J. Pijar MIPA, Vol. 16 No.4, September 2021: 442-448
DOI: 10.29303/jpm.v16i4.2603

ISSN 1907-1744 (Print)
ISSN 2460-1500 (Online)

VALIDATING METACOGNITIVE AWARENESS INVENTORY (MAI) IN CHEMISTRY LEARNING: A RASCH MODELS ANALYSIS

Jono Irawan[1], Samsul Hadi[1], Zulandri[1], Abdul Syukur[1], Jamaluddin[1], and Saprizal Hadisaputra[2]*
[1]Master of Science Education Postgraduate University of Mataram, Mataram, Indonesia
[2]Chemistry Education Programs, Faculty of Teacher Training and Science Education, University of Mataram, Indonesia
*Email: rizal@unram.ac.id

Acceped: July 31 2021. Approved: August 31 2021. Published: Sept 04 2021

Abstract: Metacognition is awareness and understanding of students' thought processes on the cognitive aspects of independent learning outcomes. This study aims to validate metacognitive instruments, including eight declarative knowledge indicators, procedural knowledge, conditional knowledge, planning, information management, monitoring, debugging, and evaluation. The validation method uses Rasch modeling (item response theory), namely Map analysis Wright (person-item map), Measure DIFF, Validity (fit statistics for the draft questionnaire), reliability Cronbach alpha (summary statistics). Results map analysis Wright shows the distribution of items and the power of discrimination is categorized as good with the logit value around 1.0. The curve Measure DIFF and the probability value indicate that the item does not have an element of distinguishing gender characteristics if the probability value is above 0.05. The validation of the assessment instrument showed an out value in-fit-of 0.5 MNSQ 1.5, a value in-fit-out of 0.5 ZSTD 1.5, and a value of 0.4 PTMEA 0.85. Furthermore, the reliability of the instrument category is highly reliable, with a value of Cronbach alpha 0.94. Based on the results of the IRT analysis, it was concluded that the MAI instrument in chemistry learning was appropriate to be used to measure the metacognitive abilities of high school students.

Keywords: Validation, Item Response Theory (IRT), Metacognitive Awareness Inventory (MAI)

INTRODUCTION

Science learning is a process of understanding natural phenomena or scientific events based on inquiry [1]. Science learning is expected to lead students to meet 21st-century competencies, according to the National Education Association (2010) is 4C skills, which include critical thinking, communication, creativity, and collaboration [2-4]. Furthermore, the skills that need to be considered to fulfill the 4C skills are students' metacognitive. Student metacognition is the level of student awareness of the academic abilities that exist within themselves [5].

Metacognition is an essential aspect of science learning, and of course, metacognitive abilities are achieved from meaningful experiences through scientific activities. Students' metacognitive skills still need to be seriously trained in learning, especially related to procedural and conditional knowledge, and skills in regulating the learning process [6-8]. Metacognitive abilities in learning should be focused on constructing knowledge through a rational relationship between the authentic experience and the material being studied [9-10]. In this case, thinking about what is thought relates to students' awareness of their ability to develop various possible ways to solve problems [11].

The process of realizing and organizing students' thinking, known as metacognition, includes thinking about how students approach the problem, choosing the strategies used to find solutions, and asking themselves about the problem [12]. The way students solve problems through the study of science in learning will foster self-awareness of their abilities. The development of metacognitive abilities in solving problems makes students more aware of their abilities [13].

Measuring the level of metacognitive ability requires a valid and reliable instrument on the area and conditions of students' understanding. The instrument was modified based on the condition and level of students' understanding of the questions in the test instrument in terms of language and ease of understanding. The test instrument used also avoids the nature of distinguishing the gender of students so that there is no gender element in influencing the results of the tests given. The assessment instrument was validated with the analysis item response test (IRT). IRT is a general framework of a special mathematical function that describes the interaction between people(person)with the item (test item) [14]. IRT method for obtaining objective, fundamental, linear measures of stochastic observations of categorical responses. IRT assumes this modeling and expresses the relationship between two variables through mathematical equations called logistic equations [15].

The analysis of the IRT method resulted in a valid, reliable assessment instrument, a good level of discrimination, and there was no influence of gender on the tests carried out. All educators and researchers can use instruments resulting from the analysis to measure students' metacognitive abilities in learning. Knowledge of the metacognitive level certainly helps educators design learning strategies

ISSN 1907-1744 (Print)
ISSN 2460-1500 (Online)

that are appropriate to the metacognitive level of students.

RESEARCH METHOD

This research is a quantitative descriptive study using a survey technique based on the metacognitive awareness inventory given to respondents. The number of samples in this study 130 respondents were high school students in class X and class XI in West Lombok, Central Lombok, and Mataram City.

The validated instrument was a metacognitive awareness instrument that had been developed, consisting of 50 statement items [16]. The indicators of the instrument consist of declarative knowledge, procedural knowledge, conditional knowledge, planning, information management, monitoring, debugging, and evaluation. Each indicator consists of many statements, which can be seen in Table 1. The data was collected by compiling MAI in a google form and then distributed through the WA group to the respondents by the chemistry teacher in high school.

Analysis of data obtained from student responses to metacognitive awareness instruments was analyzed using the Rasch model method (Item Response Theory). IRT analysis obtained data from the analysis to determine valid instruments, namely the study of Map Wright's (Person-item Map), Measure DIFF, Validity (Fit Statistics for The Draft Questionnaire), Reliability Cronbach's Alpha (Summary statistic) [17-19].

Table 1. Indicators and Statements on Metacognitive Awareness Instruments

Indicator	Item Number Statement	Number of
Declarative Knowledge	5, 10, 12, 16, 17, 20, 32, 46	8
Procedural Knowledge	3, 14, 27, 33	4
Conditional Knowledge	15, 18, 26, 29, 35	5
Planning	4, 6, 8, 22, 23, 42, 45	7
Information Management	9, 13, 30, 31, 37, 39, 41, 43, 47,	10
Monitoring	1, 2, 11, 21, 28, 34, 48	7
Debugging	25, 40, 44, 50	5
Evaluation	7, 19, 24, 36, 38, 49	6
Total		50

Table 2. Criteria for Fit Statistics for The Draft Questionnaire

Fit Statistics	Criteria for acceptance
Infit-Outfit Mean Square (MNSQ)	0.5 MNSQ 1.5
Infit-Outfit Z-Standard (ZSTD)	-2.0 ZSTD 2.0
Point Measure Correlation (PTMEA)	0.4 PTMEA 0.85

Table 3. Criteriavalue Cronbach alpha and person-item reliability

Value Cronbach alpha	Criteria	Value of person-item rebiability	Criteria
< 0.5	Poor	< 0.67	Weak
0.5 – 0.6	Poor	0.67 – 0.80	Enough
0.6 – 0.7	Enough	0.80 – 0.90	Good
0.7 – 0.8	Good	0.90 – 0.94	Very good
> 0.8	Very good	> 0.95	Special

RESULTS AND DISCUSSION

Test instruments play an essential role in determining the success of a lesson. The test instrument must have accuracy in measuring the variables due to treatment. An analysis of the test instrument was carried out first to avoid errors in measuring the variables. The analyzes carried out included map Wright's (person-item map), and the test instrument gender equality test (measure DIFF), validity test (fit statistic), and reliability test (Cronbach alpha).

Analysis of Wright's Map (Person-Item Map)

Wright's map is a map that shows the distribution of instrument items for all assessment indicators, starting from the items with the lowest ability to the highest ability. This map defines item parameters (difficulty, discrimination) regarding the degree of latent trait [20]. The MAI instrument items have an even distribution, i.e., they are below the M average, just above the M average, and above the M average (figure 1). Discrimination is estimated that the logit value is around 1.0 for the item difficulty level. A value is more significant than 1.0 means that

J. Pijar MIPA, Vol. 16 No.4, September 2021: 442-448
DOI: 10.29303/jpm.v16i4.2603
ISSN 1907-1744 (Print)
ISSN 2460-1500 (Online)

the item distinguishes between high and low abilities more than expected. A value of less than 1 means the item distinguishes between high and low skills less than expected [21]. Item discrimination shows how far a problem can distinguish high and low-ability individuals [17]. The items with the highest difficulty level are items I27 and I17. This item is only answered by students who have value logit more than one. Item I43 has the lowest level of difficulty. The logit value of items I34, I35, I40, I41, I43 is below the logit value of the students. The item has a low level of discrimination. A good assessment item is an item that, as a whole, has an even variation in the level of difficulty from low to high.

Measure DIFF

Rasch modeling provides a tool that can detect bias (DIF) based on the response given to a particular item based on respondents' demographic data. Practically any item is referred to as having a DIF (biased) when its DIF-probability value is less than 5% (0.05) [18]. DIF provides information about the level of difficulty of the items for each item based on the demographic profile of the respondents. This will be a very useful analysis to map the overall ability based on the characteristics of students. Item code questions I17, I22, I23, I27, I34, I41, I43, I44, and I49 have different responses from male and female students with probability values below 0.05 (Table 4). Question instruments that distinguish the ability to answer men and women are categorized as bad questions, so they must be removed from the assessment instrument.

The measure DIFF curve is a curve that shows students' responses to the test instrument based on gender. The curve shows the range of gender differences in response to the item questions indicated by the normal green line. The farther the range of responses of men (black) and women (red) means, the greater the item's bias about distinguishing gender characteristics. The item items received are right on and or around the normal line [17].

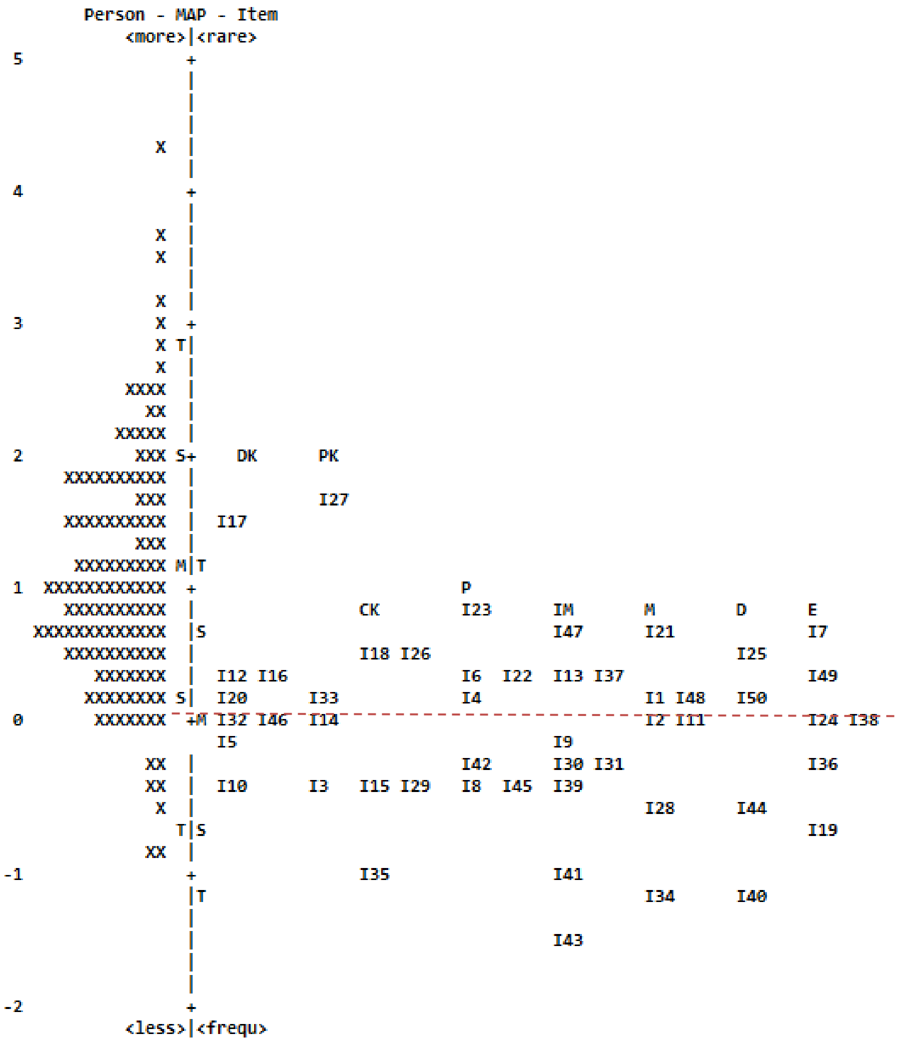

Figure 1. Distribution and Difficulty and Discrimination Power Problem Item

ISSN 1907-1744 (Print)
ISSN 2460-1500 (Online)

Table 4. Probability values measure DIFF

Item	Value Probability	Item	Value Probability	Item	Value Probability	Item	Value Probability	Item	Value Probability
I1	0.4850	I11	0.2301	I21	0, 0974	I31	1000	I41	0, 0009
I2	0.3162	I12	0.7723	I22	0, 0112	I32	0, 8294	I42	1000
I3	0.3949	I13	0.6039	I23	0, 0187	I33	0, 9410	I43	0, 0022
I4	0.2423	I14	0.2037	I24	0, 6642	I34	0, 0038	I44	0, 0275
I5	0.8637	I15	0.1486	I25	0, 8794	I35	0, 0857	I45	0, 2102
I6	0.3250	I16	1000	I26	0, 8573	I36	0, 3831	I46	0, 1157
I7	0.4716	I17	0.0458	I27	0, 0044	I37	0, 5456	I47	0, 8347
I8	1,000	I18	0.2035	I28	0, 0687	I38	0, 5215	I48	0, 1024
I9	0.0928	I19	1,000	I29	0, 2775	I39	0, 3351	I49	0, 0484
I10	0.8993	I20	0, 8373	I30	0, 6396	I40	0, 9097	I50	0, 5448

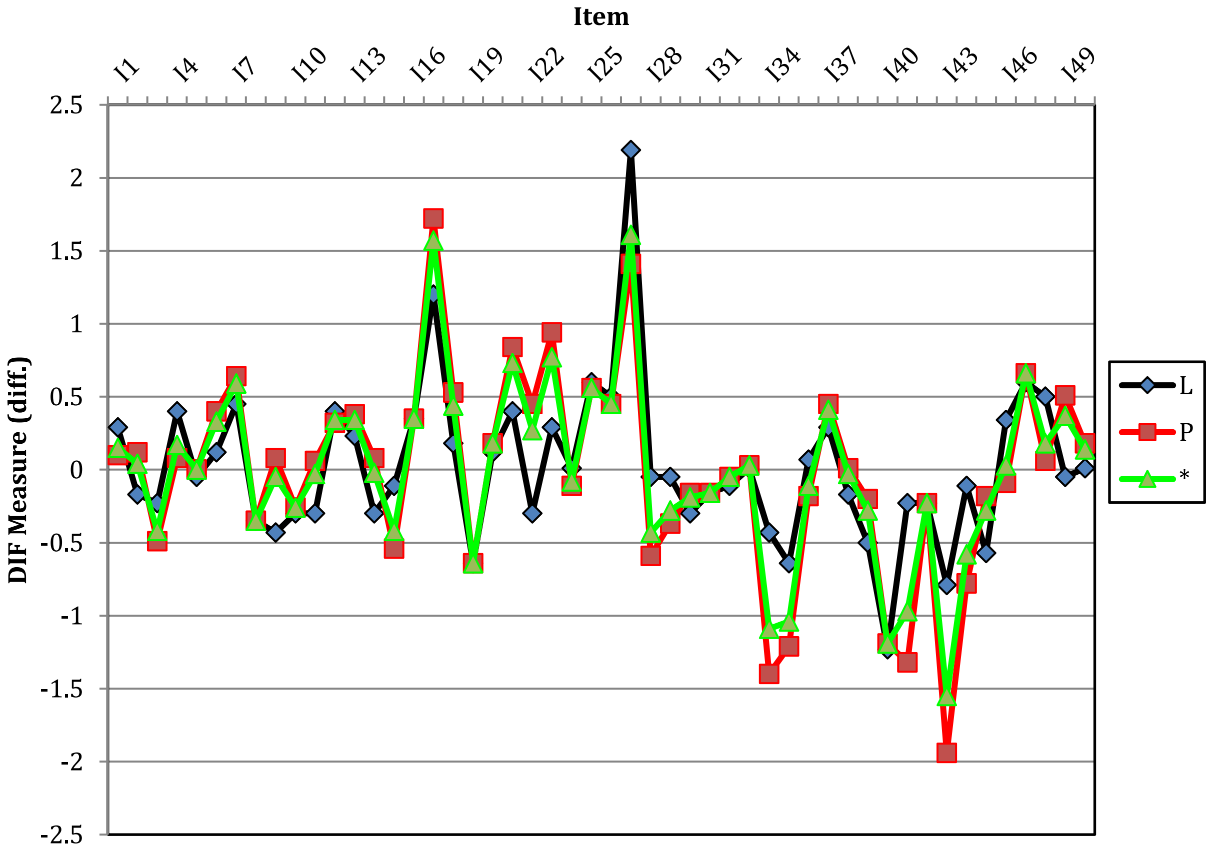

Figure 2. Graph of measure DIFF

Validity (Fit Statistics for The Draft Questionnaire)

Fit statistic is a criterion used to check the suitability of items that do not fit (misfits) [17]. Analysis of 50 items on the MAI instrument obtained 34 items that meet the valid criteria. The valid instrument in table 5 does not have biased items (misfit). Biased items can affect the results of metacognitive measurements, so they need to be removed from the item group. Valid assessment instruments can measure the measured variables appropriately. Elimination of misfit instruments does not reduce the number of indicators for assessing students' metacognitive abilities. The level of suitability of the item is determined by three criteria, as shown in table 1. If the item does not get more than one of the criteria above, the item is said to be unfit or unusable or needs to be replaced [20,22].

J. Pijar MIPA, Vol. 16 No.4, September 2021: 442-448
DOI: 10.29303/jpm.v16i4.2603
ISSN 1907-1744 (Print)
ISSN 2460-1500 (Online)

Table 5. Statistic Fit Items Metacognitive Instruments

Indicator	**Statement Items Adaptation**	**Infit MNSQ**	**Outfit MNSQ**	**Infit ZSTD**	**Outfit ZSTD**	**PTMEA**
Declarative Knowledge	I know the concepts most important to learn	0.8189	0.8461	-1.5792	-1.2392	0.5183
	I can remember course material well	0.9179	0.8963	-0.6291	-0.7191	0.5462
	I am good at managing information	0.8308	0.8442	-1.4892	-1.2792	0.5402
	I understand the strengths and weaknesses of my understanding	0.8233	0.807	-1.5292	-1.5692	0.5195
	I was able to arrange how well I learned	1.2439	1.2177	1.9912	1.7212	0.4717
	I was able to judge how far I understood the subject matter	1.054	0.9763	0.4711	-0.129	0.5318
Procedural Knowledge	I have a specific purpose for each learning strategy that I use	1.1716	1.1673	1.3912	1.2712	0.5068
	I use useful learning strategies	1.1136	1.0769	0.9311	0.6011	0.461
	0,461'mtrying to use my strategy I have ever used effectively in the past	0.8804	0.8657	-1.0391	-1.1191	0.5647
Conditional Knowledge	I know when each strategy I use is effective	0.862	0.8197	-1.2091	-1.5492	0.5692
	I can motivate myself to learn when I need it	0.8069	0.767	-1.6792	-1.9292	0.6157
	I use different strategies depending on the situation	1.0137	1.0166	0.151	0.161	0.4977
	I can learn well when I know something about the topic	0.8673	0.8288	-1.1591	-1.4692	0.5971
Planning	Before starting a task, I set a specific goal	0.8454	0.8297	-1.3492	-1.4192	0.6059
	I ask myself about the subject matter before I start studying	1.1196	1.1594	0.9311	1.0712	0.3767
Information Management	I consciously focus on important information	0.7992	0.8104	-1.8592	-1.6892	0.5726
	I use the systematics and reasoning of an article to help me in learning	0.9271	0.9388	-0.6091	-0.4891	0.4949
	I try to divide the learning process and easier	1.2394	1.2179	1.8112	1.4912	0.4016
	I ask myself whether the material I'm reading has anything to do with what I already know	1.1199	1.1157	0.9711	0.8711	0.4936
	I'm more concerned with the overall meaning than the less specific	0.856	0.8443	-1.1991	-1.1892	0.5796
	I try translating concepts new in my own words	0.8289	0.8647	-1.4492	-1.0291	0.5217
	I focus on the meaning and importance of information new	0.9535	0.911	-0.349	-0.6691	0.5853
	I use my powers of reasoning to cover up my shortcomings	0.933	0.9659	-0.5291	-0.229	0.4839
Monitoring	I ask myself have I considered all options when solving a problem	0.981	0.9422	-0.119	-0.4091	0.5684
	I find myself analyzing the benefits of various strategies when I study	1.1366	1.0941	1.1711	0.8011	0.5078
	I find myself stopping regularly to check my understanding	0.8215	0.7843	-1.5392	-1.7692	0.5794
	I ask myself how good my learning process is when I learn something new	0.9519	0.9124	-0.349	-0.6191	0.5658
Debugging	I stopped and looked back at the lesson material that was not clear	1.0955	1.0677	0.7911	0.5411	0.4596
	I changed the strategy of learning when I couldn't understand	0.886	0.8532	-0.8791	-0.9991	0.5398
Evaluation	I know how well I do some homework when I have completed the task the	0.9145	0.8687	-0.6691	-0.9691	0.535

I ask myself, is their way easier to perform a task after I finish	0.8516	0.8256	-1.2691	-1.4192	0.5555
I ask myself I myself do I have to consider all possible solutions to solve a problem once I	0.9788	0.9988	-0.139	0.031	0.5129
I asked myself about the success of my goals when I finished studying	0.9315	1.0396	-0.5491	0.351	0.5545

Reliability Cronbach alpha (Summary statistic)

values Cronbach alpha or the alpha coefficient shows the interaction between the person and the item of the assessment instrument. Cronbach's alpha was used to measure the internal consistency of the parallel instrument items developed to measure the target unidimensional outcome construct [23]. The results of the item analysis showed a special level of reliability with a value of 0.95. The person-item interaction is categorized as very good with a value Cronbach alpha of 0.94.

Table 6. Instrument Reliability (Cronbach's alpha)

Person		Item	
Real RMSE	0.26	Real RMSE	0.14
Separation	3.34	Separation	4.27
Person reliability	0.92	Item reliability	0.95
Cronbach's alpha			0.94

CONCLUSION

Based on the results of the IRT analysis of the MAI instrument, the 34 items are suitable for use with good distribution, discrimination power, and difficulty level. The results of the DIFF measure analysis show that six items have the potential for gender differences in the assessment and must be eliminated. IRT analysis obtained a valid MAI instrument with high interaction between students and items, namely with a value Cronbach alpha of 0.94.

REFERENCES

[1] Hodson, D. (2014). Learning science, learning about science, doing science: Different goals demand different learning methods. International Journal of Science Education, 36(15), 2534-2553.

[2] Astuti, A. P., Aziz, A., Sumarti, S. S., & Bharati, D. A. L. (2019, June). Preparing 21st century teachers: implementation of 4C character's pre-service teacher through teaching practice. In Journal of Physics: Conference Series (Vol. 1233, No. 1, p. 012109). IOP Publishing.

[3] Mahanani, P. (2018, December). Analysis of Challenges and Needs of Generation Behavior in 21st Century. In International Conference on Education and Technology (ICET 2018) (pp. 146-149). Atlantis Press.

[4] Oktapiani, N., & Hamdu, G. (2020). Desain Pembelajaran STEM berdasarkan Kemampuan 4C di Sekolah Dasar. Jurnal Ilmiah Pendidikan Dasar, 7(2), 99-108.

[5] Danial, M. (2016). Kesadaran metakognitif, keterampilan metakognitif, dan penguasaan konsep kimia dasar. Jurnal Ilmu Pendidikan, 17(3).

[6] Schraw, G., Olafson, L., Weibel, M., & Sewing, D. (2012). Metacognitive knowledge and field-based science learning in an outdoor environmental education program. In Metacognition in science education (pp. 57-77). Springer, Dordrecht.

[7] Zohar, A., & Barzilai, S. (2013). A review of research on metacognition in science education: Current and future directions. Studies in Science education, 49(2), 121-169.

[8] Asy'ari, M., Ikhsan, M., & Muhali, M. (2018). Validitas instrumen karakterisasi kemampuan metakognitif mahasiswa calon guru fisika. Prisma Sains: Jurnal Pengkajian Ilmu dan Pembelajaran Matematika dan IPA IKIP Mataram, 6(1), 18-26.

[9] Hadwin, A. F., Järvelä, S., & Miller, M. (2011). Self-regulated, co-regulated, and socially shared regulation of learning. Handbook of self-regulation of learning and performance, 30, 65-84.

[10] Sitzmann, T., & Ely, K. (2011). A meta-analysis of self-regulated learning in work-related training and educational attainment: what we know and where we need to go. Psychological bulletin, 137(3), 421.

[11] Anderson, T., & Kanuka, H. (1999). Using constructivism in technology-mediated learning: Constructing order out of the chaos in the literature.

[12] Anggo, M. (2011). Pemecahan masalah matematika kontekstual untuk meningkatkan kemampuan metakognitif siswa. Edumatica: Jurnal Pendidikan Matematika.

[13] Lestari, F., SARYANTONO, B., Syazali, M., Saregar, A., MADIYO, M., JAUHARIYAH, D., & Rofiqul, U. M. A. M. (2019). Cooperative Learning Application with the Method of" Network Tree Concept Map": Based on Japanese Learning System Approach. Journal for the Education of Gifted Young Scientists, 7(1), 15-32.

[14] Sumintono, B. (2018, February). Rasch Model Measurements as Tools in Assesment for Learning. In 1st International Conference on

ISSN 1907-1744 (Print)
ISSN 2460-1500 (Online)

Education Innovation (ICEI 2017) (pp. 38-42). Atlantis Press.
[15] Pasquali, L., & Primi, R. (2003). Basic theory of Item Response Theory: IRT. Avaliação Psicológica, 2(2), 99-110.
[16] Abdullah, R., & Soemantri, D. (2018). Validasi Metacognitive Awareness Inventory pada Pendidikan Dokter Tahap Akademik. Jurnal eJKI, 6(1), 15-23.
[17] Sumintono, B., & Widhiarso, W. (2015). Aplikasi pemodelan rasch pada assessment pendidikan. Trim komunikata.
[18] Sumintono, B. (2017). Rasch Model Measurement as Tools in Assessment for Learning.
[19] Saefi, M., Fauzi, A., Kristiana, E., Adi, W. C., Muchson, M., Setiawan, M., ... & Ramadhani, M. (2020). Validating of Knowledge, Attitudes, and Practices Questionnaire for Prevention of COVID-19 Infections among Undergraduate Students: A RASCH and Factor Analysis. EURASIA Journal of Mathematics, Science and Technology Education, 16(12).
[20] Vitasari, S. D. (2018). Assessment Instrument of Scientific Literacy Skills Based on Nature of Science For Middle School. Journal of Development Research, 2(2), 59-65.
[21] Whitmore, M. L., & Schumacker, R. E. (1999). A comparison of logistic regression and analysis of variance differential item functioning detection methods. Educational and Psychological Measurement, 59(6), 910-927.
[22] SE, CES-D. Item Calibration Logit Real, Infit MNSQ Infit ZSTD Outfit MNSQ, and Outfit ZSTD. "Evaluation of the CES-D in Six Countries Using Rasch Item Response Theory (IRT) Analysis." Education 11: 2-7.
[23] Heo, M., Kim, N., & Faith, M. S. (2015). Statistical power as a function of Cronbach alpha of instrument questionnaire items. BMC medical research methodology, 15(1), 1-9.