AKSIOMA: Jurnal Program Studi Pendidikan Matematika ISSN 2089-8703 (Print)
Volume 9, No. 4, 2020, 865-871 ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v9i4.3107

MATHEMATICS STUDENTS' HOTS ASSESSMENT

Benidiktus Tanujaya[1], Rully Charitas Indra Prahmana[2*], Jeinne Mumu[3]

[1,3]Universitas Papua, Manokwari, Indonesia
[2*]Universitas Ahmad Dahlan, Yogyakarta, Indonesia
*Corresponding author
E-mail: *b.tanujaya@unipa.ac.id* [1)]
rully.indra@mpmat.uad.ac.id [2*)]
j.mumu@unipa.ac.id [3)]

Received 21 September 2020; Received in revised form 28 December 2020; Accepted 29 December 2020

Abstract

Assessment is a crucial aspect of education. A critical point in the evaluation is the validity of the instruments used in conducting the assessment. However, some studies do not pay more attention to this section, which results in the invalid results of the resulting research. This study aimed to map the indicators of the Higher-Order Thinking Skills (HOTS) of mathematics students and analyze their existence as components of the instruments. The subjects were 203 senior high school students of science, Manokwari, Indonesia. Test instruments that involved five critical and four creative thinking were used to measure students' HOTS. The data was analyzed using multidimensional scaling (MDS) to map the indicators. The results showed that the five indicators of critical thinking skills form a unified distribution pattern, while the four indicators of creativity tend to spread. Therefore, each indicator used has a unique contribution in explaining the HOTS of mathematics students.

Keywords: Creative thinking; critical thinking; HOTS instrument: multidimensional scaling.

Abstrak

Penilaian atau evaluasi merupakan aspek penting dari pendidikan. Titik kritis dalam evaluasi adalah validitas instrumen yang digunakan dalam melakukan penilaian. Namun, sejumlah penelitian tidak fokus memperhatikan bagian ini, yang berakibat pada hasil penelitian yang tidak valid. Penelitian ini bertujuan untuk memetakan indikator Higher Order Thinking Skills (HOTS) matematika siswa dan menganalisis keberadaannya sebagai komponen penting pada suatu instrumen. Subjek penelitian ini adalah 203 siswa SMA IPA di Manokwari, Indonesia. Instrumen tes yang melibatkan lima indikator berpikir kritis dan empat indikator berpikir kreatif digunakan untuk mengukur HOTS siswa. Data dianalisis menggunakan multidimensional scaling (MDS) untuk memetakan seluruh indikator. Hasil penelitian menunjukkan bahwa kelima indikator keterampilan berpikir kritis membentuk pola sebaran yang menyatu, sedangkan keempat indikator kreativitas cenderung menyebar. Oleh karena itu, setiap indikator yang digunakan memiliki kontribusi unik dalam menjelaskan HOTS matematika siswa.

***Kata kunci**: Berfikir kreatif; berfikir kritis; instrumen HOTS; multidimensional scaling.*

INTRODUCTION

Mathematics is one of the most critical subjects in the education system in various countries, including Indonesia. It is indicated by including this material in several evaluation programs at the international level, such as TIMSS and PISA. Besides, as one of the scientific thinking parts, mathematics is needed for the development of students' thinking skills (Koerber, Mayer, Osterhaus, Schwippert, & Sodian, 2015), mathematical literacy abilities (Heriyadi & Prahmana, 2020), and their characters,

DOI: https://doi.org/10.24127/ajpm.v9i4.3107

such as honesty, discipline, perseverance, responsibility, and confidence (Tanujaya, 2016). Therefore, students need to have sufficient mathematical knowledge and skills to face a better future in every area of life.

Merely having mathematics knowledge is not enough; students must be able to think critically to solves mathematics problems (Peter, 2012). Consequently, students must learn mathematics with understanding. They have to construct their knowledge actively through experience and previous knowledge, and to conduct an assessment for improving the learning process.

The assessment of students' achievement is essential to the teaching and learning process (Bilgin, Karakuyu, & Ay, 2015; Keller, Neumann, & Fischer, 2017). Assessment is a process of gathering data that accurately reflects students' achievement of the curriculum expectations in a subject. Thus, there are some purposes of evaluation, although the primary purpose of assessment is basically to gather information and provides feedback to support the teaching and learning process (Tanujaya, 2017); facilitate student learning, and improve teaching practice of the teacher (Suurtaam, et al. 2016). The assessment drives the teaching and learning process.

Assessment is a crucial aspect of education, while the standard criterion for the appropriate evaluation is validity (Drijvers, Kodde-Buitenhuis, & Doorman, 2019). A critical point of the assessment is the validity of the instruments used in conducting the evaluation. Validity in education research is a principal problem because it involves the accuracy of instruments used for measurement. It means that the lack of instruments' validity can provide research results that lack validity as well. Therefore, the validity of an instrument needs to be considered in a study.

There are four groups of validity, namely statistical conclusion, internal, construct, and external or generalization. Construct validity can be translation validity or criteria related validity. Meanwhile, translation validity is further divided into face validity and content validity (Dross, 2011). Furthermore, Zamanzadeh et al. (2015) stated that content validity is essential in research, among other types.

Content validity can be represented in the phases of development and expert judgment (Yaghmale, 2003). Content validity, also known as content-related, intrinsic, relevance, representative and logical or sampling validity, can be used to measure interest variables. Therefore, content validity measures the completeness and representativeness of the scale content. It refers to the degree at which an instrument covers the content meant to be measured and can be obtained from literature, representatives of relevant populations, and experts.

At all levels of the Indonesian education system, the evaluation of the success of mathematics instruction is based on students' HOTS. Among various thinking abilities acquired during formal education, critical and creative thinking skills are two components that should be considered in learning mathematics (NCTM, 2000). In this regard, several researches noted that critical and creative thinking skills have two principal dimensions of HOTS (Wang & Wang, 2011).

Based on these theories, Tanujaya (2016) developed an instrument to measure the HOTS of mathematics students using the two dimensions of critical and creative skills. The instrument has good validity and reliability based on some phase of development, expert judgment, field trials, and then analyzed statistically using Structural Equation Modelling (SEM). It is

DOI: https://doi.org/10.24127/ajpm.v9i4.3107

a standard procedure used by some experts in developing an instrument test with some modification (Coulacoglou & Saklofske, 2018). The instrument constructed is said to be valid according to the whole process.

Structural Equation Modelling (SEM) is a multivariate quantitative technique employed to describe the relationships among observed variables. The method helps the researcher to test or validate a theoretical model for theory testing and extension (Thakkar, 2020). The technique could be view as a combination of three statistical methods, namely multiple regression, path analysis, and confirmatory factor analysis (Salkind, 2010). Therefore, SEM provides comes a higher level of complexity, requiring more excellent knowledge about the conditions and assumptions for appropriate usage. Without due consideration, the results and conclusions based on its application can be seriously flawed or invalid (Hair, Ringle, & Sarstedt, 2013). Some assumptions for valid usage of SEM, among others: endogenous variables and exogenous variables have a linear relationship, the variables should affect and cause relationship, and the sample size is generally 20 times more than the number the indicator (Thakkar, 2020). Consequently, the complexity of applying SEM results in need for another statistical method that is easy to use by presenting the same but more informative analysis results.

There are several relevant questions related to the study, such as in learning mathematics, what is the relationship between critical and creative thinking skills of high school students? Do there have a close relationship? How are these related? Could these two skills be formed at the same time, or learned separately? To answer these questions, it is necessary to analyze the relationship through the mapping of various indicators of critical and creative thinking skills of mathematics students.

Various statistical analysis methods are available to find the relationship between variables in their observations, including correlation and regression analysis (Schmidt-Catran, Fairbrother, & Andreß, 2019; Brysbaert, 2019). The two-statistical analysis produces statistical data in a numerical format, which can be evaluated in one dimension. Results of data analysis presented in image or graphic have many advantages compared to numerical form. Several research results could deduce a higher number and many kinds of conclusions by using the image or graphic format. Hence, more information could be generated from the corresponding research representing the observed populations. One of the statistical methods which produce an image or graphical format from the analysis is multidimensional scaling.

Multidimensional scaling (MDS) is a statistical technique that can be used to produce geometric models of proximities data (Jacoby & Armstrong II, 2014), or mapping the structure of objects (Davidson, Richards, & Rounds Jr., 1986). MDS represents measurements of similarity (or dissimilarity) among pairs of objects as distances between points of a low-dimension in multidimensional space. The graphical display of the correlations provided by MDS enables the researcher to analyze the data and explore its structure visually. Too often shows regularities that remain hidden when studying arrays of numbers (Borg & Groenen, 2006).

Therefore, this study aims to map the HOTS of mathematics students' indicators using the multidimensional scaling statistic method. The results of this study are used to explore the existence of various indicators of The HOTS instruments for mathematics students. These results are also expected to contribute to developing a suitable strategy in mathematics learning to improve the critical and creative thinking skills of mathematics students.

DOI: https://doi.org/10.24127/ajpm.v9i4.3107

METHOD

The object analysis of this study is the instrument used to measure the HOTS of mathematics students. The essay test was developed by Tanujaya (2016). The instrument measures both critical and creative skills and consists of nine questions representing HOTS's indicators. Critical thinking skills' indicators include prediction of impact, problem-solving, decision making, conceptual, and principles of understanding. Meanwhile, creativity's indicators consist of four items, namely working within the boundaries of competence, overcoming new challenges, having different reasoning patterns, and having lateral (imaginative) thinking.

The subjects for this study were 203 students majoring in Natural Sciences at one of state senior high school in Manokwari, Indonesia, were used as subjects for the test instrument, and it lasted for 1 hour (60 minutes). Assessment of students' work uses a holistic rubric that can evaluate three main components, namely question understanding, answer procedure, and correctness of answers. The data obtained from this assessment were students' test scores ranging from 0 to 108, which were subsequently converted from 0 to 100.

The results were statistically analyzed using MDS. As a statistical technique, it is used to reduce the complexity of a data set to permit the visual appreciation of the underlying relational structures (Hout, Papesh, & Goldinger, 2013). Therefore, this research should be able to find and visually recognize the relationships between several indicators that construct critical and creative skills using MDS.

Data analysis was performed using the MINITAB program package. The study's output was a two-dimensional graph produced by MDS, and it provided information about HOTS indicators' distribution. Based on similarity factors, indicators can be classified through their distribution. This distribution related to Hout, et al. (2016), which stated that the output of MDS is a 'map' that conveys the relationship between items, in this regard, similar elements are located proximal to one another, while different ones are proportionately further apart.

RESULTS AND DISCUSSION

The HOTS' developed instrument should be valid with a unique role. The instrument has good validity if each of these indicators must have a unique contribution to higher-order thinking skills. However, when there is an overlap among the indicators in explaining thinking skills, the instrument is not valid and should not be used. Therefore, it is necessary to conduct a study to find out the existence of indicators used to measure HOTS.

There are different types of statistical methods developed to generate data analysis results in the image or graphical format for measure HOTS's indicators. One of them is MDS which result showed that the mapping has a disperse configuration, and graphical representation's details were revealed in Figures 1, 2, and 3. The five HOTS indicators for critical thinking skills were represented in Figure 1, while the remaining four creative skills were indicated in Figure 2. Meanwhile, both critical and creative skills' distribution arrangements represent in Figure 3.

On the other hand, learning mathematics requires thinking mathematically. Mathematics thinking skills, especially Higher Order Thinking Skills (HOTS), are essential aspects of mathematics instruction (Tanujaya,

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v9i4.3107

Prahmana, & Mumu, 2017). There is a linear, positive, and strong relationship between HOTS and the performance of mathematics students, such as self-regulated learning, habit of mind, and creativity (Hodiyanto & Firdaus, 2020). Students with a high level of higher-order thinking skills tend to be more successful in their studies (Yang, 2015; Budsankom, et al. 2015). Students with HOTS can learn, improve their performance, and reduce their weaknesses (Yee, et al. 2011).

HOTS is the highest level in the hierarchy of cognitive processes. This higher-level thinking allows students to excel and achieve intellectual freedom (Limbach & Waugh, 2009). HOTS of students happen when they get new information, keep in memory and compile, link to existing knowledge, and generate this information to achieve a goal or solve a complicated situation. HOTS can challenge a person to interpret and analyze data, consequently allowing students to think critically about a lot of available data in a limited time. (Yee, et al. 2015). Therefore, to evaluate the progress of mathematics instruction, achievement should be accessed through the instrument of students' HOTS. Does the instrument use measures students' higher-order thinking skills have good content validity?

The students' ability to use mathematics concepts (CRITICAL_1), apply working principles (CRITICAL _2), predicting the impacts of both (CRITICAL_3), solving related problems (CRITICAL_4), and their decision making (CRITICAL_5) are the five critical thinking skills' indicators used for measuring HOTS. In contrast, the four creative skills' indicators are student's ability to solve mathematical problems by working at their competence limit (CREATIVE_1), trying new things (CREATIVE_2), with their divergence (CREATIVE_3), and imaginative abilities (CREATIVE_4).

Figure 1 showed that five HOTS indicators for critical thinking skills tend to disperse, and none of them has overlapping positions in a two-dimensional scatter plot. The distribution pattern explained that the indicators represent different natures of characters and could be used to generate a comprehensive information on HOTS of the study's subjects.

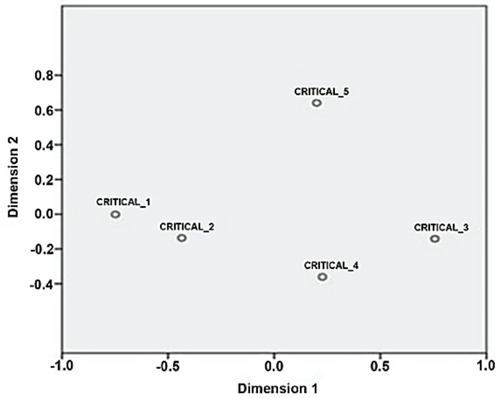

Figure 1. Distribution of HOTS indicators for critical thinking skills.

Furthermore, it appears that the indicators analyzed formed three groups based on their proximity. The first consists of CRITICAL_1 and 2, while CRITICAL_3 and CRITICAL_4 are contained in the second group. CRITICAL_5 is formed in the third group.

The existence of the first group shows that students' ability to utilize mathematics concepts has a close correlation with using the subject's principles. An idea is a set of properties linked by specific rules (Hulse, Egeth, & Deese, 1980). It is constructed by observing the features of a set of appropriate examples, while a principle is the result of a study of two or more concepts. The greater the mastery of mathematical concepts, the higher the

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v9i4.3107

ability to use its corresponding principles. Students are required to learn various interconnected concepts for mastering mathematics principles.

The principle is the result of the study of two or more mathematical concepts. Furthermore, students are expected to know more about utilizing or understanding mathematical concepts (Tanujaya, 2016). For example, when the sum of two real numbers is said to be commutative, it is one of the principles in the number of real numbers, while both are two concepts in mathematics. To understand the commutative principle, a student must first know the thoughts of addition and real numbers.

Furthermore, the second group's formation is due to the close relationship between the student's ability to predict the impact of using mathematics concepts and principles (CRITICAL_3) and solving problems (CRITICAL_4). When students can predict the effect, they can solve the problems.

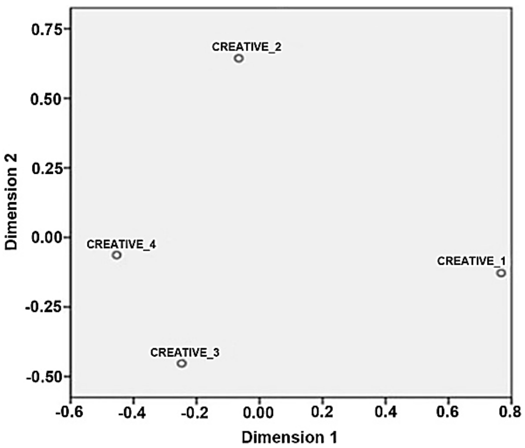

Figure 2. Distribution of HOTS indicators for creative thinking skills

Similar to Figure 1, the distribution of four HOTS indicators in Figure 2 is much the same as the first one, and it illustrated it disperse configurations with none of them showed in overlapping positions. This distribution arrangement indicated that the four indicators are accurately measured using the different features with each of them in their respective groups.

The indicators of creative thinking skills are located far apart, and it's a confirmation that there is no significant relationship among them. Students' ability to solve problems by working on the limits of their competence (CREATIVE_1) does not have a significant connection to trying new things (CREATIVE_2). Furthermore, their ability to think differently (CREATIVE_3), does not have a significant relationship with imaginative reasoning (CREATIVE_4). There is no significant correlation between two different creativity indicators as they do not have a close relationship.

The indicators of creative thinking skills differ from one another because creativity is the process of bringing new and original ideas into existence. It means thinking and acting innovatively (Ann Mean, 2008). Creativity levels vary from individuals in the same manner with actions and thoughts.

Moreover, as a skill, creative thinking can be trained and developed. It agrees with de Bono's opinion (1990), which states that the ability of human reasoning is not something that is given but can be trained and developed. Therefore, Ann Mean (2008) explained that natural creativity would remain hidden until one is put in a position to use them.

The distribution pattern was shown in a non-overlapping sequence when nine HOTS indicators were represented in one graphical illustration. The following figured the observed distribution of students in mathematics learning. The mapping provided in Figure 3 shows that the five indicators of critical thinking skills are building a

AKSIOMA: Jurnal Program Studi Pendidikan Matematika
Volume 9, No. 4, 2020, 865-871
ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v9i4.3107

more reliable and unified structure and producing independent groups. In contrast, the ones corresponding to creative skills tend to have more scattered configurations. Each creativity's indicator forms different groups because of their high variation. The scatter plot also shows that there is a high degree of similarity among critical thinking skills' indicators, but on the other hand, creativities differ. Therefore, the display in Figure 3 provided a corresponding result to what was presented in Figure 1 and Figure 2.

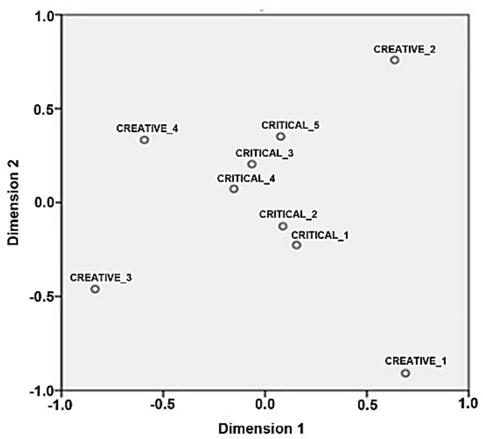

Figure 3. Distribution of HOTS indicators

Furthermore, Figure 3 presents information that there is some space among the indicators. This circumstance shows that there are dimensions that have not been used on the instrument developed. In other words, there are still different dimensions that need to be involved in measuring HOTS. This statement corresponded to the Mertens' concept (2015) who states that there are two main threats to construct validity, one of which is the construct underrepresentation. Construct under representation is a situation where the assessment to narrow and fails to include essential dimensions of the construct.

In some literature, The HOTS dimension consists of three different aspects, namely critical thinking, creative thinking, and decision making (Glassner & Schwarz, 2007; Vidergor, 2018); critical thinking, systemic thinking, and creative thinking (Teqja & Dennis Jr., 2016); critical thinking, design thinking, and systems thinking (Wang & Wang, 2011). Therefore, it can be stated that the instrument being developed has good construct validity, but less on content validity. There are still several dimensions that need to be included in the HOTS instrument.

Nevertheless, based on Figure 3, there are no overlaps among the nine indicators evaluated. On the other hand, the evaluation has been used for multiple purposes, such as providing student grades, system monitoring, determining interventions, improving teaching and learning, or providing individual feedback to students (Newton, 2007; Graham, Hebert, & Harris, 2015). Furthermore, each indicator has a unique role in explaining the HOTS of mathematics students, although some indicators need to be included on the instrument.

The distribution of indicators also confirms that as a statistical analysis tool, MDS can be used to evaluate the validity of instruments developed. Therefore, as a statistical technique, MDS can be used as an alternative to providing evidence about the validity of a measurement instrument. It's because Mohajan (2017) stated that instruments' validity plays a role in determining quality, and only a valid instrument will produce credible research.

CONCLUSION AND SUGGESTION

Indicators for critical thinking skills demonstrate higher similarities compared to that of creativity. These indicators can be arranged into one group, while those of creativity cannot

DOI: https://doi.org/10.24127/ajpm.v9i4.3107

be brought together. However, all of them have a series of contributions to the HOTS of mathematics students. Their development requires a different treatment even when they may be related. The development of critical thinking skills can be compatible with other indicators. In contrast, creativity's build-up cannot be synced with others. The results of this study confirm that MDS can be used to test the validity of measurement instruments.

Furthermore, as a suggestion, MDS also includes information about the lack of dimension used in the instrument was developed. It is essential to providing the same results with SEM in the development of an instrument. Therefore, further development instrument is needed to improve this instrument developed to include another dimension, such as Design Thinking.

REFERENCES

Ann Mean, L. (2008). *On Creativity. Awakening the Creative Mind*. Kuala Lumpur: Pelanduk Publications.

Bilgin, I., Karakuyu, Y., & Ay, Y. (2015). The effects of project based learning on undergraduate students' achievement and self-efficacy beliefs towards science teaching. *Eurasia Journal of Mathematics, Science & Technology Education*, *11*(3), 469-477. https://doi.org/10.12973/eurasia.2014.1015a

Borg, I., & Groenen, P. (2006). Modern multidimensional scaling: Theory and applications. *Journal of Educational Measurement*, *40*(3), 277-280. https://doi.org/10.1111/j.1745-3984.2003.tb01108.x

Brysbaert, M. (2019). How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. *Journal of Cognition*, *2*(1), 1-38. https://doi.org/10.5334/joc.72

Budsankom, P., Sawangboon, T., Damrongpanit, S., & Chuensirimongkol, J. (2015). Factors affecting higher order thinking skills of students: a meta-analytic structural equation modelling study. *Educational Research and Reviews*, *10*(19), 2639-2652. https://doi.org/10.5897/ERR2015.2371

Coulacoglou, C., & Saklofske, D. (2018). *Psychometrics and Psychological Assessment: Principles and Applications*. Cambridge: Academic Press.

Davison, M. L., Richards, P. S., & Rounds Jr., J. B. (1986). Multidimension scaling in counselling research and practice. *Journal of Counselling and Development*, *6*, 178-184. https://doi.org/10.1002/j.1556-6676.1986.tb01309.x

de Bono, E. (1990). *Lateral Thinking: A Text Book of Creativity*. London: Penguin Books.

Drijvers, P., Kodde-Buitenhuis, H., & Doorman, H. (2019). Assessing mathematical thinking as part of curriculum reform in the Netherlands. *Educational Studies in Mathematics*, *102*, 435–456. https://doi.org/10.1007/s10649-019-09905-7

Drost, E. A. (2011). Validity and reliability in social research. *Education Research and Perspectives*, *38*(1), 105-124. https://www.erpjournal.net/wp-

DOI: https://doi.org/10.24127/ajpm.v9i4.3107

content/uploads/2020/02/ERPV38-1.-Drost-E.-2011.-Validity-and-Reliability-in-Social-Science-Research.pdf

Glassner, A., & Schwarz, B. B. (2007). What stands and develops between creative and critical thinking? Argumentation?. *Thinking Skills and Creativity*, *2*(1), 10-18. https://doi.org/10.1016/j.tsc.2006.10.001

Graham, S., Hebert, M., & Harris, K. R. (2015). Formative assessment and writing: A meta-analysis. *The Elementary School Journal*, *115*(4), 523-547. https://doi.org/10.1086/681947

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2013). Partial least squares structural equation modeling: Rigorous applications, better results and higher acceptance. *Long range planning*, *46*(1-2), 1-12. https://ssrn.com/abstract=2233795

Heriyadi & Prahmana, R. C. I. (2020). Pengembangan lembar kegiatan siswa menggunakan pendekatan pendidikan matematika realistik. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *9*(2), 395-412. http://dx.doi.org/10.24127/ajpm.v9i2.2782

Hodiyanto & Firdaus, M. (2020). The self regulated learning, habit of mind, and creativity as high order thinking skills predictors. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *9*(1), 21-30. http://dx.doi.org/10.24127/ajpm.v9i1.2589

Hout, M. C., Godwin, H. J., Fitzsimmons, G., Robbins, A., Menneer, T., & Goldinger, S. D. (2016). Using multidimensional scaling to quantify similarity in visual search and beyond. *Attention, Perception, & Psychophysics*, *78*, 3-20. https://doi.org/10.3758/s13414-015-1010-6

Hout, M. C., Papesh, M. H., & Goldinger, S. D. (2013). Multidimensional scaling. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(1), 93-103. https://doi.org/10.1002/wcs.1203

Hulse, S. H., Egeth, H., & Deese, J. (1980). *The Psychology of Learning*. New York: McGraw-Hill.

Jacoby, W. G., & Armstrong II., D. A. (2014). Bootstrap confidence regions for multidimensional scaling solutions. *American Journal of Political Science*, *58*(1), 264-278. https://doi.org/10.1111/ajps.12056

Keller, M. M., Neumann, K., & Fischer, H. E. (2017). The impact of physics teachers' pedagogical content knowledge and motivation on students' achievement and interest. *Journal of Research in Science Teaching*, *54*(5), 586-614. https://doi.org/10.1002/tea.21378

Koerber, S., Mayer, D., Osterhaus, C., Schwippert, K., & Sodian, B. (2015). The development of scientific thinking in elementary school: A comprehensive inventory. *Child Development*, *86*(1), 327-336. https://doi.org/10.1111/cdev.12298

Limbach, B., & Waugh, W. (2009). Developing higher level thinking. *Journal of Instructional Pedagogies*, 1-9. https://www.aabri.com/manuscripts/09423.pdf

DOI: https://doi.org/10.24127/ajpm.v9i4.3107

Mertens, D. M. (2015). *Research and Evaluation in Education and Psychology*. Los Angeles: Sage Publications.

Mohajan, H. K. (2017). Two criteria for good measurement in research: Validity and reliability. *Annals of Spiru Haret University*, *17*(3), 58-82. https://mpra.ub.uni-muenchen.de/83458/1/MPRA_paper_83458.pdf

NCTM. (2000). *Principles and Standards for School Mathematics*. Reston: The National Council of Teachers of Mathematics, Inc.

Newton, P. E. (2007). Clarifying the purposes of educational assessment. *Assessment in Education: Principles, Policy & Practice*, *14*(2), 149-170. https://doi.org/10.1080/09695940701478321

Peter, E. E. (2012). Critical thinking: Essence for teaching mathematics and mathematics problem solving skills. *African Journal of Mathematics and Computer Science Research*, *5*(3), 39-43. https://doi.org/10.5897/AJMCSR11.161

Salkind, N. J. (2010). *Structural Equation Modelling*. Los Angeles: SAGE Publishing. https://doi.org/10.4135/978141296 1288.

Schmidt-Catran, A. W., Fairbrother, M., & Andreß, H. J. (2019). Multilevel models for the analysis of comparative survey data: Common problems and some solutions. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, *71*(1), 99-128. https://doi.org/10.1007/s11577-019-00607-9

Suurtamm, C., Thompson, D. R., Kim, R. Y., Moreno, L. D., Sayac, N., Schukajlow, S., Silver, E., Ufer, S., & Vos, P. (2016). *Assessment in Mathematics Education*. Cham: Springer. https://doi.org/10.1007/978-3-319-32394-7_1

Tanujaya, B. (2016). Development an instrument to measure higher order thinking skills in senior high school mathematics instruction. *Journal Education and Practice*, *7*(21), 144-148. https://www.iiste.org/Journals/index.php/JEP/article/view/31982/32852

Tanujaya, B. (2017). Application assessment as learning in mathematics instruction. *Advances in Social Science, Education, and Humanities Research*, *100*, 140-145. https://doi.org/10.2991/seadric-17.2017.30

Tanujaya, B., Prahmana, R. C. I., & Mumu, J. (2017). Mathematics instruction, problems, challenges, and opportunities: A case study in Manokwari regency, Indonesia. *World Transactions on Engineering and Technology Education*, *15*(3), 287-291. http://www.wiete.com.au/journals/WTE&TE/Pages/Vol.15,%20No.3%20(2017)/16-Tanujaya-B.pdf

Teqja, Z., & Dennis Jr., S. F. (2016). Creative thinking, critical thinking and systemic thinking-key instruments to deeply transform the higher education system in Albania: The case of landscape architecture. *Educational Alternatives*, *14*, 543-555. https://www.scientific-publications.net/get/1000021/1474994295652029.pdf

DOI: https://doi.org/10.24127/ajpm.v9i4.3107

Thakkar, J. J. (2020). *Structural Equation Modelling: Application for Research and Practice*. Cham: Springer. https://doi.org/10.1007/978-981-15-3793-6

Vidergor, H. E. (2018). Effectiveness of the multidimensional curriculum model in developing higher-order thinking skills in elementary and secondary students. *The Curriculum Journal, 29*(1), 95-115. https://doi.org/10.1080/09585176.2017.1318771

Wang, S., & Wang, H. (2011) Teaching higher order thinking in the introductory MIS course: A model-direct approach. *Journal Education for Business*, *86*(4), 208-212. https://doi.org/10.1080/08832323.2010.505254

Weintrop, D., Beheshti, E., Horn, M., Orton, K., Jona, K., Trouille, L., & Wilensky, U. (2016). Defining computational thinking for mathematics and science classrooms. *Journal of Science Education and Technology, 25*(1), 127-147. https://doi.org/10.1007/s10956-015-9581-5

Yaghmale, F. (2003). Content validity and its estimation. *Journal of Medical Education*, 1(3), 25-27. https://doi.org/10.22037/jme.v3i1.870

Yang, Y. T. C. (2015). Virtual CEOs: A blended approach to digital gaming for enhancing higher order thinking and academic achievement among vocational high school students. *Computers & Education, 81*, 281-295. https://doi.org/10.1016/j.compedu.2014.10.004

Yee, M. H, Othman, W., Yunos, J., Tee, T. K., Hassan, R., & Mohamad, M. M. (2011). The level of marzano higher order thinking skills among technical education students. International. *Journal of Social Science and Humanity, 1*(2), 121-125. http://ijssh.org/papers/20-H009.pdf

Yee, M. H., Yunos, J. M., Othman, W., Hassan, R., Tee, T. K., & Mohamad, M. M. (2015). Disparity of learning styles and higher order thinking skills among technical students. *Procedia-Social and Behavioral Sciences, 204*, 143-152. https://doi.org/10.1016/j.sbspro.2015.08.127

Zamanzadeh, V., Ghahramanian, A., Rassouli, M., Abbaszadeh, A., Alavi-Majd, H., & Nikanfar, A. (2015). Design and implementation content validity study: Development of an instrument for measuring patient-centered communication. *Journal of Caring Sciences, 4*(2), 165-178. https://doi.org/10.15171/jcs.2015.017