

Measuring of Metacognition in Problem-Posing: Standardization of the Metacognition in Problem-Posing Scale

Mesut ÖZTÜRK¹, İsmail SARIKAYA², Mustafa ALBAYRAK³

¹Department of Mathematics and Science Education, Bayburt University Faculty of Education, Bayburt, Türkiye

²Department of Primary Education, Bayburt University Faculty of Education, Bayburt, Türkiye

³Department of Mathematics and Science Education, Bayburt University Faculty of Education, Bayburt, Türkiye

What is already known on this topic?

- Metacognition has been widely studied in mathematics education, mainly in the context of problem-solving.
- Research shows that problem-posing also involves key metacognitive processes such as planning, monitoring, and evaluation.
- However, there is a lack of valid and reliable quantitative instruments specifically designed to measure metacognition in problem-posing (MPP).

What does this study add on this topic?

- This study develops and validates a new measurement instrument (the MPP Scale) to assess middle school students' MPP.
- The scale captures three core factors of metacognition: planning, monitoring, and evaluation.
- The MPP Scale provides researchers and educators with a practical and standardized tool for quantitatively examining metacognitive engagement in problem-posing activities.

Abstract

While research on metacognition in mathematics education has grown recently, it has primarily focused on problem-solving, with little attention given to problem-posing. However, problem-posing also involves complex metacognitive processes that warrant deeper investigation. This study addresses this gap by developing and validating a new instrument to assess metacognition in problem-posing among middle school students. The metacognition in problem-posing Scale is an eight-item measure encompassing three factors: planning, monitoring, and evaluation. The study was conducted using a scale-development design within a survey research framework, and data were collected through a two-stage simple random sampling. Psychometric analyses confirmed the scale's reliability and its construct validity. The metacognition in problem-posing Scale offers a promising tool for exploring students' metacognitive engagement in problem-posing tasks and has potential applications in both research and classroom settings.

Keywords: Evaluation, metacognition, monitoring, planning, problem-posing, scale

Introduction

Metacognition has become a central focus in contemporary psychological research and is gaining increasing attention in mathematics education (Schneider & Artelt, 2010; Baummanns & Rott, 2022a). While its role has been extensively examined in the context of problem-solving (Artzt & Armour-Thomas, 1992; Schoenfeld, 1992; Cozza & Oreshkina, 2013), relatively little research has explored how metacognitive processes operate during problem-posing (Baumanns & Rott, 2022a, 2022b).

Problem-posing involves planning, monitoring, and evaluating one's thinking while generating mathematical tasks. These are inherently metacognitive processes, but they have not been adequately studied. Most research in this area has been qualitative, with few experimental or correlational studies (Baumanns & Rott, 2022b). One likely reason for this lack of quantitative work is the absence of a reliable and valid instrument specifically designed to assess metacognition in the context of problem-posing.

To address this gap, the present study introduces the metacognition in problem-posing (MPP) Scale—a new tool developed to measure students' metacognitive engagement during mathematical problem-posing. The MPP Scale captures three key factors of metacognition: planning, monitoring, and evaluation. This scale provides a foundation for advancing quantitative research on MPP by offering standardized measurements. This contributes to a more complete integration of metacognitive theory into mathematics education.

Theoretical Framework

Problem-Posing

Different definitions of problem-posing have been proposed in the literature, which generally fall into two conceptual categories: The first conceptualization views problem-posing as an extension or component of problem-solving (Baumanns & Rott, 2022c). The second view treats problem-posing as a distinct activity focused on generating new problems (Baumanns & Rott, 2021a, 2022c). While the emphases differ, the two perspectives are not mutually exclusive and can be integrated. Accordingly, this study adopts a broad definition that incorporates both conceptualizations of problem-posing.

Problem-posing has also been classified into three types: free, semi-structured, and structured (Stoyanova & Ellerton, 1996). In free problem-posing, students are given no constraints and are asked to generate

Corresponding author: Mesut Öztürk
E-mail: mesutozturk@live.com

Cite this article as: Öztürk, M., Sarikaya, İ., & Albayrak, M. (2026). Measuring of metacognition in problem-posing: standardization of the metacognition in problem-posing scale. *HAYEF: Journal of Education*, 23, 0001, doi: 10.5152/hayef.2026.26001.



problems based on open-ended prompts or real-life situations (Işık & Kar, 2012). Students must independently select a domain, formulate meaningful problem content, and relate it to authentic contexts (Baumanns & Rott, 2022a). In semi-structured problem-posing, students are provided with a situation that includes predefined constraints and are asked to formulate problems within those boundaries (Işık & Kar, 2012). Within these constraints, students are expected to generate coherent and meaningful solutions. In structured problem-posing, students are presented with a complete problem and asked to modify or create new problems based on it (Baumanns & Rott, 2022b; Işık & Kar, 2012). In this context, students may reflect on the structure and intent of a given problem and transform it into new variations. Together, these different types of problem-posing illustrate how metacognition is embedded in the design and construction of mathematical tasks, as the ability to pose meaningful problems requires students to plan, monitor, and evaluate their thinking, regardless of structure.

Metacognition

Metacognition involves monitoring, regulating, and controlling one's cognitive processes while performing a task (Swanson, 1990). Metacognition has two components: knowledge of cognition and regulation of cognition. Knowledge of cognition refers to an individual's awareness of their cognitive processes. Cognition regulation refers to the way individuals regulate their cognitive processes. Cognition regulation is divided into three components: planning, monitoring, and evaluation. (i) Planning includes the skills to be performed before starting a task (Öztürk & Ada, 2022). The belief that a student can pose a problem (Baumanns & Rott, 2022b) and drawing a figure to understand the problem are skills related to planning. (ii) Monitoring includes skills for performing a task (Baumanns & Rott, 2022b). Some skills exhibited by students while performing problem-posing tasks are metacognitive. For example, students can progress by checking their writings in the problem-posing process (Baumanns & Rott, 2022b), by checking the goals they set, or by trusting that they can pose problems in this process. (iii) Evaluation includes skills to evaluate the problem-posing process or to check the established problem after completing the problem-posing process (Baumanns & Rott, 2022b). For example, after posing the problem, the student can check the problem by reading it (Baumanns & Rott, 2022b), distinguish between necessary and unnecessary information in the problem (Pelczar & Gamboa, 2009), and check the correctness of their solution (Baumanns, & Rott, 2022b).

Metacognition in Problem-Posing

While some research on problem-posing sees problem-posing as a part of problem-solving (Kontorovich, Koichu, Leikin, & Berman, 2012), some state that problem-posing involves a different process than problem-solving (Baumanns & Rott, 2022b). Studies that consider problem-solving as part of problem-posing state that MPP could be considered within the problem-solving framework. Studies that consider problem-posing as different from problem-solving have developed new theoretical frameworks for MPP. For example, Baumanns and Rott (2022b) aimed to develop a theoretical framework for MPP. They evaluated metacognition in the problem-posing process in the categories of "planning," "monitoring and controlling," and "evaluation."

The Present Study

The literature on mathematics education shows that metacognition has been studied intensively in recent years. Although some studies have examined aspects of MPP (Kontorovich et al., 2012), further research is still needed (Baumanns & Rott, 2021b). To enable generalization, quantitative research examining MPP is required. Although scales exist for general metacognition (dos Santos Kawata et al., 2021; Semerari et al., 2003; Urban & Urban, 2023) and mathematical metacognition (Yang, 2012), no scale specifically measures metacognition

in the problem-posing process itself. Therefore, this study aims to develop a measurement tool to assess middle school students' MPP.

Methods

Research Model

This study was designed as a scale-development study within a survey research framework. This study followed the scale-development stages because it aimed to develop a valid and reliable MPP Scale for middle school students. To determine the factor structure of the scale, principal component analysis (PCA) was first performed using data collected from a group of students. The compatibility of the structure obtained with the data collected from a different group was tested using confirmatory factor analysis (CFA).

Sample

The sample comprised fifth- through eighth-grade students attending public middle schools in two cities in northern Türkiye. A multi-stage sampling procedure was employed. Initially, a sampling frame of public middle schools in selected cities was established. From this frame, four schools per city were randomly selected for the study. Subsequently, one grade level (fifth, sixth, seventh, or eighth grade) was randomly selected within each school. Data were collected from students in the selected classes on a voluntary basis during regular class hours in the classroom setting. Necessary permissions were obtained from the school administration.

The first phase of this study was conducted with the participation of 316 middle school students. Of the participants, 69 (21.9%) were fifth-grade students, 73 (23.1%) were sixth-grade students, 84 (26.5%) were seventh-grade students, and 90 (28.5%) were eighth-grade students. Of the participants, 176 (55.7%) were girls and 140 (44.3%) were boys.

The second phase of the study (CFA) was conducted with the participation of 223 middle school students. Sixty (26.9%) of the participants were 5th-grade students, 63 (28.3%) were 6th-grade students, 54 (24.2%) were 7th-grade students, and 46 (20.6%) were 8th-grade students. Of the participants, 115 (51.6%) were girls and 108 (48.4%) were boys.

Preparation of the Trial Form of the Scale and the Process

The development of the measurement of the factors was guided by the Standards for Educational and Psychological Testing (AERA, APA, & NCME, 2014). Before starting the study, the researchers obtained the necessary application and ethical permissions. The study reported in this paper was conducted with the approval of the Ethics Committee of Bayburt University (Approval no: E-51694156-050.99-36868; Date: October 14, 2021). The researchers then created an item pool for the scale. While creating the item pool, the researchers examined the relevant literature. The researchers created a draft trial form comprising 20 items. Two groups of experts were consulted for suggestions and improvements of the items (AERA, APA, & NCME, 2014). The draft form was presented to five faculty members, one expert in assessment and evaluation, one in Turkish education, and three experts in mathematics education (working on problem-posing). This process aimed to obtain expert feedback on the form's language, face validity, and content validity. First, the researchers directed the form to the language expert and checked the intelligibility of the items. In line with the opinions of the language expert, the researchers made language changes to three items. Subsequently, the researchers presented the draft to the other four experts. Experts were asked to provide their opinions as "Appropriate/Not Appropriate" for each item; revision requests were received for unsuitable items. In addition, an area was reserved for the scale items that the experts wanted to add at the bottom of the form. The consensus rate for determining suitable items was set at 80%. Based on the expert feedback, eight items that did not meet this rate

were removed from the form. In addition, five new items were added to the form based on the suggestions of two experts. The new draft form, consisting of 17 items, was sent back to the linguist and then presented to the other four experts. In line with the feedback from the experts, an item that was thought to be similar was removed from the form, and minor language changes were made in two articles. Finally, a trial form consisting of 16 items was used for the factor analysis. A five-point Likert-type scale was used: Never (1), Rarely (2), Sometimes (3), Often (4), and Always (5).

A five-point rating response category was used for each question. This form was applied to 20 middle school students, 10 in the fifth grade and 10 in the eighth grade, and it was evaluated according to the students' views on intelligibility. In particular, the reason the fifth graders were selected in the pre-test was to determine whether the form was understandable by the students at the first level of middle school. After this stage, the draft form was administered to 316 middle school students, and PCA was conducted. Subsequently, the suitability of the structure obtained from the data collected from 223 middle school students was tested using CFA. The researchers collected the data face-to-face during the lesson, as deemed appropriate by the school administrators and teachers. Before the scales were filled, brief information was given to the students about the importance of data reliability, and the consent form was completed. The filling time of the 16-item scale ranged from 4 to 10 minutes.

Data Analysis

Before data analysis, missing, erroneous values, and extreme data were examined. Data were collected from 318 students for the first phase. However, two forms, the majority of which were found to be left blank, were excluded from the study; therefore, the analysis was carried out with the data obtained from 316 students. Missing data were filled using the Full Information Maximum Likelihood (FIML) method (Enders & Bandalos, 2001). Histograms, box-and-whisker plots, and stem-and-leaf plots were examined on an item basis, and the kurtosis and skewness values were checked. The reference ranges for kurtosis and skewness values were accepted as +1 and -1, respectively. The analysis concluded that the dataset had a univariate normal distribution.

To evaluate the data's suitability, R-matrix determinant coefficients, item correlation matrix, and anti-image values were examined. To determine whether there was a multicollinearity problem, the R-matrix determinant coefficients were examined, and it was observed that the relevant values were higher than 0.00001. It is suggested that in this case, the R-matrix determinant coefficients should be examined, and if the relevant values are greater than 0.00001, there is no multicollinearity problem. Field (2013) also reported that the values in the item correlation matrix should not be below 0.30 and above 0.90, and the values in the anti-image table should be 0.50 and above. The analysis showed that both assumptions were met. In addition, the Kaiser-Meyer-Olkin (KMO) measure (0.88) indicated adequate sampling adequacy, and Bartlett's test of sphericity was significant ($\chi^2(120)=1032.08$; $p < .001$), confirming that the data were suitable for factor analysis. The findings revealed that both test results were suitable for factor analysis.

Although PCA was initially employed for factor extraction, an additional analysis using Principal Axis Factoring (PAF) was conducted to address methodological considerations regarding common-factor models. The results obtained from PAF were largely consistent with the PCA findings in terms of factor structure and item loadings. Therefore, the PCA results were retained for reporting due to their interpretability and consistency.

The scree plot and factor structure matrix values were examined to determine the number of factors. The decision for item validity, item

factor structure matrix values, and common-factor variances were checked. The acceptable lower limit for common-factor variance was 0.40, and the acceptable lower limit for factor structure matrix values was 0.60. Because there is a relationship between the factors of the scale, the distribution of the items to the factors was determined by making an oblique rotation. In oblique rotation, the delta angle was set to =0.0, and the kappa coefficient was set to 4. Thus, it aimed to see the factor structure more clearly (Erkuş, 2012). While naming the factors, the structure and theoretical framework that the items aimed to measure were considered. Correlation analysis was performed to determine the direction and level of the relationship between the factors. Within the scope of reliability analysis, Spearman-Brown's two-half test reliability and Cronbach's alpha internal-consistency coefficients were examined and reported for both the factors and the entire scale. Items that did not have a sufficient level of item-total correlation, factor structure matrix value, or common-factor variance were excluded.

In the second stage of the study, the model identified was tested using CFA with data obtained from 223 students. Univariate normality was first examined, and the same procedures described above were applied to the dataset. Cook's distance, leverage values, and Mahalanobis distance coefficients were analyzed to assess multivariate normality. The results indicate that the assumption of multivariate normality was met. Model fit was evaluated using fit indices (χ^2/df , SRMR, RMSEA, AGFI, NNFI, RMR, CFI, GFI, IFI, NFI, and RFI). Univariate and multivariate normality tests and PCA were conducted using SPSS 24.0, and CFA was performed using AMOS.

Results

In this part of the study, the PCA findings for construct validity are presented. The CFA findings for the verification of the structure are then included.

After determining that the dataset was suitable for analysis, PCA was performed. In the first stage, the total correlation values of the items were examined. The items with values below the acceptable limit were removed from the scale form (items 4, 12, 14, and 16), and the analysis was repeated. With the remaining 12 items, a structure consisting of four factors was obtained, explaining 39.24% of total variance. When the common-factor variances of the items were examined, the lower limit decreased to 0.28. Therefore, items with a factor variance of less than 0.40, which is considered the lower limit in the study (Items 2, 3, 13), were excluded from the analysis, and PCA was repeated. As a result of the PCA with nine items, it was seen that item five was below 0.10 in two different factors; that is, it was an overlapping item. Therefore, the relevant item was excluded from the analysis.

The PCA was repeated with the remaining eight items. The correlation matrix, determinant coefficients, and anti-image values were rechecked to determine whether the last dataset obtained was suitable for analysis. The findings showed that the values in the item correlation matrix were between 0.10 and 0.45, and the determinant coefficients were higher than 0.01. Anti-image covariance values were between 0.72 and 0.88, and the anti-image correlation values were between 0.73 and 0.82. The KMO value of the dataset was calculated as 0.777. The Bartlett test result was significant ($\chi^2(28)=349.224$; $p < .05$). The findings show that the sample size is sufficient for the final dataset and that the sphericity assumption is met. Accordingly, the dataset was suitable for analysis. After this stage, the scree plot was examined to determine the number of factors in the scale (Figure 1).

The scree plot in Figure 1 indicates the existence of a three-factor structure. Factor characteristics, including the number of factors related to the structure, eigenvalues, explained variance rates, and cumulative variance rates, are given in Table 1.

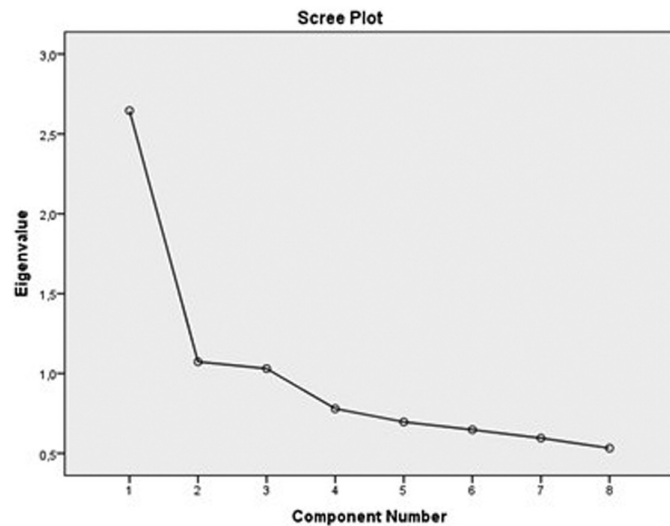


Figure 1.
Scree Plot Graph.

Table 1 shows a three-factor structure with an eigenvalue of at least 1. Factors alone explain at least about 13% of the latent variable. The first factor alone explains about 33% of the relevant structure. In addition, it is understood from the table that a structure explaining approximately 59% of the total variance was obtained. The results of the rotated factor solution performed with the oblique rotation method are presented in Table 2.

Table 2 shows the rates of explaining the variance of each item in a common factor. Common-factor variance of items: The values are between 0.47 and 0.75. The first factor, called monitoring, includes two items. It is understood that the factor structure matrix values of the items in this factor are 0.70 and 0.85, respectively. The second factor is the planning factor, which includes three items. It is seen that the factor structure matrix values of the items in the factor are between 0.61 and 0.82. The last factor of the scale is the evaluation factor, which consists of three items. The factor structure matrix values of the items in this factor vary between 0.64 and 0.79. Correlation analysis was carried out to determine whether there is a relationship between the factors and, if so, the direction and level of the relationship. The obtained results are presented in Table 3.

Table 1.
Factor Properties

Factors	Eigenvalue	Variance (%)	Total Variance (%)
1	2.65	33.07	33.07
2	1.07	13.41	46.48
3	1.03	12.88	59.36

Table 2.
Rotated Factor Loadings

Factor Name	Item No.	Item No. (New)	Communality	Factor 1	Factor 2	Factor 3
Monitoring	1	1	.54	0.70		
	10	2	.75	0.85		
Planning	6	3	.69		0.82	
	8	4	.65		0.81	
	15	5	.47		0.61	
Evaluation	7	6	.65			0.79
	9	7	.50			0.70
	11	8	.49			0.64

Table 3..
Correlation Between Factors

Factors	1	2
1. Monitoring	—	
2. Planning	0.34 **	—
3. Evaluation	0.31 *	0.37 *

Table 4.
Reliability Coefficients

	Cronbach's α	Spearman-Brown
Monitoring	0.68	0.68
Planning	0.74	0.68
Evaluation	0.73	0.74
Overall Scale	0.80	0.80

Table 3 shows the correlation values between the factors. Monitoring and planning ($r=.34$, $p < .01$); monitoring and evaluation ($r=.31$, $p < .05$); planning and evaluation ($r=.37$, $p < .05$) factors. The results indicate positive, moderate, and significant correlations between them. This finding provides evidence that there is no multicollinearity problem between the factors of the scale and that the factors are intended to measure distinct features. Cronbach's alpha internal-consistency coefficient and Spearman-Brown split-half test reliability were calculated for the whole scale and factors on the same sample. The coefficients are shown in Table 4.

Table 4 shows that both Cronbach's α and Spearman-Brown coefficients for the factors range from 0.68 to 0.74. It shows the Cronbach's α and the Spearman-Brown for the overall scale are 0.80. Findings indicate that the data of the scale are reliable.

Confirmatory factor analysis was conducted to test the suitability of the structure obtained as a result of PCA. Confirmatory factor analysis indicated "a good fit" for the suggested 3-factor model ($\chi^2(16)=32.963$, $p < .01$; CFI=0.95, AGFI=0.92, IFI=0.96, RMSEA=0.07<0.08,

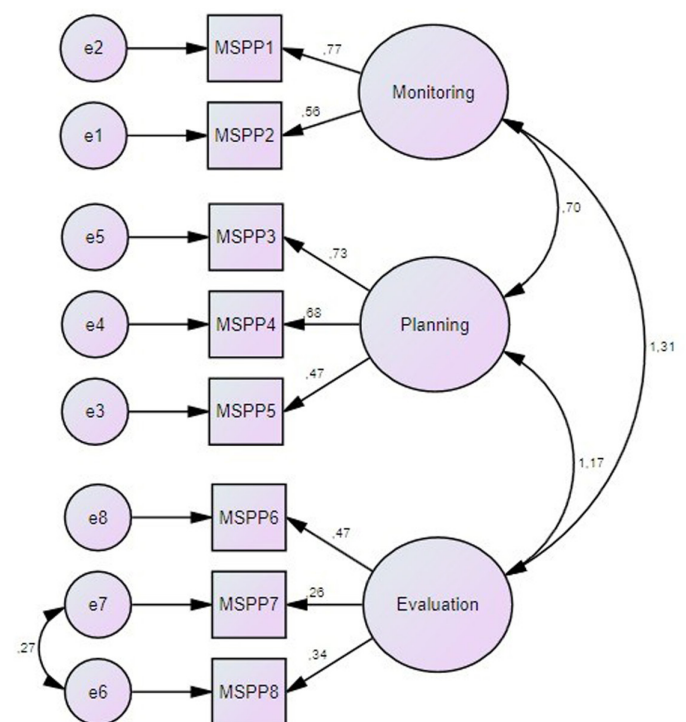


Figure 2.
Path Diagram.

and SRMR=0.07<0.08). Figure 2 presents all the standardized path coefficients.

Figure 2 pointed out that the correlation values of the items ranged from 0.26 to 0.77. These values indicate the correlation between each item and the latent variable. In addition, the relevant values indicate that the items are sufficient to distinguish the feature to be measured.

The MPP, which consists of eight items and three factors, will provide valid and reliable measurements. There is no reverse item in the scale. The minimum score obtained from the scale is eight; the maximum score is 40.

Discussion and Recommendations

The findings of this study support the existing literature while also providing new contributions. The results indicated that MPP consists of three factors: planning, monitoring, and evaluation. This structure is consistent with previous theoretical frameworks that conceptualize MPP within these categories (Baumanns & Rott, 2022b; 2022c). Thus, the factor structure of the developed scale aligns with the conceptual foundations proposed in earlier research.

Beyond confirming the three-factor structure, the study also revealed a noteworthy finding: the monitoring factor includes fewer items compared to planning and evaluation. Although factors with few items should be treated with caution, prior research indicates that they may be retained when items are strongly related and distinct (Worthington & Whittaker, 2006). In this study, the monitoring factor was retained based on both theoretical considerations and empirical support, as factor loadings indicated acceptable relationships with the construct. Moreover, metacognition in problem-posing is commonly conceptualized as comprising planning, monitoring, and evaluation, and the items in this scale clearly reflect this structure. Therefore, the monitoring factor was preserved to maintain theoretical consistency. Nevertheless, this limitation should be acknowledged, and future studies should include more items to strengthen reliability. Although the three-factor structure was supported, the results suggested limited discriminant validity, indicating that the factors are highly correlated. However, the structure was retained based on theoretical considerations.

The scale developed in this study addresses a significant gap in the literature. While previous studies on MPP have largely been qualitative (Kontorovich et al., 2012; Baumanns & Rott, 2022b; 2022c), the present study offers a quantitative measurement tool that enables broader generalization. In this respect, the study contributes methodologically to research in mathematics education.

The explained variance (approximately 60%) is consistent with commonly accepted standards in scale-development research (DeVellis, 2011; Hair et al., 2009). In addition, each factor explains more than 10% of the variance, which further supports the adequacy of the factor structure (Erkuş, 2012). Item factor loadings ranged from 0.61 to 0.85, indicating acceptable to strong loadings according to established criteria (Hinkin, 1995; Comrey & Lee, 2016). Furthermore, the removal of overlapping items strengthened the construct clarity of the scale.

Conclusion

Overall, the findings demonstrate that the MPP Scale is a valid and reliable instrument for assessing middle school students' metacognitive engagement in problem-posing. The scale provides a foundation for future quantitative and intervention-based research in this area.

In addition, the consistency between PCA and PAF results supports the stability of the factor structure. This strengthens the construct

validity of the scale and suggests that the identified factors are robust across different extraction methods.

Practical Implications and Future Research

When the literature is examined, it has been determined that the existing studies neglect the quantitative studies that will examine the problem-posing process of middle school students. This study provides a useful tool for future quantitative research by developing a measurement tool that can measure the metacognitive skills of middle school students in the problem-posing process with a small number of easy-to-apply items. In this context, future researchers can use this scale to determine the metacognitive skills of middle school students in the problem-posing process. Researchers can design relational studies using the developed scale and identify other variables related to MPP. They can also design intervention studies using this scale and examine its effects on MPP.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Artificial Intelligence Usage Statement: ChatGPT and Grammarly were used only for language editing. The authors are fully responsible for the content.

Ethics Committee Approval: All legal permits were obtained for the study. The study reported in this paper was conducted under the approval of the Ethics Committee of Bayburt University (No: E-51694156-050.99-36868 Date: October 14, 2021).

Informed Consent: Written informed consent was obtained from the participants who participated in this study.

Peer-review: Externally peer-reviewed.

Author Contributions: –Concept – M.Ö.; Design – İ.S.; Supervision M.Ö.; Resources – M.A.; Materials – M.Ö.; Data Collection and/or Processing – M.Ö.; Analysis and/or Interpretation – İ.S.; Literature Search – M.Ö.; Writing Manuscript – M.Ö.; Critical Review – M.A.

Declaration of Interests: The authors declare that they have no competing interest.

Funding: This work was financially supported by the Scientific and Technological Research Council of Turkey (TÜBİTAK) under grant number 121K859.

References

- Baumanns, L., & Rott, B. (2021a). Rethinking problem-posing situations: A review. *Investigations in Mathematics Learning*, 13(2), 59–76. [\[CrossRef\]](#)
- Baumanns, L., & Rott, B. (2021b). Metacognitive behaviour in problem posing—a case study. [Conference session]. 44th Conference of the International Group for the Psychology of Mathematics Education, Khon Kaen, Thailand.
- Baumanns, L., & Rott, B. (2022a). Developing a framework for characterising problem-posing activities: A review. *Research in Mathematics Education*, 24(1), 28–50. [\[CrossRef\]](#)
- Baumanns, L., & Rott, B. (2022b). Identifying metacognitive behavior in problem-posing processes. *International Journal of Science and Mathematics Education*, 21(5), 1381–1406. [\[CrossRef\]](#)
- Baumanns, L., & Rott, B. (2022c). The process of problem posing: Development of a descriptive phase model of problem posing. *Educational Studies in Mathematics*, 110(2), 251–269. [\[CrossRef\]](#)
- Beavers, A. S., Lounsbury, J. W., Richards, J. K., Huck, S. W., Skolits, G. J., & Esquivel, S. L. (2013). Practical considerations for using exploratory factor analysis in educational research. *Practical Assessment, Research, and Evaluation*, 18(1), 1–13. [\[CrossRef\]](#)
- Comrey, A. L., & Lee, H. B. (2016). *A first course in factor analysis* (2nd ed). Psychology Press

- Cozza, B., & Oreshkina, M. (2013). Cross-cultural study of cognitive and meta-cognitive processes during math problem solving. *School Science and Mathematics*, 113(6), 275–284. [\[CrossRef\]](#)
- DeVellis, R. F. (2011). *Scale development: Theory and applications* (3rd ed). Thousand Oaks, CA: Sage Publications.
- Dos Santos Kawata, K. H., Ueno, Y., Hashimoto, R., Yoshino, S., Ohta, K., Nishida, A., Ando, S., Nakatani, H., Kasai, K., & Koike, S. (2021). Development of metacognition in adolescence: The congruency-based metacognition scale. *Frontiers in Psychology*, 11, 565231. [\[CrossRef\]](#).
- Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 8(3), 430–457. [\[CrossRef\]](#)
- Erkuş, A. (2012). Psikolojide ölçme ve ölçek geliştirme-I: Temel kavramlar ve işlemler [Measurement and scale development in psychology-I: Basic concepts and processes]. *Pegem Akademi*.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). *Multivariate data analysis* (7th ed). London: Pearson.
- Hinkin, T. R. (1995). A review of scale development practices in the study of organizations. *Journal of Management*, 21(5), 967–988. [\[CrossRef\]](#)
- Işık, C., & Kar, T. (2012). Matematik dersinde problem kurmaya yönelik öğretmen görüşleri üzerine nitel bir çalışma [A qualitative study on teacher views of problem posing in mathematics lesson]. *Milli Eğitim Dergisi*, 194, 199–215.
- Kontorovich, I., Koichu, B., Leikin, R., & Berman, A. (2012). An exploratory framework for handling the complexity of mathematical problem posing in small groups. *Journal of Mathematical Behavior*, 31(1), 149–161. [\[CrossRef\]](#)
- Öztürk, M., & Ada, K. (2022). Problem çözme ve problem kurmada üstbilis [Metacognition in problem solving and problem posing]. In K. Özgen, T. Kar, S. Çenberci & Y. Zengin (Eds.). *Matematikte problem çözme ve problem kurma [Problem solving and problem solving in mathematics]*. Pegem Academy.
- Pelzer, I., & Gamboa, F. (2009). Problem posing: Comparison between experts and novices. In M. Tzekaki, M. Kaldrimidou & H. Sakonidis (Eds.). *Proceedings of the 33th conference of the international group for the psychology of mathematics education*, 4 (pp. 353–360). PME.
- Schneider, W., & Artelt, C. (2010). Metacognition and mathematics education. *ZDM*, 42(2), 149–161. [\[CrossRef\]](#)
- Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. In D. Grouws (Ed.). *Handbook of research on mathematics teaching and learning*. London: Macmillan.
- Semerari, A., Carcione, A., Dimaggio, G., Falcone, M., Nicolò, G., Procacci, M., & Alleva, G. (2003). How to evaluate metacognitive functioning in psychotherapy? The metacognition assessment scale and its applications. *Clinical Psychology and Psychotherapy*, 10(4), 238–261. [\[CrossRef\]](#).
- Stoyanova, E., & Ellerton, N. F. (1996). A framework for research into students' problem posing in school mathematics. In P. C. Clarkson (Ed.). *Technology in mathematics education* (pp. 518–525). Mathematics Education Research Group of Australasia. Melbourne, Australia.
- Swanson, H. L. (1990). Influence of metacognitive knowledge and aptitude on problem solving. *Journal of Educational Psychology*, 82(2), 306–314. [\[CrossRef\]](#)
- Urban, K., & Urban, M. (2023). How can we measure metacognition in creative problem-solving? Standardization of the MCPS scale. *Thinking Skills and Creativity*, 49, 101345. [\[CrossRef\]](#)
- Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *Counseling Psychologist*, 34(6), 806–838. [\[CrossRef\]](#)
- Yang, K. L. (2012). Structures of cognitive and metacognitive reading strategy use for reading comprehension of geometry proof. *Educational Studies in Mathematics*, 80(3), 307–326. [\[CrossRef\]](#).

The MPP Scale

Dear Participants,

This scale has been developed to assess to reveal the metacognition in problem-posing for middle school students. Please read the following items carefully and mark your most appropriate choice. Your time to complete the scale will be between 3 and 6 minutes. The study is carried out only with voluntary participants, and if you are not willing to participate, please inform the practitioner.

Thank you for your participation.

Items	Never	Rarely	Sometimes	Often	Always
While posing a problem, I check my writing.					
While posing problems, I ask myself questions to ensure I approach my goal.					
Before I start posing a problem, I can tell whether I can pose a problem with the given information.					
Before I start posing a problem, I decide whether I can pose a problem.					
Before I start posing a problem, I draw a diagram to help express the problem clearly.					
After I finish posing a problem, I read the problem statement several times to see if I made any mistakes.					
After I finish posing a problem, I identify which information in the problem scenario is necessary and which is not.					
After I pose a problem, I solve it to see whether it is correct.					