019; Meijer et al., 2013).

This can also be used to determine the extent to which various methods and techniques affect physical metacognition. Also, it can be used to assess metacognition by looking at differences in thinking and learning styles. In future studies, the analysis of the scale validity and reliability should be done by adapting other languages, and then re-examine the collected data.

Examining the correlation between metacognitive behavior and academic ability (Jacobse & Harskamp, 2012; Veenman, 2011) and comparing with think-aloud protocols (Mansyur et al., 2018; Schellings, Van Hout-Wolters, Veenman, & Meijer, 2013). Like other metacognition instruments that have been used before, PMI has several gaps in the scoring method (Harrison & Vallin, 2017) consequently it requires a large-scale assessment (Schellings & van Hout-Wolters, 2011).

This study has limited data which is not using a random sample, and participants are grouped in classes at an institution. Although the selected sample has learned physics but they have not yet learned about how to apply metacognition in solving physics problems specifically. Because PMI is a self-report hence there is a possibility of bias when participants fill in the instrument by choosing the best one. PMI may not be an ideal measure, but the knowledge we gain from examining its psychometrics can provide information about a proper metacognition measure.

References

Akben, N. (2018). Effects of the Problem-Posing approach on students' problem solving skills and metacognitive awareness in science education. *Research in Science Education*, *48*(1), 1–23. https://doi.org/10.1007/s11165-018-9726-7

Balta, N., Mason, A. J., & Singh, C. (2016). Surveying Turkish high school and university students' attitudes and approaches to physics problem solving. *Physical Review Physics Education Research*, *12*(1), 010129. https://doi.org/10.1103/PhysRevPhysEducRes.12.010129

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, *25*(24), 3186–3191. https://doi.org/10.1097/00007632-200012150-00014

Bowen, N. K., & Guo, S. (2013). *Structural equation modeling* (Vol. 15). New York, NY: Oxford University Press.

Brown, A. L. (1978). Knowing when, where, and how to remember: A problem of metacognition. *Advances in Instructional Psychology*, *1*(1), 225–253.

Chi, M. T. H. (2006). Two approaches to the study of experts' characteristics. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge Handbook of Expertise and Expert Performance* (pp. 21–30). New York, NY: Cambridge University Press.

Colthorpe, K., Sharifirad, T., Ainscough, L., Anderson, S., & Zimbardi, K. (2018). Prompting undergraduate students' metacognition of learning: implementing 'meta-learning' assessment tasks in the biomedical sciences. *Assessment & Evaluation in Higher Education*, *43*(2), 272–285. https://doi.org/10.1080/02602938.2017.1334872

Dinsmore, D. L., Alexander, P. A., & Loughlin, S. M. (2008). Focusing the conceptual lens on metacognition, self-regulation, and self-regulated learning. *Educational Psychology Review*, *20*(4), 391–409. https://doi.org/10.1007/s10648-008-9083-6

Flavell, J. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, *34*(10), 906–911. https://doi.org/10.1037/0003-066x.34.10.906

Geldhof, G. J., Preacher, K. J., & Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods*, *19*(1), 72–91. https://doi.org/10.1037/a0032138

Ghanizadeh, A. (2017). The interplay between reflective thinking, critical thinking, self-monitoring, and academic achievement in higher education. *Higher Education*, *74*(1), 101–114. https://doi.org/10.1007/s10734-016-0031-y

Haeruddin, Prasetyo, Z. K., & Supahar. (2019). The Development of a Metacognition Instrument for College Students to Solve Physics Problems. *International Journal of Instruction*, *13*(1), 1–16. Retrieved from http://www.e-iji.net/volumes/359-onlinefirst

Hair Jr, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis* (7th ed.). Harlow, United Kingdom: Pearson Education Limited.

Harrison, G. M., & Vallin, L. M. (2017). Evaluating the metacognitive awareness inventory using empirical factor-structure evidence. *Metacognition and Learning*, *13*(1), 15–38. https://doi.org/10.1007/s11409-017-9176-z

Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review*, *16*(3), 235–266. https://doi.org/10.1023/B:EDPR.0000034022.16470.f3

Hoyle, R. H. (2012). *Handbook of structural equation modeling*. New York, NY: The Guilford Press.

Jacobs, J. E., & Paris, S. G. (1987). Children's metacognition about reading: issues in definition, measurement, and instruction. *Educational Psychologist*, *22*(3), 255–278. https://doi.org/10.1080/00461520.1987.9653052

Jacobse, A. E., & Harskamp, E. G. (2012). Towards efficient measurement of metacognition in mathematical problem solving. *Metacognition and Learning*, *7*(2), 133–149. https://doi.org/10.1007/s11409-012-9088-x

Jahangard, Z., Soltani, A., & Alinejad, M. (2016). Exploring the relationship between metacognition and attitudes towards science of senior secondary students through a structural equation modeling analysis. *Journal of Baltic Science Education*, *15*(3), 340–349.

Joreskog, K. G., & Sorbom, D. (1993). *LISREL 8: Structural Equation Modeling With the SIMPLIS Command Language*. Uppsala, Sweden: Scientific Software International, Inc.

Kamata, A., Turhan, A., & Darandari, E. (2013, April). *Estimating reliability for multidimensional composite scale scores*. Paper presented at the annual meeting of American Educational Research Association, Florida State University, Chicago, USA.

Kipnis, M., & Hofstein, A. (2008). The inquiry laboratory as a source for development of metacognitive skills. *International Journal of Science and Mathematics Education*, *6*(3), 601–627. https://doi.org/10.1007/s10763-007-9066-y

Koch, A. (2001). Training in metacognition and comprehension of physics texts. *Science Education*, *85*(6), 758–768. https://doi.org/10.1002/sce.1037

Kryjevskaia, M., Stetzer, M. R., & Grosz, N. (2014). Answer first: Applying the heuristic-analytic theory of reasoning to examine student intuitive thinking in the context of physics. *Physical Review Special Topics - Physics Education Research*, *10*(2), 020109. https://doi.org/10.1103/PhysRevSTPER.10.020109

Mansyur, J., Lestari, W., Werdhiana, I. K., & Rizal, M. (2018). Students' Metacognition Skills in Physics Problem Solving Based on Epistemological Beliefs. In D. K. Walanda, M. Basir, N. Mery, K. Manurung, Lukman, J. Mansyur, … & Syamsuriwal (Eds.), *First Indonesian Communication Forum of Teacher Training and Education Faculty Leaders International Conference on Education 2017* (pp. 28–31). Palu, Indonesia: Tadulako University.

Meijer, J., Sleegers, P., Elshout-Mohr, M., van Daalen-Kapteijns, M., Meeus, W., & Tempelaar, D. (2013). The development of a questionnaire on metacognition for students in higher education. *Educational Research*, *55*(1), 31–52. https://doi.org/10.1080/00131881.2013.767024

Mundilarto. (2003). Kemampuan mahasiswa menggunakan pendekatan analitis kuantitatif dalam pemecahan soal fisika [Students' ability to use quantitative analytical approaches in solving physics problems]. *Journal of Mathematics and Science Education/Jurnal Pendidikan Matematika Dan Sains*, *3*(VIII), 137–142.

Ozturk, N. (2017). Assessing metacognition: Theory and practices. *International Journal of Assessment Tools in Education*, *4*(2), 134–134. https://doi.org/10.21449/ijate.298299

Pallant, J. (2011). *SPSS survival manual a step by step guide to data analysis using the SPSS program* (4th ed.). Sydney, Australia: Allen & Unwin.

Pathuddin, Budayasa, I. K., & Lukito, A. (2019). Metacognitive activity of male students: Difference field independent-dependent cognitive style. *Journal of Physics: Conference Series*, *1218*(1), 1–4. https://doi.org/10.1088/1742-6596/1218/1/012025

Patton, M. Q. (2014). *Qualitative research & evaluation method* (4th ed.). Los Angeles, CA: Sage Publications.

Retnawati, H. (2016). *Analisis kuantitatif instrumen penelitian: Panduan peneliti, mahasiswa, dan psikometrian*. [Quantitative analysis of research instruments: A guide for researchers, students and psychometrics]. Yogyakarta, Indonesia: Parama Publishing.

Sart, G. (2014). The effects of the development of metacognition on project-based learning. *Procedia - Social and Behavioral Sciences*, *152*, 131–136. https://doi.org/10.1016/j.sbspro.2014.09.169

Schellings, G., & van Hout-Wolters, B. (2011). Measuring strategy use with self-report instruments: Theoretical and empirical considerations. *Metacognition and Learning*, *6*(2), 83–90. https://doi.org/10.1007/s11409-011-9081-9

Schellings, G., Van Hout-Wolters, B. H. A. M., Veenman, M. V. J., & Meijer, J. (2013). Assessing metacognitive activities: The in-depth comparison of a task-specific questionnaire with think-aloud protocols. *European Journal of Psychology of Education*, *28*(3), 963–990. https://doi.org/10.1007/s10212-012-0149-y

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, Vol. 19, pp. 460–475. https://doi.org/10.1006/ceps.1994.1033

Schumacker, R. E., & Lomax, R. G. (2016). *A begginer's guide to structural equation modeling* (4th ed.). New York, NY and London, UK: Routledge Taylor & Francis Group.

Souza, A. C. de, Alexandre, N. M. C., & Guirardello, E. de B. (2017). Propriedades psicomericas na avaliacao de instrumentos: avaliacao da confiabilidade e da validade [Psychometric properties in instrument evaluation: reliability and validity evaluation]. *Epidemiology and Health Services: Brazilian Health System Journal/Epidemiologia e Servicos de Saude: Revista Do Sistema Unico de Saude Do Brasil*, *26*(3), 649–659. https://doi.org/10.5123/S1679-49742017000300022

Sperling, R. A., Howard, B. C., Miller, L. A., & Murphy, C. (2002). Measures of children's knowledge and regulation of cognition. *Contemporary Educational Psychology*, *27*(1), 51–79. https://doi.org/10.1006/ceps.2001.1091

Srinivasan, D. P., & Pushpam, M. A. M. (2016). Exploring the influence of metacognition and metaemotion strategies on the outcome of students of IX Std. *American Journal of Educational Research*, *4*(9), 663–668. https://doi.org/10.12691/education-4-9-3

Taasoobshirazi, G., Bailey, M., & Farley, J. (2015). Physics metacognition inventory part II: Confirmatory factor analysis and rasch analysis. *International Journal of Science Education*, *37*(17), 2769–2786. https://doi.org/10.1080/09500693.2015.1104425

Taasoobshirazi, G., & Farley, J. (2013). Construct validation of the physics metacognition inventory. *International Journal of Science Education*, *35*(3), 447–459. https://doi.org/10.1080/09500693.2012.750433

Unlu, Z. K., & Dokme, I. (2019). Adaptation of physics metacognition inventory to Turkish. *International Journal of Assessment Tools in Education*, *6*(1), 125–137. https://doi.org/10.21449/ijate.483104

Uyar, R. O., Yilmaz Genc, M. M., & Yasar, M. (2018). Prospective preschool teachers' academic achievements depending on their goal orientations, critical thinking dispositions and self-regulation skills. *European Journal of Educational Research*, *7*(3), 601–613. https://doi.org/10.12973/eu-jer.7.3.601

Veenman, M. V. J. (2011). Learning to self-monitor and self-regulate. In R. E. Mayer & P. A. Alexander (Eds.), *Handbook of Research on Learning and Instruction* (pp. 197–218). New York, NY: Routledge Taylor & Francis Group.

Veenman, M. V. J., Kok, R., & Blöte, A. W. (2005). The relation between intellectual and metacognitive skills in early adolescence. *Instructional Science*, *33*(3), 193–211. https://doi.org/10.1007/s11251-004-2274-8

Widhiarso, W., & Ravand, H. (2014). Estimating reliability coefficient for multidimensional measures: A pedagogical illustration. *Review of Psychology*, *21*(2), 111–121.

Winne, P. H., & Perry, N. E. (2000). Measuring Self-Regulated Learning. In M. Boekaerts, M. Zeidner, & P. R. Pintrich (Eds.), *Handbook of Self-Regulation* (pp. 531–566). New York, NY: Academic Press.

Winston, K. A., Van Der Vleuten, C. P. M., & Scherpbier, A. J. J. A. (2010). An investigation into the design and effectiveness of a mandatory cognitive skills programme for at-risk medical students. *Medical Teacher*, *32*(3), 236–243. https://doi.org/10.3109/01421590903197035

Yuberti, Latifah, S., Anugrah, A., Saregar, A., Misbah, & Jermsittiparsert, K. (2019). Approaching problem-solving skills of momentum and impulse phenomena using context and problem-based learning. *European Journal of Educational Research*, *8*(4), 1217–1227. https://doi.org/10.12973/eu-jer.8.4.1217

Appendix

Indonesian Version

Agar dapat memahami bagaimana Anda memecahkan masalah dalam fisika, silahkan memberikan pandangan setiap pernyataan berikut dengan cara memberi tanda silang (X) salah satu pilihan jawaban yang disediakan:

(1) Tidak pernah (2) Jarang (3) Kadang-kadang (4) Pernah (5) Selalu

No	Pernyataan	Pilihan				
1	Saya memikirkan masalah apa yang ditanyakan sebelum memulai menyelesaikan masalah fisika	1	2	3	4	5
2	Saat menyelesaikan masalah fisika, secara berkala saya bertanya pada diri sendiri apakah saya memenuhi tujuan saya	1	2	3	4	5
3	Saya meminta bantuan ketika saya tidak mengerti masalah fisika	1	2	3	4	5
4	Saya menggambar sesuatu untuk membantu saya memecahkan masalah fisika	1	2	3	4	5
5	Saya mempertimbangkan dengan baik ketika menyelesaikan masalah fisika	1	2	3	4	5
6	Ketika menyelesaikan masalah fisika, saya mengetahui bagaimana bekerja dengan baik	1	2	3	4	5
7	Ketika menyelesaikan masalah fisika, saya memiliki tujuan tertentu pada setiap strategi yang saya gunakan	1	2	3	4	5
8	Saya memeriksa kembali pekerjaan saya setelah menyelesaikan masalah fisika	1	2	3	4	5
9	Setelah menyelesaikan masalah fisika, saya periksa lagi jawabannya	1	2	3	4	5
10	Saya menggunakan gambar untuk membantu menyelesaikan masalah fisika	1	2	3	4	5
11	Saya mengetahui bagaimana menerapkan strategi untuk berhasil menyelesaikan masalah fisika	1	2	3	4	5
12	Ketika menyelesaikan masalah fisika, saya mengetahui mengapa menggunakan strategi tertentu	1	2	3	4	5
13	Ketika menyelesaikan masalah fisika, saya mengetahui kapan harus menggunakan strategi tertentu	1	2	3	4	5
14	Sebelum menyelesaikan masalah fisika, saya memikirkan nilai jawaban yang wajar	1	2	3	4	5
15	Ketika menyelesaikan masalah fisika, saya bertanya pada diri sendiri tentang seberapa baik melakukannya	1	2	3	4	5
16	Ketika menyelesaikan masalah fisika, secara berkala mengevaluasi seberapa baik saya melakukannya	1	2	3	4	5
17	Setelah menyelesaiakan masalah fisika, saya melihat kembali untuk memeriksa apakah telah melakukan prosedur yang benar	1	2	3	4	5
18	Saya mengetahui mengapa menggambar penting untuk menyelesaikan masalah fisika	1	2	3	4	5
19	Sebelum saya memulai menyelesaikan masalah fisika, saya telah mengetahui sebuah rencana untuk menyelesaikannya	1	2	3	4	5
20	Sebelum menyelesaikan masalah fisika, saya mengidentifikasi semua bagian penting	1	2	3	4	5
21	Ketika menyelesaikan masalah fisika, saya bertanya pada diri sendiri, apakah telah mencapai tujuan saya	1	2	3	4	5
22	Saya mencari bantuan ketika saya tidak mengerti masalah fisika yang sedang diselesaikan	1	2	3	4	5
23	Saya menggambar sesuatu berdasarkan masalah fisika yang saya selesaikan	1	2	3	4	5
24	Sebelum menyelesaikan masalah fisika, saya menghilangkan informasi yang tidak diperlukan	1	2	3	4	5
25	Setelah menyelesaikan masalah fisika, saya memeriksa kembali untuk melihat apakah jawaban saya masuk akal	1	2	3	4	5
26	Saya mengubah strategi, ketika saya gagal menyelesaikan masalah fisika	1	2	3	4	5