

The Development and Validation of the Metacognitive Awareness Scale for Middle School Students*

Ayşe Gönültaş Yıldırım^a, Ayşegül İşlekeller Bozca^b

Received : 18 August 2026
Revised : 6 September 2026
Accepted : 30 September 2026
DOI : 10.26822/iejee.2026.466

^a **Corresponding Author:** Ayşe Gönültaş Yıldırım, Social Services Specialist, Ministry of Interior of the Republic of Türkiye
E-mail: gonultasayse@gmail.com
ORCID: <https://orcid.org/0000-0003-4354-1723>

^b Ayşegül İşlekeller Bozca, Istanbul University-Cerrahpaşa, Department of Special Education
E-mail: islekellerbozca@iuc.edu.tr
ORCID: <https://orcid.org/0000-0003-2912-5379>

* This article was derived from the doctoral dissertation of the first author, Ayşe Gönültaş Yıldırım, under the supervision of Ayşegül İşlekeller Bozca.

Abstract

This study aimed to develop a psychometrically sound instrument to assess the metacognitive awareness of middle school students. A 5-point Likert-type instrument was developed to assess middle school students' metacognitive knowledge and regulatory processes. The study was conducted with a total of 1009 students. Exploratory factor analysis (EFA) was conducted with data from 475 students. Preliminary diagnostics supported factorability (KMO = .952; Bartlett's $\chi^2 = 8377.999$, $p < .01$). Based on statistical and theoretical criteria, 17 items were removed, yielding a unidimensional structure consisting of 29 items that explained 38.392% of the total variance. Factor loadings ranged between .469 and .644. Confirmatory factor analysis (CFA) was subsequently performed with an independent sample of 534 students. Model fit indices ($\chi^2/df = 2.770$, CFI = .904, TLI = .899, IFI = .904, NFI = .870, RFI = .860, RMSEA = .048) indicated an acceptable fit for the proposed one-factor model. Internal consistency was high (Cronbach's $\alpha = .942$), and item-total correlations ranged from .467 to .651. Test-retest reliability, assessed with 126 students from a public middle school over a three-week interval, demonstrated excellent temporal stability (ICC(3,1) = .997, 95% CI [.996, .998]). Overall, the scale demonstrated strong validity and reliability and provides a developmentally appropriate instrument for assessing metacognitive awareness in middle school students.

Keywords:

Metacognition, Middle School Students, Scale Development, Validity and Reliability, Temporal Stability

Introduction

Metacognition refers to individuals' awareness and regulation of their own cognitive processes (Flavell, 1979; Schraw & Dennison, 1994; Veenman et al., 2006). It comprises two fundamental components: metacognitive knowledge and metacognitive regulation (Baker & Brown, 1980; Pintrich, 2002; Schraw & Dennison, 1994; Schraw & Moshman, 1995; Vrugt & Oort, 2008). Although metacognition has been conceptualized in different ways, most contemporary frameworks distinguish between knowledge of cognition and regulation of cognition (Schraw & Dennison, 1994;



www.iejee.com
ISSN: 1307-9298

2026 Published by KURA Education & Publishing.
This is an open access article under the CC BY-NC-ND license. (<https://creativecommons.org/licenses/by/4.0/>)

Veenman et al., 2006). Based on these theoretical perspectives, the conceptual framework of metacognition is presented in Figure 1.

Figure 1.

Conceptual framework of metacognition synthesized from previous theoretical models.



Metacognitive knowledge refers to individuals' understanding of their own cognitive processes and includes declarative, procedural, and conditional knowledge (Cross & Paris, 1988; Schraw & Moshman, 1995). Metacognitive regulation refers to the processes through which learners plan, monitor, control, and evaluate their cognitive activities during learning (Jacobs & Paris, 1987; Schraw & Moshman, 1995; Veenman et al., 2006). Together, these complementary components form the theoretical basis of metacognitive awareness and provide the conceptual framework underlying the development of the present scale.

Students with well-developed metacognitive skills regulate their learning more effectively by planning, monitoring, and evaluating their cognitive processes (Kuhn, 2000). These competencies are particularly important during middle school, when students experience rapid cognitive development and establish increasingly independent learning habits (Karakelle & Saraç, 2010; Özsoy, 2011). Consistent with this developmental perspective, previous research has shown that higher levels of metacognitive awareness are associated with improved academic achievement, motivation, problem-solving, and self-regulated learning (Coutinho & Neuman, 2008; Desoete, 2008; Howard et al., 2000; Jahangard et al., 2016; Vrugt & Oort, 2008). Furthermore, intervention studies have demonstrated that metacognitive skills can be enhanced through appropriately designed instructional approaches, highlighting the importance of educational practices that explicitly foster metacognitive awareness (Çakıroğlu, 2007; Güner & Erbay, 2021; Owo & Ikwut, 2015; Özsoy, 2011; Young & Fry, 2008).

As students transition from concrete to formal operational thinking during middle school, assessing their metacognitive awareness becomes increasingly important for supporting independent learning and self-regulation. Reliable assessment instruments are therefore needed to help teachers identify students' metacognitive strengths and weaknesses and to design appropriate instructional interventions. A

review of the literature on metacognition reveals that several measurement instruments have been developed internationally for children (Annevirta et al., 2007; Jacobs & Paris, 1987; Lecce et al., 2010; Mokhtari & Reichard, 2002; Pereira-Laird & Deane, 1997; Schraw & Dennison, 1994; Sperling et al., 2002). However, in Türkiye, available instruments assessing metacognition remain limited (Abacı et al., 2006; Çetinkaya & Erkin, 2002; Karakelle & Saraç, 2007; Öztürk, 2012).

Among the earliest Turkish instruments, Çetinkaya and Erkin (2002) developed a scale examining metacognition in relation to reading comprehension and academic achievement. However, the instrument focused primarily on reading strategies rather than metacognition as a broader construct. Similarly, Karakelle and Saraç (2007) developed the Metacognitive Awareness Scale for Children for students aged 10–14. Although this scale represented an important contribution, it primarily assessed declarative awareness and provided limited coverage of regulatory processes such as planning, monitoring, and evaluation.

In addition to these national instruments, one of the most widely used international tools is the Metacognitive Awareness Inventory (MAI) developed by Schraw and Dennison (1994). However, the MAI was originally designed for adults and university students, and its level of abstraction and linguistic complexity may limit its suitability for middle school students. Similarly, the Junior Metacognitive Awareness Inventory (Jr. MAI), although developed for younger populations, has been reported to focus more on declarative knowledge and does not sufficiently capture the dynamic and process-oriented nature of regulatory processes such as planning, monitoring, and evaluation.

Furthermore, many existing instruments, both internationally and in the Turkish context, tend to assess metacognition either within specific domains (e.g., reading) or within limited dimensions, and may not adequately reflect the integrated and process-based nature of metacognitive functioning in contemporary learning environments. Given these limitations, there is a clear need to develop a comprehensive, up-to-date, and developmentally appropriate metacognitive awareness scale for middle school students. Such an instrument should encompass both metacognitive knowledge and control processes while accounting for students' cognitive, linguistic, and cultural characteristics.

Collectively, previous instruments either target different age groups, focus on domain-specific metacognitive skills, or fail to integrate metacognitive knowledge and regulatory processes within a single developmentally appropriate framework. Consequently, a psychometrically robust instrument specifically

designed for middle school students remains unavailable. No existing instrument simultaneously captures both metacognitive knowledge and regulatory processes in a developmentally appropriate, process-oriented, and culturally relevant manner for middle school students within the Turkish educational context. Developing this type of scale is therefore a critical priority for supporting effective learning and instructional design at the middle school level. The present study aims to address this gap by developing a metacognitive awareness scale that integrates both knowledge and control processes, reflects the dynamic and process-based nature of learning, is linguistically and cognitively appropriate for middle school students, and aligns with contemporary educational frameworks emphasizing self-regulated learning.

Development of the Scale and Significance of the Study

Middle school students are in a developmental period marked by the transition from concrete to abstract thinking. At this stage, their skills in self-expression and self-evaluation may still be limited. Therefore, the wording, conceptual level, and questioning style of metacognition scales developed for high school or university students may be insufficient in terms of clarity and validity for middle school learners. In addition, when considering the cultural context, some existing scales require reconstruction to reflect the social, cognitive, and linguistic realities of middle school students in Turkey.

This situation highlights the need for a new, valid, and reliable measurement tool that accurately reflects the fundamental components of metacognition and is suitable for the age, developmental level, and cultural context of middle school students. A scale developed for this purpose would not only help determine students' levels of metacognitive awareness but also assist teachers in guiding learning processes more effectively. Moreover, in line with the constructivist learning principles of the Türkiye Yüzyılı Maarif Model, such a tool would contribute to the literature by providing a contemporary instrument that comprehensively measures students' abilities to construct, plan, monitor, and evaluate knowledge.

The purpose of this study is to develop a valid and reliable metacognitive awareness scale for middle school students that is tailored to their cognitive and linguistic characteristics, encompassing the primary metacognitive processes: cognitive awareness, planning, monitoring, regulation, and evaluation. This study aims to address the lack of appropriate scales in the national literature and to make both theoretical and practical contributions by providing a tool that enables students to manage their own learning processes more intentionally. In addition to addressing an important measurement gap, the proposed scale

provides researchers and educators with a practical instrument for evaluating students' metacognitive awareness within contemporary educational settings.

Method

Research Design and Sample

This study employed a general survey design aimed at developing a metacognitive awareness scale. In research based on survey models, a past or existing phenomenon, event, or situation is described as it naturally occurs without any external manipulation (Büyüköztürk et al., 2009). The survey model is typically used to identify individuals' opinions or characteristics such as interests, abilities, and attitudes, and it is generally conducted with larger samples than other research designs (Büyüköztürk et al., 2009).

A purposive sampling method was used in the study. Purposive sampling is preferred in cases where a detailed investigation is necessary and allows for selecting participants who are both information-rich and capable of representing the target population, based on available resources (Büyüköztürk et al., 2009). Since the research sample was selected according to specific criteria such as district, school type, and grade level, the purposive sampling method was considered appropriate (Büyüköztürk et al., 2009). In scale development studies, determining the sample size is based on theoretical principles. It is generally recommended that the sample size be at least five times the number of items on the scale (Conway & Huffcutt, 2003). While 200 participants are considered sufficient to obtain reliable factors in factor analysis (Tabachnick & Fidell, 2007), 300 participants are regarded as good, 500 as very good, and 1000 as excellent (Kline, 1998).

The study group consisted of students enrolled in four different middle schools located in the Esenyurt, Avcılar, and Beylikdüzü districts of Istanbul during the first semester of the 2025–2026 academic year. Before data collection, ethical approval was obtained from the Ethics Committee of Istanbul University, along with permissions from the Ministry of National Education (MEB), the participating institutions, teachers, and parents. The data were collected face-to-face through paper-based questionnaires. A total of 1,067 students participated in the study; however, data from 58 students were excluded from the analysis due to various reasons. These excluded cases included 19 fifth-grade, 15 sixth-grade, 11 seventh-grade, and 13 eighth-grade students. The primary reasons for exclusion were incomplete responses, multiple selections for a single item, and answering patterns that indicated low validity (e.g., marking only "strongly agree" for all items). The analyses were conducted using valid data from 1,009 students. Of these participants, 494 (49%) were female and 515 (51%) were male. The distribution by grade level was as follows: fifth grade – 297 (29%),

sixth grade – 263 (26%), seventh grade – 231 (23%), and eighth grade – 218 (22%). In terms of school type, 577 (57%) students attended public middle schools, 229 (23%) attended imam hatip middle schools, and 203 (20%) attended private middle schools. The developed metacognitive awareness scale was administered across these diverse populations, and the validity and reliability of the scale were examined accordingly.

Table 1.

Descriptive Characteristics of Middle School Students

Descriptive Variable	Category	f	%
Gender	Female	494	49
	Male	515	51
Grade Level	5th Grade	297	29
	6th Grade	263	26
	7th Grade	231	23
	8th Grade	218	22
School Type	Imam Hatip Middle School	229	23
	Public Middle School	577	57
	Private Middle School	203	20
Total		1009	100

Additionally, data were recollected for evaluation of test–retest reliability. However, since it was not feasible to obtain data from all participants again, this stage involved only 126 students from a public middle school. Among these participants, 48.4% ($n = 61$) were female and 51.6% ($n = 65$) were male. Regarding grade distribution, 26.2% ($n = 33$) were fifth-grade students, 27% ($n = 34$) were sixth-grade students, 22.2% ($n = 28$) were seventh-grade students, and 24.6% ($n = 31$) were eighth-grade students. All students in this group attended public middle schools, with a total of 126 participants.

Development of the Metacognitive Awareness Scale for Middle School Students

Metacognition is defined as an individual's ability to recognize, monitor, evaluate, and, when necessary, regulate their own cognitive processes (Flavell, 1979). This concept indicates that learning is not limited to the acquisition of knowledge but also involves reflecting on how learning occurs. Metacognitive skills encompass the dimensions of cognitive awareness, planning, monitoring, evaluation, and regulation, enhancing students' ability to manage their learning processes more effectively (Schraw & Dennison, 1994). The literature emphasizes that metacognition is closely connected to various cognitive domains, including academic achievement, problem-solving, critical thinking, and self-regulation (Veenman et al., 2006). Therefore, supporting students' metacognitive skills in educational settings enables them to take responsibility for their own learning and achieve lasting comprehension.

In this study, a metacognitive awareness scale was developed, and its validity and reliability were

examined better to understand students' awareness of their learning processes, their ways of thinking, and how they monitor and evaluate their own learning. Students' awareness of what and how they learn is closely related to their ability to plan, monitor, and regulate the learning process. Accordingly, the developed scale aims to reveal how students apply their metacognitive skills in different learning contexts across various subjects.

This research also aims to contribute to the development of more conscious and practical approaches among educators and future researchers in understanding students' thinking processes. The validity and reliability study of the scale was conducted with 1,009 students enrolled in the 5th, 6th, 7th, and 8th grades at four different middle schools located on the European side of Istanbul during the first semester of the 2025–2026 academic year.

To evaluate the content and face validity of the items in the initial pool, expert opinions were gathered from two professors and two associate professors who have conducted research in the field of metacognition. Additionally, feedback was received from a research assistant specialized in Curriculum and Instruction, an experienced classroom teacher, and a Turkish language teacher responsible for reviewing linguistic clarity and grammar. In total, contributions were provided by six experts and one teacher.

The experts were asked to evaluate each item in the pool as "appropriate," "inappropriate," or "needs revision." Based on their assessments, necessary revisions and modifications were applied to the relevant items. The initial pool, consisting of 60 items, was reviewed by six academicians and one Turkish language teacher to ensure content and face validity. Following the expert evaluations concerning content accuracy, clarity, and conceptual consistency, 14 items were removed from the scale. Specifically, five items were eliminated due to unclear expressions, three because they duplicated other items, three for not adequately representing metacognitive concepts, and three for containing multiple ideas within a single statement. After these revisions, the scale was reduced to 46 items.

Subsequently, to assess the clarity and comprehensibility of the 46-item version, the items were read aloud to a group of 20 students from the 5th, 6th, 7th, and 8th grades. Students were asked to identify any statements they found difficult to understand. Based on their feedback, necessary adjustments were applied. It was noted that 5th-grade students had particular difficulty understanding abstract terms such as "method," "strategy," and "mnemonic." To address this issue, these terms were accompanied by parenthetical examples to enhance comprehension.

As a result of these procedures, the face and content validity of the scale were confirmed, and the instrument was finalized for implementation. The Metacognitive Awareness Scale for Middle School Students was administered during the first semester of the 2025–2026 academic year to a total of 1,009 students (grades 5–8) attending four different middle schools in Istanbul. The validity and reliability analyses of the scale were conducted using data collected from this administration. Specifically, exploratory factor analysis (EFA) was conducted on data from 475 students, and confirmatory factor analysis (CFA) was performed on data from 534 students.

Data Analysis

The Metacognitive Awareness Scale for Middle School Students, developed within the scope of this study, was administered to the designated sample group, and the collected data were analyzed to examine the scale's psychometric properties. The analysis was conducted using SPSS 22.0 statistical software. To evaluate the normality of data distribution, skewness and kurtosis values were examined. According to George and Mallery (2010), values between -1 and +1 indicate that the data are normally distributed. The analysis revealed a skewness value of 0.087 and a kurtosis value of 0.787, confirming that the dataset exhibited a normal distribution.

Within the scope of validity analyses, both content validity and construct validity of the scale were examined. Content validity was ensured through expert opinion, as detailed in the section titled Development of Data Collection Instruments. To determine construct validity, Exploratory Factor Analysis (EFA) was conducted on the first independent sample group ($n = 475$). The analysis was performed using the Principal Axis Factoring (PAF) method, which is more appropriate for identifying latent constructs as it focuses on shared variance among items rather than total variance. Given that metacognitive components are theoretically related, an oblique rotation method (Direct Oblimin) was applied. To determine the number of factors, multiple criteria were considered, including eigenvalues, scree plot, and parallel analysis. The results of the parallel analysis supported a single-factor structure. The number of factors to be retained was determined based on the Kaiser criterion (eigenvalues > 1) and a rigorous examination of the Scree Plot. The scree plot clearly indicated a single-factor structure, showing a sharp break after the first component. Before conducting the factor analysis, the Kaiser-Meyer-Olkin (KMO) coefficient and Bartlett's Test of Sphericity were calculated to assess the suitability of the data. The KMO value of 0.952 and significant Bartlett's test result ($\chi^2 = 8377.999$, $p < .01$) confirmed that the dataset was highly suitable for factor analysis. To confirm the factor structure obtained from EFA,

Confirmatory Factor Analysis (CFA) was subsequently performed with a second independent sample group ($n = 534$). In line with established guidelines, model fit was evaluated using multiple indices, including χ^2/df , CFI, TLI, IFI, and RMSEA.

Additionally, the reliability of the scale was assessed using internal consistency analysis. The analysis yielded a Cronbach's alpha coefficient of .942, indicating a very high level of internal consistency. This result suggests that the items of the scale consistently measure the same construct. Furthermore, test-retest reliability was assessed to evaluate the temporal stability of the scale. The Intraclass Correlation Coefficient (ICC) value was found to be 0.997, demonstrating excellent consistency. This finding indicates that the scale produces stable and dependable results when administered at different times. To provide additional evidence for the validity of the instrument, item-total correlation analysis was also performed. The correlation coefficients ranged from .467 to .651, indicating that all items were significantly and moderately to strongly correlated with the total scale score. This finding confirms that each item corresponds with the overall structure of the scale and contributes meaningfully to the construct being measured.

Findings

Results Related to Construct Validity

Exploratory Factor Analysis

In this study, the sample size was first examined to determine the appropriateness of the data for factor analysis. Although different views exist in the literature regarding sample size, samples smaller than 30 are generally considered weak, while samples up to 500 are regarded as sufficient for most analyses (Büyüköztürk et al., 2009). To further evaluate the data's appropriateness for factor analysis, the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett's Test of Sphericity were conducted (Barnett & Kaiser, 1978). A KMO value of 0.50 or above indicates that the data are suitable for factor analysis, while a significant Bartlett's test ($p < .05$) suggests sufficient inter-variable correlations (Tavşancıl, 2019; Williams et al., 2010). The analysis revealed a KMO value of 0.952 and a significant Bartlett's test result ($\chi^2 = 8377.999$, $p < 0.01$), confirming that the dataset was highly suitable for factor analysis.

During the Exploratory Factor Analysis (EFA), factor loadings were reviewed to determine whether the items represented a specific construct. A factor loading value of .40 or higher is generally accepted as sufficient (Büyüköztürk, 2002). It is recommended that an item's loading on its primary factor be substantially higher than its loading on other factors, with a minimum difference of .10. Items with low

communality values ($< .30$) are considered to have weak representation within the construct (Costello & Osborne, 2005). However, theoretical consistency was also taken into account alongside statistical criteria during item evaluation (Rattray & Jones, 2007).

In line with these principles, the initial 46-item form was analyzed, and 17 items were removed based on statistical and theoretical considerations. The reasons for item removal are summarized below:

- Items 2, 5, 6, 8, 13, 42, 44, 45, and 46 were removed because their factor loadings were below .40.
- Items 16, 23, 24, and 26 were excluded due to high loadings on multiple factors (loading difference $< .10$), which could reduce conceptual consistency.
- Items 27, 29, 33, and 39 were removed because of low communality values ($< .30$), indicating poor representation of the underlying construct.

Following item removal, Exploratory Factor Analysis using the Principal Axis Factoring (PAF) method with Direct Oblimin rotation was conducted. The number of factors was determined using parallel analysis as the primary and most robust criterion, supported by eigenvalues and the scree plot. All criteria consistently

indicated a single-factor structure.

As a result, a single-factor structure consisting of 29 items was obtained, explaining 38.392% of the total variance. In addition, the results of the parallel analysis provided further support for the unidimensional structure, confirming that a single-factor solution was the most appropriate representation of the data. Although metacognition is theoretically conceptualized as a multidimensional construct, the findings of the present study revealed a unidimensional structure. This result can be explained by the developmental characteristics of middle school students, for whom metacognitive processes may function as an integrated and holistic system rather than as clearly differentiated components.

These findings indicate that the scale represents metacognitive awareness as a unified construct among middle school students. Therefore, construct validity was confirmed, and the final version of the scale consisted of 29 items. To further support the validity and reliability of the obtained structure, detailed results of the factor analysis are presented in Table 2, which includes factor loadings item means, standard deviations, item-total correlation coefficients, and other relevant data.

Table 2.

Results of the Factor Analysis for the Metacognitive Awareness Scale for Middle School Students

Item No	Item Statement (English Translation)	M	SD	Item-Total Correlation	Factor Loading
m1	I know which sources (books, videos, lecture notes, etc.) I should use to learn a topic well.	4.07	1.12	.649	.688
m3	I notice the mistakes I make while studying.	3.72	1.07	.551	.589
m4	When I do not understand a topic, I try to learn it by using different methods (reading another source, asking the teacher, trying again, etc.).	3.95	1.13	.597	.633
m7	I know which subjects are easy and which are difficult for me.	3.94	1.23	.556	.596
m9	Before starting a new topic, I think about what I already know and what I need to learn.	3.57	1.21	.511	.543
m10	While learning something, I try to understand why that information is important.	3.76	1.18	.561	.595
m11	I think about how I can use what I have learned in real life.	3.69	1.25	.535	.568
m12	While learning, I ask myself, "How can I understand this topic better?"	3.72	1.14	.571	.603
m14	To avoid forgetting important information, I create reminders (taking notes, setting an alarm on my phone, etc.).	3.64	1.36	.467	.501
m15	Visuals, graphs, and tables help me understand the topic I am studying.	3.61	1.25	.488	.519
m17	Before studying, I determine which topics I need to learn.	3.77	1.23	.625	.661
m18	I plan in advance how to use my time while studying.	3.64	1.23	.508	.542
m19	I prepare the materials I will use (books, videos, notes, etc.) beforehand.	3.82	1.23	.603	.638
m20	I organize a suitable study environment (quiet place, tidy desk, etc.) to concentrate better.	4.00	1.22	.595	.635
m21	While studying, I decide which information to learn first and which to learn later.	3.55	1.27	.626	.662
m22	I know that I need to spend more time on topics that are difficult to understand.	3.84	1.13	.620	.658
m25	I check whether I understand the topic while studying.	3.88	1.18	.651	.688
m28	I assess myself to check what I have learned.	3.81	1.10	.591	.627
m30	I try to identify which parts of a topic are difficult for me.	3.87	1.13	.597	.636
m31	I underline or take notes on the important information in the text I read.	3.86	1.19	.614	.652
m32	When solving questions, I think about which information I need.	3.76	1.17	.547	.583
m34	While studying a topic, I check how much progress I have made.	3.76	1.13	.617	.655
m35	I make sure to complete the assignments given by my teacher on time.	3.97	1.13	.646	.685
m36	While learning a topic, I try to determine which information is most important for me.	3.72	1.16	.618	.655
m37	I reuse the learning methods that worked well before (reviewing my notes, practicing, using visuals, etc.).	3.93	1.13	.640	.678
m38	I review the methods I use while studying (summarizing, listening, solving questions, practicing, using visuals, etc.) and change them if necessary.	3.72	1.17	.550	.588
m40	I organize my study environment to improve my concentration.	3.94	1.18	.593	.629
m41	While studying, I decide which learning method (writing, listening, watching, etc.) helps me learn better.	3.90	1.14	.599	.636
m43	After an exam, I think about how I can correct the questions I got wrong.	3.83	1.18	.531	.567

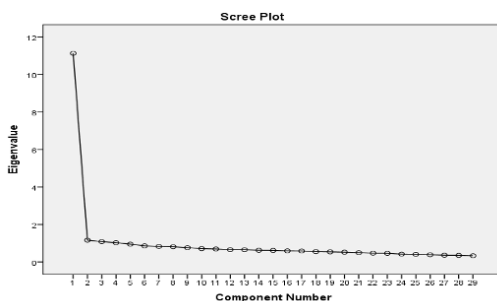
The correlation analysis conducted among the items of the scale revealed that all items were positively and statistically significantly correlated with one another ($p < .01$). This finding indicates that the items were moderately to highly interrelated. Examination of the table demonstrates that all items have significant ($p < .01$) and moderate to strong correlations with the total scale score.

These correlation values confirm that the items of the scale are valid and measure the same theoretical construct. The adequacy of the item-total correlation coefficients provides important evidence for the internal consistency and homogeneity of the scale. The examination of both item-item and item-total correlations suggest that the statements included in the scale sufficiently represent the metacognitive construct intended to be assessed.

To provide a more comprehensive evaluation of the factor structure of the scale, a scree plot illustrating the factor loadings was created (Figure 2). This visual representation supports the analytical findings and clearly assists in determining the number of factors.

Figure 2.

Scree Plot of Eigenvalues for the Metacognitive Awareness Scale for Middle School Students



Confirmatory Factor Analysis

Confirmatory Factor Analysis (CFA) is a statistical method used to test whether the data obtained from a sample fit a theoretically specified model (Sun, 2005). To evaluate model fit, several fit indices are examined. Since each index has its own strengths and limitations, it is recommended to use multiple indices when assessing model fit (Cole, 1987; Sümer, 2000). Commonly used indices include the Chi-square goodness-of-fit test (χ^2), Goodness of Fit Index (GFI), Adjusted Goodness of Fit Index (AGFI), Comparative Fit Index (CFI), Incremental Fit Index (IFI), Tucker-Lewis Index (TLI), Root Mean Square Residual (RMR), and Root Mean Square Error of Approximation (RMSEA).

In this study, a CFA was performed on the 29-item single-factor model identified through Exploratory Factor Analysis (EFA) to confirm the validity of the obtained factor structure. The CFA results demonstrated that all items loaded on a single factor. This represents the

overall structure of metacognition. In other words, all items in the scale were found to contribute to measuring metacognition as a unified construct, confirming that the scale reflects a unidimensional model of metacognitive awareness. Detailed CFA results are presented in Table 3.

Table 3.

Confirmatory Factor Analysis Results for the Metacognitive Awareness Scale for Middle School Students

Model	χ^2/df	NFI	RFI	IFI	TLI	CFI	RMSEA
Metacognitive Awareness Scale for Middle School Students	2.770	.870	.860	.904	.899	.904	.048

When evaluating the fit indices obtained from the CFA, the χ^2/df ratio was found to be 2.770, which falls within the acceptable range reported in the literature and indicates that the model fits the data adequately (Schermmelleh-Engel et al., 2003). The other fit indices were calculated as follows: NFI = .87, RFI = .86, IFI = .904, TLI = .899, CFI = .904, and RMSEA = .048. These findings demonstrate that the model tested in the CFA presents an acceptable level of fit with the observed data.

Sample size is an important factor in interpreting the results of confirmatory factor analysis; however, the literature presents differing perspectives on this issue. Some researchers argue that there should be at least 10 to 20 participants per item, whereas others suggest that the sample size should be at least 200 participants (Hox & Bechger, 1998; Rosseel, 2020). The sample size used in the present study met at least one of these criteria, supporting the validity of the analyses.

The path diagram of the confirmatory factor analysis model is presented in Figure 3, while the regression coefficients related to the model are detailed in Table 4. All relationships among the items within the model were found to be statistically significant ($p < .05$), indicating strong evidence for the scale's structural validity.

Figure 3.

Path Diagram of the Confirmatory Factor Analysis for the Metacognitive Awareness Scale for Middle School Students

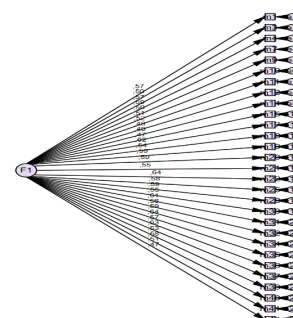


Table 4.
Regression Analysis Results for the Metacognitive Awareness Scale for Middle School Students

Item	Factor	Factor Loading	Standard Error	C.R.	p
M1	Factor 1	.572	—	—	—
M3	Factor 1	.502	.059	13.636	***
M4	Factor 1	.574	.063	15.109	***
M7	Factor 1	.594	.065	15.480	***
M9	Factor 1	.504	.064	13.682	***
M10	Factor 1	.508	.066	13.755	***
M11	Factor 1	.521	.067	14.039	***
M12	Factor 1	.482	.065	13.213	***
M14	Factor 1	.492	.068	13.417	***
M15	Factor 1	.473	.064	13.009	***
M17	Factor 1	.621	.069	15.991	***
M18	Factor 1	.542	.066	14.468	***
M19	Factor 1	.592	.064	15.454	***
M20	Factor 1	.585	.067	15.319	***
M21	Factor 1	.548	.069	14.588	***
M22	Factor 1	.644	.066	16.396	***
M25	Factor 1	.581	.066	15.230	***
M28	Factor 1	.590	.063	15.403	***
M30	Factor 1	.552	.066	14.668	***
M31	Factor 1	.606	.064	15.713	***
M32	Factor 1	.556	.064	14.739	***
M34	Factor 1	.589	.064	15.398	***
M35	Factor 1	.637	.064	16.266	***
M36	Factor 1	.571	.068	15.041	***
M37	Factor 1	.627	.066	16.100	***
M38	Factor 1	.525	.065	14.120	***
M40	Factor 1	.575	.064	15.128	***
M41	Factor 1	.513	.065	13.878	***
M43	Factor 1	.469	.062	12.912	***

Note. p < .001

As shown in Table 4, the factor loadings of the items clustered under a single factor range from .469 (M43) to .644 (M22). All factor loadings were found to be statistically significant ($p < .001$). These results indicate that each item sufficiently represents the factor it belongs to and that the overall structure of the scale aligns well with its intended purpose.

Reliability Findings

Internal consistency is a crucial indicator of reliability that demonstrates the extent to which the items within a measurement instrument represent the same construct and evaluate the internal coherence among them. It expresses the overall integrity of the test and the association of each item with the total score. The internal consistency coefficient generally ranges between 0 and 1, with higher values indicating greater similarity among the items of the scale (Bonett & Wright, 2015; Tavakol & Dennick, 2011; Taber, 2018).

The reliability of the instrument was assessed using the Cronbach's alpha (α) coefficient. In the literature, α values between .80 and 1.00 are widely accepted

as indicators of high reliability (Cortina, 1993; Bonett, 2002). Following the factor analysis conducted within the scope of the research, the 29-item scale yielded a Cronbach's alpha coefficient of .942, indicating that it is a highly reliable measurement tool. Although the Cronbach's alpha coefficient (.942) is relatively high and may suggest potential item redundancy, the items were carefully designed to represent different aspects of metacognitive processes. Therefore, the high internal consistency is interpreted as an indication of strong coherence among items rather than redundancy. Detailed results of the reliability analysis are presented in Table 5.

Table 5.
Reliability Analysis Results for the Metacognitive Awareness Scale for Middle School Students

Scale	Cronbach's Alpha (α)
Metacognitive Awareness Scale for Middle School Students	.942

To determine the stability of the scale scores over time, a test-retest reliability analysis was conducted (DeVellis, 2017). This method involves administering the same measurement instrument to the same group of participants at two different points in time and examining the correlation between the scores obtained from both administrations. It is recommended that the interval between administrations be determined based on the nature of the construct being measured (Streiner et al., 2015). The literature suggests that a time gap of two to four weeks is appropriate and that data from at least 30 participants are sufficient for this type of analysis (Tavşancıl, 2019; Seğer, 2018).

Accordingly, to assess the temporal stability of the scale, two administrations conducted three weeks apart were compared using the Intraclass Correlation Coefficient (ICC). The analysis yielded an $ICC(3,1) = 0.997$, with a 95% confidence interval [0.996, 0.998]. Based on the classification proposed by Koo and Li (2016), ICC values below 0.50 indicate poor reliability, values between 0.50 and 0.75 indicate moderate reliability, values between 0.75 and 0.90 indicate good reliability, and values above 0.90 indicate excellent reliability. Therefore, the obtained result demonstrates that the developed scale provides highly stable and precise measurements over time.

Scoring of the Scale

In Likert-type scales, an individual's total score is typically calculated by summing or averaging the responses provided to each item (Vaske et al., 2017). Such rating instruments are also known as summated rating scales. However, several studies have suggested that the inclusion of reverse-coded (negatively worded) items may negatively impact the internal consistency of the scale and lead to systematic

measurement errors in the factor structure (Cheng & Hamid, 1997; Turgut & Erden, 2013; Wong et al., 2003). For this reason, no negatively worded items were incorporated in the developed scale.

During the analysis process, it was observed that some initially designed reverse items did not load on the expected factor and reduced the internal consistency coefficient. Since these items failed to establish meaningful relationships with the overall construct, they were removed from the scale. Consequently, the final version of the scale consists of 29 positively worded items only. Each item on the scale is rated using a five-point Likert-type format, ranging from 1 (Strongly Disagree) to 5 (Strongly Agree). Accordingly, respondents indicate their degree of agreement with each statement on a continuum where 1 represents complete disagreement, 2 indicates disagreement, 3 reflects neutrality, 4 denotes agreement, and 5 signifies complete agreement. This structure allows for the measurement of participants' metacognitive awareness levels based on the extent to which they endorse each item. The lowest possible total score on the scale is 29, and the highest possible score is 145. Higher scores indicate a greater level of metacognitive awareness among middle school students.

Discussion And Conclusion

During the development of a scale to assess students' metacognitive awareness levels, a comprehensive literature review was conducted, and an initial pool of 60 items was generated. To ensure content and face validity, the items were evaluated by six field experts specializing in metacognition, and a Turkish language teacher assessed the linguistic clarity and grammatical accuracy of the items. Based on expert feedback, 14 items that were either conceptually inadequate or linguistically unclear were removed from the scale.

To test the applicability and reliability of the developed instrument, the scale was administered to 1,009 students enrolled in the 5th, 6th, 7th, and 8th grades at four different middle schools situated on the European side of Istanbul during the first semester of the 2025–2026 academic year. The large sample size, along with data collected from diverse school types and grade levels, enhanced the reliability and generalizability of the results.

Following data collection, an Exploratory Factor Analysis (EFA) was conducted on the initial 46-item version to assess the structural validity of the scale. During the EFA, items with low communalities and factors with fewer than three items were considered unstable and subsequently removed. Using the Principal Axis Factoring (PAF) method with Direct Oblimin rotation, factor loadings, communality values, and conceptual coherence were examined simultaneously. As a result,

17 items (items 2, 5, 6, 8, 13, 16, 23, 24, 26, 27, 29, 33, 39, 42, 44, 45, and 46) were eliminated due to low loadings, cross-loadings, or weak communalities. Ultimately, a single-factor structure comprising 29 items was obtained, explaining 38.392% of the total variance.

Although metacognition is theoretically conceptualized as a multidimensional construct, the unidimensional structure obtained in this study can be explained by the developmental characteristics of middle school students. At this stage, metacognitive processes may operate as an integrated and holistic system rather than as clearly differentiated components.

Although the total variance explained (38.392%) may appear relatively modest, it is considered acceptable in social sciences, particularly for complex and multifaceted constructs such as metacognition. In scale development studies, an explained variance rate of approximately 30% or higher is generally regarded as sufficient to represent the underlying construct (Hair, J. F. et al., 2019; Tabachnick, B. G. & Fidell, L. S., 2013). Previous studies have also reported similar variance levels, supporting the adequacy of this value in representing complex psychological structures (Field, A., 2024).

The structure derived from the EFA was subsequently tested through Confirmatory Factor Analysis (CFA) to verify its validity. The CFA results indicated that the proposed model demonstrated an acceptable fit to the data and was consistent with the theoretical framework. These findings provide strong evidence for the structural validity of the scale and support the adequacy of the unidimensional model. Overall, the results suggest that the scale measures a coherent and valid construct capable of representing middle school students' metacognitive awareness.

To evaluate reliability, the Cronbach's alpha coefficient was calculated, yielding a value of .942, indicating a highly reliable measurement instrument. In general, coefficients between $.60 \leq \alpha < .80$ suggest adequate reliability, whereas coefficients between $.80 \leq \alpha \leq 1.00$ indicate strong reliability. Accordingly, the developed Metacognitive Awareness Scale for Middle School Students demonstrates strong internal consistency.

Additionally, test-retest reliability was examined to assess temporal stability. The Intraclass Correlation Coefficient (ICC) calculated from two administrations conducted three weeks apart was $ICC(3,1) = .997$, 95% CI [0.996, 0.998]. According to Koo and Li (2016), ICC values above .90 indicate excellent reliability. This result confirms that the scale produces stable and consistent measurements over time.

The findings of the present study are consistent with previous research on metacognitive awareness while

also extending the existing literature by providing evidence from a developmentally appropriate instrument specifically designed for middle school students. For example, Karakelle and Saraç (2007) reported that the Children's Metacognitive Awareness Scale (CMAS-C), originally developed with a multifactorial structure for students in Grades 3–9, was subsequently interpreted as a unidimensional construct through the use of total scores. Similarly, Çetinkaya and Erkin (2002) developed a metacognitive inventory for sixth-grade students consisting of four dimensions: evaluation, self-control, awareness, and cognitive strategies. Akin et al. (2007) validated the Turkish version of Schraw and Dennison's (1994) Metacognitive Awareness Inventory (MAI) for university students. Their findings supported an eight-factor structure comprising declarative, procedural, and conditional knowledge together with planning, monitoring, information management, debugging, and evaluation. Taken together, these findings suggest that although metacognitive processes are theoretically conceptualized as distinct dimensions, they may operate as an integrated system during middle school, supporting the unidimensional structure identified in the present study.

Yıldız et al. (2009) developed a metacognition scale for students in Grades 6–8 consisting of eight dimensions representing both metacognitive knowledge and metacognitive regulation. Whereas their study supported a multidimensional structure, the present study identified a unidimensional structure based on empirical evidence obtained through exploratory and confirmatory factor analyses. Although both instruments address similar metacognitive processes, the present scale conceptualizes these processes as manifestations of a single underlying construct rather than as independent dimensions.

Namlu (2004) developed a four-factor Metacognitive Learning Strategies Scale for university students, focusing primarily on planning, monitoring, and evaluation strategies. In contrast, the present study supports a more parsimonious single-factor model that facilitates score interpretation and practical application in school settings. Similarly, the Motivational, Cognitive, and Metacognitive Competencies Scale adapted by Aktamış and Uça (2010) combines motivational, cognitive, and metacognitive competencies within a broader framework rather than measuring metacognitive awareness independently. Similarly, the Metacognitive Learning Strategies Determination Scale (MLSDDS) by Çöğenli and Güven (2014) targeted preservice teachers and identified four factors—planning, monitoring, evaluation, and affective regulation. Unlike these multidimensional instruments, the present scale conceptualizes these components as indicators of a single latent construct, resulting in a more

parsimonious measurement model that may be more practical for educational assessment. Similarly, Demir (2013) proposed a three-dimensional metacognition scale for preservice teachers, again highlighting that previous instruments have predominantly focused on older populations rather than middle school students.

Considering these studies collectively, the unidimensional Metacognitive Awareness Scale for Middle School Students developed in the present research shows theoretical and structural similarities with both single-factor (e.g., Akin et al., 2007; Demir, 2013; Karakelle & Saraç, 2007) and multidimensional but conceptually integrative scales (e.g., Aktamış & Uça, 2010; Çöğenli & Güven, 2014; Namlu, 2004; Yıldız et al., 2009). The primary contribution of the present study lies in the development of a psychometrically validated instrument specifically designed for students in Grades 5–8, using data collected from diverse school settings and addressing the limited availability of developmentally appropriate metacognitive assessment tools in Türkiye. The strong psychometric properties of the scale, together with its practical applicability, suggest that it can serve as a reliable instrument for both educational research and classroom practice. From a developmental perspective, the unidimensional structure identified in this study may indicate that middle school students do not yet differentiate metacognitive knowledge and regulatory processes as separate cognitive systems. Instead, these processes may operate as a unified construct during early adolescence.

The findings of the present study are consistent with foundational theories of metacognition, which conceptualize metacognition as involving both awareness of cognition and the regulation of cognitive processes (Flavell, 1979; Schraw & Dennison, 1994; Veenman et al., 2006). During middle school, metacognitive awareness and regulatory skills play an important role in improving learning quality, supporting self-regulated learning, and promoting academic development (Kuhn, 2000; Veenman et al., 2006). Previous studies have shown that students with higher levels of metacognition tend to exhibit more stable academic performance, employ more effective learning strategies, and demonstrate greater success in problem-solving. Moreover, metacognition supports self-regulation and independent learning by enabling learners to plan, monitor, and evaluate their own learning processes.

The Metacognitive Awareness Scale for Middle School Students developed in this study differs from earlier instruments both theoretically and in measurement scope. For example, Karakelle and Saraç (2007) primarily emphasized cognitive awareness, whereas Akin et al. (2007) adapted the MAI for university students. Consequently, neither instrument was

specifically designed to reflect the developmental and linguistic characteristics of middle school learners. Similarly, Yıldız et al. (2009) employed an eight-factor scale to differentiate metacognition into several subcomponents. In contrast, the present study supports a unidimensional representation in which cognitive awareness, planning, monitoring, evaluation, and regulation function as interrelated indicators of a single latent construct. This interpretation is supported by both the exploratory and confirmatory factor analyses conducted in the present study.

Unlike previous instruments that primarily emphasized learning strategies (Namlu, 2004; Qögenli & Güven, 2014), the present scale was specifically designed to assess metacognitive awareness as a distinct psychological construct. The Motivational, Cognitive, and Metacognitive Competencies Scale (Aktamış & Uça, 2010), though comprehensive, did not isolate metacognition as a distinct construct. In contrast, the current scale evaluates students' metacognitive awareness through observable learning behaviours, measuring not only "knowing what one knows" but also the ability to plan, monitor, regulate, and evaluate one's own learning. Accordingly, the scale reflects the practical dimension of metacognitive awareness by capturing how students consciously manage their learning processes. Moreover, it provides a developmentally appropriate and educationally relevant measure for assessing metacognitive awareness among middle school students.

The primary contribution of this study lies in establishing a developmentally appropriate, process-oriented, and culturally relevant measurement tool for assessing metacognitive awareness in middle school students. Unlike existing instruments, the present scale integrates both metacognitive knowledge and regulatory processes within a unified structure, reflecting the holistic nature of metacognitive functioning in this age group. In this respect, the present study addresses a critical gap in the literature by offering a scale specifically designed for middle school students, both in terms of developmental appropriateness and alignment with contemporary educational practices.

From a practical perspective, the developed scale provides teachers with a valid and reliable tool for monitoring students' metacognitive awareness, identifying areas requiring instructional support, and designing targeted learning interventions. In addition, school counselors and educational psychologists may employ the instrument to evaluate students' metacognitive development, while researchers may use it to investigate the relationships between metacognitive awareness and variables such as academic achievement, motivation, self-regulated learning, and critical thinking. Moreover, the scale can be used in intervention studies to evaluate the

effectiveness of instructional programs aimed at improving metacognitive skills.

Limitations and Future Research

Although the present study provides strong evidence for the validity and reliability of the Metacognitive Awareness Scale for Middle School Students, several limitations should be acknowledged. First, the sample was restricted to middle schools located in specific districts of Istanbul, which may limit the generalizability of the findings to other regions and socio-cultural contexts. Future studies should therefore replicate the scale development process with more diverse samples across different geographical and cultural settings. Second, the data relied exclusively on self-report measures, which may be subject to social desirability bias and students' limited ability to accurately reflect on their own cognitive processes. Incorporating complementary methods such as teacher observations, performance-based assessments, or mixed-method approaches would provide a more comprehensive evaluation of metacognitive awareness. Third, although metacognition is theoretically conceptualized as a multidimensional construct, the analyses revealed a unidimensional factor structure. This outcome may reflect the developmental characteristics of middle school students, yet future research should explore whether multidimensional models can be identified in different age groups or through alternative measurement approaches. Finally, the study employed a cross-sectional design, which does not allow for examination of developmental changes in metacognitive awareness over time. Longitudinal studies would be valuable in tracking how metacognitive skills evolve throughout middle school and beyond. Addressing these limitations will not only strengthen the psychometric robustness of the scale but also contribute to a deeper understanding of metacognitive development in diverse educational contexts.

Author Contributions: All authors contributed equally to the conception, design, analysis, and writing of this study.

Conflict of Interest: The authors declare that there is no conflict of interest among them.

Ethical Declaration: We declare that the rules specified in the Higher Education Institutions Scientific Research and Publication Ethics Guidelines have been followed and that no actions contrary to scientific research and publication ethics have been taken. All authors contributed to the study, and all responsibility for any potential ethical violations lies with the authors.

Ethics Committee Approval: Ethical approval for this study was obtained from the Social and Humanities

Sciences Ethics Committee of Istanbul University-Cerrahpaşa (Date: 04.03.2025, No: 2025/172).

Funding: This research has not received any funding.

Copyright: The data created or analyzed during this study are available from the authors upon reasonable request.

Data Availability Statement: The data created or analyzed during this study are available from the authors upon reasonable request.

Use of Artificial Intelligence in Writing Assistance: No artificial intelligence tools were used during the writing of this manuscript. However, artificial intelligence-based tools (Napkin AI) were used solely for the creation of Figure 1. The research design, data collection, analysis, and interpretation were carried out entirely by the authors.

References

- Abacı, R., Çetin, B., & Akin, A. (2006). Biliş ötesi farkındalık envanteri: Geçerlik ve güvenirlik çalışması. 14. Ulusal Psikoloji Kongresi Özet Kitapçığı, Hacettepe Üniversitesi, Ankara.
- Akin, A., Abacı, R., & Çetin, B. (2007). The validity and reliability study of the Turkish version of the Metacognitive Awareness Inventory. *Educational Sciences: Theory & Practice*, 7(2), 655–680.
- Aktamış, H., & Uça, S. (2010). Motivasyonel, bilişsel ve bilişüstü yeterlilikler ölçeği'nin Türkçeye uyarlanması. *İlköğretim Online*, 9(3), 980–989.
- Annevirta, T., Laakkonen, E., Kinnunen, R., & Vauras, M. (2007). Developmental dynamics of metacognitive knowledge and text comprehension skill in the first primary school years. *Metacognition and Learning*, 2, 21–39. <https://doi.org/10.1007/s11409-007-9002-8>
- Baker, L., & Brown, A. L. (1980). Metacognitive skills and reading. In P. D. Pearson (Ed.), *Handbook of reading research* (pp. 1–74). Longman.
- Barnett, M., & Kaiser, D. (1978). The relationship between intellectual-achievement responsibility attributions and performance. *Child Study Journal*, 8, 209–215.
- Bonett, D. G. (2002). Sample size requirements for testing and estimating coefficient alpha. *Journal of Educational and Behavioral Statistics*, 27(4), 335–340. <https://doi.org/10.3102/10769986027004335>
- Bonett, D. G., & Wright, T. A. (2015). Cronbach's alpha reliability: Interval estimation, hypothesis testing, and sample size planning. *Journal of Organizational Behavior*, 36(1), 3–15. <https://doi.org/10.1002/job.1960>
- Büyüköztürk, Ş. (2002). Faktör analizi: Temel kavramlar ve ölçek geliştirmede kullanımı. *Kuram ve Uygulamada Eğitim Yönetimi*, 32, 470–483.
- Büyüköztürk, Ş., Kılıç-Çakmak, E., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2009). *Bilimsel araştırma yöntemleri* (3. baskı). Pegem Akademi.
- Cheng, S. T., & Hamid, P. N. (1997). Dispositional optimism in Chinese people: What does the life orientation test measure? *International Journal of Psychology*, 32(4), 221–230.
- Cole, D. A. (1987). Utility of confirmatory factor analysis in test validation research. *Journal of Consulting and Clinical Psychology*, 55(4), 584–594. <https://doi.org/10.1037/0022-006X.55.4.584>
- Conway, J. M., & Huffcutt, A. I. (2003). A review and evaluation of exploratory factor analysis practices in organizational research. *Organizational Research Methods*, 6(2), 147–168. <https://doi.org/10.1177/1094428103251621>
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78(1), 98–104. <https://doi.org/10.1037/0021-9010.78.1.98>
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation*, 10(1), 7. <https://doi.org/10.7275/jyj1-4868>
- Coutinho, S. A., & Neuman, G. (2008). A model of metacognition, achievement goal orientation, learning style and self-efficacy. *Learning Environments Research*, 11(2), 131–151. <https://doi.org/10.1007/s10984-008-9043-8>
- Cross, D. R., & Paris, S. G. (1988). Developmental and instructional analyses of children's metacognition and reading comprehension. *Journal of educational psychology*, 80(2), 131–142.
- Çakıroğlu, A. (2007). Üstbiliş. *Türkiye Sosyal Araştırmalar Dergisi*, 11(2), 21–27.
- Çetinkaya, P., & Erkin, E. (2002). Assessment of metacognition and its relationship with reading comprehension, achievement, and aptitude. *Boğaziçi Eğitim Fakültesi Dergisi*, 19(1), 1–11.

- Çögenli, A., & Güven, M. (2014). Bilişüstü öğrenme stratejileri belirleme ölçeğinin geçerlik ve güvenirlik çalışması. *Dicle Üniversitesi Ziya Gökalp Eğitim Fakültesi Dergisi*, 22, 283–297.
- Demir, Ö. (2013). A validation and reliability study of the metacognition scale in Turkey. *Global Journal of Human Social Science*, 13(10), 26–35.
- Desoete, A. (2008). Multi-method assessment of metacognitive skills in elementary school children: How you test is what you get. *Metacognition and Learning*, 3(3), 189–206. <https://doi.org/10.1007/s11409-008-9026-0>
- DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). SAGE Publications.
- Field, A. (2024). *Discovering statistics using IBM SPSS statistics*. Sage publications limited.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34(10), 906–911. <https://doi.org/10.1037/0003-066X.34.10.906>
- Flavell, J. H. (1985). *Cognitive development*. Prentice-Hall.
- Ford, J. K., MacCallum, R. C., & Tait, M. (1986). The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel Psychology*, 39(2), 291–314. <https://doi.org/10.1111/j.1744-6570.1986.tb00583.x>
- George, D., & Mallery, M. (2010). *SPSS for Windows step by step: A simple guide and reference* (10th ed.). Pearson.
- 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>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis*.
- Howard, B. C., McGee, S., Shia, R., & Hong, N. (2000, April). *Metacognitive self-regulation and problem-solving: Expanding the theory base through factor analysis*. [Paper presentation] American Educational Research Association (AERA) Annual Meeting, New Orleans, LA, United states.
- Hox, J. J., & Bechger, T. M. (1998). An introduction to structural equation modeling. *Family Science Review*, 11(4), 354–373.
- Jacobs, J. E., & Paris, S. G. (1987). Children's metacognition about reading: Issues in definition, measurement and instruction. *Educational Psychologist*, 22, 255–278. https://doi.org/10.1207/s15326985ep2203_4
- Jahangard, Z., Soltani, A., & Alinejad, M. (2016). Exploring the relationship between metacognition and attitudes towards science of senior secondary students through a structural equation modeling analysis. *Journal of Baltic Science Education*, 15(3), 340–349. <https://doi.org/10.33225/jbse/16.15.340>
- Karakelle, S., & Saraç, S. (2007). Çocuklar için Üst Bilişsel Farkındalık Ölçeği (ÜBFÖ-Ç) A ve B formları: Geçerlik ve güvenirlik çalışması. *Türk Psikoloji Yazıları*, 10(20), 87–103.
- Karakelle, S., & Saraç, S. (2010). Üst biliş hakkında bir gözden geçirme: Üstbiliş çalışmaları mı yoksa üstbilişsel yaklaşım mı? *Türk Psikoloji Yazıları*, 13(26), 45–60.
- Kline, R. B. (1998). *Principles and practice of structural equation modeling*. Guilford Press.
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Kuhn, D. (2000). Metacognitive development. *Current Directions in Psychological Science*, 9(5), 178–181. <https://doi.org/10.1111/1467-8721.00088>
- Lecce, S., Zocchi, S., Pagnin, A., Palladino, P., & Taumoepeau, M. (2010). Reading minds: The relation between children's mental state knowledge and their metaknowledge about reading. *Child Development*, 81, 1876–1893. <https://doi.org/10.1111/j.1467-8624.2010.01517.x>
- Mokhtari, K., & Reichard, C. A. (2002). Assessing students' metacognitive awareness of reading strategies. *Journal of Educational Psychology*, 94, 249–259. <https://doi.org/10.1037/0022-0663.94.2.249>
- Namlu, A. G. (2004). Bilişötesi öğrenme stratejileri ölçme aracının geliştirilmesi: Geçerlilik ve güvenirlik çalışması. *Anadolu Üniversitesi Sosyal Bilimler Dergisi*, 2, 123–136.
- Owo, W. J., & Ikwut, E. F. (2015). Relationship between metacognition, attitude and academic achievement of secondary school chemistry students in Port Harcourt, Rivers State. *Journal of Research & Method in Education*, 5(6), 6–12.

- Özsoy, G. (2011). An investigation of the relationship between metacognition and mathematics achievement. *Asia Pacific Education Review*, 12, 227–235. <https://doi.org/10.1007/s12564-010-9129-6>
- Öztürk, E. (2012). Okuma stratejileri üstbilişsel farkındalık envanterinin Türkçe formunun geçerlik ve güvenirlik çalışması. *İlköğretim Online*, 11(2), 292–305.
- Pereira-Laird, J. A., & Deane, F. P. (1997). Development and validation of a self-report measure of reading strategy use. *Reading Psychology: An International Quarterly*, 18, 185–235. <https://doi.org/10.1080/0270271970180204>
- Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching, and assessing. *Theory into Practice*, 41(4), 219–225. https://doi.org/10.1207/s15430421tip4104_3
- Rayner, A. (2005, September). *Reflections on context based science teaching: A case study of physics students for physiotherapy* [Poster presentation]. Annual UniServe Science Blended Learning Symposium Proceedings, Sydney.
- Rattray, J., & Jones, M. C. (2007). Essential elements of questionnaire design and development. *Journal of Clinical Nursing*, 16(2), 234–243. <https://doi.org/10.1111/j.1365-2702.2006.01573.x>
- Rosseel, Y. (2020). Small sample solutions for structural equation modeling. In *Small sample size solutions* (pp. 226–238). Routledge. <https://doi.org/10.4324/9780429273872-19>
- Safran, M. (Ed.) (2019). *Dünya eğitime yön verenler* (Cilt 1-2). Millî Eğitim Bakanlığı Yayınları.
- Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, 8(2), 23–74.
- Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science*, 26(1–2), 113–125. <https://doi.org/10.1023/A:1003092417751>
- Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19, 460–475. <https://doi.org/10.1006/ceps.1994.1033>
- Schraw, G., & Moshman, D. (1995). Metacognitive theories. *Educational Psychology Review*, 7(4), 351–371. <https://doi.org/10.1007/BF02212307>
- Seğer, İ. (2018). *Psikolojik test geliştirme ve uyarılama süreci: SPSS ve LISREL uygulamaları* (2. baskı). Anı Yayıncılık.
- Sperling, R. A., Howard, B. C., Miller, L. A., & Murphy, C. (2002). Measures of children's knowledge and regulation of cognition. *Contemporary Educational Psychology*, 27, 51–79. <https://doi.org/10.1006/ceps.2001.1091>
- Streiner, D. L., Norman, G. R., & Cairney, J. (2015). *Health measurement scales: A practical guide to their development and use* (5th ed.). Oxford University Press. <https://doi.org/10.1093/med/9780199685219.001.0001>
- Sun, J. (2005). Assessing goodness of fit in confirmatory factor analysis. *Measurement and Evaluation in Counseling and Development*, 37(4), 240–256. <https://doi.org/10.1080/07481756.2005.11909764>
- Sümer, N. (2000). Yapısal eşitlik modelleri: Temel kavramlar ve örnek uygulamalar. *Türk Psikoloji Yazıları*, 3(6), 49–74.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Pearson Education.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6. Baskı). MA: Pearson.
- Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, 48(6), 1273–1296. <https://doi.org/10.1007/s11165-016-9602-2>
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53–55. <https://doi.org/10.5116/ijme.4dfb.8dfd>
- Tavşancıl, E. (2019). *Tutumların ölçülmesi ve SPSS ile veri analizi* (6. baskı). Nobel Yayıncılık.
- Turgut, T., & Erden, N. S. (2013). Olumsuz test ifadelerinin iç tutarlılığa ve faktör yapısına etkileri. *Istanbul University Journal of the School of Business Administration*, 42(2).
- Vaske, J. J., Beaman, J., & Sponarski, C. C. (2017). Rethinking internal consistency in Cronbach's alpha. *Leisure Sciences*, 39(2), 163–173. <https://doi.org/10.1080/01490400.2015.1127189>
- Veenman, M. V. J., Van Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and Learning*, 1, 3–14. <https://doi.org/10.1007/s11409-006-6893-0>

- Vrugt, A., & Oort, F. J. (2008). Metacognition, achievement goals, study strategies and academic achievement: Pathways to achievement. *Metacognition and Learning*, 3, 123–146. <https://doi.org/10.1007/s11409-008-9022-4>
- Williams, B., Onsman, A., & Brown, T. (2010). Exploratory factor analysis: A five-step guide for novices. *Australasian Journal of Paramedicine*, 8(3), 1–13. <https://doi.org/10.33151/AJP.8.3.93>
- Wong, N. Y., Rindfleisch, A., & Burroughs, J. E. (2003). Do reverse-worded items confound measures in cross-cultural consumer research? The case of material values scale. *Journal of Consumer Research*, 30(1), 72–91. <https://doi.org/10.1086/374697>
- Yıldız, E., Akpınar, E., Tatar, N., & Ergin, Ö. (2009). İlköğretim öğrencileri için geliştirilen biliş üstü ölçeği'nin açımlayıcı ve doğrulayıcı faktör analizi. *Kuram ve Uygulamada Eğitim Bilimleri*, 9(3), 1573–1604.
- Young, A., & Fry, J. D. (2008). Metacognitive awareness and academic achievement in college students. *Journal of the Scholarship of Teaching and Learning*, 8(2), 1–10.