



Araştırma Makalesi • Research Article

Mathematical Higher-Order Thinking Scale: Turkish Adaptation, Validity and Reliability Study

Cem Kurdal*, Abdullah Kaplan**

Abstract: Mathematics education aims to enhance students' computational skills and develop their critical, creative, problem-solving, and metacognitive thinking abilities. In this context, the valid and reliable assessment of mathematical higher-order thinking skills is crucial for evaluating instruction and monitoring cognitive development. This study aims to adapt the Mathematical Higher-Order Thinking Scale, developed by Zhou et al. (2024), for use in Turkish culture and to develop a valid and reliable measurement tool. The original scale consists of four main dimensions: mathematical critical thinking, creative thinking, problem solving, and metacognitive skills and nine subdimensions. The scale was translated into Turkish using the back-translation method, revised based on expert opinions, and administered to 417 high school students. The construct validity of the scale was examined through confirmatory factor analysis (CFA) and second-order CFA. Measurement invariance across gender was tested at the configural, metric, scalar, and residual levels, and the model was found to be invariant across groups. Reliability analyses were conducted, and correlations among the dimensions were examined to ensure internal consistency. The findings indicated that the Turkish version of the Mathematical Higher-Order Thinking Scale is a valid and reliable instrument for assessing students' higher-order thinking skills in mathematics.

Keywords: High school students, Mathematical higher-order thinking skills, Scale adaptation

Matematiksel Üst Düzey Düşünme Ölçeği: Türkçeye Uyarlama, Geçerlik ve Güvenirlik Çalışması

Öz: Matematik eğitimi, öğrencilerin yalnızca işlem becerilerini değil; eleştirel, yaratıcı, problem çözme ve üstbilişsel düşünme becerilerini de geliştirmeyi hedefler. Bu bağlamda, matematiksel üst düzey düşünme becerilerinin geçerli ve güvenilir biçimde ölçülmesi hem öğretimin değerlendirilmesi hem de bilişsel gelişimin izlenmesi açısından önemlidir. Bu araştırma, Zhou vd. (2024) tarafından geliştirilen Matematiksel Üst Düzey Düşünme Ölçeği'nin Türk kültürüne uyarlanmasını ve geçerli, güvenilir bir ölçme aracı elde etmeyi amaçlamaktadır. Orijinal ölçekte matematiksel eleştirel düşünme, yaratıcı düşünme, problem çözme ve üstbilişsel beceriler olmak üzere dört temel boyut ve dokuz alt boyut yer almaktadır. Geri çeviri yöntemiyle Türkçeye uyarlanan ölçek, uzman görüşleri doğrultusunda düzenlenmiş ve 417 lise öğrencisine uygulanmıştır. Ölçeğin yapı geçerliği doğrulayıcı faktör analizi (DFA) ve ikinci düzey DFA ile incelenmiş; cinsiyete göre ölçüm değişmezliği analizi yapılarak biçimsel, metrik, skalar ve hata değişmezliği düzeylerinde modelin değişmezliği doğrulanmıştır. Güvenirlik analizleri sonucunda boyutlar arası korelasyonlar incelenmiş ve ölçeğin iç tutarlılığı sağlanmıştır. Bulgular, ölçeğin Türk kültüründe geçerli ve güvenilir bir ölçme aracı olduğunu göstermektedir.

Anahtar Kelimeler: Lise öğrencileri, Matematiksel üst düzey düşünme becerileri, Ölçek uyarlama

Cite as/ Atf: Kurdal, C. & Kaplan, A. (2026). Mathematical Higher-Order Thinking Scale: Turkish Adaptation, Validity and Reliability Study. *Anemon Muş Alparslan Üniversitesi Sosyal Bilimler Dergisi*, 14(1), 316-362
<http://dx.doi.org/10.18506/anemon.1698071>

Received/Geliş: 13 May/Mayıs 2025

Accepted/Kabul: 14 Nov/Kasım 2025

Published/Yayın: 30 Apr/Nisan 2026

e-ISSN: 2149-4622. © 2026 The Author(s). This article is licensed under the Creative Commons Attribution-NonCommercial4.0 International License (CC BY-NC 4.0): <https://creativecommons.org/licenses/by-nc/4.0/>



* Arş. Gör. Dr., Bayburt Üniversitesi, Matematik ve Fen Bilimleri Eğitimi Bölümü, ORCID: 0009-0008-0880-4093
cemkurdal@gmail.com

**Prof. Dr., Atatürk Üniversitesi, Matematik ve Fen Bilimleri Eğitimi Bölümü, ORCID: 0000-0001-6743-6368
akaplan@atauni.edu.tr

Introduction

In the 21st century, changing forms of information production require individuals to access information and effectively utilize higher-level cognitive processes, such as analyzing, evaluating, and transferring the information they obtain to new situations (Kurnaz, 2019; Topal et al., 2025). Education systems, in particular, emphasize the need to instill these skills from an early age, prioritizing the development of mental competencies, such as critical thinking, creative thinking, problem-solving, and metacognitive awareness in students (Altun, 2024). Mathematics education, by its very nature, which underlies mental processes, such as abstraction, generalization, and inference, is considered a suitable discipline for developing these higher-level thinking skills (English & Halford, 2012; Lin, 2023).

Accordingly, international assessment systems, such as PISA and TIMSS, systematically evaluate students' success not only at the procedural knowledge level but also in higher-order cognitive processes, such as reasoning, judgment, and problem-solving (İncikabı et al., 2016; OECD, 2023). In the Turkish context, the updated mathematics curriculum in the 21st Century Education Model (MEB, 2024), which supports this approach, aims to develop students' mathematical thinking skills in a multidimensional way, thereby making them active individuals in the information society. In the program, the development of critical and creative thinking, the structuring of problem-solving processes, and the acquisition of metacognitive awareness are considered among the basic competencies.

It is extremely important that these mental competencies are developed in learning processes and measured in a valid and reliable manner. However, most of the scales developed in the literature for measuring these skills are either designed independently of the field of mathematics or focus only on a single cognitive dimension (e.g., only creative thinking or only problem-solving) (Gürefe & Akçakin, 2018; Korkmaz & Karakuş, 2009; Ünal & Aral, 2014). In addition, the limited number of psychometric tools that are specific to mathematical content, based on theoretical foundations, and capable of multidimensional measurement makes it difficult to evaluate higher-order thinking skills holistically in education. Indeed, students' mathematical thinking skills often require the interaction of more than one cognitive process, and these processes need to be evaluated without separating them (Schoenfeld, 1992). In this context, Zhou et al. The Mathematical Higher-Order Thinking Scale (MHOTS), developed by (2024), provides the opportunity to evaluate students' multi-layered mental processes by considering them in four basic dimensions. This scale consists of four sub-dimensions: mathematical critical thinking, mathematical creative thinking, mathematical problem solving, and mathematical metacognitive skills.

Mathematical Critical Thinking

Mathematical critical thinking is a high-level cognitive process that involves an individual's ability to generate information, make judgments, and make decisions by analyzing a mathematical situation in a multidimensional way. This dimension encompasses the ability to perform mathematical operations correctly and the ability to question, explain, and justify thought processes (Monteleone et al., 2023). Critical thinking includes skills, such as interpretation, analysis, evaluation, explanation, and creation, and the application of these skills in a mathematical context is defined as "critical mathematical thinking." Students' behaviors of questioning mathematical solution processes, making judgments, and justifying these judgments are among the key indicators of critical mathematical thinking. In this context, a student's ability to solve a mathematical problem and explain the solution method, justify the strategy used, and consider alternative solutions demonstrates their critical mathematical thinking skills (Samura & Juandi, 2020). This skill supports the individual's thinking ability and social learning processes, such as defending one's ideas in classroom discussions and critically evaluating the ideas of others (Sachdeva & Eggen, 2021).

Mathematical critical thinking is closely related to creative thinking, problem-solving, and metacognitive skills, as individuals can both observe and reconstruct their own thinking processes, thereby developing more in-depth solution strategies (Zhang et al., 2024; Monteleone et al., 2023). From this perspective, mathematical critical thinking is an individual skill and a constructive, multifaceted, and meaning-based learning process. Mathematical critical thinking allows students to question, evaluate, and justify their solution process, while reconstructing this process in different, original, and unconventional ways forms the basis of mathematical creative thinking.

Mathematical Creative Thinking

Mathematical creative thinking is a high-level cognitive process that enables an individual to approach a mathematical problem in multifaceted and original ways, going beyond conventional patterns. This skill encompasses not only arriving at the correct solution but also diversifying solution paths, developing original ideas, and demonstrating active cognitive productivity in this process. This dimension encompasses generating correct solutions to a mathematical problems and the ability to produce many valid and meaningful ideas (fluency) and to develop original and unusual ideas among them (innovation) (Suherman & Vidákovich, 2022). Fluency refers to the student's capacity to present more than one solution path to the same problem, while innovation reflects the ability to select original paths from among these and creatively restructure them (Fortes & Andrade, 2019; Sriraman, 2009). In this context, mathematical creative thinking involves not only the student producing solutions but also their ability to generate, compare, and refine these solution paths through flexible reasoning and reflective decision-making. This is a multidimensional cognitive competency that includes the ability to justify these solutions meaningfully, propose alternative approaches, and express original mathematical thought structures (Furnham et al., 2013; Kholid et al., 2024). While mathematical creative thinking supports students in generating multiple and original solution paths, the process of structuring, evaluating, and effectively applying these solutions also requires problem-solving skills.

Mathematical Problem Solving

Mathematical problem-solving is a high-level cognitive process that involves an individual understanding a mathematical situation, developing appropriate solutions, applying these solutions, and evaluating the process. This dimension encompasses performing operations correctly and skills, such as interpreting given information, generating strategies, monitoring the solution process, and evaluating the outcome (Öztürk et al., 2023). In this study, problem-solving skills are addressed in two main sub-dimensions: reasoning skills and reflective skills. Reasoning skills involve the student making inferences by establishing logical relationships between the data in the problem and planning the solution path according to these inferences (Suherman & Vidákovich, 2022). Especially in structured problems, such as ratios, proportions, and patterns, students' ability to produce correct solutions largely depends on how effectively they use their reasoning processes (Öztürk et al., 2021). Reflective skills, on the other hand, refer to the student's ability to monitor the solution process, evaluate the functionality of the strategies they have chosen, and reorganize the process when necessary (Carbonneau et al., 2020). This skill enables the student to evaluate not only the result but also the process, and significantly increases problem-solving success, mainly when supported by metacognitive awareness (Suherman & Vidákovich, 2022). In this context, mathematical problem-solving is a holistic learning skill that extends beyond simply reaching a solution; it encompasses the processes of planning, monitoring, and structuring thought. This multi-layered cognitive structure of the problem-solving process requires the individual's awareness of their own thought processes, their experiences, and their ability to manage these processes consciously; this forms the basis of mathematical metacognitive skills.

Mathematical Metacognitive Skills

Mathematical metacognitive skills are a multi-layered cognitive self-regulation system related to an individual's awareness of their own mathematical thinking processes, planning, monitoring, and restructuring these processes when necessary (Muncer et al., 2021; Suliani et al., 2024). This skill involves possessing information and understanding how, when, and why to use that information and

consciously managing this usage process (Güner & Erbay, 2021). Mathematical metacognitive skills are examined under three main components: metacognitive knowledge, metacognitive experience, and metacognitive monitoring.

Metacognitive knowledge refers to an individual's knowledge of their own cognitive characteristics, learning strategies, and the requirements of tasks (Suliani et al., 2024). This type of knowledge, metacognitive knowledge, encompasses three fundamental components: knowing what one knows (declarative), knowing how to do something (procedural), and knowing when and in what situation to use a particular strategy (contingent knowledge) (Güner & Erbay, 2021). Students with well-developed metacognitive knowledge can more realistically assess their own competencies and develop more effective learning processes by choosing strategies appropriate to the type of problem they encounter (Muncer et al., 2021).

Metacognitive experience is the student's awareness of the emotions, thoughts, and judgments they experience during a mathematical task, as well as their ability to guide the process based on these internal experiences (Yüksel et al., 2021). This experience is crucial in reflecting what an individual feels, knows, thinks, and how confident they are while solving a particular problem (Suliani et al., 2024). In particular, feelings of difficulty or success experienced during problem-solving can cause individuals to change their strategy choices, which is effective in consciously managing the process (Şahinkaya et al., 2022). Metacognitive monitoring is the ability of an individual to observe their mathematical thinking process in real time, control the solution process, and reorganize it by intervening when necessary (Güner & Erbay, 2021). This monitoring skill to identify their current position during the solution process, recognize their mistakes, and adjust their strategies to steer the process in the right direction (Suliani et al., 2024). Research indicates that developing metacognitive skill enables students to successfully manage the outcome and the process, and significantly enhances their mathematical achievement (Aydın & Öztürk, 2024). In this context, mathematical metacognitive skills are a fundamental cognitive system that reflects how students access information, structure and evaluate information in the mathematical learning process; they are integrated with self-awareness and support in-depth learning (Suliani et al., 2024; Muncer et al., 2021).

In recent years in Türkiye, various scale development and adaptation studies have been carried out regarding metacognitive skills, creative and critical thinking, and mathematical problem-solving processes. However, the number of measurement tools that directly link metacognitive awareness to the problem-solving process is quite limited. Most scales developed in Türkiye either measure general metacognitive awareness or address problem-solving processes indirectly. The majority of these studies address general metacognitive constructs or establish a limited relationship with the mathematical context. For example, the scales developed by Fırat Durdukoca and Arıbaş (2019) and Kaplan and Duran (2016) aim to measure the general level of metacognitive awareness. They are not structured specifically for mathematical content. The scales developed by Çanakçı and Özdemir (2011) and Hacıömeroğlu (2011), respectively, aim to measure attitudes and beliefs regarding mathematical problem solving and exclude higher-level cognitive processes. Furthermore, Suliani et al. (2024), Muncer et al. (2021), and Yüksel et al. (2021), highlight the significance of metacognitive skills in individuals' mathematical thinking processes, but state that there is an insufficient number of scales that directly measure these constructs.

In line with this general picture, most measurement tools developed and adapted in mathematical problem-solving in Türkiye focus on attitudes, skills, beliefs, and achievement in terms of the areas they evaluate. For example, attitude scales aim to measure students' tendencies and motivations towards problem-solving (Çanakçı, 2008; Uğurluoğlu, 2008). Skill scales are geared towards evaluating individuals' problem-solving performance (Çömlekoğlu, 2001; Özsoy, 2008; Uysal, 2007; Kızılkaya & Aşkar, 2009; Özpinar, 2012; Muyo, 2015; Çelik, 2017; Atlıhan, 2021; Gürsoy & Çeliköz, 2021). However, metacognitive regulatory dimensions of the problem-solving process, such as planning, monitoring, control, and evaluation, are mostly overlooked in these tools.

Similarly, belief scales have measured students' epistemological beliefs and self-efficacy perceptions regarding problem-solving (Kayan, 2008; Hacıömeroğlu, 2011; Gümüş & Umay, 2017). These scales assess subjective beliefs about problem-solving rather than cognitive awareness in the process. Achievement tests (Sadık, 2006; Güneş, 2022; Süzer, 2022; Yaşar, 2022) measure problem-solving performance but do not directly assess individuals' metacognitive awareness of their own problem-solving processes. An examination of the study groups reveals that the developed scales are applied across a wide range of levels, from second grade of primary school to university level (Değirmenci & Deringöl, 2024; Çanakçı, 2008; Uysal, 2007; Çelik, 2017; Özpınar, 2012; Atlıhan, 2021; Özsoy, 2008; Sadık, 2006; Süzer, 2022; Yaşar, 2022; Hacıömeroğlu, 2011; Kayan, 2008; Muyo, 2015; Çömlekoğlu, 2001; Gümüş & Umay, 2017; Gürsoy & Çeliköz, 2021). This situation highlights that despite extensive studies on measuring problem-solving skills at various ages and educational levels, remains lack of valid and reliable measurement tool to evaluate metacognitive awareness in the mathematical problem-solving process of high school students, along with their creative and critical thinking skills. In this context, adapting the scale developed by Zhou et al. (2024), which addresses the dimensions of critical thinking, creative thinking, problem-solving, and metacognitive thinking holistically, into Turkish makes an important contribution to meeting the multidimensional measurement needs in the field.

Therefore, rather than developing a new scale, the present study aims to adapt the Mathematical Higher-Order Thinking Scale (MHOTS) developed by Zhou et al. (2024), which has a solid theoretical foundation and a multidimensional structure, for use in Turkish. This approach enabled the benefits of the international validity and comprehensive factor structure of the original scale, while also facilitating cultural and linguistic adaptations suitable for the Turkish educational context. The adapted scale has a unique structure that enables the evaluation of students' mathematical thinking processes in a multidimensional manner. The scale developed by Zhou et al. (2024) aims to measure the cognitive, affective, and metacognitive aspects of mathematical thinking simultaneously addressing the dimensions of critical thinking, creative thinking, problem-solving, and metacognitive thinking within a holistic framework. Considering that the existing tools for measuring mathematical thinking skills in Türkiye are generally unidimensional, the translation of this scale into Turkish fills an important gap in the field. Furthermore, this study is consistent with national education policies as, it contributes to the assessment of higher-order thinking and problem-solving skills emphasized in the Turkish Century Education Model (MEB, 2024). Therefore, the Turkish adaptation of the scale can make a significant contribution to the field of mathematics education at both theoretical and practical levels.

Purpose

This study aims to fill a significant gap in the literature and provide an original and functional contribution by adapting the MHOTS (Mathematical Higher-Order Thinking Scale) developed by Zhou et al. (2024) into Turkish, presenting the validity and reliability evidence of the scale in a Turkish sample, and examining the psychometric properties of the measurement instrument. In this context, the study investigated whether the four-dimensional structure of the scale remains valid within the Turkish cultural context.

The adapted scale is an important tool for academic research and teachers to identify, monitor, and plan strategies for developing students' higher-order thinking skills in their classroom practices. For researchers, the scale enables a holistic analysis of the cognitive, affective, and metacognitive dimensions of mathematical thinking; for policymakers, it provides a concrete measurement tool for assessing 21st-century skills and the higher-order thinking achievements emphasized in the Turkish Century Education Model.

The findings of this study can make significant contributions to both monitoring the effectiveness of curricula and developing evidence-based practices that support higher-order thinking skills in teacher training processes. It is anticipated that the results of this research will guide studies on measuring, monitoring, and developing students' higher-order thinking skills in mathematics education. This

adapted scale allows for a multidimensional and holistic assessment of students' cognitive, metacognitive, creative, and critical thinking skills in a mathematical context, filling a significant gap in the literature and offering a unique and functional contribution to the field.

Method

This research aims to adapt the Mathematical Higher-order Thinking Scale (MHOTS), originally developed by Zhou et al. (2024), to assess the mathematical higher-order thinking skills of high school students, to Turkish culture. Accordingly, a survey model, a quantitative research method, was chosen for this study. The survey model is a research design that aims to describe an existing situation, phenomenon, or characteristic as it is and within its natural environment (Karasar, 2000). In this study, validity and reliability analyses were conducted as part of the adaptation process of the MHOTS scale into Turkish; the construct validity of the scale was examined using confirmatory factor analyses. Furthermore, the internal reliability coefficients of the scale were calculated to evaluate the consistency of each dimension, and the correlations between the dimensions were examined to support the structural integrity of the scale.

Translation Study

The adaptation process was carried out in accordance with criteria accepted in national and international literature (Hambleton et al., 2005). In the first stage, the 45 items of the original scale were independently translated into Turkish by three foreign language experts and three mathematics education experts with proficiency in the English-Turkish language. Then, each translation was combined, and a consensus was reached. This translated text was then re-translated into English using the back-translation method by three foreign language educators who were different from the expert team that had done the first translation. This process was conducted to assess whether the translations retained their conceptual and semantic integrity (Brislin, 1970). The back-translations were compared with the original scale, and it was found that they maintained semantic integrity.

Following expert opinions, a Turkish version of the scale was created through final revisions, and a pilot application was conducted with a group of 30 high school students. In this study, the scale form was evaluated in terms of students' understanding of the statements, and minor simplifications were made to the item structure accordingly. Based on the feedback obtained from the pilot application, the scale was finalized and unanimously approved by the expert team, making it ready for implementation. In the Turkish adapted scale, the four core dimensions (Mathematical Critical Thinking, Mathematical Creative Thinking, Mathematical Problem Solving, and Mathematical Metacognitive Skills) and their nine sub-dimensions were retained, as in the original form. The scale is structured as a 5-point Likert scale (1 = Strongly disagree, 5 = Strongly agree), as in the original study. In this context, the translation and adaptation process focused on ensuring cultural and linguistic equivalence at a qualitative level; based on expert opinions and the results of the pilot application, it was observed that the scale has been adapted to be understandable and contextually appropriate for Turkish high school students. However, it should be noted that since linguistic equivalence was examined only qualitatively in this process, the findings offer a limited linguistic evaluation.

Research Group

The study's research group consisted of high school students studying in a province located in Northeast Türkiye, studying during the 2024-2025 academic year. The research group comprised a total of 417 high school students, including 210 females (50.36%) and 207 males (49.64%). The distribution of participants by grade level was as follows: 87 (20.9%) 9th-grade students, 107 (25.7%) 10th-grade students, 109 (26.1%) 11th-grade students, and 114 (27.3%) 12th-grade students. In the literature, the recommended lower limit for valid factor analysis results is 400 (Goretzko et al., 2021; Güvendir et al., 2025). Accordingly, the sample size of this study appears to be sufficient.

Data Collection Tool

The Mathematical Metacognitive Thinking Scale (MHOTS) was developed by Zhou et al. (2024) to assess the higher-order mathematical thinking skills of high school students. During the scale development process, 53 items were identified with the contributions of experts in mathematics education and assessment; these were then reduced to 45 items based on expert opinions and analyses. The scale includes four core dimensions related to mathematical thinking: mathematical critical thinking, mathematical creative thinking, mathematical problem-solving, and mathematical metacognitive skills. Each core dimension is further divided into sub-dimensions, resulting in nine sub-factors. The scale is structured using a 5-point Likert scale (1=Strongly disagree, 5=Strongly agree), and data regarding students' gender and class variables were also collected during the application process.

The first dimension of the scale, Mathematical Critical Thinking, includes the sub-dimensions of Curiosity and Systematic Thinking. This dimension aims to measure students' tendencies to question mathematical situations deeply and think systematically. Examples include statements. For example, "When faced with a mathematical problem, I ask questions to find different solutions," and "When solving a problem, I arrange the steps I follow in a logical order." The second dimension, Mathematical Creative Thinking, consists of the sub-dimensions of Flexibility and Innovation. This dimension evaluates students' ability to solve mathematical problems using unconventional approaches. For example, statements "I try different methods to solve a mathematical problem," and "I enjoy generating original ideas in mathematics," fall under this dimension. The third dimension, Mathematical Problem Solving, encompasses the sub-dimensions of Reasoning Skills and Reflective Skills. This dimension aims to measure students' abilities in logical reasoning, strategy formation, and evaluating the solution process. An example of this is, "When solving a mathematical problem, I verify the result I reach in different ways." Examples of this dimension include statements. For example, "After solving a problem, I evaluate whether the method I used was appropriate." The fourth dimension, Mathematical Metacognitive Skills, consists of the sub-dimensions of Knowledge, Experience, and Monitoring. This dimension measures students' ability to plan, monitor, and organize their own learning processes. For example, statements like "Before solving a problem, I think about what approach I will take" and "I change my solution strategy by learning from the mistakes I make in mathematics" are examples of this dimension.

Data Collection and Analysis Process

This research began with obtaining the necessary permission via email from the responsible author who developed the scale. After obtaining ethical permissions, the process of adapting the scale into Turkish commenced. Following the adaptation process, the Turkish version of the scale was administered to high school students. Participants answered the demographic questions in the first part of the form and then indicated their opinions on the scale items using a 5-point Likert scale. The application was conducted on a voluntary basis, and necessary permissions were obtained from the school administration and teachers. The data collection process was carried out in a manner that did not disrupt class operations.

In the analysis process, the Kaiser-Meyer-Olkin (KMO) coefficient and Bartlett's sphericity test were first examined to determine the suitability of the data for factor analysis. The results of these tests showed that the data were suitable for factor analysis, and the Confirmatory Factor Analysis (CFA) process was initiated. SPSS, Jamovi, and AMOS programs were used together to test the structural validity of the scale and the model's fit to the sample. Since the data type is at the ordinal level, the WLSMV (Weighted Least Squares Mean and Variance Adjusted) estimation method and polychoric correlation matrix were preferred in the CFA analysis. To evaluate model fit, the χ^2/df , CFI, NFI, GFI, AGFI, and RMSEA fit indices were reported. The findings indicate that the model provides good fit values ($\chi^2/df < 3$, RMSEA $< .08$, CFI, NFI $\geq .90$, GFI, AGFI $\geq .85$).

The analysis revealed that the four core dimensions (Mathematical Critical Thinking, Mathematical Creative Thinking, Mathematical Problem Solving, and Mathematical Metacognitive

Skills) and their nine sub-dimensions were validated in accordance with the original structure. Standardized factor loadings ranging from .51 to .91 indicated that the items adequately represented their respective dimensions. Furthermore, the high and significant correlation coefficients between the core dimensions support the construct validity of the model. In addition, the measurement invariance of the scale with respect to the gender variable was examined. In this context, structural equation modeling (SEM) was used to test the configural, metric (factor loadings), scalar (cross-sectional), and strict (error variances) levels of invariance, respectively. The fact that the obtained ΔCFI , ΔTLI , and $\Delta RMSEA$ values remained within acceptable limits ($\Delta CFI \leq .01$, $\Delta TLI \leq .01$, $\Delta RMSEA \leq .015$) indicates that the scale functions equivalently for female and male students and that measurement invariance has been ensured.

As part of the reliability analyses of the scale, Cronbach's alpha, McDonald's Omega, Guttman Split-Half, and Spearman-Brown coefficients were calculated; coefficients above .80 for all dimensions indicated that the scale has high internal consistency. In conclusion, the analyses conducted revealed that the MHOTS can be used validly and reliably in Turkish high school students.

Findings

In the process of adapting the scale into Turkish, the suitability of the data for factor analysis was examined first. Accordingly, kurtosis and skewness values were calculated for each sub-dimension in the dataset, and the results obtained are presented in Table 1.

Table 1. Descriptive Statistics for Sub-Dimensions

| Sub-Dimensions | N | \bar{X} | SS | Kurtosis | Skewness |
|-----------------------------------|-----|-----------|-----|----------|----------|
| Mathematical Critical Thinking | 417 | 40.15 | .72 | -.94 | .31 |
| Mathematical Creative Thinking | 417 | 28.08 | .63 | -.17 | .04 |
| Mathematical Problem Solving | 417 | 34.90 | .74 | -1.05 | .28 |
| Mathematical Metacognitive Skills | 417 | 33.55 | .57 | -1.15 | .30 |

As shown in Table 1, the kurtosis and skewness values for all sub-dimensions were in the range of -1.5 to +1.5. Büyüköztürk (2007) and Tabachnick and Fidell (2013) state that values within this range ensure normality. Accordingly, it is assumed that the data set meets the assumption of normal distribution. Furthermore, it was determined that the sample size is sufficient and there were no missing or outlier values. To test the suitability of the data set for factor analysis, the Kaiser-Meyer-Olkin (KMO) and Bartlett's Sphericity Tests were applied, and the findings are presented in Table 2.

Table 2. KMO and Bartlett Sphericity Test Results for Each Subdimension

| | | | |
|-----------------------------------|--------------------------|--------------------|---------|
| Mathematical Critical Thinking | KMO Value | .97 | |
| | Bartlett Sphericity Test | Chi-square | 5774.84 |
| | | Degrees of freedom | 78 |
| | | p | .000* |
| Mathematical Creative Thinking | KMO Value | .87 | |
| | Bartlett Sphericity Test | Chi-square | 927.39 |
| | | Degrees of freedom | 45 |
| | | p | .000* |
| Mathematical Problem Solving | KMO Value | .96 | |
| | Bartlett Sphericity Test | Chi-square | 5202.31 |
| | | Degrees of freedom | 55 |
| | | p | .000* |
| Mathematical Metacognitive Skills | KMO Value | .95 | |
| | Bartlett Sphericity Test | Chi-square | 4449.80 |
| | | Degrees of freedom | 55 |
| | | p | .000* |

p* < .05

According to Table 2, the KMO values for each sub-dimension ranged between .87 and .97, and all values were above .60. Furthermore, the Bartlett Sphericity Test was statistically significant for each sub-dimension ($p < .05$). According to Büyüköztürk (2006), for factor analysis to be possible, the KMO value must be greater than .60 and the Bartlett test must be statistically significant. In this context, the findings indicate that the dataset is highly suitable for factor analysis. During the adaptation process, the structural steps in the scale's development phase were followed, and the factor structure and reliability were evaluated separately for each dimension.

Mathematical Critical Thinking Dimension CFA Results

To test the validity of the two-factor structure of the mathematical critical thinking dimension, a confirmatory factor analysis (CFA) was applied. This analysis examined the extent to which previous findings were confirmed in the sample and tested the validity of the structure. The CFA was performed using the AMOS program, and the resulting model fit indicators are visualized in Figure 1. As seen in Figure 1, the confirmatory factor analysis of the mathematical critical thinking dimension confirmed the two-factor structure of the model. The factor loadings of the items related to the "Curiosity" factor ranged from .86 to .90, while the factor loadings of the items related to the "Systematicity" factor ranged from .86 to .89. These values indicated that all items strongly represent their respective factors. Furthermore, the correlation coefficient of .77 between the two factors reveals that the Curiosity and Systematicity dimensions were highly related to each other, but did not measure the same structure. When Figure 1 is examined, the model's fit indices were calculated as follows: $\chi^2 = 117.072$, $df = 64$, $p = .007$, $\chi^2/df = 1.829$, CFI = .99, GFI = .96, NFI = .98, AGFI = .94, and RMSEA = .045. Information regarding the evaluation of these values is presented in detail in Table 3.

Table 3. Compatibility Indices related to the Dimension of Mathematical Critical Thinking

| Compatibility Indices | Zhou vd. | The values obtained within the scope of this study. | Acceptable range of compliance | Perfect fit range | Decision (Kline, 2019) |
|-----------------------|----------|---|--------------------------------|-------------------|------------------------|
| χ^2 / df | 2.670 | 1.829 | 3- 5 | 0- 3 | Perfect Fit |
| RMSEA | 0.072 | .045 | $\leq .08$ | $\leq .06$ | Perfect Fit |
| NFI | 0.923 | .98 | $\geq .90$ | $\geq .95$ | Perfect Fit |
| CFI | 0.950 | .99 | $\geq .90$ | $\geq .95$ | Perfect Fit |
| GFI | 0.931 | .96 | $\geq .85$ | $\geq .90$ | Perfect Fit |
| AGFI | 0.917 | .94 | $\geq .85$ | $\geq .90$ | Perfect Fit |

When Table 3 is examined, the χ^2/df value being below 3 indicates that the model is at an acceptable level. The RMSEA value being below .06 (.037) shows that the model has a low error rate and exhibits excellent fit. All other fit indices, NFI, CFI, GFI, and AGFI, are above .95 or .90, indicating that the model has a high level of structural validity.

Based on the fit criteria proposed by Kline (2019), all indices obtained in this study had an excellent fit level. Furthermore, the fact that these results coincide with the structural model reached by Zhou et al. (2024), who developed the original scale, indicates that the factor structure in the adapted scale has been preserved and validated in accordance with the original structure. Accordingly, it can be stated that the two-factor structure related to the dimension of mathematical critical thinking is consistently maintained based on both theoretical grounds and empirical findings.

Following confirmatory factor analysis, convergent and divergent validity values for the Mathematical Critical Thinking dimension were examined. In this context, the Mean Explained Variance (AVE) values for the Curiosity and Systematics sub-dimensions were calculated as .77 and .76, respectively, and the Composite Reliability (CR) values were calculated as .95 and .96. These values met the criteria of $AVE \geq .50$ and $CR \geq .70$. This finding indicates that both sub-dimensions have

convergent validity and that the items adequately represent the factor to which they belong (Fornell & Larcker, 1981; Hair et al., 2019). Regarding divergent validity, it was determined that the AVE values for the Mathematical Critical Thinking dimension (.77 and .76) were higher than the squares of the correlation coefficients (r^2) of this dimension with the other dimensions [MED–MPCÇ = $(.743)^2 = .552$; MED–MÜB = $(.766)^2 = .587$]. This result showed that the Curiosity and Systematics sub-dimensions were statistically distinguishable from other constructs in the scale. Therefore, the obtained AVE and CR values supported both the theoretical consistency and psychometric robustness of the Mathematical Critical Thinking dimension.

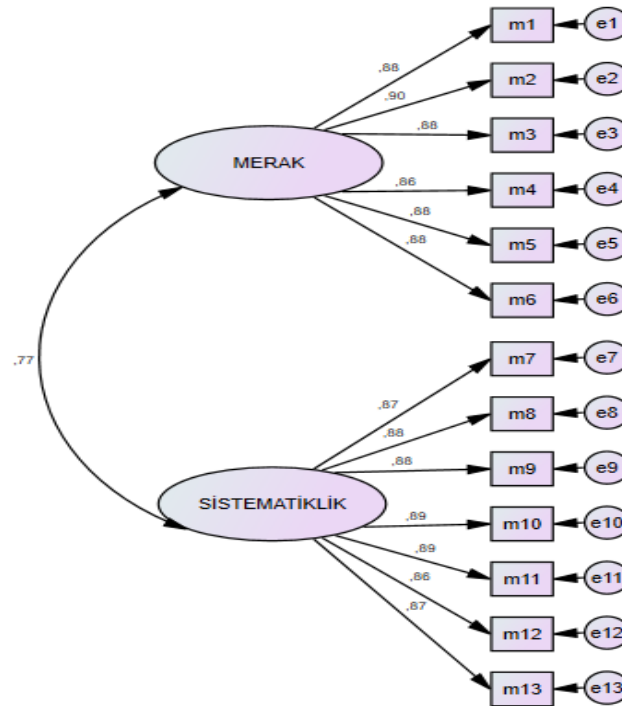


Figure 1. CFA Results regarding the Dimension of Mathematical Critical Thinking (Merak: Curiosity, Sistemantiklik: Systematics)

Mathematical Creative Thinking Dimension CFA Results

To test the validity of the two-factor structure of the mathematical creative thinking dimension, CFA was applied. This analysis examined the extent to which previous findings were confirmed in the sample and tested the validity of the structure. CFA was performed using the AMOS program, and the resulting model fit indicators are visualized in Figure 2. Figure 2 shows that the model exhibits an acceptable level of fit and that the measurement instrument maintains its construct validity. When examining the standardized factor loadings, it is observed that the loadings of the items in the Flexibility dimension range from .48 to .67, and the loadings of the items in the Innovation dimension range from .50 to .61. The fact that all items have factor loadings above .40 indicates that the items adequately represent their respective dimensions. Furthermore, the observed correlation of .85 between the two factors reveals that the dimensions have a significant and strong relationship with each other. These results generally show that the model fits the data well and that the two-dimensional structure of the scale is confirmed. When Figure 2 is examined, the model's fit indices were calculated as follows: $\chi^2 = 97.354$, $df = 34$, $p = .000$, $\chi^2/df = 2.863$, $CFI = .93$, $GFI = .96$, $NFI = .91$, $AGFI = .93$, and $RMSEA = .067$. Information regarding the evaluation of these values is presented in detail in Table 4.

Table 4. Compatibility Indices related to the Mathematical Creative Thinking Dimension

| Compatibility Indices | Zhou vd. | The values obtained within the scope of this study. | Acceptable range of compliance | Perfect fit range | Decision (Kline, 2019) |
|-----------------------|----------|---|--------------------------------|-------------------|------------------------|
| X2/df | 3.599 | 2.863 | 3- 5 | 0- 3 | Perfect Fit |
| RMSEA | .057 | .067 | ≤.08 | ≤.06 | Perfect Fit |
| NFI | .91 | .90 | ≥.90 | ≥.95 | Perfect Fit |
| CFI | .95 | .93 | ≥.90 | ≥.95 | Acceptable fit |
| GFI | 0.930 | .96 | ≥.85 | ≥.90 | Perfect Fit |
| AGFI | 0.907 | .93 | ≥.85 | ≥.90 | Perfect Fit |

When Table 4 is examined, a χ^2/df value below 3 indicates that the model is at an acceptable level. An RMSEA value below .07 (.067) shows that the model has a low error rate and exhibits an excellent fit. All other fit indices, NFI (.896), CFI (.929), GFI (.956), and AGFI (.929), are at an acceptable level, indicating good fit and structural validity.

Based on the fit criteria proposed by Kline (2019), all indices obtained in this study have an excellent fit. Furthermore, the congruence of these results with the structural model developed by Zhou et al. (2024), who developed the original scale, reveals that the factor structure in the adapted scale has been preserved and validated in accordance with the original structure. Accordingly, it can be stated that the two-factor structure related to the dimension of mathematical creative thinking is consistently maintained, based on both theoretical grounds and empirical findings.

In terms of convergent validity, the AVE values for the Flexibility and Innovation sub-dimensions were calculated as .34 and .33, respectively, and the CR values were .72 and .71. Although the AVE values are below .50, the CR values being above .70 indicates that the construct has an acceptable level of convergent validity (Fornell & Larcker, 1981; Hair et al., 2019). This result shows that the relevant items are significantly related to the factor and that the internal consistency of the construct is ensured. In the divergent validity analysis, it was observed that the AVE value of the Mathematical Creative Thinking dimension (.58) is higher than the squares of the correlation coefficients with the other dimensions [MRT–MCT = $(.127)^2 = .016$; MRT–MMS = $(.135)^2 = .018$]. This indicates that the Flexibility and Innovation sub-dimensions are distinguishable from other constructs within the scale. Therefore, the obtained AVE and CR values support the theoretical integrity of the Mathematical Creative Thinking dimension and strengthen the overall validity of the scale.

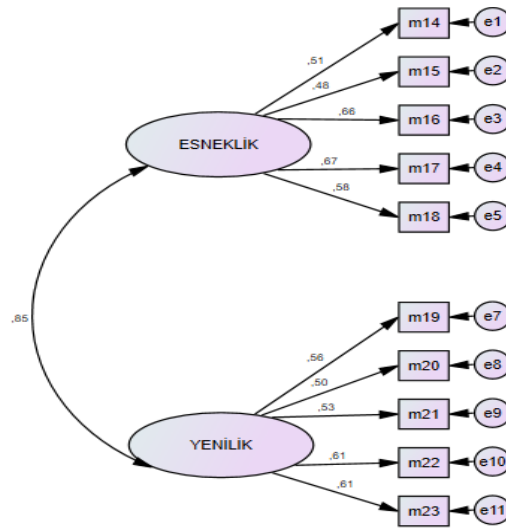


Figure 2. CFA Results regarding the Dimension of Mathematical Creative Thinking (Esneklik: Flexibility, Yenilik: Innovation)

Mathematical Problem Solving Dimension CFA Results

To test the validity of the two-factor structure of the mathematical problem-solving dimension, Confirmatory Factor Analysis (CFA) was applied. This analysis examined the extent to which previous findings were confirmed on the sample and tested the validity of the structure. CFA was performed using the AMOS program, and the resulting model fit indicators are visualized in Figure 3. Figure 3 shows that the model exhibits an excellent fit and strongly ensures the structural validity of the measurement instrument. When the standardized factor loadings were examined, it was determined that the loadings of the items in the Reasoning dimension ranged from .88 to .92, and the loadings of the items in the Reflective dimension ranged from .88 to .91. The fact that the factor loadings of all items are well above .40 indicates that the items represent their respective dimensions at a high level. Furthermore, the observed correlation of .77 between the two factors reveals that the dimensions have a significant and strong relationship with each other. When Figure 3 is examined, the model's fit indices were calculated as follows: $\chi^2 = 65.613$, $df = 34$, $p = .01$, $\chi^2/df = 1.526$, $CFI = .99$, $GFI = .97$, $NFI = .99$, $AGFI = .96$, and $RMSEA = .036$. Information regarding the evaluation of these values is presented in detail in Table 5.

Table 5. Compatibility Indices Related to the Mathematical Problem Solving Dimension

| Compatibility Indices | Zhou vd. | The values obtained within the scope of this study. | Acceptable range of compliance | Perfect fit range | Decision (Kline, 2019) |
|-----------------------|----------|---|--------------------------------|-------------------|------------------------|
| X2/df | 3.368 | 1.526 | 3- 5 | 0- 3 | Perfect Fit |
| RMSEA | 0.076 | .036 | ≤.08 | ≤.06 | Perfect Fit |
| NFI | 0.928 | .99 | ≥.90 | ≥.95 | Perfect Fit |
| CFI | 0.948 | .99 | ≥.90 | ≥.95 | Perfect Fit |
| GFI | 0.929 | .97 | ≥.85 | ≥.90 | Perfect Fit |
| AGFI | 0.906 | .96 | ≥.85 | ≥.90 | Perfect Fit |

Table 5 shows that a χ^2/df value below 3 indicates that the model is at an acceptable level. An RMSEA value below .06 (.036) shows that the model has a low error rate and exhibits excellent fit. All other fit indices, NFI, CFI, GFI, and AGFI, are .95 or higher, indicating a high level of structural validity for the model.

Based on the fit criteria proposed by Kline (2019), all indices obtained in this study are found to have an excellent fit level. Furthermore, the fact that these results coincide with the structural model reached by Zhou et al. (2024), who developed the original scale, reveals that the factor structure in the adapted scale has been preserved and validated in accordance with the original structure. Accordingly, it can be stated that the two-factor structure related to the mathematical problem-solving dimension is consistently maintained based on both theoretical grounds and empirical findings.

In the convergent validity analysis, the AVE values for the Reasoning Skills and Reflective Skills sub-dimensions were calculated as .79 and .80, respectively, and the CR values were .96 for both dimensions. The satisfaction of the conditions $AVE \geq .50$ and $CR \geq .70$ indicated that the items represent the factors at a high level and that convergent validity is achieved. In terms of divergent validity, it was found that the AVE value for the Mathematical Problem Solving dimension (.80) is higher than the squares of the correlation coefficients of this dimension with the other dimensions [e.g., $MPS-MCT = (.743)^2 = .552$; $MPS-MMS = (.784)^2 = .614$]. This finding shows that the Reasoning and Reflective components are distinguishable from other constructs in the scale; thus, it strongly supports the overall construct validity of the scale (Fornell & Larcker, 1981; Hair et al., 2019).

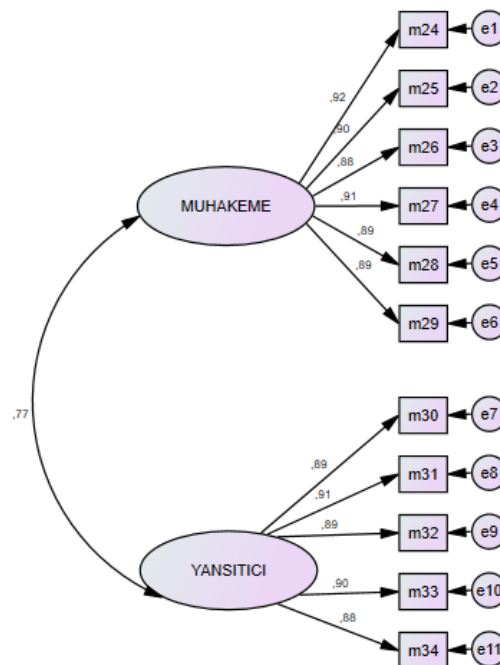


Figure 3. CFA Results Regarding the Problem-Solving Dimension (Muhakeme: Reasoning Skills, Yansitici: Reflective skills)

Mathematical Metacognitive Skills Dimension CFA Results

To test the validity of the three-factor structure of the mathematical metacognitive skills dimension, CFA was applied. This analysis examined the extent to which previous findings were confirmed on the sample and tested the validity of the structure. CFA was performed using the AMOS program, and the resulting model fit indicators are visualized in Figure 4. Figure 4 shows that the model exhibits an excellent fit and strongly ensures the structural validity of the measurement instrument. When examining the standardized factor loadings, it was determined that the loadings of the items in the Knowledge dimension ranged from .81 to .97, those in the Experience dimension from .55 to .57, and those in the Monitoring dimension from .55 to .61. The fact that the factor loadings of all items were well above .40 indicates that the items adequately represent their respective dimensions. Furthermore, correlations between the factors ranging from .45 to .57 revealed that the dimensions are significantly and moderately related to each other. When Figure 4 is examined, the model's fit indices were calculated

as $\chi^2 = 49.55$, $df = 41$, $p = .168$, $\chi^2/df = 1.208$, $CFI = 1.00$, $GFI = .98$, $NFI = .99$, $AGFI = .96$, and $RMSEA = .024$, respectively. Information regarding the evaluation of these values is presented in detail in Table 6.

Table 6. Fit Indices for the Mathematical Metacognitive Skills Dimension

| Compatibility Indices | Zhou vd. | The values obtained within the scope of this study. | Acceptable range of compliance | Perfect fit range | Decision (Kline, 2019) |
|-----------------------|----------|---|--------------------------------|-------------------|------------------------|
| χ^2/df | 3.003 | 1.208 | 3- 5 | 0- 3 | Perfect Fit |
| RMSEA | 0.071 | 0.024 | $\leq .08$ | $\leq .06$ | Perfect Fit |
| NFI | 0.905 | .99 | $\geq .90$ | $\geq .95$ | Perfect Fit |
| CFI | 0.933 | 1.00 | $\geq .90$ | $\geq .95$ | Perfect Fit |
| GFI | 0.933 | .98 | $\geq .85$ | $\geq .90$ | Perfect Fit |
| AGFI | 0.902 | .96 | $\geq .85$ | $\geq .90$ | Perfect Fit |

Table 6 shows that a χ^2/df value below 3 indicates that the model is at an acceptable level. An RMSEA value below .06 (.024) shows that the model has a low error rate and exhibits excellent fit. All other fit indices, NFI, CFI, GFI, and AGFI, are .95 or higher, indicating a high level of structural validity for the model.

Based on the fit criteria proposed by Kline (2019), all indices obtained in this study have an excellent fit level. Furthermore, the fact that these results coincide with the structural model reached by Zhou et al. (2024), who developed the original scale, reveals that the factor structure in the adapted scale has been preserved and validated in accordance with the original structure. Accordingly, it can be stated that the three-factor structure related to the mathematical metacognitive skills dimension is consistently maintained based on both theoretical grounds and empirical findings.

In terms of convergent validity, the AVE values for the three sub-dimensions were calculated as .76, .78, and .77, respectively, and the CR values were .93, .93, and .93. All values being above the acceptable thresholds ($AVE \geq .50$; $CR \geq .70$) indicates that each sub-dimension has convergent validity. In the divergent validity analysis, it was determined that the AVE values for the Mathematical Metacognitive Skills dimension (.77 and above) were higher than the squares of the correlation coefficients with the other dimensions (e.g., $MMS-MCT = (.766)^2 = .587$; $MMS-MPS = (.784)^2 = .614$). This indicated that the Metacognitive Knowledge, Experience, and Monitoring sub-dimensions were statistically distinguishable from each other and from other constructs. Therefore, the findings indicated that the Mathematical Metacognitive Skills dimension demonstrated both convergent and divergent validity, thereby strengthening the overall psychometric structure of the scale.

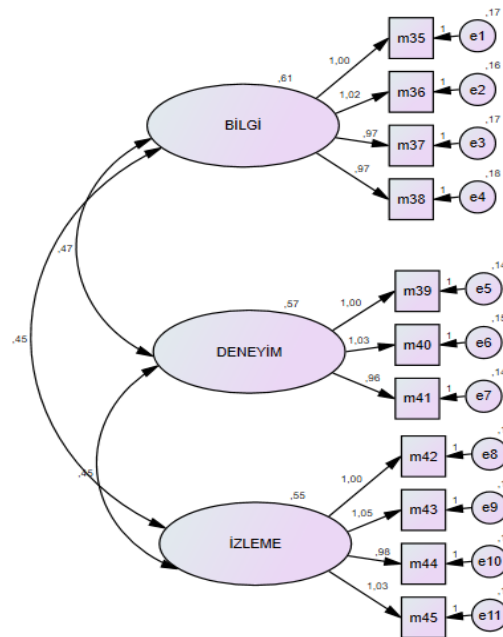


Figure 4. CFA Results Regarding the Mathematical Metacognitive Skills Dimension (Bilgi: M. Knowledge, Deneyim: M. Experience, İzleme: M. Monitoring)

CFA Results for the Entire Scale

This section of the study examines the extent to which previous findings are confirmed on the sample and tests the validity of the construct. DFA was performed using AMOS and JAMOVI programs, and the resulting model fit indicators are visualized in Figure 5. Figure 5 shows that the model exhibits an acceptable level of fit and that the measurement instrument maintains its construct validity.

When standardized factor loadings were examined, it was observed that the loadings of the items in the Mathematical Critical Thinking (Curiosity and Systematics) dimension ranged from .87 to .88; those in the Mathematical Creative Thinking (Flexibility and Innovation) dimension range from .51 to .73; those in the Mathematical Problem Solving (Reasoning Skills and Reflective Skills) dimension ranged from .86 to .89; and those in the Mathematical Metacognitive Skills (Knowledge, Experience, and Observation) dimension ranged from .87 to .91. The fact that all items had factor loadings above .50 indicated that the items adequately represented their respective sub-dimensions.

At the second level, the standardized loadings of the MHOTS, the higher-level factor representing the four sub-dimensions, ranged from .19 to 1.01. These values indicated that the sub-dimensions significantly predicted the overall construct. Furthermore, the correlation coefficients between the sub-dimensions ranged from .21 to .97. The fact that most of these values were above .50 suggests that the dimensions were related but conceptually distinct, thereby supporting the construct validity of the scale.

The overall fit indices of the model also support these results. The fit indices obtained from the analysis were $\chi^2 = 1405.019$, $df = 930$, $\chi^2/df = 1.511$, $CFI = .973$, $NFI = .925$, $GFI = .874$, $AGFI = .860$, and $RMSEA = .035$. These values indicate that the model shows excellent fit in terms of CFI and RMSEA, and that the χ^2/df ratio being below 2 indicates a high level of fit to the data. Information regarding the evaluation of these values is presented in detail in Table 11. The results obtained in the Jamovi analysis also show parallelism with these findings. The confirmatory factor analysis, performed using the WLSMV estimation method yielded $CFI = .998$, $TLI = .998$, $SRMR = .035$, and $RMSEA = .022$, indicating that the model maintains a very good fit in a second sample as well.

Table 11. Fit Indices of the Mathematical Higher-Order Thinking Scale

| Compatibility Indices | Zhou vd. | The values obtained within the scope of this study. | Acceptable range of compliance | Perfect fit range | Decision (Kline, 2019) |
|-----------------------|----------|---|--------------------------------|-------------------|------------------------|
| X ² /df | 3.003 | 1.511 | 3- 5 | 0- 3 | Perfect Fit |
| RMSEA | 0.071 | 0.035 | ≤.08 | ≤.06 | Perfect Fit |
| NFI | 0.905 | .93 | ≥.90 | ≥.95 | Acceptable fit |
| CFI | 0.933 | .97 | ≥.90 | ≥.95 | Perfect Fit |
| GFI | 0.933 | .87 | ≥.85 | ≥.90 | Acceptable fit |
| AGFI | 0.902 | .86 | ≥.85 | ≥.90 | Acceptable fit |

Table 11 shows that the fit indices obtained from the confirmatory factor analysis (CFA) of the MHOTS are generally within acceptable and excellent fit ranges. The X²/df ratio of 1.511 indicates that the model shows an excellent fit. The RMSEA value (.035) also supports the model having a good level of fit. While the CFI (.97) value reflects excellent fit, the NFI (.93), GFI (.87), and AGFI (.86) values are at an acceptable fit level. These results show that the validated factor structure of the Turkish form of the scale is consistent with the original model and provides a good fit to the data.

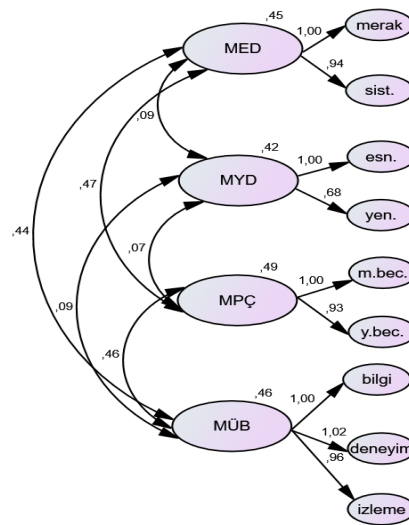


Figure 5. First-Order Confirmatory Factor Analysis Model of the Mathematical Higher-Order Thinking Scale (MED: Mathematical Critical Thinking, MYD: Mathematical Creative Thinking, MPC: Mathematical Problem Solving, MÜB: Mathematical Metacognitive Skills, Merak: Curiosity, Sist.: Systematics, Esn.: Flexibility, Yen.: Innovation, M. Bec.: Reasoning Skills, Y. Bec.: Reflective Skills, Bilgi: M. Knowledge, Deneyim: M. Experience, İzleme: M. Monitoring)

This section of the study examines the extent to which previous findings are confirmed on the sample, and the validity of the construct is tested using second-order confirmatory factor analysis (2nd-order CFA). The analysis was performed using AMOS and Jamovi software, and the resulting model fit indicators are presented in Figure 6. Figure 6 shows that the second-order structure of the model exhibits an acceptable level of fit, and the measurement instrument maintains its construct validity at the overall factor level.

When standardized factor loadings are examined, it is observed that the loadings of the items in the Mathematical Critical Thinking (Curiosity and Systematics) dimension range from .88 to .88; those in the Mathematical Creative Thinking (Flexibility and Innovation) dimension range from .73 to .16; those in the Mathematical Problem Solving (Reasoning Skills and Reflective Skills) dimension range from .86 to .89; and those in the Mathematical Metacognitive Skills (Knowledge, Experience, and Observation) dimension range from .87 to .91. The fact that all items have factor loadings above .50 indicates that the items adequately represent their respective sub-dimensions.

The second-level DFA results showed that the standardized loadings of the MHOTS, representing the higher-level factor that encompasses the four sub-dimensions, ranged from .19 to 1.01. These values

indicate that the sub-dimensions significantly predict the overall construct. Furthermore, the correlation coefficients between the sub-dimensions were observed to range from .19 to .99. These results reveal that the sub-dimensions are related but conceptually distinct, thus supporting the construct validity of the scale.

The overall fit indices of the model also support these results. The fit indices obtained from the analysis were $\chi^2 = 1411.478$, $df = 932$, $\chi^2/df = 1.514$, CFI = .973, TLI = .971, NFI = .925, GFI = .873, AGFI = .859, and RMSEA = .035. These values indicate that the model exhibits an excellent fit in terms of CFI, TLI, and RMSEA, and the χ^2/df ratio being below 2 indicates that the model provides a high level of fit to the data.

The results obtained in the Jamovi analysis are also consistent with these findings. The confirmatory factor analysis performed using the WLSMV estimation method yielded CFI = .998, TLI = .998, SRMR = .035, and RMSEA = .022, indicating that the model provides a very good fit in a second sample as well. Detailed information regarding the evaluation of these values is presented in Table 12.

Table 12. Level 2 CFA Fit Indices for the Mathematical Higher-Order Thinking Scale

| Compatibility Indices | Zhou vd. | The values obtained within the scope of this study. | Acceptable range of compliance | Perfect fit range | Decision (Kline, 2019) |
|-----------------------|----------|---|--------------------------------|-------------------|------------------------|
| X^2/df | 3.003 | 1.514 | 3- 5 | 0- 3 | Perfect Fit |
| RMSEA | 0.071 | 0.035 | $\leq .08$ | $\leq .06$ | Perfect Fit |
| NFI | 0.905 | .93 | $\geq .90$ | $\geq .95$ | Acceptable fit |
| CFI | 0.933 | .97 | $\geq .90$ | $\geq .95$ | Perfect Fit |
| GFI | 0.933 | .87 | $\geq .85$ | $\geq .90$ | Acceptable fit |
| AGFI | 0.902 | .86 | $\geq .85$ | $\geq .90$ | Acceptable fit |

Table 12 presents the results of the second-order CFA for the Mathematical Meta-Order Thinking Scale. The analysis findings show that the model fits the data very well. The X^2/df ratio of 1.514 indicates that the model has an excellent fit. The RMSEA value of .035 also supports this result. While the CFI (.97) value is at an excellent fit level, the NFI (.93), GFI (.87), and AGFI (.86) values are within acceptable fit ranges. These results confirm the statistical validity of the second-order factor structure and that Mathematical Meta-Order Thinking skills are represented holistically.

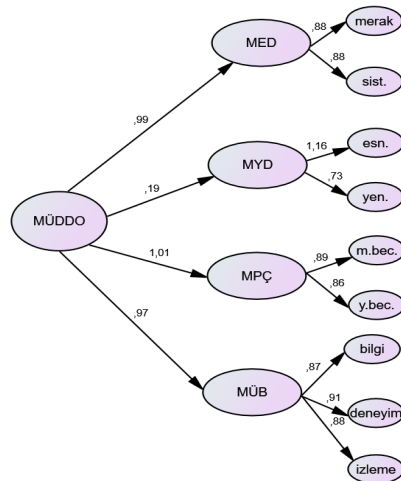


Figure 6. Second-Order CFA Model of the Mathematical Meta-Order Thinking Scale (MÜDDÖ: MHOTS, MED: Mathematical Critical Thinking, MYD: Mathematical Creative Thinking, MPÇ: Mathematical Problem Solving, MÜB: Mathematical Metacognitive Skills, Merak: Curiosity, Sist.: Systematics, Esn.: Flexibility, Yen.: Innovation, M. Bec.: Reasoning Skills, Y. Bec.: Reflective Skills, Bilgi: M. Knowledge, Deneyim: M. Experience, İzleme: M. Monitoring)

Findings of Measurement Constancy Analysis

In this part of the research process, a measurement invariance analysis of the scale was performed based on the gender variable. The analysis aimed to test the configural, metric, scalar, and strict invariance levels based on confirmatory factor analysis. The fit indices of the models created to determine whether the model performed the same way among gender groups are presented in Table 13.

Table 13. Results of the Fit Index of Multiple Group CFA Analysis by Gender

| Types of Invariance | χ^2 | sd | CFI | RMSEA | $\Delta\chi^2$ | Δ sd | Sig. | Δ CFI |
|------------------------------|----------|------|-----|-------|----------------|-------------|------|--------------|
| Configural Invariance | 2481.72 | 1864 | .97 | .028 | - | - | - | - |
| Metric Invariance | 2517.48 | 1900 | .97 | .028 | 35.76 | 36 | .00 | .000 |
| Scalar Invariance | 2530.19 | 1908 | .97 | .028 | 12.71 | 8 | .00 | .001 |
| Residual / Strict Invariance | 2604.91 | 1967 | .96 | .028 | 74.72 | 59 | .00 | .002 |

As shown in Table 13, it was seen that the formal model showed a good level of fit (CFI = .97, RMSEA = .028). Then, in the metric invariance model, the values Δ CFI = .000 and $\Delta\chi^2(36) = 35.76$, $p < .001$, showed that the factor loadings did not differ significantly between the gender groups. Similarly, in the scalar invariance model, Δ CFI = .001 was found and no significant deterioration was observed in the fit indices. Finally, in the model where the error variances were equalized, a value of Δ CFI = .002 was obtained. The fact that the Δ CFI values remain below .010 when transitioning from one model to another shows that the scale provides complete measurement invariance in terms of gender. For this reason, it was observed that this difference remained within acceptable limits. These findings demonstrated that the scale exhibited complete measurement invariance with respect to the gender variable, meaning that the scale measured the same construct in both male and female groups. In other words, scale scores were comparable between groups.

Reliability Findings of the Scale

In this part of the research process, reliability analyses were performed for the Turkish adaptation of the Mathematical Higher-Order Thinking Scale and its sub-dimensions, and the results were reported. In this context, Cronbach's alpha, Guttman's λ coefficient, McDonald's Omega, and Spearman-Brown reliability coefficients, as well as AVE and CR values, were calculated. The results are presented in Table 14.

Table 14. Reliability Coefficients for the Sub-Dimensions and the Whole Scale

| Ana Boyut | Alt Boyut | Cronbach alfa | Guttman's | Omega | Spearman-Brown | AVE | CR |
|-----------------------------------|-------------------|---------------|-----------|-------|----------------|-----|-----|
| Mathematical Critical Thinking | Curiosity | .95 | .95 | .95 | .96 | .77 | .95 |
| | Systematics | .96 | .95 | .95 | .95 | .76 | .96 |
| Mathematical Creative Thinking | Flexibility | .70 | .70 | .70 | .67 | .34 | .72 |
| | Innovation | .70 | .70 | .70 | .72 | .33 | .71 |
| Mathematical Problem Solving | Reasoning Skills | .96 | .96 | .96 | .96 | .79 | .96 |
| | Reflective skills | .95 | .95 | .95 | .95 | .80 | .96 |
| Mathematical Metacognitive Skills | M. Knowledge | .93 | .93 | .93 | .93 | .76 | .93 |

| | | | | | | | |
|-------|-----------------------------------|-----|-----|-----|-----|-----|-----|
| | M. Experience | .91 | .91 | .92 | .90 | .78 | .93 |
| | M. Monitoring | .93 | .93 | .93 | .93 | .77 | .93 |
| MHOTS | Mathematical Critical Thinking | .97 | .96 | .96 | .88 | .77 | .97 |
| | Mathematical Creative Thinking | .81 | .81 | .81 | .78 | .58 | .96 |
| | Mathematical Problem Solving | .96 | .96 | .96 | .96 | .80 | .98 |
| | Mathematical Metacognitive Skills | .95 | .95 | .95 | .90 | .77 | .97 |

As shown in Table 14, the internal consistency coefficients (Cronbach alpha, Guttman's Lambda, Omega, and Spearman-Brown) of all sub-dimensions of the scale were above .70, indicating that the scale has a high level of overall reliability. Specifically, these coefficients are above .90 in the dimensions of "Mathematical Critical Thinking," "Mathematical Problem Solving", and "Mathematical Metacognitive Skills." Furthermore, the fact that most AVE values were above .50 and all CR values were above .70 revealed that the construct had sufficient construct validity and high composite reliability. Only the sub-dimensions of the primary dimension of Mathematical Creative Thinking showed low AVE values. Literature indicates that construct validity is considered acceptable when AVE values are below .50, and CR values are above .70 (Fornell & Larcker, 1981; Hair et al., 2019). These values indicate an acceptable level of reliability and demonstrate that these sub-dimensions have sufficient consistency in terms of measurement (DeVellis, 2017).

In this phase of the study, the relationships between the four sub-dimensions of the scale were examined using Pearson correlation analysis. The obtained correlation coefficients and significance levels are presented in Table 15.

Table 15. Factor Correlations for the MHOTS Scale

| | MCT | MRT | MPS | MMS |
|-----|--------|-------|--------|-----|
| MCT | 1 | | | |
| MRT | .178* | 1 | | |
| MPS | .743** | .228* | 1 | |
| MMS | .765** | .245* | .774** | 1 |

* $p < .05$, ** $p < .01$ (MCT: Mathematical Critical Thinking, MRT: Mathematical Reflective Thinking, MPS: Mathematical Problem Solving, MMS: Mathematical Metacognitive Skills)

Table 15 shows that the correlations between MCT and MMS ($r = .765$, $p < .01$) and MPS ($r = .743$, $p < .01$) are robust. This indicated that critical thinking was highly related to both metacognitive awareness and problem-solving processes. Furthermore, the correlations of the MRT sub-dimension with other variables were relatively lower. For example, while the relationship between MRT and MPS ($r = .228$, $p < .05$) and MÜB ($r = .245$, $p < .05$) was weak, it was statistically significant. These results suggest that the MRT sub-dimension has a more independent structure and weaker connections with other concepts in the scale. In addition, the high correlation between MPS and MMS was noteworthy ($r = .774$, $p < .01$). These findings revealed that students' mathematical problem-solving skills were closely related to their metacognitive competencies and higher-order thinking abilities. Overall, the relationships between the sub-dimensions of the scale are reliable in terms of internal consistency. Therefore, it can be said that the scale can validly and reliably measure the mathematical higher-order thinking skills of high school students.

Discussion

In this study, the Mathematical Meta-Order Thinking Scale, adapted into Turkish, stands out as a theoretically structured measurement tool designed to assess students' mathematical thinking skills in a multidimensional way. The scale's four core dimensions, critical thinking, creative thinking, problem-

solving, and metacognitive skills, encompass the operational aspects of mathematical thinking and its analytical, creative, and self-regulatory dimensions. This structure is consistent with previous research that highlights the need for measurement tools appropriate to the nature of mathematical thinking processes (Schoenfeld, 1992; English & Halford, 2012; Zhang et al., 2024). Indeed, a large portion of the scales in the existing literature focus only on specific cognitive domains; for example, some studies address only metacognitive skills (Güner & Erbay, 2021; Kaplan & Duran, 2016), while others address only creative thinking (Furnham et al., 2013; Kholid et al., 2024). However, due to the nature of cognitive processes, these skills often work in interaction and need to be evaluated holistically within a mathematical context. In this regard, the scale used in this study, which can measure four different mental dimensions simultaneously and in relation to each other, fills an important gap in the literature and offers a unique contribution to measurement and evaluation.

In this context, the fact that the scale used in this study addresses four different cognitive processes together fills an important gap in the literature. The action research conducted by Altun (2024) also emphasizes the need for holistic approaches, especially in developing higher-order thinking skills of prospective teachers. Similarly, the meta-analysis study conducted by Aydın and Öztürk (2024) shows that metacognitive awareness trainings have significant and strong effects on both problem-solving and general mathematics achievement. These findings demonstrate that the Mathematical Higher-Order Thinking Scale is a tool that can be used not only for diagnosis but also for evaluating the effectiveness of instructional interventions.

Furthermore, international and national assessment reports, such as those by the OECD (2023) and İncikabı et al. (2016) point to the importance of students having the ability to analyze, structure, and transfer information to new situations, rather than simply accessing information. The critical thinking, creative thinking, and metacognitive components included in this scale contain elements that directly measure these 21st-century skills. In this respect, the scale is also consistent with the goals emphasized in the Turkish Century of Education Model (MEB, 2024) regarding the development of multifaceted mathematical thinking and higher-order mental competencies.

In conclusion, the Mathematical Higher-Order Thinking Scale, adapted into Turkish in this study, stands out as a robust measurement tool both theoretically and statistically, providing a valuable contribution to the assessment of students' higher-order cognitive skills in the field of mathematics education. The scale is a valuable tool for researchers, teachers, and curriculum developers to monitor students' thinking skills in a multifaceted manner. Further studies could broaden the scale's scope of application by conducting validity analyses for different age groups, cultural contexts, or teaching levels.

Conclusion

In this study, the validity and reliability of the Mathematical Higher-Order Thinking Scale developed by Zhou et al. (2024) were examined in a Turkish sample. In validity studies, exploratory and confirmatory factor analyses were conducted, and for reliability analyses, internal consistency and split-half reliability coefficients were calculated, and as well as interdimensional correlations. In the process of adapting the scale into Turkish, first, the 45 items in the original form were translated into Turkish by three independent foreign language experts and three mathematics educators. These translations were then combined to create a common text. Subsequently, this text was translated into English using the back-translation method by three different foreign language educators, and its conceptual and semantic equivalence was verified according to Brislin's (1970) method. Following final revisions based on expert opinions, the resulting Turkish form was piloted with 30 high school students, and minor linguistic simplifications were made according to the students' level of understanding. The final version of the scale retains the four main dimensions and nine sub-dimensions from the original form and is structured as a 5-point Likert scale.

The results of the DFA conducted in the study showed that the structure determined by exploratory analyses was validated on the sample data. When evaluated in terms of key indicators, such as model fit indices (χ^2/df , RMSEA, CFI, GFI, AGFI, and NFI), it was observed that the models for all

sub-dimensions had an excellent fit level (Kline, 2019). In particular, the RMSEA values being below .06 and the CFI, GFI, and NFI values being close to or above .95 indicate that the factor structures in the scale have been empirically validated. In addition, as a result of the analyses performed to determine the measurement invariance of the scale with respect to the gender variable, all formal, metric, scalar, and error invariance models showed acceptable fit. The Δ CFI values remaining below .010 at all stages reveal that the scale provides complete measurement invariance with respect to gender. This finding demonstrates that the Mathematical Higher-Order Thinking Scale is a valid measure of the same construct in both male and female groups, and that scale scores are comparable across gender groups.

Reliability analyses of the scale revealed that Cronbach's α , Guttman's λ , McDonald's Omega, and Spearman-Brown coefficients for both the sub-dimensions and the overall scale mainly were above .90. This indicates that the scale has high internal consistency and provides reliable measurements. The reliability coefficients for the "Flexibility" and "Innovation" components within the "Mathematical Creative Thinking" sub-dimension, which are around .70, show that these sub-dimensions are also acceptably reliable (DeVellis, 2017). When all these findings are considered together, it can be said that the Turkish adapted form of the Mathematical Higher-Order Thinking Scale is both a valid and reliable measurement tool. The scale provides a functional structure for researchers and practitioners to evaluate the mathematical thinking skills of high school students in a multidimensional manner.

Recommendations

This study makes significant contributions to the adaptation of the Mathematical Higher-Order Thinking Scale into Turkish and to the performance of its validity and reliability analyses. However, it also has some limitations. First, this study was conducted only with high school students. This limits the generalizability of the validity and reliability findings of the scale to different age groups. In future studies, it is recommended that the scale be applied to different age and education groups, such as middle school, university, and adult individuals, and that comparative analyses be conducted according to age and developmental level. Second, the data obtained in this study were collected through self-reports. Data obtained based on participants' self-reports may be subject to effects, such as social desirability bias. Therefore, it is recommended that multiple data sources, such as teacher observations, portfolios, or performance evaluations based on problem-solving processes, be used in future research. Third, this study focused only on the psychometric properties and did not investigate the relationships between the dimensions included in the scale and students' academic achievement or other cognitive/affective variables. Future research could strengthen criterion validity by examining the relationship between higher-order mathematical thinking skills and variables, such as academic achievement, math anxiety, self-regulation, problem-solving competence, creativity, and critical thinking. Finally, cross-cultural comparisons could be conducted across different socio-cultural regions to assess the relationship between the constructs included in the scale and the cultural context. Such studies would provide important data for determining the cross-cultural validity level of the scale.

Declarations

- 1. Author Contributions Statement:** This study was jointly conducted by two authors. The research design, data collection process, data analysis, manuscript drafting, and scale adaptation procedures were undertaken by Author 1. The literature review, development of the theoretical framework, interpretation of findings, and revision of the manuscript in terms of academic language were carried out by Author 2. The final review, formatting, and submission of the article were completed through the joint contribution of both authors. All authors have reviewed and approved every stage of this study, including the final version.
- 2. Conflict of Interest:** The authors declare that there is no conflict of interest.
- 3. Ethics Approval:** Ethics approval for this study was granted by the Ethics Committee of Bayburt University Rectorate with the decision dated 20.02.2025, numbered 43, 2nd session (Document Verification Code: BSFBL18BSC), confirming unanimously that the research proposal complies with ethical principles.
- 4. Research Design:** This study is a quantitatively designed research article employing a survey model to examine high school students' mathematical higher-order thinking skills. In the survey model, the existing situation is defined and analyzed as it naturally occurs within a predetermined group (Karasar, 2000). The scale adaptation, validity, and reliability procedures were planned within this framework. To ensure the linguistic and conceptual equivalence of the scale, Brislin's (1970) back-translation technique was used, and linguistic simplifications were made based on a pilot study conducted with 30 students. The suitability of the data set for factor analysis was assessed using the KMO and Bartlett tests and skewness–kurtosis indices to evaluate normality. Construct validity was examined using exploratory factor analysis in SPSS (principal components + direct oblimin) and confirmatory factor analysis in LISREL (χ^2/df , RMSEA, CFI, GFI, AGFI, NFI). Internal consistency was evaluated using Cronbach's α , Guttman's λ , McDonald's omega, and Spearman–Brown coefficients; structural adequacy using AVE and CR values; and the coherence among subdimensions through correlation analysis. This methodological framework confirms that the scale is a valid and reliable measurement instrument.

TÜRKÇE SÜRÜM

Giriş

21. yüzyılda değişen bilgi üretim biçimleri, bireylerden yalnızca bilgiye ulaşmalarını değil, ulaştıkları bilgiyi analiz etme, değerlendirme ve yeni durumlara transfer etme gibi üst düzey bilişsel süreçleri de etkili bir şekilde kullanmalarını beklemektedir (Kurnaz, 2019; Topal vd., 2025). Özellikle eğitim sistemleri, bu becerilerin erken yaşlardan itibaren kazandırılması gerektiğini vurgulamakta; öğrencilerin eleştirel düşünme, yaratıcı düşünme, problem çözme ve üstbilişsel farkındalık gibi zihinsel yeterlikleri geliştirmelerine öncelik vermektedir (Altun, 2024). Matematik eğitimi ise, soyutlama, genelleme ve çıkarım yapma gibi zihinsel süreçlerin temelinde yer alan doğası gereği, bu üst düzey düşünme becerilerinin geliştirilebilmesi için uygun bir disiplin olarak kabul edilmektedir (English & Halford, 2012; Lin, 2023).

Bu doğrultuda, uluslararası ölçme değerlendirme sistemleri olan PISA ve TIMSS gibi programlarda da öğrencilerin sadece işlemsel bilgi düzeyinde değil, aynı zamanda akıl yürütme, muhakeme ve problem çözme gibi üst düzey bilişsel süreçlerde ne derece başarılı oldukları sistematik olarak değerlendirilmektedir (İncikabı vd., 2016; OECD, 2023). Türkiye bağlamında da bu yaklaşımı destekleyen biçimde güncellenen Türkiye Yüzyılı Maarif Modelinde matematik dersi öğretim programı (MEB, 2024); öğrencilerin matematiksel düşünme becerilerini çok boyutlu olarak geliştirmeyi, bu sayede onları bilgi toplumunun aktif bireyleri hâline getirmeyi amaçlamaktadır. Programda özellikle eleştirel ve yaratıcı düşünmenin geliştirilmesi, problem çözme süreçlerinin yapılandırılması ve üstbilişsel farkındalığın kazandırılması, temel yeterlilikler arasında sayılmaktadır.

Bu zihinsel yeterliklerin öğrenme süreçlerinde geliştirilebilmesi kadar, geçerli ve güvenilir biçimde ölçülmesi de son derece önemlidir. Ancak literatürde bu becerilerin ölçümüne yönelik geliştirilen ölçeklerin çoğu ya matematik alanından bağımsız kurgulanmış ya da yalnızca tek bir bilişsel boyuta (örneğin sadece yaratıcı düşünme ya da yalnızca problem çözme) odaklanmıştır (Gürefe & Akçakin, 2018; Korkmaz & Karakuş, 2009; Ünal & Aral, 2014). Buna ek olarak, matematiksel içeriğe özgü, kuramsal temele dayalı ve çok boyutlu ölçüm yapabilen psikometrik araçların sayısının sınırlı olması, eğitimde üst düzey düşünme becerilerinin bütüncül olarak değerlendirilmesini zorlaştırmaktadır. Nitekim öğrencilerin matematiksel düşünme becerileri, çoğu zaman birden fazla bilişsel sürecin etkileşim hâlinde yürütülmesini gerektirmekte ve bu süreçlerin birbirinden ayrıştırılmadan değerlendirilmesi gerekmektedir (Schoenfeld, 1992). Bu bağlamda, Zhou vd. (2024) tarafından geliştirilen Matematiksel Üst Düzey Düşünme Ölçeği (MÜDDÖ), öğrencilerin bu çok katmanlı zihinsel süreçlerini dört temel boyutta ele alarak değerlendirme imkânı sunmaktadır. Söz konusu ölçek; matematiksel eleştirel düşünme, matematiksel yaratıcı düşünme, matematiksel problem çözme ve matematiksel üstbilişsel beceriler olmak üzere dört alt boyuttan oluşmaktadır.

Matematiksel Eleştirel Düşünme

Matematiksel eleştirel düşünme, bireyin matematiksel bir durumu çok boyutlu biçimde analiz ederek bilgi üretme, yargıya varma ve karar verme becerisini içeren üst düzey bilişsel bir süreçtir. Bu boyut, yalnızca matematiksel işlemleri doğru yapmayı değil, aynı zamanda düşünme süreçlerini sorgulama, açıklama ve gerekçelendirme becerisini de kapsar (Monteleone vd., 2023). Eleştirel düşünme; yorumlama, analiz etme, değerlendirme, açıklama ve yaratma gibi becerileri içerdiği gibi, bu becerilerin matematiksel bağlamda uygulanmasıyla "eleştirel matematiksel düşünme" olarak tanımlanır. Öğrencilerin matematiksel çözüm süreçlerini sorgulama, yargı üretme ve bu yargıları gerekçelendirme davranışları, eleştirel matematiksel düşünmenin temel göstergeleri arasında yer almaktadır. Bu bağlamda, öğrencinin bir matematiksel problemi çözmekle kalmayıp, çözüm yolunu açıklaması, kullanılan stratejiyi gerekçelendirmesi ve alternatif çözüm yollarını ele alabilmesi, onun eleştirel matematiksel düşünme becerisine sahip olduğunu gösterir (Samura & Juandi, 2020). Bu beceri yalnızca bireyin bireysel düşünme yetisini değil, aynı zamanda sınıf içi tartışmalarda düşünce savunma ve

başkalarının düşüncelerini eleştirel biçimde değerlendirme gibi sosyal öğrenme süreçlerini de desteklemektedir (Sachdeva & Eggen, 2021).

Matematiksel eleştirel düşünme, yaratıcı düşünme, problem çözme ve üstbilişsel becerilerle yakından ilişkilidir; çünkü birey, kendi düşünme sürecini hem izler hem de yeniden yapılandırarak daha derinlemesine çözüm stratejileri geliştirebilir (Zhang vd., 2024; Monteleone vd., 2023). Bu açıdan bakıldığında, matematiksel eleştirel düşünme, sadece bireysel bir beceri değil; aynı zamanda yapıcı, çok yönlü ve anlam temelli bir öğrenme sürecidir. Bu yönüyle, matematiksel eleştirel düşünme öğrencinin çözüm sürecini sorgulamasına, değerlendirmesine ve gerekçelendirmesine imkân tanırken; bu sürecin farklı, özgün ve alışılmadık yollarla yeniden kurgulanması ise matematiksel yaratıcı düşünmenin temelini oluşturmaktadır.

Matematiksel Yaratıcı Düşünme

Matematiksel yaratıcı düşünme, bireyin matematiksel bir problemi alışılmış kalıpların ötesine geçerek çok yönlü ve özgün biçimlerde ele almasını sağlayan üst düzey bir bilişsel süreçtir. Bu beceri, yalnızca doğru çözüme ulaşmayı değil; çözüm yollarını çeşitlendirme, özgün düşünce geliştirme ve bu süreçte aktif bilişsel üretkenlik sergilemeyi de kapsar. Bu boyut, yalnızca matematiksel bir probleme doğru çözümler üretmeyi değil; aynı zamanda çok sayıda geçerli ve anlamlı fikir üretebilme (akıcılık) ile bu fikirler arasında özgün ve sıra dışı olanları geliştirebilme (yenilik) yetilerini birlikte içermektedir (Suherman & Vidákovich, 2022). Akıcılık, öğrencinin aynı probleme yönelik birden fazla çözüm yolu ortaya koyabilme kapasitesini ifade ederken; yenilik, bu yollar arasından özgün olanları seçebilme ve yaratıcı biçimde yeniden yapılandırabilme becerisini yansıtır (Fortes & Andrade, 2019; Sriraman, 2009). Bu bağlamda, matematiksel yaratıcı düşünme, öğrencinin yalnızca çözüm üretmesini değil; bu çözümleri anlamlı biçimde gerekçelendirmesini, alternatif yollar önermesini ve özgün matematiksel düşünce yapılarını ifade edebilmesini de kapsayan çok boyutlu bir bilişsel yeterlilik (Furnham vd., 2013; Kholid vd., 2024). Matematiksel yaratıcı düşünme, öğrencilerin çoklu ve özgün çözüm yolları üretmesini desteklerken; bu çözüm yollarının yapılandırılması, değerlendirilmesi ve etkili biçimde uygulanması süreci problem çözme becerisini de gerekli kılmaktadır.

Matematiksel Problem Çözme

Matematiksel problem çözme, bireyin karşılaştığı bir matematiksel durumu anlayarak uygun çözüm yolları geliştirmesi, bu yolları uygulaması ve süreci değerlendirmesiyle oluşan üst düzey bir bilişsel süreçtir. Bu boyut, yalnızca işlemleri doğru yapmayı değil; verilenleri yorumlama, strateji üretme, çözüm sürecini izleme ve değerlendirme gibi becerileri de kapsar (Öztürk vd., 2023). Problem çözme becerisi, bu çalışmada muhakeme becerileri ve yansıtıcı beceriler olmak üzere iki temel alt boyutta ele alınmaktadır. Muhakeme becerileri, öğrencinin problemdeki veriler arasında mantıksal ilişkiler kurarak çıkarım yapmasını ve çözüm yolunu bu çıkarımlara göre planlamasını içerir (Suherman & Vidákovich, 2022). Özellikle oran-orantı ve örüntü gibi yapılandırılmış problemlerde, öğrencilerin doğru çözüm üretmesi, büyük ölçüde muhakeme süreçlerini ne kadar etkili kullandıklarına bağlıdır (Öztürk vd., 2021). Yansıtıcı beceriler ise öğrencinin çözüm sürecini izlemesi, seçtiği stratejilerin işlevselliğini değerlendirmesi ve gerektiğinde süreci yeniden düzenleyebilmesini ifade eder (Carbonneau vd., 2020). Bu beceri, öğrencinin yalnızca sonucu değil, süreci de değerlendirmesini sağlar ve özellikle üstbilişsel farkındalıkla desteklendiğinde, problem çözme başarısını anlamlı biçimde artırır (Suherman & Vidákovich, 2022). Bu bağlamda matematiksel problem çözme, yalnızca sonuca ulaşmakla sınırlı olmayan; düşünmeyi planlama, izleme ve yapılandırma süreçlerini içeren bütüncül bir öğrenme becerisidir. Bu bağlamda matematiksel problem çözme, yalnızca sonuca ulaşmakla sınırlı olmayan; düşünmeyi planlama, izleme ve yapılandırma süreçlerini içeren bütüncül bir öğrenme becerisidir. Problem çözme sürecinin bu bilişsel çok katmanlı yapısı, bireyin kendi düşünme süreçlerine ilişkin farkındalığını, deneyimlerini ve bu süreçleri bilinçli biçimde yönetebilmesini gerektirir; bu ise matematiksel üstbilişsel becerilerin temelini oluşturmaktadır.

Matematiksel Üstbilişsel Beceriler

Matematiksel üstbilişsel beceriler, bireyin kendi matematiksel düşünme süreçlerinin farkında olması, bu süreçleri planlaması, izlemesi ve gerektiğinde yeniden yapılandırmasıyla ilişkili çok katmanlı bir bilişsel özdüzenleme sistemidir (Muncer vd., 2021; Suliani vd., 2024). Bu beceri, yalnızca bilgiye sahip olmayı değil; bilginin nasıl, ne zaman ve neden kullanılacağını anlamayı ve bu kullanım sürecini bilinçli biçimde yönetmeyi içerir (Güner & Erbay, 2021). Matematiksel üstbilişsel beceriler, üç temel bileşen altında incelenmektedir: üstbilişsel bilgi, üstbilişsel deneyim ve üstbilişsel izleme.

Üstbilişsel bilgi, bireyin kendi bilişsel özelliklerine, öğrenme stratejilerine ve görevlerin gerekliliklerine dair bilgi sahibi olmasını ifade eder (Suliani vd., 2024). Bu bilgi türü; neyi bildiğini bilme (bildirimsel), nasıl yapılacağını bilme (yordam bilgisi) ve ne zaman-hangi durumda nasıl bir strateji kullanılacağını bilme (koşul bilgisi) olmak üzere üç temel bileşeni kapsar (Güner & Erbay, 2021). Üstbilişsel bilgisi gelişmiş öğrenciler, kendi yeterliliklerini daha gerçekçi değerlendirebilir ve karşılaştıkları problem türüne uygun stratejileri seçerek daha etkili öğrenme süreçleri geliştirebilir (Muncer vd., 2021).

Üstbilişsel deneyim, öğrencinin bir matematiksel görev sırasında yaşadığı duygu, düşünce ve yargıların farkında olması ve bu içsel yaşantılara göre sürece yön vermesidir (Yüksel vd., 2021). Bu deneyim, bireyin belirli bir problemi çözerken ne hissettiğini ne bildiğini düşündüğünü ve ne kadar emin olduğunu yansıması açısından oldukça kritiktir (Suliani vd., 2024). Özellikle problem çözme sırasında hissedilen zorluk ya da başarı duyguları, bireyin strateji tercihlerinde değişiklik yapmasına neden olabilir ve bu durum, sürecin bilinçli yönetilmesinde etkili olur (Şahinkaya vd., 2022).

Üstbilişsel izleme ise bireyin matematiksel düşünme sürecini gerçek zamanlı olarak gözlemlemesi, çözüm sürecini kontrol etmesi ve gerektiğinde müdahale ederek süreci yeniden düzenlemesidir (Güner & Erbay, 2021). Bu izleme becerisi, öğrencinin çözüm sırasında nerede olduğunu fark etmesini, hatalarını tanımasını ve strateji değişikliği yaparak süreci doğru yöne taşımalarını sağlar (Suliani vd., 2024). Araştırmalar, bu becerinin gelişmiş olmasının öğrencinin sadece sonucu değil, süreci de başarıyla yönetmesini sağladığını ve matematik başarılarını anlamlı biçimde artırdığını göstermektedir (Aydın & Öztürk, 2024). Bu bağlamda matematiksel üstbilişsel beceriler, öğrencinin matematiksel öğrenme sürecinde bilgiye nasıl eriştiğini, bu bilgiyi nasıl yapılandırdığını ve süreci nasıl değerlendirdiğini yansıtan; öz farkındalıkla bütünleşmiş, derinlemesine öğrenmeyi destekleyen temel bir bilişsel sistemdir (Suliani vd., 2024; Muncer vd., 2021).

Türkiye’de son yıllarda üstbilişsel beceriler, yaratıcı ve eleştirel düşünme ile matematiksel problem çözme süreçlerine ilişkin çeşitli ölçek geliştirme ve uyarlama çalışmaları gerçekleştirilmiştir. Bununla birlikte, üstbilişsel farkındalığın problem çözme süreciyle doğrudan ilişkilendirildiği ölçme araçlarının sayısı oldukça sınırlıdır. Türkiye’de geliştirilen ölçeklerin çoğu ya genel üstbilişsel farkındalığı ölçmekte ya da problem çözme süreçlerini dolaylı biçimde ele almaktadır. Bu çalışmaların büyük bölümü, genel üstbilişsel yapıları ele almakta veya matematiksel bağlamla sınırlı düzeyde ilişki kurmaktadır. Örneğin, Fırat Durdukoca ve Arıbaş (2019) ile Kaplan ve Duran (2016) tarafından geliştirilen ölçekler, genel üstbilişsel farkındalık düzeyini ölçmeye yönelik olup, matematiksel içeriğe özgü yapılandırılmamıştır. Çanakçı ve Özdemir (2011) ile Hacıömeroğlu (2011) tarafından geliştirilen ölçekler ise sırasıyla matematik problem çözmeye ilişkin tutum ve inançları ölçmeye yönelik olup, üst düzey bilişsel süreçleri kapsam dışı bırakmıştır. Ayrıca Suliani vd. (2024), Muncer vd. (2021) ve Yüksel vd. (2021) gibi çalışmalar, bireylerin matematiksel düşünme süreçlerinde üstbilişsel becerilerin önemini vurgulamakla birlikte, bu yapıları doğrudan ölçen ölçeklerin sayıca yetersiz olduğunu ifade etmektedir.

Bu genel tabloya paralel biçimde, Türkiye’de matematiksel problem çözme alanında geliştirilen ve uyarlanan ölçme araçlarının büyük bölümü, değerlendirdikleri alan bakımından tutum, beceri, inanç ve başarı odaklıdır. Örneğin tutum ölçekleri öğrencilerin problem çözmeye yönelik eğilim ve motivasyonlarını ölçmeyi amaçlamıştır (Çanakçı, 2008; Uğurluoğlu, 2008). Beceri ölçekleri, bireylerin problem çözme performansını değerlendirmeye yöneliktir (Çömlekoğlu, 2001; Özsoy, 2008; Uysal, 2007; Kızılkaya & Aşkar, 2009; Özpinar, 2012; Muyo, 2015; Çelik, 2017; Atlıhan, 2021; Gürsoy &

Çeliköz, 2021). Ancak bu araçlarda problem çözme sürecinin planlama, izleme, kontrol ve değerlendirme gibi üstbilişsel düzenleme boyutları çoğunlukla göz ardı edilmiştir.

Benzer şekilde, inanç ölçekleri öğrencilerin problem çözmeye yönelik epistemolojik inançlarını ve öz yeterlik algılarını ölçmüştür (Kayan, 2008; Hacıömeroğlu, 2011; Gümüş & Umay, 2017). Bu ölçekler süreçteki bilişsel farkındalık yerine, problem çözme hakkındaki öznel inançları değerlendirmektedir. Başarı testleri (Sadık, 2006; Güneş, 2022; Süzer, 2022; Yaşar, 2022) ise problem çözme performansını ölçmekte, fakat bireylerin kendi çözüm süreçlerine ilişkin üstbilişsel farkındalıklarını doğrudan değerlendirmemektedir. Çalışma grupları bakımından yapılan incelemede, geliştirilen ölçeklerin ilkökul ikinci sınıftan üniversite düzeyine kadar geniş bir yelpazede uygulandığı görülmektedir (Değirmenci & Deringöl, 2024; Çanakçı, 2008; Uysal, 2007; Çelik, 2017; Özpınar, 2012; Atlıhan, 2021; Özsoy, 2008; Sadık, 2006; Süzer, 2022; Yaşar, 2022; Hacıömeroğlu, 2011; Kayan, 2008; Muyo, 2015; Çömlekoğlu, 2001; Gümüş & Umay, 2017; Gürsoy & Çeliköz, 2021). Bu durum, farklı yaş ve eğitim düzeylerinde problem çözme becerilerinin ölçülmesine yönelik kapsamlı çalışmalar yapılmış olmasına karşın, özellikle lise öğrencilerinin matematiksel problem çözme sürecindeki üstbilişsel farkındalıklarını, yaratıcı ve eleştirel düşünme becerileriyle birlikte değerlendiren geçerli ve güvenilir bir ölçme aracının henüz bulunmadığını göstermektedir. Bu bağlamda, eleştirel düşünme, yaratıcı düşünme, problem çözme ve üstbilişsel düşünme boyutlarını bütüncül biçimde ele alan Zhou vd. (2024) tarafından geliştirilen ölçeğin Türkçeye uyarlanması, alandaki bu çok boyutlu ölçme ihtiyacını karşılaması bakımından önemli bir katkı sunmaktadır.

Bu nedenle, mevcut çalışma yeni bir ölçek geliştirmekten ziyade, kuramsal temelleri sağlam ve çok boyutlu bir yapıya sahip olan Zhou vd. (2024) tarafından geliştirilen Mathematical Higher-Order Thinking Scale (MHOTS) ölçeğinin Türkçeye uyarlanmasını amaçlamaktadır. Bu yaklaşım, özgün ölçeğin uluslararası geçerliliği ve kapsamlı faktör yapısından yararlanmayı, aynı zamanda Türk eğitim bağlamına uygun kültürel ve dilsel uyarlamaların yapılmasını mümkün kılmıştır. Uyarlanan ölçek, öğrencilerin matematiksel düşünme süreçlerini çok boyutlu biçimde değerlendirebilme açısından özgün bir yapıya sahiptir. Zhou vd. (2024) tarafından geliştirilen ölçek; eleştirel düşünme, yaratıcı düşünme, problem çözme ve üstbilişsel düşünme boyutlarını bütüncül bir çerçevede ele alarak, matematiksel düşünmenin bilişsel, duyuşsal ve üstbilişsel yönlerini birlikte ölçmeyi amaçlamaktadır. Türkiye’de matematiksel düşünme becerilerini ölçmeye yönelik mevcut araçların genellikle tek boyutlu olduğu göz önüne alındığında, bu ölçeğin Türkçeye kazandırılması alandaki önemli bir boşluğu doldurmaktadır. Ayrıca, Türkiye Yüzyılı Maarif Modeli’nde vurgulanan üst düzey düşünme ve problem çözme becerilerinin değerlendirilmesine katkı sunması bakımından, çalışma ulusal eğitim politikalarıyla da doğrudan örtüşmektedir (MEB, 2024). Dolayısıyla, ölçeğin Türkçe uyarlanması hem kuramsal hem de uygulamalı düzeyde matematik eğitimi alanına anlamlı bir katkı sağlamaktadır.

Amaç

Bu çalışmada Türkçeye uyarlanan MÜDDÖ, öğrencilerin matematiksel bağlamdaki bilişsel, üstbilişsel, yaratıcı ve eleştirel düşünme becerilerini çok boyutlu ve bütüncül bir yaklaşımla değerlendirme olanağı sunarak alanyazındaki önemli bir boşluğu doldurmayı ve özgün, işlevsel bir katkı sağlamayı amaçlamaktadır. Araştırmanın temel amacı, Zhou vd. (2024) tarafından geliştirilen ölçeği Türkçeye uyarlamak, ölçeğin Türk örnekleminde geçerlik ve güvenilirlik kanıtlarını ortaya koymak ve ölçme aracının psikometrik özelliklerini incelemektir. Bu kapsamda, ölçeğin dört boyutlu yapısının Türk kültürel bağlamında da geçerliliğini koruyup korumadığı araştırılmıştır.

Uyarlanan ölçek, yalnızca akademik araştırmalar için değil; aynı zamanda öğretmenlerin sınıf içi uygulamalarında öğrencilerin üst düzey düşünme becerilerini tanılaması, izleyebilmesi ve geliştirmeye yönelik stratejiler planlaması açısından da önemli bir araç niteliğindedir. Araştırmacılar için ölçek, matematiksel düşünmenin bilişsel, duyuşsal ve üstbilişsel boyutlarını bütüncül biçimde analiz etmeye imkân tanımakta; politika yapıcılar açısından ise 21. yüzyıl becerilerini ve Türkiye Yüzyılı Maarif Modeli’nde vurgulanan üst düzey düşünme kazanımlarını somut biçimde değerlendirebilecek bir ölçme aracı sunmaktadır.

Bu çalışmanın bulguları, hem öğretim programlarının etkililiğinin izlenmesine hem de öğretmen yetiştirme süreçlerinde üst düzey düşünme becerilerini destekleyen kanıta dayalı uygulamaların geliştirilmesine önemli katkılar sunmaktadır. Araştırmadan elde edilen sonuçların, matematik eğitimi alanında öğrencilerin üst düzey düşünme becerilerinin ölçülmesi, izlenmesi ve geliştirilmesine yönelik çalışmalara yön vereceği öngörülmektedir. Uyarlanan bu ölçek, öğrencilerin matematiksel bağlamda bilişsel, üstbilişsel, yaratıcı ve eleştirel düşünme becerilerini çok boyutlu ve bütüncül bir biçimde değerlendirmeye olanak tanıyarak alanyazında var olan önemli bir boşluğu doldurmada ve alana özgün, işlevsel bir katkı sunmaktadır.

Yöntem

Bu araştırma, lise öğrencilerinin matematiksel üst düzey düşünme becerilerini değerlendirmek amacıyla Zhou vd. (2024) tarafından geliştirilen ve özgün adı "Mathematical Higher-order Thinking Scale (MHOTS)" olan Matematiksel Üst Düzey Düşünme Ölçeğini (MÜDDÖ) Türk kültürüne uyarlamayı amaçlamaktadır. Bu doğrultuda, araştırmada nicel araştırma yöntemlerinden tarama modeli tercih edilmiştir. Tarama modeli, mevcut bir durumu, olguyu ya da özelliği, olduğu şekliyle ve doğal ortamı içinde tanımlamayı amaçlayan bir araştırma desendir (Karasar, 2000). Çalışmada, MÜDDÖ ölçeğinin Türkçeye uyarlanması süreci kapsamında geçerlik ve güvenilirlik analizleri yapılmış; ölçeğin yapı geçerliği doğrulayıcı faktör analizleriyle incelenmiştir. Ayrıca, ölçeğin iç güvenilirlik katsayıları hesaplanarak her bir boyutun tutarlılığı değerlendirilmiş ve boyutlar arasındaki korelasyonlar incelenerek ölçeğin yapısal bütünlüğü desteklenmiştir.

Çeviri Çalışması

Uyarlama süreci, ulusal ve uluslararası literatürde kabul görmüş ölçütler doğrultusunda gerçekleştirilmiştir (Hambleton vd., 2005). İlk aşamada, özgün ölçekte yer alan 45 madde, İngilizce-Türkçe dil yeterliliği olan üç yabancı dil uzmanı ve üç matematik eğitimi uzmanı tarafından bağımsız olarak Türkçeye çevrilmiştir. Daha sonra, her çeviri birleştirilerek uzlaşa sağlanmış ve bu çeviri metni, alanında uzman ilk çeviriyi yapan uzman ekibinden farklı üç yabancı dil eğitimcisi tarafından geriye çeviri (back-translation) yöntemiyle İngilizceye yeniden çevrilmiştir. Bu süreç, çevirilerin kavramsal ve anlamsal bütünlüğünü koruyup korumadığını değerlendirmek amacıyla yapılmıştır (Brislin, 1970). Geriye çeviriler ile özgün ölçek karşılaştırılmış ve anlam bütünlüğü taşıdığı görülmüştür.

Uzman görüşleri doğrultusunda yapılan son düzenlemelerle Türkçe form oluşturulmuş ve ölçeğin pilot uygulaması, 30 kişilik bir lise öğrencisi grubuna uygulanmıştır. Bu çalışmada ölçek formu, öğrencilerin ifadeleri anlama düzeyleri açısından değerlendirilmiş ve bu doğrultuda madde yapısında küçük sadeleştirmeler yapılmıştır. Pilot uygulamadan elde edilen geri bildirimlere göre ölçek, nihai hâline getirilerek uzman ekibin oy birliğiyle uygulamaya hazır duruma getirilmiştir. Türkçeye uyarlanan ölçekte, özgün formda olduğu gibi dört temel boyut (Matematiksel Eleştirel Düşünme, Matematiksel Yaratıcı Düşünme, Matematiksel Problem Çözme ve Matematiksel Üstbilişsel Beceriler) ile bu boyutların altında yer alan dokuz alt boyut korunmuştur. Ölçek, özgün çalışmada olduğu gibi 5'li Likert tipi (1 = Hiç katılmıyorum, 5 = Tamamen katılıyorum) olarak yapılandırılmıştır. Bu bağlamda, gerçekleştirilen çeviri ve uyarlama süreci nitel düzeyde kültürel ve dilsel eşdeğerliğin sağlanmasına odaklanmış; uzman görüşleri ve pilot uygulama sonuçları doğrultusunda ölçeğin Türk lise öğrencileri için anlaşılabilir ve bağlama uygun bir biçime kavuştuğu görülmüştür. Ancak, bu süreçte dilsel eşdeğerlik yalnızca nitel olarak incelendiğinden, elde edilen bulguların dilsel açıdan sınırlı bir değerlendirme sunduğu dikkate alınmalıdır.

Araştırma grubu

Çalışmanın araştırma grubunu, 2024-2025 eğitim öğretim yılında Türkiye'nin Kuzeydoğusunda yer alan bir ilde öğrenim görmekte olan lise öğrencileri oluşturmaktadır. Araştırma grubunda, 210 kadın (%50,36) ve 207 erkek (%49,64) olmak üzere toplam 417 lise öğrencisi bulunmaktadır. Katılımcıların sınıf seviyelerine göre dağılımı ise, 87 (%20,9) lise 1. sınıf, 107 (%25,7) lise 2. sınıf, 109 (%26,1) lise 3. sınıf ve 114 (%27,3) lise 4. sınıf öğrencisidir. Alanyazında faktör analizlerinin geçerli sonuç

verebilmesi için önerilen alt sınır 400 olarak kabul edilmektedir (Goretzko vd., 2021; Güvendir, Vd., 2025). Bu doğrultuda araştırmanın örneklem büyüklüğünün yeterli olduğu görülmektedir.

Veri Toplama Aracı

MÜDDÖ, Zhou vd. (2024) tarafından lise öğrencilerinin matematiksel üst düzey düşünme becerilerini değerlendirmek amacıyla geliştirilmiştir. Ölçek geliştirme sürecinde, matematik eğitimi ve ölçme değerlendirme alanında uzman kişilerin katkılarıyla toplam 53 madde belirlenmiş; bu maddelerden uzman görüşü ve analizler doğrultusunda 45 maddeye indirgenmiştir. Ölçek, matematiksel düşünme ile ilişkili dört temel boyut içermektedir: matematiksel eleştirel düşünme, matematiksel yaratıcı düşünme, matematiksel problem çözme ve matematiksel üstbilişsel beceriler. Her temel boyut, kendi içerisinde alt boyutlara ayrılarak toplam dokuz alt faktör halinde yapılandırılmıştır. Ölçek 5’li Likert tipinde (1=Hiç katılmıyorum, 5=Tamamen katılıyorum) düzenlenmiş olup, uygulama sürecinde öğrencilerin cinsiyet ve sınıf değişkenine ilişkin veriler de toplanmıştır.

Ölçeğin birinci boyutu olan Matematiksel Eleştirel Düşünme, *Merak* ve *Sistematiklik* alt boyutlarını içermektedir. Bu boyut, öğrencilerin matematiksel durumları derinlemesine sorgulama ve sistematik düşünme eğilimlerini ölçmeyi amaçlamaktadır. Örnek maddeler arasında “Bir matematik problemiyle karşılaştığımda farklı çözüm yolları bulmak için sorular sorarım.” ve “Bir problemi çözerken izlediğim adımları mantıksal bir sırayla düzenlerim.” ifadeleri yer almaktadır. İkinci boyut olan Matematiksel Yaratıcı Düşünme, *Esneklik* ve *Yenilik* alt boyutlarından oluşmaktadır. Bu boyut, öğrencilerin matematiksel problemleri alışılmışın dışında yaklaşımlarla çözebilme becerilerini değerlendirmektedir. Örneğin, “Bir matematiksel problemi çözmek için farklı yöntemler denerim.” ve “Matematikte özgün fikirler üretmekten hoşlanırım.” gibi ifadeler bu boyut kapsamında yer almaktadır. Üçüncü boyut olan Matematiksel Problem Çözme, *Muhakeme Becerileri* ve *Yansıtıcı Beceriler* alt boyutlarını kapsamaktadır. Bu boyut, öğrencilerin mantıksal akıl yürütme, strateji oluşturma ve çözüm sürecini değerlendirme yetilerini ölçmeyi amaçlamaktadır. Bu kapsamda “Bir matematik problemini çözerken ulaştığım sonucu farklı yollarla doğrularım.” ve “Problemi çözdükten sonra kullandığım yöntemin uygun olup olmadığını değerlendiririm.” örnek maddeleri yer almaktadır. Dördüncü boyut olan Matematiksel Üstbilişsel Beceriler ise *Bilgi*, *Deneyim* ve *İzleme* alt boyutlarından oluşmaktadır. Bu boyut, öğrencilerin kendi öğrenme süreçlerini planlama, izleme ve düzenleme becerilerini ölçmektedir. Örneğin, “Bir problemi çözmeden önce nasıl bir yol izleyeceğimi düşünürüm.” ve “Matematikte yaptığım hatalardan ders çıkararak çözüm stratejimi değiştiririm.” gibi ifadeler bu boyuta örnektir.

Veri Toplama ve Analiz Süreci

Araştırmaya, ölçeği geliştiren sorumlu yazardan e-posta yoluyla gerekli izin alınarak başlanmıştır. Etik izinler de alındıktan sonra ölçeğin Türkçeye uyarılma sürecine geçilmiştir. Uyarılma süreci tamamlandıktan sonra, ölçeğin Türkçe formu lise öğrencilerine uygulanmıştır. Katılımcılar, formun ilk kısmında yer alan demografik soruları yanıtladıktan sonra, ölçek maddelerine ilişkin görüşlerini 5’li Likert tipi derecelendirme üzerinden belirtmişlerdir. Uygulama gönüllülük esasına dayalı olarak gerçekleştirilmiş; okul yönetimi ve öğretmenlerden gerekli izinler alınmış, veri toplama süreci ders işleyişini aksatmayacak biçimde yürütülmüştür.

Analiz sürecinde, öncelikle verilerin faktör analizine uygunluğunu belirlemek için Kaiser-Meyer-Olkin (KMO) katsayısı ve Bartlett küresellik testi incelenmiştir. Bu testlerin sonucunda verilerin faktör analizine uygun olduğu görülmüş ve DFA sürecine geçilmiştir. Ölçeğin yapı geçerliğini ve modelin örneklem üzerindeki uyumunu test etmek amacıyla SPSS, Jamovi ve AMOS programları birlikte kullanılmıştır. Veri tipi ordinal düzeyde olduğu için, DFA analizinde WLSMV (Weighted Least Squares Mean and Variance Adjusted) kestirim yöntemi ve polikorik korelasyon matrisi tercih edilmiştir. Model uyumunu değerlendirmek üzere χ^2/df , CFI, NFI, GFI, AGFI ve RMSEA uyum indeksleri raporlanmıştır. Elde edilen bulgular, modelin iyi uyum değerleri sağladığını göstermektedir ($\chi^2/df < 3$, RMSEA $< .08$, CFI, NFI $\geq .90$, GFI, AGFI $\geq .85$).

Analiz sonucunda, dört temel boyut (Matematiksel Eleştirel Düşünme, Matematiksel Yaratıcı Düşünme, Matematiksel Problem Çözme ve Matematiksel Üstbilişsel Beceriler) ile bunların altında yer alan dokuz alt boyutun orijinal yapıya uygun şekilde doğrulandığı görülmüştür. Standardize edilmiş faktör yüklerinin .51 ile .91 arasında değişmesi, maddelerin kendi boyutlarını yeterli düzeyde temsil ettiğini ortaya koymuştur. Ayrıca temel boyutlar arasındaki korelasyon katsayılarının yüksek düzeyde ve anlamlı olması, modelin yapı geçerliliğini desteklemektedir. Bununla birlikte, ölçeğin cinsiyet değişkenine göre ölçüm değişmezliği incelenmiştir. Bu kapsamda yapısal eşitlik modellemesi (SEM) yaklaşımıyla sırasıyla configural (biçimsel), metric (faktör yükleri), scalar (kesit) ve strict (hata varyansları) değişmezlik düzeyleri test edilmiştir. Elde edilen ΔCFI , ΔTLI ve $\Delta RMSEA$ değerlerinin kabul edilebilir sınırlar içinde ($\Delta CFI \leq .01$, $\Delta TLI \leq .01$, $\Delta RMSEA \leq .015$) kalması, ölçeğin kadın ve erkek öğrenciler için eşdeğer biçimde işlediğini ve ölçüm değişmezliğinin sağlandığını göstermiştir.

Ölçeğin güvenilirlik analizleri kapsamında Cronbach's alpha, McDonald's Omega, Guttman Split-Half ve Spearman-Brown katsayıları hesaplanmış; tüm boyutlar için katsayıların .80'in üzerinde olması, ölçeğin iç tutarlılığının yüksek olduğunu göstermiştir. Sonuç olarak, yürütülen analizler, MÜDDÖ'nün Türk lise öğrencilerinde geçerli ve güvenilir biçimde kullanılabileceğini ortaya koymuştur.

Bulgular

Ölçeğin Türkçeye uyarlanma sürecinde öncelikle verilerin faktör analizine uygunluğu incelenmiştir. Bu doğrultuda veri setindeki her alt boyuta ilişkin basıklık ve çarpıklık değerleri hesaplanmış ve elde edilen sonuçlar Tablo 1'de sunulmuştur.

Tablo 8. Alt Boyutlara Ait Tanımlayıcı İstatistikler

| Boyut | N | \bar{X} | SS | Basıklık | Çarpıklık |
|------------------------------------|-----|-----------|-----|----------|-----------|
| Matematiksel Eleştirel Düşünme | 417 | 40.15 | .72 | -.94 | .31 |
| Matematiksel Yaratıcı Düşünme | 417 | 28.08 | .63 | -.17 | .04 |
| Matematiksel Problem Çözme | 417 | 34.90 | .74 | -1.05 | .28 |
| Matematiksel Üstbilişsel Beceriler | 417 | 33.55 | .57 | -1.15 | .30 |

Tablo 1 incelendiğinde, tüm alt boyutlara ilişkin basıklık ve çarpıklık değerlerinin -1.5 ile +1.5 aralığında olduğu görülmektedir. Büyüköztürk (2007) ve Tabachnick ve Fidell (2013) bu aralık içinde kalan değerlerin normalliği sağladığını ifade etmektedir. Bu doğrultuda, veri setinin normal dağılım varsayımını karşıladığı kabul edilmiştir. Ayrıca örneklem büyüklüğünün yeterli olduğu ve kayıp ya da uç değer bulunmadığı da belirlenmiştir. Veri setinin faktör analizine uygunluğunu test etmek amacıyla Kaiser-Meyer-Olkin (KMO) ve Bartlett Küresellik Testi uygulanmış, elde edilen bulgular Tablo 2'de verilmiştir.

Tablo 9. Her Bir Alt Boyuta Ait KMO ve Bartlett Küresellik Testi Sonuçları

| | | | |
|------------------------------------|---------------------------|---------------------|---------|
| Matematiksel Eleştirel Düşünme | KMO Değeri | | .97 |
| | Bartlett Küresellik Testi | Ki-kare | 5774,84 |
| | | Serbestlik derecesi | 78 |
| | | p | .000* |
| Matematiksel Yaratıcı Düşünme | KMO Değeri | | .87 |
| | Bartlett Küresellik Testi | Ki-kare | 927.39 |
| | | Serbestlik derecesi | 45 |
| | | p | .000* |
| Matematiksel Problem Çözme | KMO Değeri | | .96 |
| | Bartlett Küresellik Testi | Ki-kare | 5202.31 |
| | | Serbestlik derecesi | 55 |
| | | p | .000* |
| Matematiksel Üstbilişsel Beceriler | KMO Değeri | | .95 |
| | Bartlett Küresellik Testi | Ki-kare | 4449.80 |
| | | Serbestlik derecesi | 55 |
| | | | |

p .000*

p* < .05

Tablo 2'ye göre her bir alt boyut için KMO değerlerinin .87 ile .97 arasında değiştiği ve tüm değerlerin .60'ın üzerinde olduğu görülmektedir. Ayrıca Bartlett Küresellik Testi her bir alt boyutta istatistiksel olarak anlamlıdır ($p < .05$). Büyüköztürk (2006)'e göre faktör analizine geçilebilmesi için KMO değerinin .60'tan büyük, Bartlett testinin ise anlamlı düzeyde olması gerekmektedir. Bu bağlamda elde edilen bulgular, veri setinin faktör analizine son derece uygun olduğunu göstermektedir. Uyarılama sürecinde, ölçeğin geliştirme aşamasındaki yapısal adımlar izlenmiş ve her boyut özelinde faktör yapısı ile güvenilirliği ayrı ayrı değerlendirilmiştir.

Matematiksel Eleştirel Düşünme Boyutu DFA Sonuçları

Matematiksel eleştirel düşünme boyutunun iki faktörlü yapısının doğruluğunu test etmek amacıyla, DFA uygulanmıştır. Bu analizde, daha önceki bulguların örneklem üzerinde ne derece doğrulandığı incelenmiş ve yapının geçerliliği test edilmiştir. DFA, AMOS programı aracılığıyla gerçekleştirilmiş olup, elde edilen model uygunluk göstergeleri Şekil 1'de görselleştirilmiştir. Şekil 1'de görüldüğü üzere, Matematiksel Eleştirel Düşünme boyutuna ilişkin doğrulayıcı faktör analizi sonucunda modelin iki faktörlü yapısı doğrulanmıştır. “Merak” faktörüne ilişkin maddelerin faktör yükleri .86 ile .90 arasında değişirken, “Sistematiklik” faktörüne ilişkin maddelerin faktör yükleri .86 ile .89 arasında bulunmuştur. Bu değerler, tüm maddelerin ilgili faktörlerini güçlü biçimde temsil ettiğini göstermektedir. Ayrıca iki faktör arasındaki korelasyon katsayısının .77 olması, Merak ve Sistematiklik boyutlarının birbirleriyle yüksek düzeyde ilişkili olduğunu, ancak aynı yapıyı ölçmediklerini ortaya koymaktadır. Şekil 1 incelendiğinde, modelin uyum indeksleri sırasıyla $\chi^2 = 117.072$, $df = 64$, $p = .007$, $\chi^2/df = 1.829$, CFI = .99, GFI = .96, NFI = .98, AGFI = .94 ve RMSEA = .045 olarak hesaplanmıştır. Bu değerlerin değerlendirilmesine yönelik bilgiler Tablo 3'te ayrıntılı şekilde sunulmuştur.

Tablo 10. Matematiksel Eleştirel Düşünme Boyutuna Ait Uyum İndeksleri

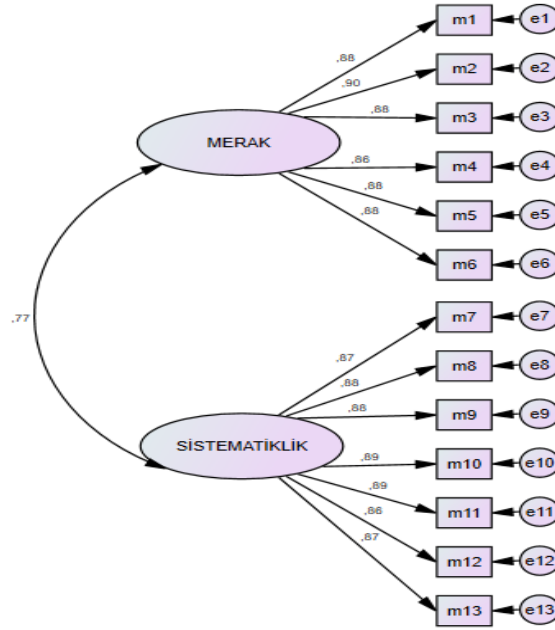
| Uyum İndeksleri | Zhou vd. | Bu çalışma kapsamında ulaşılan değerler | Kabul edilebilir uyum aralığı | Mükemmel uyum aralığı | Sonuç (Kline, 2019) |
|-----------------|----------|---|-------------------------------|-----------------------|---------------------|
| χ^2 / df | 2.670 | 1.829 | 3- 5 | 0- 3 | Mükemmel Uyum |
| RMSEA | 0.072 | .045 | $\leq .08$ | $\leq .06$ | Mükemmel Uyum |
| NFI | 0.923 | .98 | $\geq .90$ | $\geq .95$ | Mükemmel Uyum |
| CFI | 0.950 | .99 | $\geq .90$ | $\geq .95$ | Mükemmel Uyum |
| GFI | 0.931 | .96 | $\geq .85$ | $\geq .90$ | Mükemmel Uyum |
| AGFI | 0.917 | .94 | $\geq .85$ | $\geq .90$ | Mükemmel Uyum |

Tablo 3 incelendiğinde, χ^2/df değerinin 3'ün altında olması modelin kabul edilebilir düzeyde olduğunu göstermektedir. RMSEA değerinin .06'nın altında (.037) olması, modelin hata oranının düşük olduğunu ve yapının mükemmel uyum sergilediğini göstermektedir. Diğer uyum indekslerinden NFI, CFI, GFI ve AGFI değerlerinin tümü .95'in ya da .90'ın üzerinde yer almakta olup, bu da modelin yüksek düzeyde yapı geçerliliğine sahip olduğunu ortaya koymaktadır.

Kline (2019) tarafından önerilen uyum ölçütleri temel alındığında, bu çalışmada elde edilen tüm indekslerin mükemmel uyum düzeyinde olduğu görülmektedir. Ayrıca bu sonuçların, ölçeğin orijinalini geliştiren Zhou vd.(2024)'nin ulaştığı yapısal modelle örtüşmesi, uyarlanan ölçekteki faktör yapısının özgün yapıya uygun biçimde korunduğunu ve doğrulandığını ortaya koymaktadır. Bu doğrultuda, *matematiksel eleştirel düşünme* boyutuna ilişkin iki faktörlü yapının hem teorik zemine hem de ampirik bulgulara dayalı olarak tutarlı biçimde sürdürüldüğü ifade edilebilir.

Doğrulayıcı faktör analizi sonrasında, Matematiksel Eleştirel Düşünme boyutuna ait yakınsak ve ıraksak geçerlik değerleri incelenmiştir. Bu kapsamda Merak ve Sistematiklik alt boyutlarına ilişkin Ortalama Açıklanan Varyans (AVE) değerleri sırasıyla .77 ve .76, Bileşik Güvenirlik (CR) değerleri ise .95 ve .96 olarak hesaplanmıştır. Bu değerlerin $AVE \geq .50$ ve $CR \geq .70$ ölçütlerini karşıladığı görülmektedir. Bu bulgu, her iki alt boyutun da yakınsak geçerliğe sahip olduğunu ve maddelerin ait

oldukları faktörü yeterli düzeyde temsil ettiğini göstermektedir (Fornell & Larcker, 1981; Hair vd., 2019). İraksak geçerlik kapsamında, Matematiksel Eleştirel Düşünme boyutuna ait AVE değerlerinin (.77 ve .76), bu boyutun diğer boyutlarla olan korelasyon katsayılarının karelerinden (r^2) daha yüksek olduğu belirlenmiştir [MED-MPÇ = $(.743)^2 = .552$; MED-MÜB = $(.766)^2 = .587$]. Bu sonuç, Merak ve Sistematiçlik alt boyutlarının ölçekte yer alan diğer yapılardan istatistiksel olarak ayırt edilebilir olduğunu göstermektedir. Dolayısıyla elde edilen AVE ve CR değerleri, Matematiksel Eleştirel Düşünme boyutunun hem kuramsal tutarlılığını hem de psikometrik sağlamlığını desteklemektedir.



Şekil 7. Matematiksel Eleştirel Düşünme Boyutuna İlişkin DFA Sonuçları

Matematiksel Yaratıcı Düşünme Boyutu DFA Sonuçları

Matematiksel yaratıcı düşünme boyutunun iki faktörlü yapısının doğruluğunu test etmek amacıyla, DFA uygulanmıştır. Bu analizde, daha önceki bulguların örneklem üzerinde ne derece doğrulandığı incelenmiş ve yapının geçerliliği test edilmiştir. DFA, AMOS programı aracılığıyla gerçekleştirilmiş olup, elde edilen model uygunluk göstergeleri Şekil 2'de görselleştirilmiştir. Şekil 2 incelendiğinde, modelin kabul edilebilir düzeyde bir uyum gösterdiğini ve ölçme aracının yapı geçerliliğini sağladığını göstermektedir. Standardize edilmiş faktör yükleri incelendiğinde, *Esneklik* boyutundaki maddelerin yüklerinin .48 ile .67 arasında, *Yenilik* boyutundaki maddelerin yüklerinin ise .50 ile .61 arasında değiştiği görülmektedir. Tüm maddelerin faktör yüklerinin .40'ın üzerinde olması, ilgili maddelerin kendi boyutlarını yeterli düzeyde temsil ettiğini göstermektedir. Ayrıca iki faktör arasında gözlenen .85'lik korelasyon, boyutların birbirleriyle anlamlı ve güçlü bir ilişki içerisinde olduğunu ortaya koymaktadır. Bu sonuçlar genel olarak, modelin veriye iyi uyum sağladığını ve ölçeğin iki boyutlu yapısının doğrulandığını göstermektedir. Şekil 2 incelendiğinde, modelin uyum indeksleri sırasıyla $\chi^2 = 97.354$, $df = 34$, $p = .000$, $\chi^2/df = 2.863$, CFI = .93, GFI = .96, NFI = .91, AGFI = .93 ve RMSEA = .067 olarak hesaplanmıştır. Bu değerlerin değerlendirilmesine yönelik bilgiler Tablo 4'te ayrıntılı şekilde sunulmuştur.

Tablo 11. Matematiksel Yaratıcı Düşünme Boyutuna Ait Uyum İndeksleri

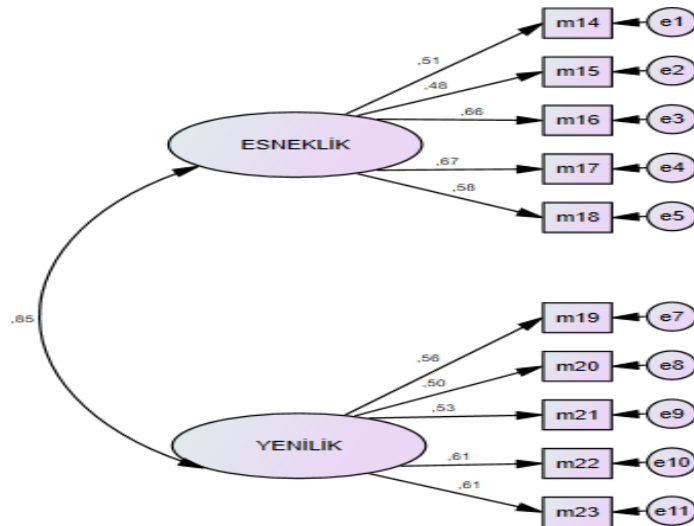
| Uyum İndeksleri | Zhou vd. | Bu çalışma kapsamında ulaşılan değerler | Kabul edilebilir uyum aralığı | Mükemmel uyum aralığı | Sonuç (Kline, 2019) |
|--------------------|----------|---|-------------------------------|-----------------------|---------------------|
| X ² /df | 3.599 | 2.863 | 3- 5 | 0- 3 | Mükemmel uyum |

| | | | | | |
|-------|-------|------|------|------|-----------------------|
| RMSEA | .057 | .067 | ≤.08 | ≤.06 | Mükemmel uyum |
| NFI | .91 | .90 | ≥.90 | ≥.95 | Mükemmel uyum |
| CFI | .95 | .93 | ≥.90 | ≥.95 | Kabul edilebilir uyum |
| GFI | 0.930 | .96 | ≥.85 | ≥.90 | Mükemmel uyum |
| AGFI | 0.907 | .93 | ≥.85 | ≥.90 | Mükemmel uyum |

Tablo 4 incelendiğinde, χ^2/df değerinin 3'ün altında olması modelin kabul edilebilir düzeyde olduğunu göstermektedir. RMSEA değerinin .07'nin altında (.067) olması, modelin hata oranının düşük olduğunu ve yapının mükemmel uyum sergilediğini göstermektedir. Diğer uyum indekslerinden NFI (.896), CFI (.929), GFI (.956) ve AGFI (.929) değerlerinin tamamı kabul edilebilir düzeyde olup, modelin iyi uyum gösterdiğini ve yapı geçerliliğinin sağlandığını ortaya koymaktadır.

Kline (2019) tarafından önerilen uyum ölçütleri temel alındığında, bu çalışmada elde edilen tüm indekslerin mükemmel uyum düzeyinde olduğu görülmektedir. Ayrıca bu sonuçların, ölçeğin orijinalini geliştiren Zhou vd. (2024)'nin ulaştığı yapısal modelle örtüşmesi, uyarlanan ölçekteki faktör yapısının özgün yapıya uygun biçimde korunduğunu ve doğrulandığını ortaya koymaktadır. Bu doğrultuda, *matematiksel yaratıcı düşünme* boyutuna ilişkin iki faktörlü yapının hem teorik zemine hem de ampirik bulgulara dayalı olarak tutarlı biçimde sürdürüldüğü ifade edilebilir.

Yakınsak geçerlik kapsamında, Esneklik ve Yenilik alt boyutlarına ilişkin AVE değerleri sırasıyla .34 ve .33, CR değerleri ise .72 ve .71 olarak hesaplanmıştır. AVE değerleri .50'nin altında olsa da, CR değerlerinin .70'in üzerinde olması yapının kabul edilebilir düzeyde yakınsak geçerliğe sahip olduğunu göstermektedir (Fornell & Larcker, 1981; Hair vd., 2019). Bu sonuç, ilgili maddelerin faktörle anlamlı biçimde ilişkili olduğunu ve yapının iç tutarlılığının sağlandığını göstermektedir. İraksak geçerlik analizinde, Matematiksel Yaratıcı Düşünme boyutunun AVE değerinin (.58), diğer boyutlarla olan korelasyon katsayılarının karelerinden daha yüksek olduğu görülmüştür [$MYD-MED = (.127)^2 = .016$; $MYD-MÜB = (.135)^2 = .018$]. Bu durum, Esneklik ve Yenilik alt boyutlarının ölçek içerisindeki diğer yapılardan ayırt edilebilir nitelikte olduğunu göstermektedir. Dolayısıyla elde edilen AVE ve CR değerleri, Matematiksel Yaratıcı Düşünme boyutunun kuramsal bütünlüğünü desteklemekte ve ölçeğin genel geçerlik yapısını güçlendirmektedir.



Şekil 8. Matematiksel Yaratıcı Düşünme Boyutuna İlişkin DFA Sonuçları

Matematiksel Problem Çözme Boyutu DFA Sonuçları

Matematiksel problem çözme boyutunun iki faktörlü yapısının doğruluğunu test etmek amacıyla, DFA uygulanmıştır. Bu analizde, daha önceki bulguların örneklem üzerinde ne derece doğrulandığı incelenmiş ve yapının geçerliliği test edilmiştir. DFA, AMOS programı aracılığıyla gerçekleştirilmiş

olup, elde edilen model uygunluk göstergeleri Şekil 3'te görselleştirilmiştir. Şekil 3 incelendiğinde, modelin mükemmel düzeyde bir uyum gösterdiği ve ölçme aracının yapı geçerliliğini güçlü biçimde sağladığı görülmektedir. Standardize edilmiş faktör yükleri incelendiğinde, Muhakeme boyutundaki maddelerin yüklerinin .88 ile .92 arasında, Yansıtıcı boyutundaki maddelerin yüklerinin ise .88 ile .91 arasında değiştiği belirlenmiştir. Tüm maddelerin faktör yüklerinin .40'ın oldukça üzerinde olması, ilgili maddelerin kendi boyutlarını yüksek düzeyde temsil ettiğini göstermektedir. Ayrıca iki faktör arasında gözlenen .77'lik korelasyon, boyutların birbirleriyle anlamlı ve güçlü bir ilişki içerisinde olduğunu ortaya koymaktadır. Şekil 3 incelendiğinde, modelin uyum indeksleri sırasıyla $\chi^2 = 65.613$, $df = 34$, $p = .01$, $\chi^2/df = 1.526$, CFI = .99, GFI = .97, NFI = .99, AGFI = .96 ve RMSEA = .036 olarak hesaplanmıştır. Bu değerlerin değerlendirilmesine yönelik bilgiler Tablo 5'te ayrıntılı şekilde sunulmuştur.

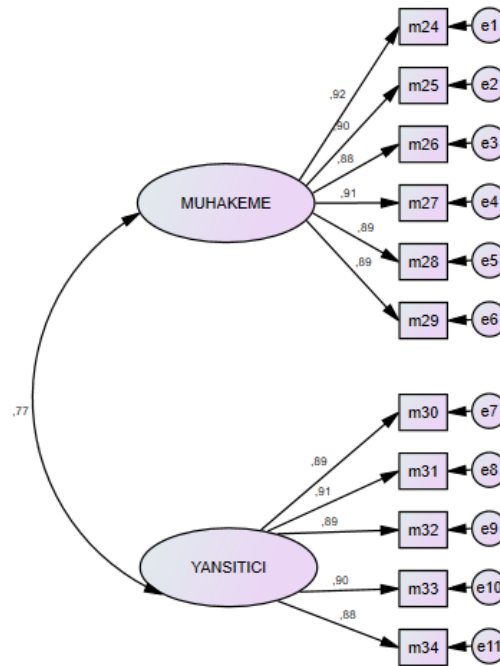
Tablo 12. Matematiksel Problem Çözme Boyutuna Ait Uyum İndeksleri

| Uyum İndeksleri | Zhou vd. | Bu çalışma kapsamında ulaşılan değerler | Kabul edilebilir uyum aralığı | Mükemmel uyum aralığı | Sonuç (Kline, 2019) |
|--------------------|----------|---|-------------------------------|-----------------------|---------------------|
| X ² /df | 3.368 | 1.526 | 3- 5 | 0- 3 | Mükemmel uyum |
| RMSEA | 0.076 | .036 | ≤.08 | ≤.06 | Mükemmel uyum |
| NFI | 0.928 | .99 | ≥.90 | ≥.95 | Mükemmel uyum |
| CFI | 0.948 | .99 | ≥.90 | ≥.95 | Mükemmel uyum |
| GFI | 0.929 | .97 | ≥.85 | ≥.90 | Mükemmel uyum |
| AGFI | 0.906 | .96 | ≥.85 | ≥.90 | Mükemmel uyum |

Tablo 5 incelendiğinde, χ^2/df değerinin 3'ün altında olması modelin kabul edilebilir düzeyde olduğunu göstermektedir. RMSEA değerinin .06'nın altında (.036) olması, modelin hata oranının düşük olduğunu ve yapının mükemmel uyum sergilediğini göstermektedir. Diğer uyum indekslerinden NFI, CFI, GFI ve AGFI değerlerinin tümü .95 ve üzerinde yer almakta olup, bu da modelin yüksek düzeyde yapı geçerliliğine sahip olduğunu ortaya koymaktadır.

Kline (2019) tarafından önerilen uyum ölçütleri temel alındığında, bu çalışmada elde edilen tüm indekslerin mükemmel uyum düzeyinde olduğu görülmektedir. Ayrıca bu sonuçların, ölçeğin orijinalini geliştiren Zhou vd.(2024)'nin ulaştığı yapısal modelle örtüşmesi, uyarlanan ölçekteki faktör yapısının özgün yapıya uygun biçimde korunduğunu ve doğrulandığını ortaya koymaktadır. Bu doğrultuda, matematiksel problem çözme boyutuna ilişkin iki faktörlü yapının hem teorik zemine hem de ampirik bulgulara dayalı olarak tutarlı biçimde sürdürüldüğü ifade edilebilir.

Yakınsak geçerlik analizinde, Muhakeme Becerileri ve Yansıtıcı Beceriler alt boyutlarına ait AVE değerleri sırasıyla .79 ve .80, CR değerleri ise her iki boyut için .96 olarak hesaplanmıştır. $AVE \geq .50$ ve $CR \geq .70$ koşullarının sağlanması, maddelerin faktörleri yüksek düzeyde temsil ettiğini ve yakınsak geçerliğin sağlandığını göstermektedir. İraksak geçerlik açısından, Matematiksel Problem Çözme boyutuna ait AVE değerinin (.80), bu boyutun diğer boyutlarla olan korelasyon katsayılarının karelerinden daha yüksek olduğu tespit edilmiştir [örneğin, $MPÇ-MED = (.743)^2 = .552$; $MPÇ-MÜB = (.784)^2 = .614$]. Bu bulgu, Muhakeme ve Yansıtıcı bileşenlerin ölçekteki diğer yapılardan ayırt edilebilir olduğunu göstermekte; böylece ölçeğin genel yapı geçerliğine güçlü destek sağlamaktadır (Fornell & Larcker, 1981; Hair vd., 2019).



Şekil 9. Problem Çözme Boyutuna İlişkin DFA Sonuçları

Matematiksel Üstbilişsel Beceriler Boyutu DFA Sonuçları

Matematiksel üstbilişsel beceriler boyutunun üç faktörlü yapısının doğruluğunu test etmek amacıyla, DFA uygulanmıştır. Bu analizde, daha önceki bulguların örneklem üzerinde ne derece doğrulandığı incelenmiş ve yapının geçerliliği test edilmiştir. DFA, AMOS programı aracılığıyla gerçekleştirilmiş olup, elde edilen model uygunluk göstergeleri Şekil 4'te görselleştirilmiştir. Şekil 4 incelendiğinde, modelin mükemmel düzeyde bir uyum gösterdiği ve ölçme aracının yapı geçerliliğini güçlü biçimde sağladığı görülmektedir. Standardize edilmiş faktör yükleri incelendiğinde, Bilgi boyutundaki maddelerin yüklerinin .81 ile .97 arasında, Deneyim boyutundaki maddelerin yüklerinin .55 ile .57 arasında, İzleme boyutundaki maddelerin yüklerinin ise .55 ile .61 arasında değiştiği belirlenmiştir. Tüm maddelerin faktör yüklerinin .40'ın oldukça üzerinde olması, ilgili maddelerin kendi boyutlarını yeterli düzeyde temsil ettiğini göstermektedir. Ayrıca faktörler arasındaki .45 ile .57 arasındaki korelasyonlar, boyutların birbirleriyle anlamlı ve orta düzeyde ilişkili olduğunu ortaya koymaktadır. Şekil 4 incelendiğinde, modelin uyum indeksleri sırasıyla $\chi^2 = 49.55$, $df = 41$, $p = .168$, $\chi^2/df = 1.208$, CFI = 1.00, GFI = .98, NFI = .99, AGFI = .96 ve RMSEA = .024 olarak hesaplanmıştır. Bu değerlerin değerlendirilmesine yönelik bilgiler Tablo 6'da ayrıntılı şekilde sunulmuştur.

Tablo 13. Matematiksel Üstbilişsel Beceriler Boyutuna Ait Uyum İndeksleri

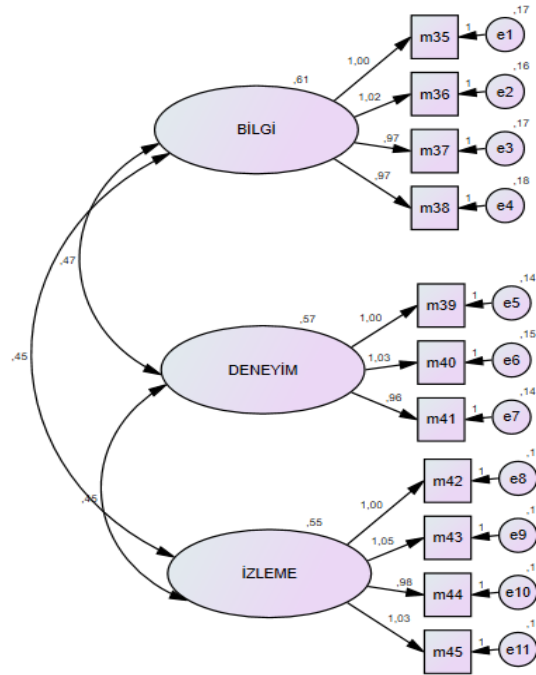
| Uyum İndeksleri | Zhou vd. | Bu çalışma kapsamında ulaşılan değerler | Kabul edilebilir uyum aralığı | Mükemmel uyum aralığı | Sonuç (Kline, 2019) |
|-----------------|----------|---|-------------------------------|-----------------------|---------------------|
| χ^2/df | 3.003 | 1.208 | 3- 5 | 0- 3 | Mükemmel uyum |
| RMSEA | 0.071 | 0.024 | $\leq .08$ | $\leq .06$ | Mükemmel uyum |
| NFI | 0.905 | .99 | $\geq .90$ | $\geq .95$ | Mükemmel uyum |
| CFI | 0.933 | 1.00 | $\geq .90$ | $\geq .95$ | Mükemmel uyum |
| GFI | 0.933 | .98 | $\geq .85$ | $\geq .90$ | Mükemmel uyum |
| AGFI | 0.902 | .96 | $\geq .85$ | $\geq .90$ | Mükemmel uyum |

Tablo 6 incelendiğinde, χ^2/df değerinin 3'ün altında olması modelin kabul edilebilir düzeyde olduğunu göstermektedir. RMSEA değerinin .06'nın altında (.024) olması, modelin hata oranının düşük olduğunu ve yapının mükemmel uyum sergilediğini göstermektedir. Diğer uyum indekslerinden NFI,

CFI, GFI ve AGFI değerlerinin tümü .95 ve üzerinde yer almakta olup, bu da modelin yüksek düzeyde yapı geçerliliğine sahip olduğunu ortaya koymaktadır.

Kline (2019) tarafından önerilen uyum ölçütleri temel alındığında, bu çalışmada elde edilen tüm indekslerin mükemmel uyum düzeyinde olduğu görülmektedir. Ayrıca bu sonuçların, ölçeğin orijinalini geliştiren Zhou vd.(2024)'nın ulaştığı yapısal modelle örtüşmesi, uyarlanan ölçekteki faktör yapısının özgün yapıya uygun biçimde korunduğunu ve doğrulandığını ortaya koymaktadır. Bu doğrultuda, *matematiksel üstbilişsel beceriler* boyutuna ilişkin üç faktörlü yapının hem teorik zemine hem de ampirik bulgulara dayalı olarak tutarlı biçimde sürdürüldüğü ifade edilebilir.

Yakınsak geçerlik kapsamında, üç alt boyuta ait AVE değerleri sırasıyla .76, .78 ve .77, CR değerleri ise .93, .93 ve .93 olarak hesaplanmıştır. Tüm değerlerin kabul edilebilir eşiklerin üzerinde olması ($AVE \geq .50$; $CR \geq .70$), her bir alt boyutun yakınsak geçerliğe sahip olduğunu göstermektedir. İraksak geçerlik analizinde, Matematiksel Üstbilişsel Beceriler boyutuna ait AVE değerlerinin (.77 ve üzeri), diğer boyutlarla olan korelasyon katsayılarının karelerinden (örneğin, MÜB–MED = $(.766)^2 = .587$; MÜB–MPÇ = $(.784)^2 = .614$) daha yüksek olduğu belirlenmiştir. Bu durum, Üstbilişsel Bilgi, Deneyim ve İzleme alt boyutlarının birbirinden ve diğer yapılardan istatistiksel olarak ayırt edilebilir olduğunu göstermektedir. Dolayısıyla elde edilen bulgular, Matematiksel Üstbilişsel Beceriler boyutunun hem yakınsak hem de ıraksak geçerliği sağladığını ve ölçeğin genel psikometrik yapısını güçlendirdiğini ortaya koymaktadır.



Şekil 10. Matematiksel Üstbilişsel Beceriler Boyutuna İlişkin DFA Sonuçları

Ölçeğin Tamamına Ait DFA Sonuçları

Çalışmanın bu bölümünde, daha önceki bulguların örneklem üzerinde ne derece doğrulandığı incelenmiş ve yapının geçerliliği test edilmiştir. DFA, AMOS ve JAMOVI programları aracılığıyla gerçekleştirilmiş olup, elde edilen model uygunluk göstergeleri Şekil 5'te görselleştirilmiştir. Şekil 5 incelendiğinde, modelin kabul edilebilir düzeyde bir uyum gösterdiği ve ölçme aracının yapı geçerliliğini sağladığı görülmektedir.

Standardize edilmiş faktör yükleri incelendiğinde, Matematiksel Eleştirel Düşünme (Merak ve Sistematiçlik) boyutundaki maddelerin yüklerinin .87 ile .88 arasında; Matematiksel Yaratıcı Düşünme (Esneklik ve Yenilik) boyutundaki maddelerin yüklerinin .51 ile .73 arasında; Matematiksel Problem Çözme (Muhakeme Becerileri ve Yansıtıcı Beceriler) boyutundaki maddelerin yüklerinin .86 ile .89 arasında; Matematiksel Üstbilişsel Beceriler (Bilgi, Deneyim ve İzleme) boyutundaki maddelerin yüklerinin ise .87 ile .91 arasında deęiştii görülmektedir. Tüm maddelerin faktör yüklerinin .50'nin üzerinde olması, ilgili maddelerin kendi alt boyutlarını yeterli düzeyde temsil ettiğini göstermektedir.

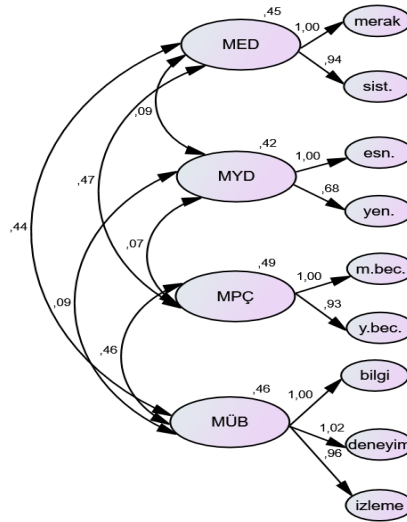
İkinci düzey düzeyde, dört alt boyutu temsil eden üst düzey faktör olan MÜDDÖ' ye ait standardize edilmiş yüklerin .19 ile 1.01 arasında deęiştii görülmektedir. Bu deęerler, alt boyutların genel yapıyı anlamlı biçimde yordadığını göstermektedir. Ayrıca alt boyutlar arasındaki korelasyon katsayılarının .21 ile .97 arasında deęiştii belirlenmiştir. Bu deęerlerin çoğunun .50'nin üzerinde olması, boyutların birbiriyle ilişkili fakat kavramsal olarak ayrıştığını; dolayısıyla ölçeğin yapı geçerliliğini desteklediğini göstermektedir.

Modelin genel uyum indeksleri de bu sonuçları desteklemektedir. Analiz sonucunda elde edilen uyum indeksleri $\chi^2 = 1405.019$, $df = 930$, $\chi^2/df = 1.511$, $CFI = .973$, $NFI = .925$, $GFI = .874$, $AGFI = .860$, ve $RMSEA = .035$ olarak bulunmuştur. Bu deęerler, CFI ve RMSEA açısından modelin mükemmel uyum gösterdiğini, χ^2/df oranının 2'nin altında olmasının ise veriye yüksek düzeyde uyum sağlandığını ortaya koymaktadır. Bu deęerlerin deęerlendirilmesine yönelik bilgiler Tablo 11'de ayrıntılı şekilde sunulmuştur. Jamovi analizinde elde edilen sonuçlar da bu bulgularla paralellik göstermektedir. WLSMV tahmin yöntemiyle yapılan doğrulayıcı faktör analizinde elde edilen $CFI = .998$, $TLI = .998$, $SRMR = .035$ ve $RMSEA = .022$ deęerleri, modelin ikinci bir örneklem üzerinde de oldukça iyi uyum sağladığını göstermektedir.

Tablo 11. Matematiksel Üst Düzey Düşünme Ölçeğine Ait Uyum İndeksleri

| Uyum İndeksleri | Zhou vd. | Bu çalışma kapsamında ulaşılan deęerler | Kabul edilebilir uyum aralığı | Mükemmel uyum aralığı | Sonuç (Kline, 2019) |
|--------------------|----------|---|-------------------------------|-----------------------|-------------------------------|
| X ² /df | 3.003 | 1.511 | 3- 5 | 0- 3 | Mükemmel uyum aralığı |
| RMSEA | 0.071 | 0.035 | ≤.08 | ≤.06 | Mükemmel uyum aralığı |
| NFI | 0.905 | .93 | ≥.90 | ≥.95 | Kabul edilebilir uyum aralığı |
| CFI | 0.933 | .97 | ≥.90 | ≥.95 | Mükemmel uyum aralığı |
| GFI | 0.933 | .87 | ≥.85 | ≥.90 | Kabul edilebilir uyum aralığı |
| AGFI | 0.902 | .86 | ≥.85 | ≥.90 | Kabul edilebilir uyum aralığı |

Tablo 11 incelendiğinde, MÜDDÖ'ye ilişkin doğrulayıcı faktör analizi (DFA) sonucunda elde edilen uyum indekslerinin genel olarak kabul edilebilir ve mükemmel uyum aralıkları içerisinde yer aldığı görülmektedir. X²/df oranının 1.511 olması modelin mükemmel uyum gösterdiğini ortaya koymaktadır. RMSEA deęeri (.035) de modelin iyi düzeyde bir uyuma sahip olduğunu desteklemektedir. CFI (.97) deęeri mükemmel uyumu yansıtırken, NFI (.93), GFI (.87) ve AGFI (.86) deęerleri kabul edilebilir uyum düzeyinde yer almaktadır. Bu sonuçlar, ölçeğin Türkçe formunun doğrulanmış faktör yapısının orijinal modele uygun ve veriye iyi uyum sağladığını göstermektedir.



Şekil 11. Matematiksel Üst Düzey Düşünme Ölçeği'nin Birinci Düzey Doğrulayıcı Faktör Analizi Modeli (MED: Matematiksel Eleştirel Düşünme, MYD: Matematiksel Yaratıcı Düşünme, MPÇ: Matematiksel Problem Çözme, MÜB: Matematiksel Üstbilişsel Beceriler, Sist.:Sistematiklik, Esn.: Esneklik, Yen.: Yenilik, M. Bec.: Muhakeme Becerileri, Y.Bec: Yansıtıcı Beceriler)

Çalışmanın bu bölümünde, daha önceki bulguların örneklem üzerinde ne derece doğrulandığı incelenmiş ve yapının geçerliliği ikinci düzey doğrulayıcı faktör analizi (2. Düzey DFA) ile test edilmiştir. Analiz, AMOS ve Jamovi programları aracılığıyla gerçekleştirilmiş olup, elde edilen model uygunluk göstergeleri Şekil 6'da sunulmuştur. Şekil 6 incelendiğinde, modelin ikinci düzey yapısının kabul edilebilir düzeyde bir uyum gösterdiği ve ölçme aracının genel faktör düzeyinde yapı geçerliliğini sağladığı görülmektedir.

Standardize edilmiş faktör yükleri incelendiğinde, Matematiksel Eleştirel Düşünme (Merak ve Sistematiklik) boyutundaki maddelerin yüklerinin .88 ile .88 arasında; Matematiksel Yaratıcı Düşünme (Esneklik ve Yenilik) boyutundaki maddelerin yüklerinin .73 ile 1.16 arasında; Matematiksel Problem Çözme (Muhakeme Becerileri ve Yansıtıcı Beceriler) boyutundaki maddelerin yüklerinin .86 ile .89 arasında; Matematiksel Üstbilişsel Beceriler (Bilgi, Deneyim ve İzleme) boyutundaki maddelerin yüklerinin ise .87 ile .91 arasında değiştiği görülmektedir. Tüm maddelerin faktör yüklerinin .50'nin üzerinde olması, ilgili maddelerin kendi alt boyutlarını yeterli düzeyde temsil ettiğini göstermektedir.

İkinci düzey DFA sonucunda, dört alt boyutu temsil eden üst düzey faktör olan MÜDDÖ'ye ait standardize edilmiş yüklerin .19 ile 1.01 arasında değiştiği belirlenmiştir. Bu değerler, alt boyutların genel yapıyı anlamlı biçimde yordadığını göstermektedir. Ayrıca alt boyutlar arasındaki korelasyon katsayılarının .19 ile .99 arasında değiştiği gözlemlenmiştir. Bu sonuçlar, alt boyutların birbiriyle ilişkili ancak kavramsal olarak ayrıştığını; dolayısıyla ölçeğin yapı geçerliliğini desteklediğini ortaya koymaktadır.

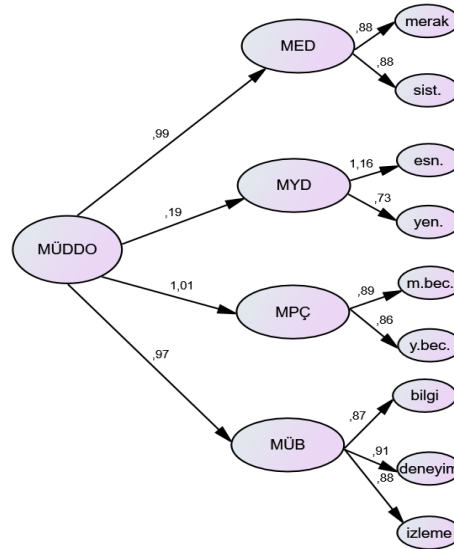
Modelin genel uyum indeksleri de bu sonuçları desteklemektedir. Analiz sonucunda elde edilen uyum indeksleri $\chi^2 = 1411.478$, $df = 932$, $\chi^2/df = 1.514$, CFI = .973, TLI = .971, NFI = .925, GFI = .873, AGFI = .859, ve RMSEA = .035 olarak bulunmuştur. Bu değerler, CFI, TLI ve RMSEA açısından modelin mükemmel uyum gösterdiğini; χ^2/df oranının 2'nin altında olmasının ise modelin veriye yüksek düzeyde uyum sağladığını göstermektedir.

Jamovi analizinde elde edilen sonuçlar da bu bulgularla paralellik göstermektedir. WLSMV tahmin yöntemiyle yapılan doğrulayıcı faktör analizinde elde edilen CFI = .998, TLI = .998, SRMR = .035 ve RMSEA = .022 değerleri, modelin ikinci bir örneklem üzerinde de oldukça iyi uyum sağladığını ortaya koymaktadır. Bu değerlerin değerlendirilmesine yönelik ayrıntılı bilgiler Tablo 12'de sunulmuştur.

Tablo 12. Matematiksel Üst Düzey Düşünme Ölçeğine Ait 2. Düzey DFA Uyum İndeksleri

| Uyum İndeksleri | Zhou vd. | Bu çalışma kapsamında ulaşılan değerler | Kabul edilebilir uyum aralığı | Mükemmel uyum aralığı | Sonuç (Kline, 2019) |
|--------------------|----------|---|-------------------------------|-----------------------|-------------------------------|
| X ² /df | 3.003 | 1.514 | 3- 5 | 0- 3 | Mükemmel uyum aralığı |
| RMSEA | 0.071 | 0.035 | ≤.08 | ≤.06 | Mükemmel uyum aralığı |
| NFI | 0.905 | .93 | ≥.90 | ≥.95 | Kabul edilebilir uyum aralığı |
| CFI | 0.933 | .97 | ≥.90 | ≥.95 | Mükemmel uyum aralığı |
| GFI | 0.933 | .87 | ≥.85 | ≥.90 | Kabul edilebilir uyum aralığı |
| AGFI | 0.902 | .86 | ≥.85 | ≥.90 | Kabul edilebilir uyum aralığı |

Tablo 12’de Matematiksel Üst Düzey Düşünme Ölçeği’ne ilişkin ikinci düzey doğrulayıcı faktör analizi (DFA) sonuçları verilmiştir. Analiz bulguları, modelin veriye oldukça iyi uyum sağladığını göstermektedir. X²/df oranının 1.514 olması, modelin mükemmel uyum düzeyinde olduğunu ortaya koymaktadır. RMSEA değerinin .035 olması da bu sonucu desteklemektedir. CFI (.97) değeri mükemmel uyum düzeyinde yer alırken, NFI (.93), GFI (.87) ve AGFI (.86) değerleri kabul edilebilir uyum aralıklarındadır. Bu sonuçlar, ölçeğin ikinci düzey faktör yapısının istatistiksel olarak doğrulandığını ve Matematiksel Üst Düzey Düşünme becerilerinin bütüncül bir yapıda temsil edildiğini göstermektedir.

**Şekil 12.** Matematiksel Üst Düzey Düşünme Ölçeği’nin İkinci Düzey Doğrulayıcı Faktör Analizi Modeli

Ölçüm Değişmezliği Analizini Bulguları

Araştırma ürecinin bu kısmında ölçeğin cinsiyet değişkenine göre ölçüm değişmezliği analizi gerçekleştirilmiştir. Analiz, doğrulayıcı faktör analizine dayalı olarak configural, metric, scalar ve strict değişmezlik düzeylerini test etmeye yöneliktir. Cinsiyet grupları arasında modelin aynı şekilde işleyip işlemediğini belirlemek için oluşturulan modellerin uyum indeksleri Tablo 13’te sunulmuştur.

Tablo 13. Cinsiyete Göre Çoklu Grup DFA Analizinin Uyum İndeks Sonuçları

| Değişmezlik Türleri | χ^2 | sd | CFI | RMSEA | $\Delta\chi^2$ | Δ sd | Sig. | Δ CFI |
|---------------------|----------|------|-----|-------|----------------|-------------|------|--------------|
| Biçimsel D. | 2481.72 | 1864 | .97 | .028 | - | - | - | - |
| Metrik D. | 2517.48 | 1900 | .97 | .028 | 35.76 | 36 | .00 | .000 |
| Skalar D. | 2530.19 | 1908 | .97 | .028 | 12.71 | 8 | .00 | .001 |
| Hata D. | 2604.91 | 1967 | .96 | .028 | 74.72 | 59 | .00 | .002 |

Tablo 13 incelendiğinde, biçimsel modelin iyi düzeyde bir uyum gösterdiği görülmektedir (CFI = .97, RMSEA = .028). Ardından gerçekleştirilen metrik değişmezlik modelinde, Δ CFI = .000 ve $\Delta\chi^2(36) = 35.76$, $p < .001$ değerleri, faktör yüklerinin cinsiyet grupları arasında anlamlı biçimde farklılaşmadığını göstermektedir. Skalar değişmezlik modelinde de benzer biçimde Δ CFI = .001 olarak bulunmuş ve uyum indekslerinde anlamlı bir bozulma gözlenmemiştir. Son olarak hata varyanslarının eşitlendiği modelde Δ CFI = .002 değeri elde edilmiştir. Bir modelden diğerine geçişte Δ CFI değerlerinin .010'un altında kalması, ölçeğin cinsiyet açısından tam ölçüm değişmezliği sağladığını göstermektedir. Bu nedenden dolayı bu farkın kabul edilebilir sınırlar içinde kaldığı görülmüştür. Bu bulgular, ölçeğin cinsiyet değişkenine göre tam ölçüm değişmezliği sağladığını, yani ölçeğin hem erkek hem de kadın gruplarında aynı yapıyı ölçtüğünü göstermektedir. Başka bir ifadeyle, ölçek puanları gruplar arasında karşılaştırılabilir niteliktedir.

Ölçeğe Ait Güvenirlik Bulguları

Araştırma sürecinin bu kısmında Türkçeye uyarlanan *Matematiksel Üst Düzey Düşünme Ölçeği* ve bu ölçeğin alt boyutları için güvenirlik analizleri yapılmış ve elde edilen sonuçlar raporlanmıştır. Bu bağlamda Cronbach's alfa, Guttman'ın λ katsayısı, McDonald's Omega ve Spearman-Brown güvenirlik katsayıları ile AVE ve CR değerleri hesaplanmıştır. Elde edilen sonuçlar Tablo 14'te sunulmuştur.

Tablo 14. Ölçeğin Alt Boyutlarına ve Tamamına Ait Güvenirlik Katsayıları

| Ana Boyut | Alt Boyut | Cronbach alfa | Guttman's | Omega | Spearman-Brown | AVE | CR |
|---------------------------------------|--------------------------------|---------------|-----------|-------|----------------|-----|-----|
| Matematiksel Eleştirel Düşünme | Merak | .95 | .95 | .95 | .96 | .77 | .95 |
| | Sistematiklik | .96 | .95 | .95 | .95 | .76 | .96 |
| Matematiksel Yaratıcı Düşünme | Esneklik | .70 | .70 | .70 | .67 | .34 | .72 |
| | Yenilik | .70 | .70 | .70 | .72 | .33 | .71 |
| Matematiksel Problem Çözme | Muhakeme Becerileri | .96 | .96 | .96 | .96 | .79 | .96 |
| | Yansıtıcı beceriler | .95 | .95 | .95 | .95 | .80 | .96 |
| Matematiksel Üstbilişsel Beceriler | Üstbilişsel Bilgi | .93 | .93 | .93 | .93 | .76 | .93 |
| | Üstbilişsel Deneyim | .91 | .91 | .92 | .90 | .78 | .93 |
| | Üstbilişsel İzleme | .93 | .93 | .93 | .93 | .77 | .93 |
| Matematiksel Üst Düzey Düşünme Ölçeği | Matematiksel Eleştirel Düşünme | .97 | .96 | .96 | .88 | .77 | .97 |
| | Matematiksel Yaratıcı Düşünme | .81 | .81 | .81 | .78 | .58 | .96 |
| | Matematiksel Problem Çözme | .96 | .96 | .96 | .96 | .80 | .98 |
| | Matematiksel | .95 | .95 | .95 | .90 | .77 | .97 |
| | Üstbilişsel Beceriler | | | | | | |

Tablo 14'te görüldüğü üzere, ölçeğin tüm alt boyutlarında iç tutarlılık katsayıları (Cronbach's alfa, Guttman's Lambda, Omega ve Spearman-Brown) .70'in üzerinde çıkmış ve bu durum, ölçeğin genel olarak yüksek düzeyde güvenirliliğe sahip olduğunu göstermektedir. Özellikle "Matematiksel Eleştirel Düşünme", "Matematiksel Problem Çözme" ve "Matematiksel Üstbilişsel Beceriler" boyutlarında bu katsayılar .90'in üzerindedir. Ayrıca, AVE değerlerinin çoğunlukla .50'nin üzerinde ve CR değerlerinin tamamının .70'in üzerinde olması, yapının yeterli yapı geçerliliğine ve yüksek bileşik güvenirliliğe sahip olduğunu ortaya koymaktadır. Sadece Matematiksel Yaratıcı Düşünme ana boyutunun alt boyutlarında AVE değerleri düşük çıkmıştır. AVE değerlerinin .50'nin altında olması durumunda, CR değerinin .70'in üzerinde olması halinde yapı geçerliliğinin kabul edilebilir olduğu literatürde belirtilmektedir (Fornell & Larcker, 1981; Hair vd., 2019). Bu değerler, kabul edilebilir güvenilirlik

düzeyine işaret etmekte olup, bu alt boyutların da ölçme açısından yeterli tutarlılığa sahip olduğunu göstermektedir (DeVellis, 2017).

Çalışmanın bu aşamasında, ölçeğin dört alt boyutu arasında yer alan ilişkiler Pearson korelasyon analizi ile incelenmiştir. Elde edilen korelasyon katsayıları ve anlamlılık düzeyleri Tablo 15'te sunulmuştur.

Tablo 15. MÜDDÖ Ölçeği İçin Faktör Korelasyonları

| | MED | MYD | MPC | MÜB |
|-----|--------|-------|--------|-----|
| MED | 1 | | | |
| MYD | .178* | 1 | | |
| MPC | .743** | .228* | 1 | |
| MÜB | .765** | .245* | .774** | 1 |

* $p < .05$, ** $p < .01$ (MED: Matematiksel Eleştirel Düşünme, MYD: Matematiksel Yansıtıcı Düşünme, MPC: Matematiksel Problem Çözme, MÜB: Matematiksel Üstbilişsel Beceriler)

Tablo 15 incelendiğinde, özellikle MED ile MÜB ($r = .765$, $p < .01$) ve MPC ($r = .743$, $p < .01$) arasındaki korelasyonlar oldukça güçlü olduğu görülmektedir. Bu durum, eleştirel düşünmenin hem üstbilişsel farkındalıkla hem de problem çözme süreçleriyle yüksek düzeyde ilişkili olduğunu göstermektedir. Bunun yanı sıra, MYD alt boyutunun diğer değişkenlerle olan korelasyonlarının görece daha düşük düzeyde kaldığı gözlemlenmiştir. Örneğin, MYD'nin MPC ($r = .228$, $p < .05$) ve MÜB ($r = .245$, $p < .05$) ile ilişkisi zayıf düzeyde olsa da istatistiksel olarak anlamlıdır. Bu sonuçlar, MYD alt boyutunun daha bağımsız bir yapıda olduğunu ve ölçekte yer alan diğer kavramlarla daha zayıf bağlar kurduğunu düşündürmektedir. Bunun yanı sıra MPC ile MÜB arasındaki yüksek korelasyon da dikkat çekicidir ($r = .774$, $p < .01$). Bu bulgular, öğrencilerin matematiksel problem çözme becerilerinin, üstbilişsel yetkinlikleri ve üst düzey düşünme yeterlikleriyle yakından ilişkili olduğunu ortaya koymaktadır. Genel olarak değerlendirildiğinde ölçeğin alt boyutları arasındaki ilişkiler iç tutarlılık bakımından güvenilirdir. Bu nedenle ölçeğin, lise öğrencilerinin matematiksel üst düzey düşünme becerilerini geçerli ve güvenilir şekilde ölçebileceği söylenebilir.

Tartışma

Bu çalışmada Türkçeye uyarlanan Matematiksel Üst Düzey Düşünme Ölçeği, öğrencilerin matematiksel düşünme becerilerini çok boyutlu biçimde değerlendirme amacıyla yapılandırılmış kuramsal bir ölçme aracı olarak öne çıkmaktadır. Ölçeğin dört temel boyutu; eleştirel düşünme, yaratıcı düşünme, problem çözme ve üstbilişsel beceriler, matematiksel düşünmenin yalnızca işlemsel değil; analitik, yaratıcı ve özdüzenleyici yönlerini de kapsamaktadır. Bu yapı, matematiksel düşünme süreçlerinin doğasına uygun ölçme araçlarına duyulan gereksinimi vurgulayan önceki araştırmalarla örtüşmektedir (Schoenfeld, 1992; English & Halford, 2012; Zhang vd., 2024). Nitekim mevcut literatürdeki ölçeklerin büyük bir kısmı yalnızca belirli bilişsel alanlara odaklanmakta; örneğin kimi çalışmalar yalnızca üstbilişsel becerileri (Güner & Erbay, 2021; Kaplan & Duran, 2016), kimileri ise yalnızca yaratıcı düşünmeyi (Furnham vd., 2013; Kholid vd., 2024) ele almaktadır. Ancak bilişsel süreçlerin doğası gereği bu beceriler çoğu zaman etkileşim hâlinde çalışmakta ve matematiksel bağlam içinde bütüncül biçimde değerlendirilmesi gerekmektedir. Bu bağlamda, bu çalışmada kullanılan ölçeğin dört farklı zihinsel boyutu aynı anda ve birbirleriyle ilişkili biçimde ölçebilmesi, alanyazında önemli bir boşluğu doldurmada ve ölçme-değerlendirme açısından özgün bir katkı sunmaktadır.

Bu bağlamda, bu çalışmada kullanılan ölçeğin dört farklı bilişsel süreci birlikte ele alması, literatürdeki önemli bir boşluğu doldurmaktadır. Altun (2024) tarafından yürütülen eylem araştırması da, özellikle öğretmen adaylarının üst düzey düşünme becerilerinin geliştirilmesinde bütüncül yaklaşımlara olan ihtiyacı vurgulamaktadır. Benzer şekilde, Aydın ve Öztürk (2024) tarafından gerçekleştirilen meta-analiz çalışması, üstbilişsel farkındalık eğitimlerinin hem problem çözme hem de genel matematik başarısı üzerinde anlamlı ve güçlü etkiler yarattığını göstermektedir. Bu bulgular, Matematiksel Üst Düzey Düşünme Ölçeği'nin yalnızca tanılayıcı değil, aynı zamanda öğretimsel müdahalelerin etkililiğini değerlendirmede de kullanılabilir bir araç olduğunu göstermektedir.

Ayrıca OECD (2023) ve İncikabı ve arkadaşları (2016) gibi uluslararası ve ulusal değerlendirme raporları da, öğrencilerin bilgiye erişiminden ziyade bilgiyi analiz edebilme, yapılandırabilme ve yeni durumlara transfer edebilme becerilerine sahip olmasının önemine işaret etmektedir. Bu ölçekte yer alan eleştirel düşünme, yaratıcı düşünme ve üstbilişsel bileşenler, tam da bu 21. yüzyıl becerilerini doğrudan ölçmeye yönelik unsurlar içermektedir. Bu yönüyle, ölçek aynı zamanda Türkiye Yüzyılı Maarif Modeli'nde (MEB, 2024) vurgulanan çok yönlü matematiksel düşünme ve üst düzey zihinsel yeterliklerin kazandırılmasına yönelik hedeflerle de tutarlıdır.

Sonuç olarak, bu çalışmada Türkçeye uyarlanan Matematiksel Üst Düzey Düşünme Ölçeği, hem kuramsal hem de istatistiksel açıdan güçlü bir ölçme aracı olarak öne çıkmakta; matematik eğitimi alanında öğrencilerin yüksek düzey bilişsel becerilerinin değerlendirilmesine yönelik nitelikli bir katkı sunmaktadır. Ölçek, araştırmacıların yanı sıra öğretmenlerin ve program geliştiricilerin de öğrencilerin düşünme becerilerini çok yönlü biçimde izleyebilmeleri açısından önemli bir araç niteliğindedir. Gelecek çalışmalarda, bu ölçeğin farklı yaş gruplarına, kültürel bağlamlara veya öğretim düzeylerine göre geçerlik analizlerinin yapılması, ölçeğin kullanım alanını genişletebilir.

Sonuç

Bu çalışmada, Zhou vd. (2024) tarafından geliştirilen Matematiksel Üst Düzey Düşünme Ölçeği'nin Türk örneklemindeki yapısının geçerlik ve güvenilirliği incelenmiştir. Geçerlik çalışmalarında açıklayıcı ve doğrulayıcı faktör analizi, güvenilirlik analizleri için ise iç tutarlılık ve iki yarı test güvenilirliği katsayıları hesaplanmış ve boyutlar arası korelasyonlar hesaplanmıştır.

Ölçeğin Türkçeye uyarlanma sürecinde ilk olarak, özgün formda yer alan 45 madde üç bağımsız yabancı dil uzmanı ve üç matematik eğitimcisi tarafından Türkçeye çevrilmiş, ardından bu çeviriler birleştirilerek ortak bir metin oluşturulmuştur. Daha sonra, bu metin farklı üç yabancı dil eğitimcisi tarafından geriye çeviri (back-translation) yöntemiyle İngilizceye çevrilmiş ve kavramsal-anlamsal eşdeğerliliği sağladığı Brislin (1970) yöntemi doğrultusunda doğrulanmıştır. Uzman görüşlerine dayalı son düzenlemelerin ardından oluşturulan Türkçe form, 30 lise öğrencisiyle pilot olarak uygulanmış, öğrencilerin anlama düzeyine göre küçük dilsel sadeleştirmeler yapılmıştır. Nihai hâli verilen ölçek, özgün formdaki dört temel boyut ve dokuz alt boyut korunarak 5'li Likert tipi şeklinde yapılandırılmıştır.

Çalışmada yapılan DFA sonuçları açıklayıcı analizlerle belirlenen yapının örneklem verileri üzerinde doğrulandığını göstermiştir. Model uyum indeksleri χ^2/df , RMSEA, CFI, GFI, AGFI ve NFI gibi temel göstergeler açısından değerlendirildiğinde, tüm alt boyutlara ait modellerin mükemmel uyum düzeyinde olduğu görülmüştür (Kline, 2019). Özellikle RMSEA değerlerinin .06'nın altında olması ve CFI, GFI, NFI gibi değerlerin .95'e yakın veya üzerinde olması, ölçekteki faktör yapıların ampirik olarak doğrulandığını göstermektedir. Ayrıca, ölçeğin cinsiyet değişkenine göre ölçüm değişmezliğini belirlemek amacıyla gerçekleştirilen analizler sonucunda, biçimsel, metrik, skalar ve hata değişmezliği modellerinin tümü kabul edilebilir uyum göstermiştir. ΔCFI değerlerinin tüm aşamalarda .010'un altında kalması, ölçeğin cinsiyet açısından tam ölçüm değişmezliği sağladığını ortaya koymaktadır. Bu bulgu, Matematiksel Üst Düzey Düşünme Ölçeği'nin hem erkek hem de kadın gruplarında aynı yapıyı geçerli biçimde ölçtüğünü ve ölçek puanlarının cinsiyet grupları arasında karşılaştırılabilir olduğunu göstermektedir.

Ölçeğin güvenilirlik analizleri sonucunda hem alt boyutlara hem de genel ölçeğe ait Cronbach's α , Guttman's λ , McDonald's Omega ve Spearman-Brown katsayılarının çoğunlukla .90'ın üzerinde olduğu tespit edilmiştir. Bu durum, ölçeğin iç tutarlılığının yüksek olduğunu ve güvenilir ölçümler sunduğunu göstermektedir. "Matematiksel Yaratıcı Düşünme" alt boyutunda yer alan "Esneklik" ve "Yenilik" bileşenlerine ait güvenilirlik katsayılarının .70 civarında olması, bu alt boyutların da kabul edilebilir düzeyde güvenilir olduğunu ortaya koymaktadır (DeVellis, 2017). Tüm bu bulgular bir arada değerlendirildiğinde, Matematiksel Üst Düzey Düşünme Ölçeği'nin Türkçeye uyarlanmış formunun hem geçerli hem de güvenilir bir ölçme aracı olduğu söylenebilir. Ölçek, lise öğrencilerinin

matematiksel düşünme becerilerini çok boyutlu olarak değerlendirme konusunda araştırmacılar ve uygulayıcılar için işlevsel bir yapı sunmaktadır.

Öneriler

Bu çalışma, Matematiksel Üst Düzey Düşünme Ölçeği'nin Türkçeye uyarlanması ve geçerlik-güvenirlik analizlerinin gerçekleştirilmesi açısından önemli katkılar sunmakla birlikte, bazı sınırlılıklar da içermektedir. İlk olarak, araştırma yalnızca lise düzeyindeki öğrencilerle yürütülmüştür. Bu durum, ölçeğin geçerlik ve güvenilirlik bulgularının farklı yaş gruplarına genellenebilirliğini kısıtlamaktadır. Gelecek çalışmalarda, ortaokul, üniversite ve yetişkin bireyler gibi farklı yaş ve eğitim gruplarına uygulanarak yaşa ve gelişim düzeyine göre karşılaştırmalı analizler yapılması önerilmektedir.

İkinci olarak, çalışmada elde edilen veriler, öz bildirim yoluyla toplanmıştır. Katılımcıların kendi beyanlarına dayalı olarak elde edilen verilerde sosyal beğenirlik yanlılığı gibi etkiler söz konusu olabilir. Bu nedenle, ilerleyen araştırmalarda, öğretmen gözlemleri, ürün dosyaları veya problem çözme süreçlerine dayalı performans değerlendirmeleri gibi çoklu veri kaynaklarının kullanılması önerilmektedir. Üçüncü olarak, bu çalışmada yalnızca psikometrik özellikler üzerinde durulmuş ve ölçekte yer alan boyutlar ile öğrencilerin akademik başarıları ya da diğer bilişsel/duyuşsal değişkenler arasındaki ilişkiler test edilmemiştir. Gelecek araştırmalarda, matematiksel üst düzey düşünme becerilerinin akademik başarı, matematik kaygısı, öz düzenleme, problem çözme yetkinliği, yaratıcılık ve eleştirel düşünme gibi değişkenlerle ilişkisi incelenerek ölçüt geçerliği güçlendirilebilir. Son olarak, ölçekte yer alan yapıların kültürel bağlamla olan ilişkisini değerlendirmek adına farklı sosyo-kültürel bölgelerde çapraz kültürel karşılaştırmalar yapılabilir. Bu tür çalışmalar, ölçeğin kültürler arası geçerlik düzeyini belirlemek açısından önemli veriler sağlayacaktır.

Beyan ve Açıklamalar

- 1. Araştırmacıların katkı oranı beyanı:** Bu çalışma iki yazar tarafından ortak olarak yürütülmüştür. Araştırma tasarımı, veri toplama süreci, veri analizi, yazım ve ölçek uyarlama süreci Yazar 1 tarafından üstlenilmiştir. Literatür taraması, kuramsal çerçevenin oluşturulması, bulguların yorumlanması ve metnin akademik dil açısından revize edilmesi ise Yazar 2 tarafından gerçekleştirilmiştir. Çalışmanın son kontrolü, biçimsel düzenlemeleri ve dergiye gönderim süreci her iki yazarın ortak katkısıyla tamamlanmıştır. Her iki yazar da çalışmanın tüm aşamalarını gözden geçirmiş ve nihai onayı vermiştir.
- 2. Çıkar çatışması:** Herhangi bir çıkar çatışması bulunmamaktadır.
- 3. Etik Raporu:** Bu çalışmanın etik kurul onayı, Bayburt Üniversitesi Rektörlüğü Etik Kurulu'nun 20.02.2025 tarihli, 43 sayılı ve 2. oturum kararı (Belge Doğrulama Kodu: BSFBL18BSC) ile, araştırma önerisinin etik ilkelere uygun olduğunun oy birliğiyle kabul edilmesiyle sağlanmıştır.
- 4. Araştırmanın Modeli:** Bu çalışma, lise öğrencilerinin matematiksel üst düzey düşünme becerilerini incelemek amacıyla yürütülen nicel bir tarama modeli olarak tasarlanan bir araştırma makalesidir. Tarama modelinde mevcut durum, önceden belirlenmiş grup üzerinde olduğu gibi ve doğal ortamında tanımlanır ve analiz edilir (Karasar, 2000). Ölçek uyarlama ve geçerlik-güvenirlik aşamaları bu model çerçevesinde planlanmıştır. Bu çalışmada, ölçeğin dilsel ve kavramsal eşdeğerliğini sağlamak üzere Brislin'in (1970) geri çeviri tekniği uygulanmış ve 30 öğrenciyle yapılan pilot uygulama sonucunda dilsel sadeleştirmeler yapılmıştır. Veri setinin faktör analizine uygunluğu KMO ve Bartlett testleri ile, normallik basıklık-çarpıklık değerleriyle sınanmış; yapısal geçerlik için SPSS'te açımlayıcı faktör analizi (temel bileşenler + direct oblimin) ve Lisrel'de doğrulayıcı faktör analizi (χ^2/df , RMSEA, CFI, GFI, AGFI, NFI) gerçekleştirilmiştir. Ölçeğin iç tutarlılığı Cronbach's α , Guttman's λ , McDonald's Omega ve Spearman-Brown katsayıları; yapısal yeterliliği AVE ve CR değerleriyle; alt boyutlar arası uyum ise korelasyon analiziyle değerlendirilmiştir. Bu yöntemsel çerçeve, ölçeğin geçerli ve güvenilir bir araç olduğunu göstermektedir.

References

Acar Güvendir, M., Kılıç, A. F. ve Özer Özkan, Y. (2025). Pursuit for scale selection: The journey to finding the right scale. *Pamukkale University Journal of Education*, 64, 338-373. <https://doi.org/10.9779/pauefd.1501330>

Altun, H. İ. (2024). *An action research on the development of algorithmic thinking skills of pre-service preschool teachers*. Unpublished master's thesis (Master's thesis - Kırşehir Ahi Evran University) was obtained from the National Thesis Center of the Council of Higher Education. (Thesis No. 909541).

Atlıhan, E. (2021). *The investigation of the relationship between grade 5 students' problem solving skills and problem posing skills in the context of creativity in patterns thesis* Unpublished master's thesis (Master's thesis – Ağrı İbrahim Çeçen University) was obtained from the National Thesis Center of the Council of Higher Education. (Thesis No. 704753).

Aydın, Ü. & Öztürk, M. (2024). The effect of metacognitive training on mathematics achievement and problem solving: a metaanalysis research. *Manisa Celal Bayar University Journal of The Faculty of Education*, 12(2), 286-304. <https://doi.org/10.52826/mcbuefd.1450124>

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, 1(3), 185-216. <https://doi.org/10.1177/135910457000100301>

Büyüköztürk, Ş. (2007). *Handbook of data analysis for social sciences: Statistics, research design, SPSS applications and interpretation* (Seventh Edition). PEGEM A Publishing.

Carbonneau, K. J., Wong, R. M., & Borysenko, N. (2020). The influence of perceptually rich manipulatives and collaboration on mathematic problem-solving and perseverance. *Contemporary Educational Psychology*, 61, 101846. <https://doi.org/10.1016/j.cedpsych.2020.101846>

Çanakçı, O. (2008). *The development and the evaluation of mathematics problem solving attitude scale* Unpublished doctoral thesis (Doctoral thesis – Marmara University) was obtained from the National Thesis Center of the Council of Higher Education. (Thesis No. 231804).

Çanakçı, O., & Özdemir, A. Ş. (2011). The construction of mathematics problem solving attitude scale. *Abant İzzet Baysal University Faculty of Education Journal*. 11(1), 119136.

Çelik, E. (2017). Metacognitive Self-Regulation Scale: The reliability and validity study of the Turkish Form. *Studies in Psychology*, 37(1), 53-71.

Çömlekoğlu, G. (2001). *The effect of calculators on prospective teachers problem solving skills* Unpublished master's thesis (Master's thesis – Balıkesir University) was obtained from the National Thesis Center of the Council of Higher Education. (Thesis No. 112628).

Değirmenci, T., & Deringöl, Y. (2024). Validity and Reliability Study of the Problem Solving Scale for Primary Schools. *National Education*, 53(244), 1849-1872. <https://doi.org/10.37669/milliegitim.1262404>

DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). Sage Publications.

Durdukoca, Ş. F., & Arıbaş, S. (2019). Development of metacognitive awareness scale for teacher candidates. *Electronic Journal of Social Sciences*, (72), 1541-1557. <https://doi.org/10.17755/esosder.474601>

English, L. D., & Halford, G. S. (2012). *Mathematics education: Models and processes*. Routledge.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39-50. <https://doi.org/10.1177/002224378101800104>

Fortes, E. C., & Andrade, R. R. (2019). Mathematical creativity in solving non-routine problems. *The Normal Lights*, 13(1), 108-135. <https://doi.org/10.56278/tnl.v13i1.1237>

Furnham, A., Hughes, D. J., & Marshall, E. (2013). Creativity, OCD, narcissism and the Big Five. *Thinking Skills and Creativity*, 10, 91-98. <https://doi.org/10.1016/j.tsc.2013.05.003>

Goretzko, D., Pham, T. T. H., & Bühner, M. (2021). Exploratory factor analysis: Current use, methodological developments and recommendations for good practice. *Current Psychology*, 40(7), 3510-3521. <https://doi.org/10.1007/s12144-019-00300-2>

Gümüş, F. Ö., ve Umay, A. (2017). Effect of problem solving strategies instruction on preservice elementary mathematics teachers' conceptual /procedural solution preferences and problem solving performance. *Elementary Education Online*, 16(2), 746-764. <https://doi.org/10.17051/ilkonline.2017.304731>

Güner, P., & Erbay, H. N. (2021). Metacognitive skills and problem-solving. *International Journal of Research in Education and Science*, 7(3), 715–734. <https://doi.org/10.46328/ijres.1594>

Güneş, A. M. (2022). The relationship between problem-solving skills, burnout levels and self-efficacy beliefs of school principals. *International Journal of Contemporary Educational Research*, 9(3), 590-602. <https://doi.org/10.33200/ijcer.1080663>

Gürefe, N., & Akçakin, V. (2018). The Turkish Adaptation of the Mathematical Resilience Scale: Validity and Reliability Study. *Journal of Education and Training Studies*, 6(4), 38-47. <https://doi.org/10.11114/jets.v6i4.2992>

Gürsoy, S. E., ve Çeliköz, N. (2021). Developing a problem solving scale for second grade primary school students: a validity and reliability study. *Mediterranean Journal of Educational Research*, 15(36), 128-163. <http://doi.org/10.29329/mjer.2020.367.8>

Hacıömeroğlu, G. (2011). Turkish adaptation of beliefs about mathematical problem solving instrument. *Dicle University Ziya Gökalp Faculty of Education Journal*, (17), 119-132.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage Learning.

Hambleton, R. K., Merenda, P. F., & Spielberger, C. D. (Eds.). (2005). *Adapting educational and psychological tests for cross-cultural assessment*. Lawrence Erlbaum Associates Publishers.

İncikabı, L., Mercimek, O., Ayanoğlu, P., Aliustaoğlu, F., & Tekin, N. (2016). An evaluation of middle school mathematics teaching programs based on tumss cognitive domains. *Elementary Education Online*, 15(4). <https://doi.org/10.17051/ieo.2016.54792>

Kaplan, A., & Duran, M. (2016). Mathematical metacognition awareness inventory towards middle school students: validity and reliability study. *Ataturk University Kazim Karabekir Faculty of Education Journal*, (32), 1-17.

Karasar, N. (2000). *Scientific research method*. Nobel Publications.

Kayan, F. (2008). *A study on preservice elementary mathematics teachers' mathematical problem solving beliefs*. Unpublished master's thesis (Master's thesis – Middle East Technical University) was obtained from the National Thesis Center of the Council of Higher Education. (Thesis No. 218040).

Kholid, M. N., Mahmudah, M. H., Ishartono, N., Putra, F. G., & Forthmann, B. (2024). Classification of students' creative thinking for non-routine mathematical problems. *Cogent Education*, 11(1), 2394738. <https://doi.org/10.1080/2331186X.2024.2394738>

Kızılkaya, G., ve Aşkar, P. (2009). The development of a reflective thinking skill scale towards problem solving. *Education and Science*, 34(154), 82-92. <https://doi.org/10.15390/ES.2009.839>

Kline, R. B. (2019). *Principles and practice of structural equation modeling*. Guilford publications.

Korkmaz, Ö. & Karakus, U. (2009). The effects of education faculties on students' critical thinking dispositions and levels. *Turkish Educational Sciences Journal*, 7(4), 879-902.

Kurnaz, A. (2019). *Planning, implementing and evaluating critical thinking teaching activities*. Education Publishing.

Lin, C. (2023). *Intellectual development and mathematics learning*. Springer.

Millî Eğitim Bakanlığı [MEB]. (2024). Türkiye Century Education Model: Secondary School Mathematics Course Curriculum (Grades 9–12). Board of Education and Training. <https://tymm.meb.gov.tr/ogretim-programlari>

Monteleone, C., Miller, J., & Warren, E. (2023). Conceptualising critical mathematical thinking in young students. *Mathematics Education Research Journal*, 35(2), 339–359. <https://doi.org/10.1007/s13394-023-00445-1>

Muncer, G., Higham, P. A., Gosling, C. J., Cortese, S., Wood-Downie, H., & Hadwin, J. A. (2021). A meta-analysis investigating the association between metacognition and math performance in adolescence. *Educational Psychology Review*, 34, 301–334. <https://doi.org/10.1007/s10648-021-09620-x>

Muyo, M. (2015). *Improving mathematical literacy and problem solving abilities of students in Prizren Faculty of Education*. Unpublished doctoral thesis (Doctoral thesis – Balikesir University) was obtained from the National Thesis Center of the Council of Higher Education. (Thesis No. 421402).

OECD. (2023). PISA 2022 Results (Volume I): *The State of Learning and Equity in Education*. OECD Publishing. <https://doi.org/10.1787/53f23881-en>

Özpinar, İ. (2012). *Study of developing scales to measure the skills in 6th - 8th grade mathematics curriculum*. Unpublished doctoral thesis (Doctoral thesis – Black Sea Technical University) was obtained from the National Thesis Center of the Council of Higher Education. (Thesis No. 344458).

Özsoy, G. (2008). Metacognition. *The Journal of Turkish Educational Sciences*, 6(4), 713-740.

Öztürk, M., Demir, Ü., & Akkan, Y. (2021). Investigation of proportional reasoning problem solving processes of seventh grade students: A mixed method research. *International Journal on Social and Education Sciences (IJonSES)*, 3(1), 48–67. <https://doi.org/10.46328/ijonSES.66>

Öztürk, M., Sarıkaya, İ., & Yıldız, K. A. (2023). Middle school students' problem solving performance: Identifying the factors that influence it. *International Journal of Science and Mathematics Education*, 22, 1363–1379. <https://doi.org/10.1007/s10763-023-10423-5>

Sachdeva, S., & Eggen, P. O. (2021). Learners' critical thinking about learning mathematics. *International Electronic Journal of Mathematics Education*, 16(3), em0644. <https://doi.org/10.29333/iejme/11003>

Sadık, R. (2006). The Comparison of success levels in four operations & problem-solving in natural numbers in primary school grade 4 and 5 students who know to play chess and who do not Unpublished master's thesis (Master's thesis – Bolu Abant İzzet Baysal University) was obtained from the National Thesis Center of the Council of Higher Education. (Thesis No. 190239).

Samura, A. O. & Juandi, D. (2020). Improving mathematical critical thinking skills through problem-based learning. *Journal of Physics: Conference Series*, 1521(3), 032102. <https://doi.org/10.1088/1742-6596/1521/3/032102>

Schoenfeld, A. H. (1992). *Learning to think mathematically: Problem solving, metacognition, and sense-making in mathematics*. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 334–370). Macmillan.

Sriraman, B. (2009). The characteristics of mathematical creativity. *ZDM Mathematics Education*, 41, 13–27. <https://doi.org/10.1007/s11858-008-0114-z>

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>

Suliani, M., Juniati, D., & Lukito, A. (2024). The influence of student's mathematical beliefs on metacognitive skills in solving mathematical problems. *International Journal of Evaluation and Research in Education*, 13(3), 1481–1491. <https://doi.org/10.11591/ijere.v13i3.27117>

Süzer, A. (2022). *Problem solving skills and processes of primary school students* Unpublished master's thesis (Master's thesis – Bartın University) was obtained from the National Thesis Center of the Council of Higher Education. (Thesis No. 713516).

Şahinkaya, T., Öztürk, M., & Albayrak, M. (2022). The effects of metacognitive IMPROVE technique on teaching ratio-proportion and developing proportional reasoning skills. *Kocaeli University Education Journal*, 5(2), 495–516. <https://doi.org/10.33400/kuje.1137016>

Tabachnick, B. G. & Fidell, L. S. (2013). *Using Multivariate Statistics* (6th ed.). Allyn and Bacon.

Topal, M., Topal, N., Görgel, A., Kama, H., & Yağız, N. (2025). Problem based learning and its theoretical foundations: a new approach to the learning process. *Socrates Journal of Interdisciplinary Social Studies*, 11(49), 46-64. <https://doi.org/10.5281/zenodo.14632143>

Uğurluoğlu, E. (2008). *An investigation of elementary students' beliefs and attitudes towards mathematics and mathematical problem solving according to certain variables* Unpublished master's thesis (Master's thesis – Eskişehir Osmangazi University) was obtained from the National Thesis Center of the Council of Higher Education. (Thesis No. 177267).

Uysal, O. (2007). *Evaluation of relation between problem solving skills, anxiety and attitudes of secondary education grade ii students directed towards mathematics course* Unpublished doctoral thesis (Doctoral thesis – Dokuz Eylül University) was obtained from the National Thesis Center of the Council of Higher Education. (Thesis No. 211583).

Ünal, M., & Aral, N. (2014). Development of problem solving scale in science education: Validity and reliability studies. *Education and Science*, 39(176), 267-278. <http://dx.doi.org/10.15390.EB.2014.3585>

Yaşar, S. (2022). *The effect of storyline-based mathematics problem solving practices on problem solving success, skills and attitudes* Unpublished master's thesis (Master's thesis – Bolu Abant İzzet Baysal University) was obtained from the National Thesis Center of the Council of Higher Education. (Thesis No. 769493).

Yüksel, M. Y., Tekin, Ö. E., & Kaplaner, K. (2021). The research of the relationship between the problem solving skills & metacognitive awareness of middle school students and the social emotional learning. *Çukurova University Faculty of Education Journal*, 50(1), 487-506.

Zhang, J., Zhou, Y., Jing, B., Pi, Z., & Ma, H. (2024). Metacognition and mathematical modeling skills: The mediating roles of computational thinking in high school students. *Journal of Intelligence*, 12(6), 55. <https://doi.org/10.3390/jintelligence12060055>

Zhou, Y., Ning, Y., Chen, J., Zhang, W., & Wijaya, T. T. (2024). Development and validation of Mathematical Higher-Order Thinking Scale for high school students. *Psychology in the Schools*, 61(8), 3160-3192. <https://doi.org/10.1002/pits.23213>