MODESTUM

EURASIA Journal of Mathematics, Science and Technology Education, 2024, 20(6), em2462
ISSN:1305-8223 (online)

OPEN ACCESS

Research Paper

https://doi.org/10.29333/ejmste/14661

Design and content validity of mathematics creative problem-solving ability instrument for junior high school students

Himmatul Ulya [1,2*], Sugiman [1], Raden Rosnawati [1]

[1] Universitas Negeri Yogyakarta, Sleman, INDONESIA
[2] Universitas Muria Kudus, Kudus, INDONESIA

Received 29 December 2023 ▪ Accepted 17 April 2024

Abstract

This research aims to design and validate the contents of the mathematical creative problem-solving (MCPS) instrument to measure creative problem-solving skills. This instrument consists of four aspects, each with related items representing latent variables for promoting students' MCPS. An instrument construct that contains aspects of divergent and convergent thinking is presented. The stimulus provided is real problems in everyday life (real-world situations) with tiered levels of structure of the observed learning outcome (SOLO) taxonomy. The question structure developed is in the form of near-transfer and far-transfer questions. The content validity process is presented by seven expert validators experienced in mathematics, mathematics education, and measurement. The level of agreement between expert assessments is determined using the Aiken formula. Aiken coefficient for all items is above the good threshold. The instrument developed has demonstrated strong content validity and is recommended for measuring MCPS skills of junior high school students.

Keywords: SOLO taxonomy, design & content validity, MPCS instrument, near transfer & far transfer

INTRODUCTION

The 21st century has brought changes in all areas of life. This change requires many new skills or changes existing skills that each individual has (Erdem, 2020). Reliable human resources need creative thinking and problem-solving strategies to complete non-routine and unpredictable work tasks (Leminger, 2020). Creativity and creative problem-solving (CPS) are competencies that today's youth need to improve themselves (Kasemsap, 2017; OECD, 2019; Ulya et al., 2024), so creativity and problem-solving are essential in the 21st century.

CPS is the ability to overcome problems to obtain new solutions (Arp, 2008; Karamustafaoglu & Pektas, 2023; Proctor, 2005). Other researchers explain that CPS is solving problems that are unstructured, complex, and related to the real world (Gizzi et al., 2022; Kasemsap, 2017). Other experts add that CPS is the ability to solve problems in a way that is unique and valued by someone (Tan & Maker, 2020; Teseo, 2019; Yuliani et al., 2019). The evaluation of CPS program for international student assessment 2012 measures students' capacity to engage in cognitive processing to understand and resolve problems, where the solution method is not immediately apparent (OECD, 2012).

When assessing students' abilities, teachers need to use quality instruments to ensure accurate measurements. In measuring students' CPS skills, teachers can use non-routine problems that require students to solve creative problems, including problems that require high-level thinking. Tests of CPS skills in mathematics must be considered with open-ended problems (Tan & Maker, 2020; van Hooijdonk et al., 2020). Non-routine questions can be story questions that have a tiered stimulus. Structure of the observed learning outcome (SOLO) taxonomy distinguishes between five levels indicated by the verb hierarchy according to cognitive complexity, namely pre-structural, unistructural, multi structural, and relational. The structure of questions, from simple to complex questions (involving several concepts), is expected to

✉ himmatul2pasca.2021@student.uny.ac.id **(*Correspondence)** ✉ sugiman@uny.ac.id ✉ rosnawati@uny.ac.id

Contribution to the literature

- The instrument for measuring mathematical creative problem-solving (MCPS) ability measures aspects of divergent and convergent thinking.
- The stimulus provided is in the form of real-world situation problems with tiered levels of SOLO taxonomy.
- The question structure developed is in the form of near-transfer and far-transfer questions. Content validity by experts in mathematics, mathematics education, and measurement is calculated using Aiken formula.

help students think in a more structured way (Biggs & Collis, 1982; Svensäter & Rohlin, 2022).

Finding solutions to non-routine problems requires developing techniques and challenges a person to think to understand the concepts involved (Johnny et al., 2017; Kozakli Ulger & Yazgan, 2021). Research results show that non-routine problem-solving experiences will benefit students, both for current learning and their future education and careers (Chirove et al., 2022; Klymchuk et al., 2020). In solving problems, a process called 'transfer' is required. The transfer process is the ability to extend what has been learned in one context to a new one (Cukurova et al., 2018; Daniels et al., 2022). Near-transfer occurs when students apply their knowledge and skills in situations and contexts like those in which learning occurs, whereas far-transfer occurs when a skill is performed in a context that is very different from the context in which it was learned (Orón & Lizasoain, 2023; Zhang et al., 2023). Thus, problem-solving questions can be categorized into two, namely near-transfer and far-transfer. Solving the near-transfer problem requires only low-level problem-solving skills, whereas far-transfer already involves high-level problem-solving skills.

Teachers have not optimized MCPS skills. This was proven in initial research with mathematics teachers in Indonesia. The interview results show that teachers have different perceptions regarding CPS. According to teachers, MCPS is solving non-routine problems, where there is a complex process involved in applying the knowledge they have, to find various solutions according to the results of their thinking. According to teachers, the characteristics of CPS questions are almost the same as those of other problem-solving but have a slight difference, namely if the questions contain open problems and various solutions emerge from the student's thinking. This aligns with research results that main difference between CPS and other problem-solving is divergent and convergent thinking (Isrok'atun, 2014; Tan & Maker, 2020; Yuliani et al., 2019).

The information indicates that some teachers already understand CPS, but teachers rarely or even never assess MCPS because it has some challenges. Teachers' challenges include students not being used to questions like this, and when they work on questions, they tend to answer according to the textbook. This means that students do not provide unique answers to the problems given. Apart from that, students are less motivated to work on questions, and many students have difficulty solving questions because they require several steps to complete. This study is in line with research results that show that students are less interested in solving mathematical problems because the length and complexity of the questions reduce student motivation (Novriania & Surya, 2017; Phonapichat et al., 2014).

The development of instruments to measure MCPS is still not much. The weakness of the instrument developed by Kim et al. (2003) is that the problems given are structured so that it is not suitable for measuring students' creativity in solving various problems. The research results recommend that the instruments developed be unstructured and unclear so students can respond with various solutions. Lee et al. (2003) complements the research of Kim et al. (2003) that MCPS instrument is developed to measure MCPS skills of talented students in mathematics and regular students. Learners with mathematics talent refer to learners with outstanding talents in understanding mathematics concepts, mathematics problem-solving ability, and high-order mathematics creativity. Experts explain that learners with mathematics talent are also learners with excellent problem-solving, metacognitive, mathematics creative thinking, and high-performance abilities in solving mathematics problems (Leikin et al., 2017; Rotigel & Fello, 2004; Yazgan-Sag, 2022). These learners could understand the complex mathematics ideas immediately; compete; and express their mathematics interests. Their learning achievements in the classroom and mathematics competitions are outstanding compared to their peers. On the other hand, regular learners to understand mathematics have average mathematics ability. They could understand the mathematics concept adequately, but they needed a longer time and put more effort into understanding any sustainable topics than the talented learners. Their problem-solving abilities are excellent although they have no mathematics talent.

The differences between talented and regular learners of mathematics include learning speed, comprehension, problem-solving, and competitive achievement (Lee et al., 2020; Ryan et al., 2022). Despite these matters, both learners have some similarities like

(1) mathematics interest found in both talented and regular learners about mathematics interest with different intensities,

(2) the development potency and the unrevealed potency for the regular learners, and

(3) the implementation of the available resources for both groups to gain the same access on textbooks, online tutorials, and supporting teachers to improve their mathematics abilities (Arthur et al., 2017).

Lee et al. (2003) suggest that tests of CPS skills in mathematics should be considered with open-ended problems. Cho and Hwang (2007) developed a MCPS skill test and state that the weakness of the instrument is that it requires students to produce as many responses as possible in a limited time to measure fluency so that students' ability to produce original solutions (originality) cannot be done optimally. Developing MCPS instruments will differentiate scoring between fluency and originality to overcome this weakness.

There is still minimal research in Indonesia to develop assessment models to measure MCPS. This is confirmed by research by Cho and Hwang (2007), which states that no standard MCPS skill test exists outside Korea. Several studies in Indonesia have used MCPS test to measure students' MCPS skills. However, these studies use a process or stage of CPS approach, i.e., consisting of mess finding, data finding, problem finding, idea finding, solution finding, and acceptance finding (Isrok'atun, 2014; Yuliani et al., 2019).

This research focuses on development of assessment as learning, emphasizing students' CPS skills. Student assessment results will be reported based on indicators of MCPS, SOLO taxonomy classification, and CPS transfer process classification. A brief description of MCPS assessment highlights some of the essential concepts implied. First, to generate practical creative ideas for solving mathematical problems, it seems very important that ideas generated through divergent thinking are also evaluated and selected through convergent thinking (Brophy, 2001; de Vink et al., 2022). Second, non-routine questions like story questions with tiered stimuli help students think in a more structured way (Jaiswal, 2019; Svensäter & Rohlin, 2022). Third, the transfer process extends what students have learned in one context to a new context (Byrnes, 1996; Rebello et al., 2007). Lastly, non-routine problem-solving experiences will benefit students, both for current learning and for their future education and careers (Gavaz et al., 2021; Klymchuk et al., 2020).

MATERIALS AND METHODS

This research was descriptive with knowledge formed by data, evidence, and logical considerations (Creswell & Poth, 2018). In practice, information was collected as data and empirical evidence to develop designs or products based on logical considerations. Next, an expert assessment was carried out to prove the validity of the content. When designing the instrument, apart from considering the criteria that emerge from the literature review, conditions that occur in the field were also considered, including the context and characteristics of students. Figure 1 shows a literature review diagram of the construction of the instrument being developed.

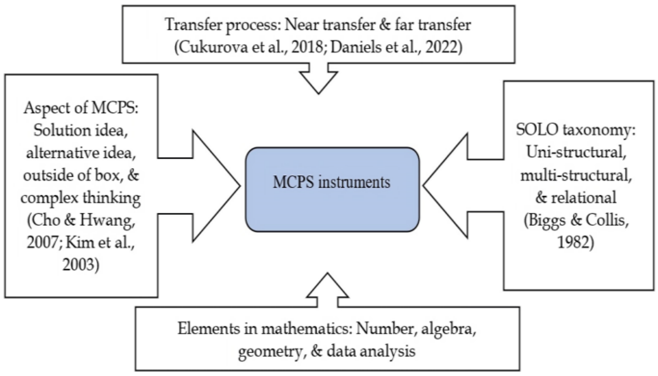

Figure 1. Theoretical framework of MCPS instrument (Source: Authors' own elaboration)

Content validity is the first that must be proven in test development to test the quality of the test items subjectively by experts (Allen & Yen, 1979). The process of assessing content validity in this research used the Delphi technique. This technique was carried out to survey and collect opinions from experts and aimed to obtain recommendations or suggestions for improving the instrument, which the instrument developer had to follow up.

The instrument validation process to prove the validity of the content in this research was carried out using expert judgment on the instrument. This research involved seven experts in mathematics, mathematics education, and measurement to assess 16 questions. Expert assessment was carried out to assess the suitability between essential indicators and question items. Experts could give scores in categories one to five (one is very irrelevant, two is irrelevant, three is less relevant, four is relevant, and five is very relevant) accompanied by an explanation of why the item does not meet the specified criteria or what must be added so that the criteria are fully met. In addition, experts could provide input related to material/substance, construction, and language. To calculate the agreement between the experts' assessments, the experts' quantitative assessments of the instrument items were analyzed using the validity formula (Aiken, 1985). Valid items have a calculated V value at least the same as the table V value. Aiken sets a lower limit for the index calculation results depending on the number of experts and the criteria used. This research involved seven experts using an assessment of five criteria. The expert assessment data was concluded using the Aiken formula

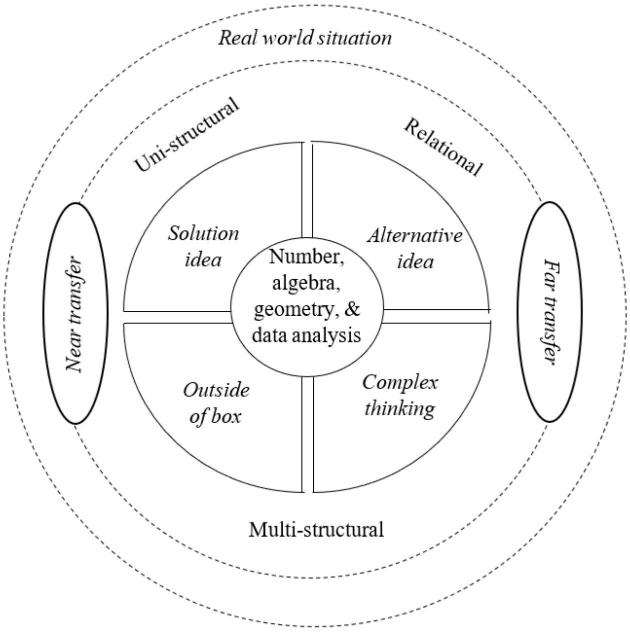

Figure 2. Developed instrument construct (Source: Authors' own elaboration)

$V = \frac{\sum s}{[n(c-1)]}$, where V is the agreement of the assessors, s=r-lo, lo is the lowest assessment score for each category, r is the score given by the assessor, c is the highest assessment score, and n is the number of assessors.

RESULTS

Instrument Construction Design

MCPS instrument IS developed to measure CPS skills of junior high school students. The question construct contains aspects of divergent and convergent thinking. Divergent thinking refers to aspects of solution ideas, alternative ideas, and outside of the box. The form of questions developed uses descriptive questions so that alternative answers given by students can be seen. The stimulus provided is in the form of real problems in everyday life (real word situations) with tiered levels of SOLO taxonomy, namely unistructural, multi-structural and relational. The question structure developed is in the form of near-transfer and far-transfer questions. Near-transfer is a non-routine problem studied in the learning process, while far-transfer is a non-routine problem using previous knowledge to solve new problems. **Figure 2** shows the constructed framework of MCPS instrument with the elements used.

The instrument construct developed facilitates students to provide various correct answers to a problem, pays attention to students' transfer abilities and the context of elements in mathematics, and applies various knowledge to solve problems. Instruments for measuring students' MCPS skills use non-routine questions with aspects of solution ideas, alternative ideas, outside-of-the-box, and complex thinking. Non-routine questions can be story questions that have a tiered stimulus. Thus, the levels in SOLO taxonomy are used in the construct of the instrument being developed. SOLO taxonomy levels used in the construct of this instrument include unistructural, multi-structural, and relational. The first level, the pre-structural level, is an irrelevant response so it is not used. Solving problems requires a transfer process. Transfer of problem-solving occurs when a student can use what he or she has learned to solve a problem different from the one presented during learning. The instrument construct developed uses two transfer categories: near transfer (close transfer) and far transfer (long transfer).

Assessment instruments for MCPS contain an outline of assessment instruments, test questions, and scoring guidelines. The assessment instrument outline contains aspects of CPS, SOLO taxonomy levels, transfer processes, elements, question indicators, and distribution of instrument items. The test questions contain items used to measure aspects of MCPS instruments. **Table 1** shows the instrument outline containing aspects, elements, question indicators, taxonomic levels, and detailed item numbers.

Instrument Content Validity

The items that have been prepared are then reviewed again before being assessed by experts. This expert assessment is intended to test the content validity of the instrument being developed. In this study, the instrument was validated by seven experts, consisting of experts in mathematics education and psychometrics. The expert validation results were analyzed using the Aiken formula, and the results are presented in **Table 2**.

Items 1, 2, 5, 7, 9, and 13 contain questions at the unistructural SOLO taxonomy level. Item 1 and item 2 measure the solution idea aspect. Item 1 contains material content on data and uncertainty, obtaining a coefficient value of 0.86, while item 2 contains the material content of numbers, obtaining a coefficient value of 0.86. Item 5 and item 7 measure alternative idea aspect. Item 5 contains material content on geometry, obtaining a coefficient value of 0.82, while item 7 contains material content on data and uncertainty, obtaining a coefficient value of 0.96. Item 9, measuring outside-of-the-box aspect with material data on data and uncertainty, obtaining a coefficient value of 0.96, while item 13 measures complex thinking aspect with material content on algebra, obtaining a coefficient value of 0.82.

Items 3, 6, 10, and 14 are questions with multi-structural levels in SOLO taxonomy. These four items measure aspects of solution ideas, alternative ideas, outside-of-the-box, and complex thinking. Item 3 and item 14 contain material content on geometry, obtaining coefficient values of 0.89 and 0.93, while items 6 and 10 contain material content on number, obtaining a coefficient value of 0.89 and 0.93.

Table 1. MCPS instrument outlines

Aspect	Element	Question indicator	SOLO taxonomy & transfer process	Item
Solution idea	Data & uncertainty	Organize data until it meets certain conditions	Unistructural–near transfer	1
	Number	Arrange consecutive integers that have a certain number of digits	Unistructural–far transfer	2
	Geometry	Create geometric shapes on specific patterns	Multi structural–near transfer	3
	Algebra	Create mathematical expressions from everyday life problems	Relational–far transfer	4
Alternative idea	Geometry	Estimate area of a flat shape through a sketch	Unistructural–near transfer	5
	Number	Arrange numbers on a certain pattern	Multi structural–far transfer	6
	Data & uncertainty	Interpret data from a graph	Unistructural–far transfer	7
	Algebra	Design algebraic forms so that system of linear equations can be solved	Relational–far transfer	8
Outside of box	Data & uncertainty	Arrange possible events from everyday problems	Unistructural–far transfer	9
	Number	Create procedures for estimating volume	Multi structural–near transfer	10
	Algebra	Solve comparison problems in everyday life	Relational–far transfer	11
	Geometry	Create a combination of geometric shapes by paying attention to their sizes	Relational–near transfer	12
Complex thinking	Algebra	Solve comparison problems in everyday life situations	Unistructural–near transfer	13
	Geometry	Determine ratio of areas of flat shapes	Multi structural–near transfer	14
	Data & uncertainty	Predict a datum that meets certain conditions	Relational–near transfer	15
	Number	Solve problems from a number pattern	Relational–far transfer	16

Table 2. Results of content validity using Aiken formula MCPS instrument outlines

No	Item	Rater							V
		1	2	3	4	5	6	7	
1	Item 1	4	4	5	5	4	5	4	0.86
2	Item 2	4	4	5	5	4	5	4	0.86
3	Item 3	4	4	5	5	4	5	5	0.89
4	Item 4	5	5	5	5	4	4	5	0.93
5	Item 5	4	5	4	5	4	4	4	0.82
6	Item 6	5	5	5	4	4	5	4	0.89
7	Item 7	5	5	5	5	4	5	5	0.96
8	Item 8	5	5	4	5	4	5	5	0.93
9	Item 9	5	5	5	5	4	5	5	0.96
10	Item 10	5	5	4	5	4	5	5	0.93
11	Item 11	5	4	5	5	4	5	5	0.93
12	Item 12	5	5	5	5	4	5	5	0.96
13	Item 13	4	5	4	4	4	4	5	0.82
14	Item 14	5	4	5	5	4	5	5	0.93
15	Item 15	5	4	5	5	4	5	5	0.93
16	Item 16	5	5	5	5	4	5	5	0.96
Average									0.91

The six items that use relational problem levels are items 4, 8, 11, 12, 15, and 16. Items 4, 8, and 11 contain material content on algebra, obtaining a coefficient value of 0.93. Item 12 contains material content on geometry, obtaining a coefficient value of 0.96. Item 15 contains material content on data and uncertainty, obtaining a coefficient value of 0.93. Item 16 contains material content on number, obtaining a coefficient value of 0.96.

The 16 items received expert assessment in the valid category because the V value calculated by the Aiken formula for each item has reached the V limit value, which can be seen from the Aiken table.

The table V value has a V value of more than 0.75. The results of proving content validity show that all items from MCPS assessed by experts are valid because all calculated V values are more than the V table. The Aiken coefficient value assessed by seven raters using five categories has exceeded 0.75. The average Aiken coefficient across items is 0.91.

DISCUSSION

The results of this research developed an MCPS instrument design to measure junior high school students' CPS skills. This assessment instrument is used as an assessment as a learning tool. Through this assessment design, students are given focused feedback for reflection on learning. This aligns with research results that teacher feedback on portfolios as assessment as learning instruments contributes to motivating and improving assessment as learning and taking responsibility for student learning and monitoring future directions (Fernandes et al., 2022). Students can reflect by thinking about the state of their learning and improving their learning. This is in accordance with the findings that assessment as learning is intended more as a learning tool, so that assessment is intended to provide students with learning experiences and can foster student self-monitoring of learning (Ghorbanpour et al., 2021; McMillan, 2018; Rosaini et al., 2019).

The novelty of MCPS instrument developed is that it is unstructured and unclear so that students can respond with various solutions. The instrument will be used to measure MCPS skills of talented students in mathematics and regular students. Therefore, since the differences between talented and regular learners of mathematics, the assessment of CPS ability must have a design to accommodate the differences and provide fair evaluation for all learners. The focus on creativity, problem-solving approach variety, fair evaluation criteria, and problem-solving process could relieve the impacts of various mathematics talent levels on the assessment results.

Questions are developed in the form of descriptions with open-ended problems to facilitate a diversity of answers from students. This follows research results that tests of CPS skills in mathematics must be considered open-ended problems (Munahefi et al., 2021; Tan & Maker, 2020; van Hooijdonk et al., 2020). MCPS instrument development differentiates scoring between solution ideas and outside-of-the-box aspects. To produce effective creative ideas for solving mathematical problems, the ideas generated need to involve divergent and convergent thinking (Suherman & Vidákovich, 2022; Yuliani et al., 2019). Non-routine questions, like story questions with tiered stimuli, help students think in a more structured way (Risnanosanti et al., 2019; Svensäter & Rohlin, 2022). The transfer process extends what students have learned in one context to a new context (Daniels et al., 2022; Zhang et al., 2023). Experience in solving non-routine problems will benefit students, both for current learning and their future education and careers (Kasemsap, 2017; Klymchuk et al., 2020).

MCPS instrument construct is developed as non-routine questions that can help students provide various correct answers to a problem. This finding is in accordance with research results and recommendations that tests of CPS skills in mathematics should be considered with open-ended problems (Maker, 2020; Tan & Maker, 2020). In line with this, research results show that open problem-solving assessment is effective for measuring students' mathematical creativity regarding the results (product) and process (Rahayuningsih et al., 2021; Ulya et al., 2019).

The aspects measured in MCPS instrument are aspects of divergent thinking, including solution ideas, alternative ideas, outside of the box, and aspects of convergent thinking, namely complex thinking. The aspects used in this instrument's construct follow research results that CPS can be measured through divergent and convergent thinking (Keles, 2022; Oraklibel et al., 2018; Yuliani et al., 2019). In contrast to this, the results of other research in measuring students' MCPS skills only use divergent aspects, including fluency (number of ideas produced), flexibility (number of different types or categories of ideas), originality (uniqueness of responses), elaboration (expansion of ideas through details addition), completeness, practicality, and general skills (Incebacak & Ersoy, 2018; Bicer et al., 2019; Chamberlin et al., 2022; Chen et al., 2016; Lee et al., 2003; Lee & Cha, 2016; Maker, 2020; Rubenstein et al., 2020; Sipayung et al., 2021; Tan & Maker, 2020; van Hooijdonk et al., 2020, 2022).

The design of MCPS instrument developed concerns students' transfer skills. Solving near-transfer problems requires only low-level problem-solving skills, whereas far-transfer already involves high-level problem-solving skills. This instrument construct, which contains near transfer and far transfer problems, facilitates students starting from non-routine problems with situations or contexts like learning to the highest cognitive treatments involving advanced abilities or declarative knowledge. This follows the findings that in solving near-transfer problems, students apply knowledge and skills in situations and contexts like the learning they have been through. In contrast, in solving far transfer, students need skills in a very different context from the context in which the skills were learned (Gilligan et al., 2020; Sala et al., 2019).

SOLO taxonomy levels used in the construct of this instrument include uni-structural, multi-structural, and relational. The research results show that the structure of questions, from simple to complex questions (involving several concepts), helps students think in a more structured way (Svensäter & Rohlin, 2022). Not all levels in SOLO taxonomy are used in creating the construct of this instrument. Lim et al. (2010) promote five levels in SOLO taxonomy: pre-structural, unistructural, multi-structural, and relational. Research findings show that the pre-structural level is intended to complete tasks and responses that require no or little working memory (Hook, 2014), in contrast to the construct of the instrument being developed. MCPS instrument developed aims to promote students' higher-level thinking skills.

According to previous research findings, MCPS instrument was developed only to identify talented children in mathematics (Kim et al., 2003). Lee et al. (2003) complement the research results of Kim et al. (2003) so that MCPS instrument developed can be used to measure MCPS skills of talented students in mathematics and regular students. In line with the results of this, Kasemsap (2017) advocates CPS skills to improve educational performance and achieve strategic goals in global education, so future research directions should broaden perspectives in promoting CPS skills in modern learning environments. As times change and technology develops rapidly, CPS skills are seen as an orientation or disposition towards mathematical activities that can be widely fostered in public schools. This is very relevant if these abilities are developed for all students in connection with the urgency of the 21st century.

After the construct design has been successfully developed and reviewed by the supervisor, the instrument is then content validated through expert judgment. This is in line with the fact that all items written according to the indicators will be processed for quality testing through a series of reviews by experts (Buitrago et al., 2023; del Pozo-Herce et al., 2023). The content validation process in this research was carried out by seven expert validators experienced in mathematics education and measurement. Content validation helps evaluate expert assessments of the instruments being developed (Aktas & Tabak, 2018; Amalina & Vidákovich, 2022; Kahveci et al., 2016). The findings regarding content validation align with research results that show that quantitative analysis of data through expert assessment helps reveal the strengths and weaknesses of the instrument. Adjustments are made based on the results, as well as qualitative observations and recommendations issued by experts, especially regarding the clarity and accuracy of items (Salfate et al., 2023). After evaluation and input by experts, several instrument items are adjusted to increase clarity so that instrument improvements are made.

This research's findings indicate that the Aiken value of each item has met the minimum threshold. The level of agreement between expert assessments is determined using the Aiken formula. The content validity coefficient value of seven raters with five answer categories is said to be good if a minimum validity coefficient of 0.75 is obtained (Aiken, 1985). The calculation results show that the overall content validity of each item obtained is more than 0.75, with an overall average of 0.91. Design and validation of content with expert assessment of instruments that display aspects and indicators that are appropriate to the abilities to be measured, namely MCPS skills. This is in line with the fact that content validity is a function of how well the dimensions and elements of a concept have been described (Almanasreh et al., 2019; Ghazali et al., 2018; Roebianto et al., 2023). This validation process aims to design an instrument to observe important aspects of a competency (Amalina & Vidákovich, 2023; Authié et al., 2023; Hrnjičić et al., 2022).

Based on the description, MCPS instrument developed has demonstrated strong content validity and is recommended for further empirical verification. The strength of this research is promoting assessment as learning in junior high school mathematics learning in the form of MCPS instrument to measure CPS skills. Instrument design originates from a dialectical process between theory and observation. Implementing this instrument requires creativity for teachers to score student answers. Based on this, the development of this instrument will continue to guide teachers in providing scoring, carrying out assessments, analyzing results, and providing feedback.

CONCLUSIONS

The findings of this research are design and evidence related to content as a source of instrument validation, which functions as a guide for assessing and analyzing students' MCPS skills. Design starts by examining field conditions and reviewing relevant theories so that a draft instrument can be developed. The initial evaluation is done through an internal trial process, where the supervisor reviews the aspects and indicators to be measured. Next, an expert assessment is carried out to prove validation regarding the content. Validation by experts and quantitative analysis substantially contributed to improving the instrument. All MCPS test items developed have good content validity in terms of the suitability of the items to the indicators. By providing content-related evidence as a source of validation, this research contributes to research in assessment as learning in mathematics learning. Furthermore, it is hoped that this MCPS instrument can contribute to educational policies related to improving mathematics learning through assessment as learning in assessing CPS skills, which are currently needed in the 21st century. At the same time, this research impacts improving mathematics learning due to the assessment design. Future research needs to be enriched by providing other sources of validity, such as construct validity. This research focuses on providing feedback as a reflection of learning by promoting CPS skills. This research is only focused on providing content-related evidence. In particular, the research provides evidence that the aspects, indicators, and items in the instrument are sufficient, coherent, relevant, and straightforward.

Author contributions: HU: conceptualization, design, data collection, analysis, writing, & funding support; **S:** conceptualization, providing technical support, & supervision; & **RR:** conceptualization, design, providing technical support, & supervision. All authors have agreed with the results and conclusions.

Funding: This study was supported by Ministry of Education, Culture, Research, and Technology, Center for Higher Education Funding, & Indonesia Endowment Funds for Education.

Ethical statement: The authors stated that the study was approved by the ethics committee at Directorate of Research and Community Service Universitas Negeri Yogyakarta, Indonesia (Approval number: B/48/UN34.9/KP.06.07/2023)

Declaration of interest: No conflict of interest is declared by the authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

Aiken, L. R. (1985). Three coefficients to analyze the reliability and validity of rating. *Educational and Psychological Measurement, 45*, 131-142. https://doi.org/10.1177/0013164485451012

Aktas, M. C., & Tabak, S. (2018). Turkish adaptation of math and me survey: A validity and reliability study. *European Journal of Educational Research, 7*(3), 707-714. https://doi.org/10.12973/eu-jer.7.3.707

Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory*. Brooks/Cole Publishing Company.

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

Amalina, I. K., & Vidákovich, T. (2022). An integrated STEM-based mathematical problem-solving test: Developing and reporting psychometric evidence. *Journal on Mathematics Education, 13*(4), 587-604. https://doi.org/10.22342/jme.v13i4.pp587-604

Amalina, I. K., & Vidákovich, T. (2023). Assessment of domain-specific prior knowledge: A development and validation of mathematical problem-solving test. *International Journal of Evaluation and Research in Education, 12*(1), 468-476. https://doi.org/10.11591/ijere.v12i1.23831

Arp, R. (2008). Scenario visualization: An evolutionary account of creative problem solving. *American Journal of Human Biology, 21*(1). https://doi.org/10.1002/ajhb.20848

Arthur, C., Badertscher, E., Goldenberg, E. P., Moeller, B., McLeod, M., Nikula, J., & Reed, K. (2017). *Strategies to improve all students' mathematics learning and achievement*. Education Development Center, Inc.

Authié, C. N., Poujade, M., Talebi, A., Defer, A., Zenouda, A., Coen, C., Mohand-Said, S., Chaumet-Riffaud, P., Audo, I., & Sahel, J.-A. (2023). Development and validation of a novel mobility test for rod-cone dystrophies, from reality to virtual reality. *American Journal of Ophthalmology, 258*, 43-54. https://doi.org/10.1016/j.ajo.2023.06.028

Bicer, A., Lee, Y., Capraro, R. M., Capraro, M. M., Barroso, L. R., & Rugh, M. (2019). Examining the effects of STEM PBL on students' divergent thinking attitudes related to creative problem solving. In *Proceedings of the 2019 IEEE Frontiers in Education Conference* (pp. 1-6). IEEE. https://doi.org/10.1109/FIE43999.2019.9028431

Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy (structure of the observed learning outcome)*. Academic Press. https://doi.org/10.1177/000494418302700311

Brophy, D. R. (2001). Comparing the attributes, activities, and performance of divergent, convergent, and combination thinkers. *Creativity Research Journal, 13*(3-4), 439-455. https://doi.org/10.1207/S15326934CRJ1334_20

Buitrago, R., Salinas, J., & Boude, O. (2023). Validation of a model for the formalization of personal learning pathways through expert judgment. *Journal of Higher Education Theory and Practice, 23*(12), 224-241. https://doi.org/10.33423/jhetp.v23i12.6283

Byrnes, J. P. (1996). *Cognitive development and learning in instructional contexts*. Allyn and Bacon.

Chamberlin, S. A., Liljedahl, P., & Savić, M. (2022). *Mathematical creativity: A developmental perspective*. Springer. https://doi.org/10.1007/978-3-031-14474-5

Chen, Y. W., Chang, W. H., & Kuo, C. C. (2016). A comparative study of the divergent problem solving abilities of mathematically and scientifically talented students and nongifted students. *Thinking Skills and Creativity, 22*, 247-255. https://doi.org/10.1016/j.tsc.2016.10.009

Chirove, M., Mogari, D., & Ugorji, O. (2022). Students' mathematics-related belief systems and their strategies for solving non-routine mathematical problems. *Waikato Journal of Education, 27*(3), 101-121. https://doi.org/10.15663/wje.v27i3.822

Cho, S., & Hwang, D.-J. (2007). Math creative problem solving ability test for identification of the mathematically gifted middle school students. *Journal of Gifted/Talented Education, 17*(1), 1-26. https://doi.org/10.9722/JGTE.2021.31.1.1

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry & research design: Choosing among five approaches*. SAGE. https://doi.org/10.13187/rjs.2017.1.30

Cukurova, M., Bennett, J., & Abrahams, I. (2018). Students' knowledge acquisition and ability to apply knowledge into different science contexts in two different independent learning settings. *Research in Science and Technological Education, 36*(1), 17-34. https://doi.org/10.1080/02635143.2017.1336709

Daniels, J. S., Moreau, D., & Macnamara, B. N. (2022). Learning and transfer in problem solving progressions. *Journal of Intelligence, 10*(4), 85. https://doi.org/10.3390/jintelligence10040085

de Vink, I. C., Willemsen, R. H., Lazonder, A. W., & Kroesbergen, E. H. (2022). Creativity in mathematics performance: The role of divergent and convergent thinking. *British Journal of Educational Psychology, 92*(2), 484-501. https://doi.org/10.1111/bjep.12459

del Pozo-Herce, P., Martínez-Sabater, A., Chover-Sierra, E., Gea-Caballero, V., Satústegui-Dordá, P. J., Saus-Ortega, C., Tejada-Garrido, C. I., Sánchez-Barba, M., Pérez, J., Juárez-Vela, R., Santolalla-Arnedo, I., & Baca-García, E. (2023). Application of the Delphi method for content validity analysis of a questionnaire to determine the risk factors of the Chemsex. *Healthcare, 11*(21), 2905. https://doi.org/10.3390/healthcare11212905

Erdem, C. (2020). Introduction to 21st century skills and education. In *21st century skills and education*. Cambridge Scholars Publishing.

Espinoza Salfate, L., Guerrero, G., Barbé Farré, J., & Márquez Salinas, F. (2023). Design and validation of a classroom observation instrument to evaluate the quality of mathematical activity from a gender perspective. *Education Sciences, 13*(3), 266. https://doi.org/10.3390/educsci13030266

Fernandes, S., Abelha, M., & Albuquerque, A. S. (2022). The student journey in PBL: Using individual portfolios to promote self-reflection and assessment as learning. *International Symposium on Project Approaches in Engineering Education, 12*, 207-211. https://doi.org/10.5281/zenodo.7058095

Gavaz, H. O., Yazgan, Y., & Arslan, C. (2021). Non-routine problem solving and strategy flexibility: A quasi-experimental study. *Journal of Pedagogical Research, 5*(3), 40-54. https://doi.org/10.33902/jpr.2021370581

Ghazali, N., Sahari Nordin, M., Hashim, S., & Hussein, S. (2018). Measuring content validity: Students self-efficacy and meaningful learning in massive open online course (MOOC) scale. In *Proceedings of the 3rd International Conferences on Education in Muslim Society* (pp. 128-133). https://doi.org/10.2991/icems-17.2018.25

Ghorbanpour, E., Abbasian, G. R., & Mohseni, A. (2021). Assessment alternatives in developing L2 listening ability: Assessment FOR, OF, AS learning or integration? Assessment *x* approach. *International Journal of Language Testing, 11*(1), 36-57.

Gilligan, K. A., Thomas, M. S. C., & Farran, E. K. (2020). First demonstration of effective spatial training for near transfer to spatial performance and far transfer to a range of mathematics skills at 8 years. *Developmental Science, 23*(4), e12909. https://doi.org/10.1111/desc.12909

Gizzi, E., Nair, L., Chernova, S., & Sinapov, J. (2022). Creative problem solving in artificially intelligent agents: A survey and framework. *International Joint Conference on Artificial Intelligence, 75*(2022), 857-911. https://doi.org/10.24963/ijcai.2023/777

Hook, P. (2014). *About SOLO taxonomy*. https://leadinglearnerdotme.files.wordpress.com/2014/09/about-solo-taxonomy-by-pam-hook-pdf.pdf

Hrnjičić, A., Alihodžić, A., Čunjalo, F., & Hamzić, D. K. (2022). Development of an item bank for measuring students' conceptual understanding of real functions. *European Journal of Science and Mathematics Education, 10*(4), 455-470. https://doi.org/10.30935/scimath/12222

Incebacak, B., & Ersoy, E. (2018). Ortaokul öğrencilerinin yaratıcı problem çözme becerileri [Creative problem solving skills of middle school students]. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi* [*Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education*], *12*(1), 1-24. https://doi.org/10.17522/balikesirnef.437352

Isrok'atun. (2014). *Situation-based learning untuk meningkatkan kemampuan creative problem solving matematis siswa* [*Situation-based learning to improve students' creative mathematical problem solving abilities*]. Universitas Pendidikan Indonesia.

Jaiswal, P. (2019). Using constructive alignment to foster teaching learning processes. *English Language Teaching, 12*(6), 10. https://doi.org/10.5539/elt.v12n6p10

Johnny, J., Abdullah, A. H., Abu, M. S., Mokhtar, M., & Atan, N. A. (2017). Difficulties in reasoning among high achievers when doing problem solving in mathematics. *Man in India, 97*(12), 61-70.

Kahveci, M., Kahveci, A., Mansour, N., & Alarfaj, M. M. (2016). Construct validity and reliability measures of scores from the science teachers' pedagogical discontentment (STPD) scale. *EURASIA Journal of Mathematics, Science and Technology Education, 12*(3), 549-558. https://doi.org/10.12973/eurasia.2015.1417a

Karamustafaoğlu, O., & Pektaş, H. M. (2023). Developing students' creative problem solving skills with inquiry-based STEM activity in an out-of-school learning environment. *Education and Information Technologies, 28*, 7651-7669. https://doi.org/10.1007/s10639-022-11496-5

Kasemsap, K. (2017). Advocating problem-based learning and creative problem-solving skills in global education. In C. Zhou (Ed.), *Handbook of research on creative problem-solving skill development in higher education* (pp. 351-377). IGI Global. https://doi.org/10.4018/978-1-5225-0643-0.ch016

Keles, T. (2022). Investigation of high school students' creative problem-solving attributes. *Journal of Pedagogical Research, 6*(4), 66-83. https://doi.org/10.33902/JPR.202215433

Kim, H., Cho, S., & Ahn, D. (2003). Development of mathematical creative problem solving ability test for identification of the gifted in math. *Gifted Education International, 18*, 164-174. https://doi.org/10.1177/026142940301800206

Klymchuk, S., Thomas, M., Alexander, J., Evans, T., Gulyaev, S., Liu, W., Novak, J., Murphy, P., Stephens, J., & Zaliwski, A. (2020). Investigating the impact of non-routine problem-solving on creativity, engagement, and intuition of STEM tertiary students. *Teaching and Learning Research Initiative*. https://eric.ed.gov/?id=ED608044

Kozakli Ulger, T., & Yazgan, Y. (2021). Non-routine problem-posing skills of prospective mathematics teachers. *Eurasian Journal of Educational Research, 21*(94), 147-168. https://doi.org/10.14689/ejer.2021.94.7

Lee, K. H., Kim, Y., & Lim, W. (2020). Risks of aiming to kill two birds with one stone: The effect of mathematically gifted and talented students in the dual realities of special schooling. *Mathematical Thinking and Learning, 23*(4), 271-290. https://doi.org/10.1080/10986065.2020.1784696

Lee, K. S., Hwang, D., & Seo, J. J. (2003). A development of the test for mathematical creative problem solving ability. *Journal of the Korea Society of Mathematical Education, Series D: Research in Mathematical Education, 7*(3), 163-189.

Lee, M., & Cha, D. (2016). A comparison of generalizability theory and many facet Rasch measurement in an analysis of mathematics creative problem solving test. *Journal of Curriculum and Evaluation, 19*(2), 251-279. https://doi.org/10.29221/jce.2016.19.2.251

Leikin, R., Leikin, M., Paz-Baruch, N., Waisman, I., & Lev, M. (2017). On the four types of characteristics of super mathematically gifted students. *High Ability Studies, 28*(1), 107-125. https://doi.org/10.1080/13598139.2017.1305330

Leminger, E. A. (2020). *Human resource professional's attitudes towards creativity* [Master's project, State University of New York].

Lim, H. L., Wun, T. Y., & Idris, N. (2010). Superitem test: As an alternative assessment tool. *International Journal for Mathematics Teaching and Learning*.

Maker, C. J. (2020). Identifying exceptional talent in science, technology, engineering, and mathematics: Increasing diversity and assessing creative problem-solving. *Journal of Advanced Academics, 31*(3), 161-210. https://doi.org/10.1177/1932202X20918203

McMillan, J. H. (2018). *Classroom assessment: Principles and practice that enhance student learning and motivation*. Pearson.

Munahefi, D. N., Khoirunnisa, K., Dwijanto, Mulyono, Fariz, R., & Noverianto, B. (2021). Development of open ended project based learning model assisted by GeoGebra for mathematical creative thinking ability. In *Proceedings of the 8th International Conference on Mathematics, Science and Education*. https://doi.org/10.1063/5.0126736

Novriania, M. R., & Surya, E. (2017). Analysis of student difficulties in mathematics problem solving ability at MTs SWASTA IRA Medan. *International Journal of Sciences: Basic and Applied Research, 33*(3), 63-75.

OECD. (2012). *PISA 2012 creative problem solving*. https://www.oecd.org/pisa/innovation/creative-problem-solving/

OECD. (2019). *PISA 2021 creative thinking framework*. OECD Publishing.

Oraklibel, R. D., Ulkebas, S. D., & Oygur, I. (2018). Creative problem-solving assessment and product design education. In *Proceedings of the International Conference on Engineering and Product Design Education*.

Orón, J. V., & Lizasoain, I. (2023). Achieving transfer from mathematics learning. *Education Sciences, 13*(2), 161. https://doi.org/10.3390/educsci13020161

Phonapichat, P., Wongwanich, S., & Sujiva, S. (2014). An analysis of elementary school students' difficulties in mathematical problem solving. *Procedia-Social and Behavioral Sciences, 116*, 3169-3174. https://doi.org/10.1016/j.sbspro.2014.01.728

Proctor, T. (2005). *Creative problem solving for managers–Developing skills for decision making and innovation*. Routledge. https://doi.org/10.1177/1350507601321014

Rahayuningsih, S., Sirajuddin, S., & Ikram, M. (2021). Using open-ended problem-solving tests to identify students' mathematical creative thinking ability. *Participatory Educational Research, 8*(3), 285-299. https://doi.org/10.17275/per.21.66.8.3

Rebello, N. S., Cui, L., Bennett, A. G., Zollman, D. A., & Ozimek, D. J. (2007). Transfer of learning in problem solving in the context of mathematics and physics. In D. H. Jonassen (Ed.), *Learning to solve complex scientific problems*. Routledge. https://doi.org/10.4324/9781315091938-10

Risnanosanti, Susyla, D., & Syofiana, M. (2019). Developing students critical thinking ability through lesson study. *Journal of Physics: Conference Series, 1320*, 012005. https://doi.org/10.1088/1742-6596/1320/1/012005

Roebianto, A., Savitri, S. I., Aulia, I., Suciyana, A., & Mubarokah, L. (2023). Content validity: Definition and procedure of content validation in psychological research. *TPM-Testing, Psychometrics, Methodology in Applied Psychology, 30*(1), 5-18. https://doi.org/10.4473/TPM30.1.1

Rosaini, R., Budiyono, B., & Pratiwi, H. (2019). Mathematics teacher supporting higher order thinking skill of students through assessment as learning in instructional model. *Journal of Physics: Conference Series, 1157*(3), 032076. https://doi.org/10.1088/1742-6596/1157/3/032076

Rotigel, J. V., & Fello, S. (2004). Mathematically gifted students: How can we meet their needs? *Gifted Child Today, 27*(4), 46-51. https://doi.org/10.4219/gct-2004-150

Rubenstein, L. D. V., Callan, G. L., Speirs Neumeister, K., Ridgley, L. M., & Hernández Finch, M. (2020). How problem identification strategies influence creativity outcomes. *Contemporary Educational Psychology, 60*, 101840. https://doi.org/10.1016/j.cedpsych.2020.101840

Ryan, V., Fitzmaurice, O., & O'donoghue, J. (2022). Student interest and engagement in mathematics after the first year of secondary education. *European Journal of Science and Mathematics Education, 10*(4), 436-454. https://doi.org/10.30935/SCIMATH/12180

Sala, G., Deniz Aksayli, N., Semir Tatlidil, K., Tatsumi, T., Gondo, Y., & Gobet, F. (2019). Near and far transfer in cognitive training: A second-order meta-analysis. *Collabra: Psychology, 5*(1), 18. https://doi.org/10.1525/collabra.203

Sipayung, T. N., Imelda, Siswono, T. Y. E., & Masriyah. (2021). The preliminary study of students' creative problem-solving ability. *Journal of Physics: Conference Series, 1836*, 012052. https://doi.org/10.1088/1742-6596/1836/1/012052

Suherman, S., & Vidákovich, T. (2022). Assessment of mathematical creative thinking: A systematic review. *Thinking Skills and Creativity, 44*, 101019. https://doi.org/10.1016/j.tsc.2022.101019

Svensäter, G., & Rohlin, M. (2022). Assessment model blending formative and summative assessments using the SOLO taxonomy. *European Journal of Dental Education, 27*(1), 149-157. https://doi.org/10.1111/eje.12787

Tan, S., & Maker, C. J. (2020). Assessing creative problem solving ability in mathematics: The discover mathematics assessment. *Gifted and Talented International, 35*(1), 58-71. https://doi.org/10.1080/15332276.2020.1793702

Teseo, R. F. (2019). *Analyses of attribute patterns of mathematical creative problem-solving ability in 6th grade students*. St. John's University.

Ulya, H., Rahayu, R., & Riyono, A. (2019). Integration of products assessment in mind mapping learning to enhance mathematical communication. *Journal of Physics: Conference Series, 1175*, 012142. https://doi.org/10.1088/1742-6596/1175/1/012142

Ulya, H., Sugiman, Rosnawati, R., & Retnawati, H. (2024). Technology-based learning interventions on mathematical problem-solving: A meta-analysis of research in Indonesia. *International Journal of Evaluation and Research in Education, 13*(1), 292-301. https://doi.org/10.11591/ijere.v13i1.26380

van Hooijdonk, M., Mainhard, T., Kroesbergen, E. H., & van Tartwijk, J. (2020). Creative problem solving in primary education: Exploring the role of fact finding, problem finding, and solution finding across tasks. *Thinking Skills and Creativity, 37*), 100665. https://doi.org/10.1016/j.tsc.2020.100665

van Hooijdonk, M., Mainhard, T., Kroesbergen, E. H., & Van Tartwijk, J. (2022). Examining the assessment of creativity with generalizability theory: An analysis of creative problem solving assessment tasks. *Thinking Skills and Creativity, 43*, 100994. https://doi.org/10.1016/j.tsc.2021.100994

Yazgan-Sağ, G. (2022). Views on mathematical giftedness and characteristics of mathematically gifted students-the case of prospective primary mathematics teachers. *Mathematics Teaching-Research Journal, 14*(5), 128-140.

Yuliani, A., Kusumah, Y. S., & Sumarmo, U. (2019). Mathematical creative problem solving ability and self-efficacy: A survey with eight grade students. *Journal of Physics: Conference Series, 1157*(3), 032097. https://doi.org/10.1088/1742-6596/1157/3/032097

Zhang, Q., Zhang, Q., Zhang, X., & Gao, C. (2023). Effects of rule variant reasoning in far transfer problem solving. *Acta Psychologica Sinica, 55*(1), 117-128. https://doi.org/10.3724/SP.J.1041.2023.00117