International Journal of Multicultural and Multireligious Understanding

http://ijmmu.com
editor@ijmmu.com
ISSN 2364-5369
Volume 9, Issue 3
March, 2022
Pages: 780-790

Content Validation through Expert Judgement of an Instrument on the Self-Assessment of Mathematics Education Student Competency

Gusti Adillah; Achmad Ridwan; Wardani Rahayu

Department of Educational Research and Evaluation, Universitas Negeri Jakarta, Indonesia

http://dx.doi.org/10.18415/ijmmu.v9i3.3738

Abstract

The purpose of this study was to determine the validity of the content based on expert judgment on the design of the self-assessment instrument on the competence of mathematics education students. The research methodology uses descriptive qualitative, with the data collection method using online observation sheets via Google Form. Quality standard checks use member-checking based on informant sources and data analysis by triangulating sources with the Males and Huberman approach. Informants consist of three experts, namely: experts in the fields of instrument development, language, and mathematical content. To determine expert agreement using Fleiss Kappa analysis. The results show that the level of expert agreement is 0.743 in the substantial agreement category. This level of agreement is satisfactory. However, the most important thing is the suggestions and suggestions for improvement recommended by the expert for the finalization of the proposed instrument. The final modification is carried out according to the expert's instructions, taking into account the purpose of the indicator, its relationship with math content, and the statement item sentences that are arranged using active sentences and are consistent and understandable.

Keywords: *Validation; Self-Assessment; Mathematics Education*

Introduction

Competencies and skills required in the twenty-first century, specifically identifying the competencies required by students in dealing with life, the working world, and citizenship, are emphasized in seven skills, namely critical thinking skills, problem-solving, and communication (Wagner, 2010). In addition, the ability to analyse, and synthesize information (Chu, Reynolds, Tavares, Notari & Lee, 2017). These abilities can be integrated into online mathematics courses.

Based on the framework and theory that have been developed to analyse the teaching and learning process and difficulties related to other math tasks (Freudenthal, 2006; Schoenfeld, 2016; Sriraman & English, 2010). Usodo (2012) states that mathematics is not just arithmetic, but can also be used to prove the truth of an idea and solve problems through a logical and structured way of thinking. Mathematics is considered a very important science and is taught at almost all levels of education, starting from elementary school, secondary school, and college (Rosmaiyadi, 2018).

Measurement of mathematics ability in Indonesia based on the PISA survey shows that the level of students' mathematical ability is still low compared to other countries. In the PISA 2018 results for the math category, Indonesia is ranked 7th from the bottom (73), with an average score of 379 (OECD, 2018). Therefore, there needs to be an effort to improve the ability of students in Indonesia in the field of mathematics. One way is by increasing students' mathematical competence.

Various instruments have been developed as measuring tools to evaluate the learning environment, including What Is Happening In This Class (WIHIC) and Student Perception Opportunity Competence Development Instruments (SPOCD; Rahayu et al., 2020), with five competencies, namely: thinking, managing self, using language, symbols, and text, participating, and contributing, which are used as an evaluation tool to measure students' perceptions of development opportunities regarding mathematical competence. There is a growing need for the use of measurement instruments related to Student Perception Opportunity Competence Development (SPOCD). Instrument adaptation is not well known among educational professionals, which may explain the existence of incomplete instruments derived from existing instruments in the field. Therefore, several relevant skills and guidelines must be acquired and applied dynamically and sustainably, so it is necessary to validate an instrument (Lajoie et al., 2015).

A reliable instrument is required, as well as instruments that have been validated using construct validity (Beyazit et al., 2020). Validity and reliability are two quality criteria that must be met by every measuring instrument. All of these are important in determining the quality of the measurements produced by the instrument (Alzubaidi et al., 2016). Content validity is defined as the extent to which the elements of an assessment instrument are relevant to the construct targeted for the assessment (Hung & Yang, 2017). In this sense, the usual way to assess the quality of an instrument is to consult an expert, which consists of evaluating an instrument using a procedure known as expert judgment (Lange, 2017).

Content validation through expert judgment is defined by Almanasreh et al. (2019) as an informed opinion from an individual with a track record in the field who is considered by others to be a qualified expert and who can provide information, proof, and assessment. Evaluation through expert judgment consists of asking several individuals to make an assessment of an instrument or to express their opinion on a particular aspect (Leite et al., 2021). Content validation is generally carried out either during instrument design or for translation validation and standardization of instruments for use in different cultures. The role of the expert is fundamental in clarifying, adding, and/or modifying the necessary aspects (Nurrohmah et al., 2018).

This study aims to determine the validity of content based on expert judgment on the design of a self-assessment instrument to test the competence of mathematics education students. At the end of this research, it is hoped that the quality of the proposed instrument will be known to perfect the instrument by finalizing it.

Methods

The methodology of this research uses descriptive, qualitative, and quantitative data collection methods using observation sheets. quality standard with member-checking and data analysis using source triangulation from Miles and Huberman to analyze content validity through expert judgment (Nowell et al., 2017). The instrument was validated by three experts who have expertise in instrument development, language, and mathematics. The selection of experts is based on a procedure where the expert who mastered the subject matter is good because of their academic background, work experience, or recognition in the community. The validation instruments are arranged in a draft text containing the following five columns:

Table 1
Grid of Self-Assessment Instruments

Dimension	Indicator	Item Number	Number of Items
Thinking Ability	Able to think critically and creatively	1,2,3,4	4
	Able to solve math problems	6	1
	Able to relate mathematical ideas	5,7,8	3
Ability to Self-Management	Confidence in competence	9	1
	Able to manage time well	14	1
	Independence in learning	12	1
	Motivation in achieving learning goals	13	1
	Able to develop knowledge	10,11,15	3
Ability to Connect with Others	Able to interact in various contexts	16,19	2
	Able to be open	17,18	2
	Able to provide emotional support to others	20	1
Ability to Contribute and Participate	Able to contribute and work with groups	21, 22, 24	3
	Actively involved in groups	23	1
Ability to use ICT	Skilled in using ICT in the learning process	26,28	2
	Able to adapt to new technology	25,27,29	3
Ability to Plan and Learning Implementation	Able to arrange lesson plans	30,31,32,33	4
	Able to determine the technique of assessment and evaluation of the learning process	34,35,36	3

The instrument proposed for expert validation is composed of 36 items and gives qualitative advice and input. Data analysis in this study is descriptive and qualitative by checking qualitative data using triangulation, as well as other supporting data analysis using Kappa Fleiss statistics to assess the level of expert agreement among three or more raters who independently assess a series of items using a certain number of instruments with ordinal categories (Xie et al., 2018). To measure the level of approval of multiple raters in several categories (Hassan et al., 2019).

$$\kappa = \frac{\Pr(\alpha) - \Pr(e)}{1 - \Pr(e)} \qquad \text{...} (1)$$

The formula is used to measure the difference between the level of agreement $\Pr(\alpha)$ and a coincidence deal $\Pr(e)$.

Based on Brennan and Prediger, a Kappa value of 0.7 indicates that the level of agreement is reliable (Gwet, 2016; Xie et al., 2018). In addition to the interpretation of Brennan and Prediger, another common interpretation is defined by Landis and Koch (Nurrohmah et al., 2018; Warrens, 2015) as shown below.

Table 2

Kappa Value Interpretation

Value	Interpretation
$\kappa \leq 0$	No agreement
κ between 0.01 and 0.20	Slight agreement
κ between 0.21 and 0.40	Fair agreement
κ between 0.41 and 0.60	Moderate agreement
κ between 0.61 and 0.80	Substantial agreement
κ between 0.81 and 1.00	Almost perfect agreement

The minimum value assumed by this coefficient is 0 and the maximum value is 1. The scale generated by Landis and Koch (1977) quantitatively expresses the agreement of power between observers and is used for the interpretation of Fleiss values. In this research, if the Kappa score meets the agreement criteria and has reached 0.7, it means that the expert assessment has ended.

Finding

The initial development stage of the instrument is carried out through content validation assessments by three experts to measure the value of Fleiss Kappa, where expert responses are organized to result in agreement on each dimension, indicator, and statement item. The deal rate is calculated using Fleiss Kappa Statistics, with 36 items (N = 36), three experts (n = 3), and three rating values (k = 4). By calculating the probability of agreeing $\Pr(\alpha)$ and the chance of agreeing by chance $\Pr(e)$. The probability of agreement indicates the degree of agreement between experts regardless of whether the agreement occurred by chance (Hassan et al., 2019; Xie et al., 2018). By applying the calculation $\Pr(\alpha)$, get value $\Pr(\alpha)$ the obtained value is 0.986. Besides that, measure the degree of agreement that might be obtained by chance. In the calculation $\Pr(e)$, get value $\Pr(e)$ the obtained value was 0.946. The two degrees of the agreement are needed to measure the level of agreement between experts by applying the Fleiss Kappa value (1) as follows:

$$\kappa = \frac{0.986 - 0.946}{1 - 0.946} = 0.743 \quad \text{..............................} \quad (1)$$

The degree of agreement between experts, indicated by the Kappa value, is 0.743. This value satisfies the Brennan and Prediger benchmarks of 0.70 as well as the substantial agreement with Landis and Koch (Nurrohmah et al., 2018). Finally, the author states that the level of expert agreement on the assessment of the proposed instrument is in the substantial agreement category (Akhanova et al., 2021; Hepworth & Rowe, 2018). However, there are some suggestions and suggestions for improvement from experts that need to be finalized.

Discussion

The results of the assessment of the proposed instrument show that all experts have a substantial agreement. However, to improve instrument quality, the author removes irrelevant statement items and modifies the addition of several items based on suggestions and recommendations from experts.

The results of expert discussions in the assessment of the proposed instrument have not been fully agreed upon by the experts, because most of the experts disagree, namely: expert 1 and expert 3, while only one expert, namely, expert 2, indicates agreement with the proposed instrument item. Table 2 below presents indicators along with items that still require further identification, so that modifications can be made as recommended by experts.

Table 3

The Indicators that Need Modification

Indicator	Items before modification	Items after expert modification
Able to think critically and creatively	1. I discuss with my classmates different ways of thinking in solving problems 2. When I make a mistake, I think of a solution to the problem. 3. When I make a mistake, I can think of a solution to the problem 4. I can think of something that is known to solve an unknown problem 6. I can think of something that is known to solve an unknown problem	1. I can solve math problems in several ways 2. When I make a mistake, I can think of a solution to the problem 3. When I make a mistake, I learn from the problem 6. I can think how things happen to understand mathematics learning
Able to relate mathematical ideas	5. In learning mathematics, I pay more attention to the process than the result 7. In learning mathematics, I can express the results of my thoughts as the answers I give 8. In learning mathematics, I can explain the meaning of the answers given	5. I can pay attention to the process of learning mathematics 7. I can express the results of my thoughts as answers given in learning mathematics 8. I can explain the meaning of the answers given in learning mathematics
Confidence in competence	9. During mathematics learning I understand why studying the material	9. I'm sure I can learn math material
Able to manage time well	14. I am ready to take math lessons	14. I can manage time to study math
Independent learning	12. When the teacher explains to other groups, I still study independently	12. I can learn independently by studying mathematics
Motivation in achieving learning goals	13. When the teacher explains to other groups, I still study independently	13. I am active in achieving the goal of learning mathematics
Able to develop	10. I think about what I have learned	10. I can develop the mathematical

knowledge	and what to do 11. I discuss what I have learned and what to do 15. Due to clear instructions, I follow what to do when studying math	knowledge that has been learned 11. I can develop the results of discussions about mathematics 15. I can do things in learning math
Able to interact in various contexts	16. Due to clear instructions, I follow what to do when studying math 19. My friends and I take turns being the leader of the discussion group	16. While studying mathematics, my friends and I share opinions 19. My friends and I take turns being the leader of the discussion group
Able to be open	17. The teacher is proud of me when I pay close attention when other students share opinions 18. In learning mathematics, I give feedback to other students	17. I can share my opinion with other group friends 18. I can give feedback to other students
Able to provide emotional support to others	20 I provide support to my group mates when understanding the learning material	20. I can provide support to group mates when understanding
Able to contribute and work with groups	21. I like to ask questions to understand the learning material 22. I like to express my opinion to understand the learning material 24. My friends and I discussed how to apply mathematics in everyday life	21. I like to ask questions to understand the learning material 22. I can express opinions to understand the learning material 24. My friends and I discussed how to apply mathematics in everyday life
Able to arrange lesson plans	30. I know how to formulate the learning objectives to be achieved 31. I know how to set learning goals to be achieved 32. I can coordinate learning components 33. I use media in learning activities	30. I can formulate the learning to be achieved 31. I can determine the learning objectives to be achieved 32. I can coordinate learning components 33. I can use media in learning activities

Dimensions of Thinking Ability, which consists of indicators "Able to Critical Thinking and Creative," where all experts assessed disagreeing with the proposed item because the item was deemed unable to measure the context of the indicator in depth. Therefore, experts advise that the item does not indicate the ability to think but rather the habit of discussing. So some items need to be revised and lead to the ability of each item. Some theories by Mutohir et al. (2019); Shively et al. (2018); Yazar Soyadı (2015) state that critical and creative thinking requires individuals to be able to analyze and evaluate thoughts with the view that they can improve thinking based on a goal. Research from Hitchcock (2017); Lo and Feng (2020) state that individuals do not just accept arguments and conclusions for granted, but also question the validity of these arguments and conclusions.

The indicator "Able to Connect Mathematical Ideas", given a linguistic assessment of the editorial items of the statement, should use direct sentences and be consistent in the use of words. So the expert concluded that there were still some items that did not show indicators of ability with the use of active sentences. This is stated by Klette et al. (2017); Tieng and Eu (2014) that if students are trained to predict problem-solving with problems involving mathematical content, students will have good abilities without realizing it.

Dimensions of Self-Management Ability are measured using five indicators: belief in competence, ability to manage time effectively, learning independence, motivation to achieve learning goals, and ability to develop knowledge; where overall, assessed that there are still many items that have not shown indicators, such as item editorials with statement sentences that have not been straightforward, and some item sentences are not clear with the purpose of indicators. Basically according to Danciu (2010); Pellegrino and Hilton (2013), self-management ability is when a person can control his thoughts, feelings, and actions. Besides that, prioritize goals, decide what to do, and take responsibility for completing the necessary actions.

Dimensions of the ability to relate to others include three indicators: the ability to interact in a variety of contexts, openness, and the ability to provide emotional support to others. The composition item consists of five items, where the results of expert assessments show that there are still some items that do not have measure indicators, some items with unclear sentences because they are measured by the teacher and not by you. Other research by Lase (2016), states that individuals have abilities or skills that enable a person to relate to other people, such as being able to communicate both verbally and non-verbally.

Dimensions of Ability to Contribute and Participate with two indicators, namely being able to contribute and work together with groups and being actively involved in groups. The constituent points consist of four items. The results of the expert assessment show that there are still items that are not by the indicators. While item sentences mostly only measure contributions, sentences must be consistent and have clear directions with indicators. Some experts state Solari (2014), that contribution is part of participation, where individuals work together to meet their needs (Bauer & Booth, 2019). Besides that, participation can be interpreted as an activity to arouse feelings and be included in a certain activity (Decaigny, 2014).

Dimensions of the Ability to Use Information and Communication Technology (ICT) include two indicators: proficiency with ICT in learning and the ability to adapt to new technologies. The constituent items consist of five items, with the results of expert assessments showing that the items are by the objectives of the indicators being measured. It's just that the item sentence can be emphasized by directing the ability. The essence of these dimensions directs individuals to understand the quality of learning (Ratheeswari, 2018), access to learning (Tieng & Eu, 2014), visualize ideas and make it easier to understand the material (Guiller et al., 2008).

Dimensions of Ability to Plan and Learning Implementation with two indicators, namely: being able to develop learning plans and being able to determine the assessment and evaluation techniques of the learning process. There are seven constituent items, with the results of expert assessments showing some items are still not by the indicators. The item sentence has not yet been conveyed to the purpose of the indicator, and the language used in the item sentences must be clear and communicative. This is the opinion of Suffian and Nachiappan (2019), that a teacher's learning planning is expected to be able to plan learning activities effectively so that they can demonstrate their ability to design learning activities such as formulating goals, choosing teaching materials and methods, and determining evaluations in learning (Zukhairina et al., 2020).

The results of the proposed instrument improvement based on the expert agreement are in Table 2. It produces 36 statement items. The addition of items is evident in most of the indicators. The final result of this expert assessment is deemed sufficient to provide suggestions, input, and recommendations for the improvement of the instrument as a whole.

Conclusion

Agreement in the expert assessment of the self-assessment instrument on the competence of mathematics education students proposed in the substantial category of Agreement. This characteristic shows that the expert gives a fair assessment and not because of coercion, or undue influence by one expert on another and that each expert acts voluntarily and of his own free will.

Finalization of instruments in this first stage, recommended by experts to modify the addition of statement items on each indicator, in addition, language corrections are carried out by experts on most of the editorial statement items.

References

Akhanova, G., Nadeem, A., Kim, J. R., Azhar, S., & Khalfan, M. (2021). *Building Information Modeling Based Building Sustainability Assessment Framework for Kazakhstan*. 1–25.

Almanasreh, E., Moles, R., & Chen, T. F. (2019). Evaluation of methods used for estimating content validity. *Research in Social and Administrative Pharmacy*, *15*(2), 214–221. https://doi.org/10.1016/j.sapharm.2018.03.066

Alzubaidi, E., Aldridge, J. M., & Khine, M. S. (2016). Learning English as a second language at the university level in Jordan: motivation, self-regulation and learning environment perceptions. *Learning Environments Research*, *19*(1), 133–152. https://doi.org/10.1007/s10984-014-9169-7

B Usodo. (2012). Karakteristik intuisi siswa sma dalam memecahkan masalah matematika ditinjau dari kemampuan matematika dan perbedaan gender. *Aksioma*, *11*(1).

Bauer, J. R., & Booth, A. E. (2019). Exploring potential cognitive foundations of scientific literacy in preschoolers: Causal reasoning and executive function. *Early Childhood Research Quarterly*, *46*, 275–284. https://doi.org/10.1016/j.ecresq.2018.09.007

Beyazit, U., Yurdakul, Y., & Ayhan, A. B. (2020). The Psychometric Properties of the Turkish Version of the Trait Emotional Intelligence Questionnaire–Child Form. *SAGE Open*, *10*(2). https://doi.org/10.1177/2158244020922904

Chu, S. K. W., Reynolds, R. B., Tavares, N. J., Notari, M., & Lee, C. W. Y. (2017). *Twenty-first-century skills and global education roadmaps. In the 21st-century skills development through inquiry-based learning* (hal. 17–32). Springer.

Danciu, E. L. (2010). Methods of developing children's emotional intelligence. *Procedia - Social and Behavioral Sciences*, *5*, 2227–2233. https://doi.org/10.1016/j.sbspro.2010.07.440

Decaigny, T. (2014). Inquisitorial and Adversarial Expert Examinations in the Case Law of the European Court of Human Rights. *New Journal of European Criminal Law*, *5*(2), 149–166.

https://doi.org/10.1177/203228441400500203

Freudenthal, H. (2006). *Revisiting mathematics education: China lectures* (Vol. 9). Springer Science & Business Media.

Galle-Tessonneau, M., & Gana, K. (2019). Development and validation of the school refusal evaluation ScaLe1 for adolescents. *Journal of Pediatric Psychology*, *44*(2), 153–163. https://doi.org/10.1093/jpepsy/jsy061

Guiller, J., Durndell, A., & Ross, A. (2008). Peer interaction and critical thinking: Face-to-face or online discussion? *Learning and Instruction*, *18*(2), 187–200. https://doi.org/10.1016/j.learninstruc.2007.03.001

Gwet, K. L. (2016). Testing the Difference of Correlated Agreement Coefficients for Statistical Significance. *Educational and Psychological Measurement*, *76*(4), 609–637. https://doi.org/10.1177/0013164415596420

Hassan, N. F., Puteh, S., & Sanusi, A. M. (2019). Fleiss's Kappa: Assessing the concept of technology-enabled active learning (TEAL). *Journal of Technical Education and Training*, *11*(1), 109–118. https://doi.org/10.30880/jtet.2019.11.01.14

Hepworth, L. R., & Rowe, F. J. (2018). Using Delphi methodology in the development of a new patient-reported outcome measure for stroke survivors with visual impairment. *Brain and Behavior*, *8*(2), 1–9. https://doi.org/10.1002/brb3.898

Hitchcock, D. (2017). Critical Thinking as an Educational Ideal. *Argumentation Library*, *30*(April), 477–497. https://doi.org/10.1007/978-3-319-53562-3_30

Hung, J.-L., & Yang, D. (2017). The Validation of an Instrument for Evaluating the Effectiveness of Professional Development Program on Teaching Online. *Journal of Educational Technology Development and Exchange*, *8*(1). https://doi.org/10.18785/jetde.0801.02

Klette, K., Blikstad-Balas, M., & Roe, A. (2017). Linking Instruction and Student Achievement. A research design for a new generation of classroom studies. *Acta Didactica Norge*, *11*(3), 19. https://doi.org/10.5617/adno.4729

Lajoie, S. P., Poitras, E. G., Doleck, T., & Jarrell, A. (2015). Modeling metacognitive activities in medical problem-solving with bio world. *Intelligent Systems Reference Library*, *76*(October), 323–343. https://doi.org/10.1007/978-3-319-11062-2_13

Landis, J. R., & Koch, G. G. (1977). Landis and Koch1977_agreement of categorical data. *Biometrics*, *33*(1), 159–174.

Lange, R. (2017). Rasch scaling and cumulative theory-building in consciousness research. *Psychology of Consciousness: Theory, Research, and Practice*, *4*(1), 135–160. https://doi.org/10.1037/cns0000118

Lase, D. (2016). Jurnal sundermann. *Journal Sunderman*, *1*(1), 28–43.

Leite, M., Infante, V., & Andrade, A. R. (2021). Using expert judgment techniques to assess reliability for long service-life components: An application to railway wheelsets. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*. https://doi.org/10.1177/1748006X211034650

Lo, C. O., & Feng, L. C. (2020). Teaching higher-order thinking skills to gifted students: A meta-analysis. *Gifted Education International*, *36*(2), 196–217. https://doi.org/10.1177/0261429420917854

Mutohir, T. C., Kusnanik, N. W., Hidayati, S., & Mukminin, A. (2019). *Critical Thinking Skills Based – Physical Activities Learning Model for Early Childhood*. *5*(22), 738–755.

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic Analysis: Striving to Meet the Trustworthiness Criteria. *International Journal of Qualitative Methods*, *16*(1), 1–13. https://doi.org/10.1177/1609406917733847

Nurrohmah, I., Sensuse, D. I., & Santoso, H. B. (2018). The expert-judgment validation and finalization of proposed interaction design process maturity instrument: Case study: E-commerce in Indonesia. *Proceedings of the 2nd International Conference on Informatics and Computing*, *ICIC 2017*, *2018-Janua*, 1–6. https://doi.org/10.1109/IAC.2017.8280647

OECD. (2018). *PISA 2018 results*.

Pellegrino, J. W., & Hilton, M. L. (2013). Education for life and work: Developing transferable knowledge and skills in the 21st century. In *Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century*. https://doi.org/10.17226/13398

Rahayu, W., Putra, M. D. K., Iriyadi, D., Rahmawati, Y., & Koul, R. B. (2020). A Rasch and factor analysis of an Indonesian version of the Student Perception of Opportunity Competence Development (SPOCD) questionnaire. *Cogent Education*, *7*(1), 1–18. https://doi.org/10.1080/2331186X .2020.1721633

Ratheeswari, K. (2018). Information Communication Technology in Education. *Journal of Applied and Advanced Research*, *3*(S1), 45. https://doi.org/10.21839/jaar.2018.v3is1.169

Rosmaiyadi. (2018). *ANALISIS KESALAHAN PENYELESAIAN SOAL ALJABAR PADA MAHASISWA PROGRAM STUDI PENDIDIKAN MATEMATIKA STKIP SINGKAWANG*. *12*(1), 59–70.

Schmid, M., Brianza, E., & Petko, D. (2020). Self-Reported Technological Pedagogical Content Knowledge (TPACK) of Pre-Service Teachers about Digital Technology Use in Lesson Plans. *Computers in Human Behavior*, *115* (September 2020), 106586. https://doi.org/10.1016/ j.chb.2020.106586

Schoenfeld, A. H. (2016). Learning to think mathematically: Problem-solving, metacognition, and sense-making in mathematics (Reprint). *Journal of education*, *196*(2), 1–38.

Shively, K., Stith, K. M., & Rubenstein, L. D. V. (2018). Measuring What Matters: Assessing Creativity, Critical Thinking, and the Design Process. *Gifted Child Today*. https://doi.org/10.1177/ 1076217518768361

Solari, E. (2014). Longitudinal prediction of 1st and 2nd grade English oral reading fluency in ELL. *Journal of adolescence*, *74*(4), 274–283. https://doi.org/10.1002/pits

Sriraman, B., & English, L. D. (2010). *Theories of mathematics education: Seeking new frontiers*. Springer.

Suffian, S., & Nachiappan, S. (2019). Analysis of Teacher Readiness towards Higher Order Thinking Skills (HOTS) Integration in Preschool Teaching and Learning (TNL). *International Journal of Academic Research in Business and Social Sciences*, *9*(7), 417–423. https://doi.org/10.6007 /ijarbss/v9-i7/6134

Tieng, P. G., & Eu, L. K. (2014). Improving Students' Van Hiele Level of Geometric Thinking Using Geometer's Sketchpad. *Malaysia Online Journal of Educational Technology*, *2*(3), 20–31.

Wagner, T. (2010). *Overcoming The Global Achievement Gap (online)*. Mass., Harvard University.

Warrens, M. J. (2015). Five Ways to Look at CohenÃ¢Â€ Â™s Kappa. *Journal of Psychology & Psychotherapy*, *05*(04), 8–11. https://doi.org/10.4172/2161-0487.1000197

Werner, C., Bedford, T., Cooke, R. M., Hanea, A. M., & Morales-Nápoles, O. (2017). Expert judgment for dependence in probabilistic modeling: A systematic literature review and future research directions. *European Journal of Operational Research*, *258*(3), 801–819. https://doi.org/10.1016/j.ejor.2016.10.018

Xie, Z., Gadepalli, C., & Cheetham, B. (2018). A Study of Chance-Corrected Agreement Coefficients for the Measurement of Multi-Rater Consistency. *International journal of simulation: systems, science & technology*, 1–9. https://doi.org/10.5013/ijssst.a.19.02.10

Yazar Soyadı, B. B. (2015). Creative and Critical Thinking Skills in Problem-based Learning Environments. *Journal of Gifted Education and Creativity*, *2*(2), 71–71. https://doi.org/10.18200 /jgedc.2015214253

Zukhairina, Wujiati, Sari, S. R., Yennizar, Zulqarnain, & Pujiarto, P. (2020). Parenting practices of early childhood education teachers in developing early childhood communication and critical thinking skills: Case studies. *Journal of Critical Reviews*, *7*(7), 125–132. https://doi.org/10.31838/jcr.07.07.20