Available online at ijci.wcci-international.org

International Journal of Curriculum and Instruction 17(2) (2025) 673–688

IJCI
International Journal of Curriculum and Instruction

A Science Curiosity Scale for Middle School Students: A Validity and Reliability Study

Fazilet Seçil Gök[a] *, Alev Doğan [b]

[a] *Gazi University Graduate School of Educational Sciences, Ankara, Türkiye*

[b] *Gazi University, Gazi Faculty of Education, Ankara, Türkiye*

Abstract

The primary objective of this study was to develop a valid and reliable scale capable of measuring the multi-dimensional curiosity levels of middle school students in the context of science education. The scale's development commenced with an initial pool of 36 items, which were formulated based on a comprehensive literature review and expert consultations. Following content validity assessments and a pilot administration, the item count was refined to 21. The finalized scale was administered to a total of 668 middle school students during the 2024–2025 academic year. The collected data were analyzed using both exploratory and confirmatory factor analyses. The exploratory factor analysis identified a robust three-factor structure for the scale: Curiosity for Discovering Scientific Knowledge, Curiosity for Using Science in Daily Life, and Experience-Based Scientific Curiosity. These three factors collectively accounted for 45.89% of the total variance. A confirmatory factor analysis further validated this three-factor structure, demonstrating a strong model fit (CFI = .96; TLI = .98; RMSEA = .019). The overall Cronbach's α reliability coefficient for the entire scale was calculated as .92, with the sub-dimensions exhibiting reliability coefficients ranging from .82 to .88. These findings suggest that the newly developed instrument is a comprehensive, valid, and reliable tool for assessing the cognitive, behavioral, and experiential facets of curiosity within science education. The scale is poised to be a valuable resource in the literature for researchers and educators seeking to quantitatively assess the impact of various instructional activities on students' curiosity levels.

Keywords: Science Education, Curiosity Scale, Middle school students

1. Introduction

Curiosity is a foundational human trait that motivates individuals' desire for new information, discovery, and comprehension. This characteristic, which is observable from early childhood, is a vital component of the learning process. The ability to determine students' curiosity levels is essential for understanding their approach to learning and

* Fazilet Seçil GÖK. ORCID ID.: https://orcid.org/0000-0002-9916-9978
E-mail address: fsecilgok@gmail.com

the degree to which they engage in scientific thought processes. Within the domain of science education, students' scientific curiosity significantly contributes to the development of positive learning attitudes, active classroom participation, the application of scientific process skills, and the cultivation of lifelong learning habits. It also serves as a fundamental source of motivation, prompting students to question, explain, and make sense of their surrounding world. Consequently, the valid and reliable measurement of students' science-related curiosity is of paramount importance for both educational researchers and practitioners.

The international literature contains numerous scales designed to measure the construct of "curiosity." While some of these tools, which are frequently used in psychology and education, assess general curiosity, others incorporate dimensions specifically related to scientific or academic achievement. For instance, the "Science Curiosity Scale" (SCS) is a single-dimensional instrument developed to evaluate the level of engagement with science among adults. In contrast, the "Science Curiosity in Learning Environments Scale (SCILE)" was created to measure students' scientific curiosity across various contexts, such as school, home, and museums, and possesses a strong structure in terms of contextual validity. For younger age groups, the "Children's Science Curiosity Scale" is a basic tool that measures students' interest and desire to learn about scientific phenomena. However, it has been critiqued for its single-dimensional design, which is said to not adequately capture the multi-faceted nature of scientific curiosity. The "Five-Dimensional Curiosity Scale" (5DCR), conversely, offers a multi-dimensional framework that explores individuals' curiosity tendencies across five sub-dimensions, including social, emotional, and cognitive aspects. The "Curiosity and Exploration Inventory-II" (CEI-II), developed by Kashdan et al. (2009), is another prominent scale for assessing an individual's drive to seek out new information and experiences. Adaptations of the "Curiosity and Exploration Inventory-II," the "Science Curiosity in Learning Environments Scale (SCILE)," and the "Children's Science Curiosity Scale" have been translated into Turkish, enriching the national literature.

Beyond these adaptations, the few curiosity scales developed domestically in Turkey are primarily focused on fields like language education, literature, mathematics, sports, and social sciences. Scales specifically designed to measure curiosity toward science are limited. For example, the "Early Childhood Science Curiosity Scale," developed by Sarışan Tungaç and Yaman (2023), provides a multi-dimensional assessment of scientific curiosity in children aged five to six. Similarly, the "Science Curiosity Scale," which directly aims to measure science-specific curiosity, is a Turkish adaptation of the single-dimensional scale originally developed by Harty and Beall (1984). While these instruments are valuable contributions, there remains a notable gap in the literature: the absence of a multi-dimensional scale, tailored to the Turkish educational system and cultural context, that specifically addresses science curiosity at the middle school level.

2. Introduction

2.1. Research Model

The study was designed to develop a scale capable of multi-dimensionally assessing middle school students' curiosity levels toward science. To achieve this, the research was structured around the survey model, a quantitative research approach commonly used to construct valid and reliable measurement tools for a specific construct.

2.2. Study Group

The research sample consisted of 668 voluntary middle school students attending public schools with varying academic achievement levels in Ankara, Turkey, during the 2024–2025 academic year. Participants were selected from the 5th, 6th, 7th, and 8th grades using a stratified random sampling method. The students' ages ranged from 10 to 14, with 51% female and 49% male. This sampling technique ensured that representative individuals were selected from each stratum, which were divided by different achievement and grade levels. This approach aimed to enhance the sample's representativeness and to ensure that the developed scale would possess more robust and generalizable psychometric properties by testing it on student groups with different developmental and academic levels. In line with the literature, two separate datasets were used for the Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). Consequently, data from 403 students were used for the EFA, and data from the remaining 265 students were used for the CFA.

2.3. Scale Development Process

In the initial phase, a preliminary item pool of 45 items was created based on a review of the literature on science education, scientific curiosity, scientific inquiry, and individual motivation. These items were drafted to measure students' science curiosity across multiple sub-dimensions. The item pool was then submitted for evaluation to a panel of 7 experts, which included 1 academician in science education, 4 experienced science teachers with doctoral degrees, 1 expert in measurement and evaluation, and 1 expert in the Turkish language. Seeking expert opinion is considered a crucial step for establishing content validity, as it ensures that the items are appropriate in terms of language, scope, and alignment with learning objectives. The inclusion of experts from diverse fields ensured that the scale's structure was well-balanced and consistent in its scientific content, linguistic clarity, and measurement adequacy. For educational measurement tools, a collaborative review by both subject matter and language/measurement experts is a recommended approach to guarantee both content validity and suitability for the target audience.

Following the experts' recommendations, a total of 6 items that were semantically similar or did not align with the scale's purpose were removed, leaving 39 items. A pilot study was then conducted with 42 middle school students outside of the main study group to test the scale's comprehensibility. The pilot data were subjected to a preliminary analysis of the scale items, and potentially problematic items were identified by examining item-total correlations and item variances. Based on student feedback, some item phrasings were simplified, and grammatical adjustments were made. Once the pilot phase was completed, the main data collection process commenced. The scale was administered during class hours at the approved middle schools. During the administration, the researcher and the classroom teacher provided necessary explanations and clarified any unclear phrases. Students filled out the anonymous scale voluntarily, and no identifying information was collected. The administration took an average of 15 to 20 minutes. The completed scale forms were coded and entered into a digital environment for statistical analysis.

Before the data analysis, a double data entry method was employed to ensure the accuracy of the data entry process. Inconsistencies between the two entries were carefully compared and corrected. This method is an effective data validation technique aimed at minimizing typing and coding errors in the dataset, and it is particularly recommended in scale development studies. After these checks, a consistent and clean dataset was prepared for analysis.

3. Findings

The findings are organized under four subheadings: findings related to Exploratory Factor Analysis, findings related to Validity, findings related to Confirmatory Factor Analysis, and findings related to the reliability of the scale.

3.1. Findings on Exploratory Factor Analysis (EFA)

To determine the construct validity, item-total correlation, exploratory factor analysis, and confirmatory factor analysis were conducted. First, however, the normality of the data was examined. The data from the normality analysis are presented in Table 1.

Table 1. Normality Test Results for the Data

Variable	Kolmogorov-Smirnov		
	Statistic	sd	p
Curiosity scale	0.13	672	.000

As shown in the table, the data do not exhibit a normal distribution, with a p-value of .000\<.05. However, the skewness and kurtosis coefficients for the data were within the acceptable range of +1.5 to -1.5 (-.114 and -1.190), and the mode, median, and mean values were quite similar (126, 106, 106). Since the skewness and kurtosis values fell within the –1.5 to +1.5 range, the distribution was considered sufficiently close to normal. Moreover, given the large sample size, the data were accepted as normally distributed, in accordance with the Central Limit Theorem.

Following the determination that the data set approximated a normal distribution, its suitability for factor analysis was evaluated using the Kaiser-Meyer-Olkin (KMO) and Bartlett's Test of Sphericity. The analysis yielded a KMO value of 0.972, and Bartlett's Test of Sphericity was found to be statistically significant (chi2(210)=4405.45; p\<.001). A KMO value above 0.90 and a significant Bartlett's test indicate that the data are exceptionally suitable for factor analysis. The findings confirm that the sample size is adequate and the data set is suitable for factor analysis.

3.2. Findings on the Validity of the Scale

Exploratory Factor Analysis (EFA) was performed to assess the scale's structural validity. During the EFA, Principal Component Analysis (PCA) was used to maximize the contribution of eigenvalues and simplify factor loadings for easier interpretation. The extracted factors were rotated using the Varimax orthogonal rotation technique to more clearly define the relationships between the variables. Varimax is a commonly preferred method in educational research because it enhances interpretability and assumes independence between factors. A factor loading threshold of 0.40 was applied. Items with values below this threshold or with high loadings on more than one factor were eliminated. The Varimax rotation facilitated the dominant loading of items onto a single factor, which made the factor structure conceptually clearer. As a result, a total of 9 items were removed due to either significant cross-loadings on multiple factors or low factor loadings, leading to a final three-factor structure with the remaining 30 items. This resulting structure, both statistically and theoretically, adequately represents the science curiosity construct the scale was designed to measure.

To determine the number of factors, multiple criteria were considered: factors with an eigenvalue above 1 (Kaiser criterion), the cumulative explained variance, and the scree plot. These criteria are widely used to ascertain the statistical significance of factors and whether they present a meaningful model for construct validity. The scree plot is presented in Figure 1.

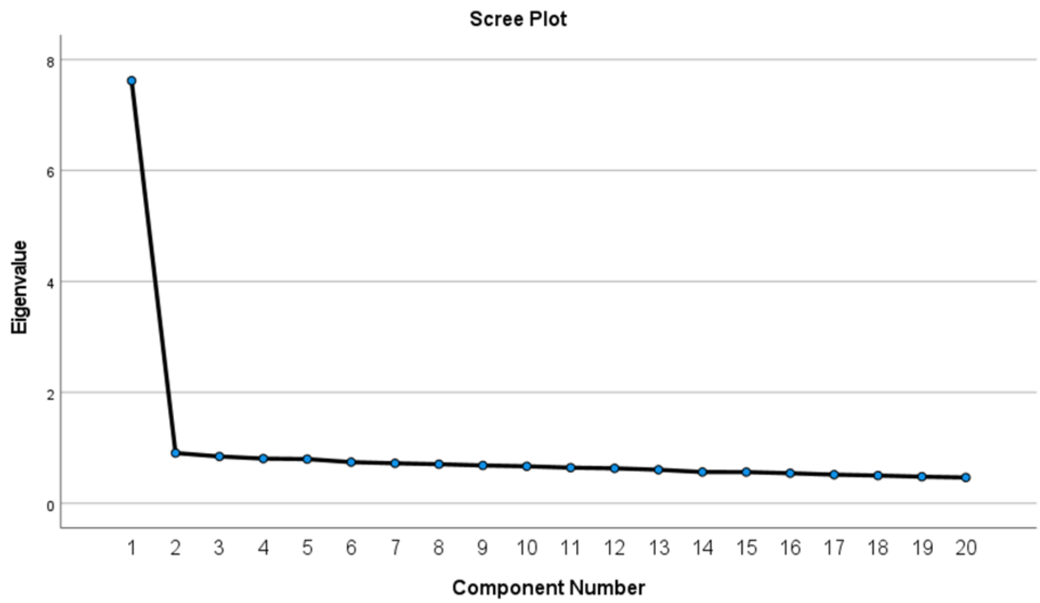

Figure 1. Scree Plot

The initial eigenvalue table showed that only one factor had an eigenvalue above 1 (8.755; explaining 38.07% of the total variance). However, a careful inspection of the scree plot and the scale's theoretical underpinnings suggested that a three-factor model would be more meaningful. Therefore, the analysis was re-run with a three-factor Varimax rotation. During this process, 9 items with factor loadings below 0.40 were sequentially removed, culminating in the final 21-item scale. In its final form:

Factor 1: eigenvalue = 3.864, explained variance = 16.80%

Factor 2: eigenvalue = 3.625, explained variance = 15.76%

Factor 3: eigenvalue = 3.065, explained variance = 13.33%

The three factors collectively explained 45.89% of the total variance. In social sciences, explaining over 40% of the variance is considered a sufficient representation of the underlying structure. The factor loadings for the items in the scale ranged from .56 to .64. Since these values are all above the .40 threshold, they demonstrate a significant relationship between the items and their respective factors, indicating a strong contribution to the measured construct. This range of factor loadings provides evidence for the scale's structural validity and the consistent grouping of items based on their content. The findings further indicate that the scale's items successfully represent the intended constructs and that the instrument can comprehensively measure the multi-dimensional nature of curiosity.

Based on the content analysis of the items, these three factors were named "Curiosity for Discovering Scientific Knowledge," "Curiosity for Using Science in Daily Life," and

"Experience-Based Scientific Curiosity," respectively. The factors, their corresponding items, and factor loadings are detailed in Table 2.

Table 2. Science Curiosity Scale Factor Analysis Results

Item	h²	Factor 1	Factor 2	Factor 3
M29	.404	.569		
M13	.451	.622		
M27	.562	.615		
M18	.451	.618		
M14	.453	.573		
M9	.410	.599		
M7	.475	.611		
M6	.465	.618		
M5	.410	.596		
M1	.552		.610	
M2	.512		.585	
M4	.451		.569	
M20	.565		.603	
M21	.401		.602	
M26	.461		.572	
M17	.465		.638	
M30	.481		.626	
M24	.598			.598
M22	.563			.563
M25	.561			.561
M28	.635			.635

As shown in Table 2, the factor loadings for the nine items in Factor 1 (M5, M6, M7, M9, M13, M14, M18, M27, M29) ranged from .569 to .622. Similarly, the loadings for the eight items in Factor 2 (M1, M2, M4, M17, M20, M21, M26, M30) were between .569 and .638, while the four items in Factor 3 (M22, M24, M25, M28) had loadings between .561 and .635. These values are all well above the conventional .40 threshold for social sciences, indicating that each item loads significantly onto its corresponding factor and successfully represents that dimension. The consistent distribution of factor loadings is a key finding that supports the scale’s validity in terms of both content similarity and structural homogeneity among the items.

In addition to factor loadings, the common variance values (h^2) presented in the table provide insights into the items' representational power within the structure. All common variance coefficients are between .404 and .635, which are above the recommended thresholds. This finding suggests that each item contributes meaningfully to its intended factor and that the structure is well-formed and homogeneous. These findings collectively demonstrate that the three-factor structure of the Science Curiosity Scale is statistically sound and theoretically valid. The fact that the items load significantly onto their intended sub-dimensions is particularly important as it aligns with the scale's theoretical framework.

3.3. Findings on Confirmatory Factor Analysis (CFA)

In the confirmatory factor analysis (CFA) conducted to test construct validity, various fit indices were calculated to determine how well the measurement model aligned with the empirical data. While some different approaches exist in the literature regarding which fit indices should be prioritized, researchers generally advise evaluating multiple measures together rather than relying on a single index. This is because each index assesses model fit from a distinct perspective, and their combined interpretation offers a more reliable understanding of the model's overall validity.

To this end, the following common fit statistics were computed: Chi-square/degrees of freedom (chi2/df), Normed Fit Index (NFI), Tucker–Lewis Index (TLI), Incremental Fit Index (IFI), Comparative Fit Index (CFI), Goodness-of-Fit Index (GFI), Adjusted Goodness-of-Fit Index (AGFI), Root Mean Square Residual (RMR), and Root Mean Square Error of Approximation (RMSEA). The resulting values were evaluated against the "good fit" and "acceptable fit" thresholds suggested in the literature. For instance, a chi2/df ratio of 2 or less is considered excellent, while a value below 5 is acceptable. Values above .90 for comparative indices like CFI, TLI, and IFI indicate a good fit, and an RMSEA value below .08 is considered acceptable, with a value below .05 indicating an excellent fit. Detailed values and comparative interpretations regarding the model's goodness of fit are presented in Table 3.

Table 3. Fit Indices

Index	Good Fit Criterion	Excellent Fit Criterion	Excellent Fit Criterion	Excellent Fit Criterion
x2/Sd	3 ≤ x2/Sd ≤ 5	0 ≤ x2/Sd ≤ 3	*1.24*	Excellent fit
RMSEA	0,05≤ RMSEA ≤ 0,08	0 ≤ RMSEA ≤ 0,05	*.019*	Excellent fit
GFI	0,90 ≤ GFI ≤ 0,95	0,95 ≤ GFI ≤ 1,00	*.969*	Excellent fit
AGFI	0,90 ≤ AGFI ≤ 0,95	0,95 ≤ AGFI ≤ 1,00	*.961*	Excellent fit
RMR	0,05≤ RMR ≤ 0,08	0 ≤ RMR ≤ 0,05	.015	Excellent fit
CFI	0,90 ≤ CFI ≤ 0,95	0,95 ≤ CFI ≤ 1,00	*.961*	Excellent fit
IFI	0,90 ≤ IFI ≤ 0,95	0,95 ≤ IFI ≤ 1,00	*.989*	Excellent fit
TLI	0,90 ≤ TLI ≤ 0,95	0,95 ≤ TLI ≤ 1,00	*.989*	Excellent fit

In light of the findings in Table 3, it can be concluded that the scale's measurement model demonstrates a general fit that aligns with the established threshold values and adequately reflects its theoretical structure. CFA is a widely used analytical technique for testing the degree to which a theoretical model, previously identified through exploratory factor analysis, corresponds with empirical data and for evaluating the overall fit of the measurement model. The evaluation of this model adopted a multi-dimensional approach based on common goodness-of-fit indices.

The analysis revealed that the chi-square/degrees of freedom ratio (chi2/df) for the three-factor, 21-item model was 1.243. This value, being below 2, indicates an exceptionally high level of fit between the model and the observed data. The Root Mean Square Error of Approximation (RMSEA) was found to be .019. This value is significantly lower than the .05 limit accepted in the literature, demonstrating an excellent level of model fit. Furthermore, other goodness-of-fit indices—GFI = .969, AGFI = .961, CFI = .990, IFI = .990, and TLI = .989—indicate that the model possesses a very strong structure in terms of both absolute and comparative fit. The fact that the CFI and IFI values are above .95, in particular, signifies a very good level of model fit. Moreover, a Tucker–Lewis Index (TLI) of .989 suggests that the model is also very efficiently structured in terms of parsimony.

All these findings confirm that the developed three-factor structure is well-supported by confirmatory analysis and that the scale possesses a high degree of construct validity. The attainment of excellent fit levels in sensitive indices like RMSEA and chi2/df shows that the model is built on solid statistical and theoretical foundations. The structural path diagram for the CFA is presented in Figure 2.

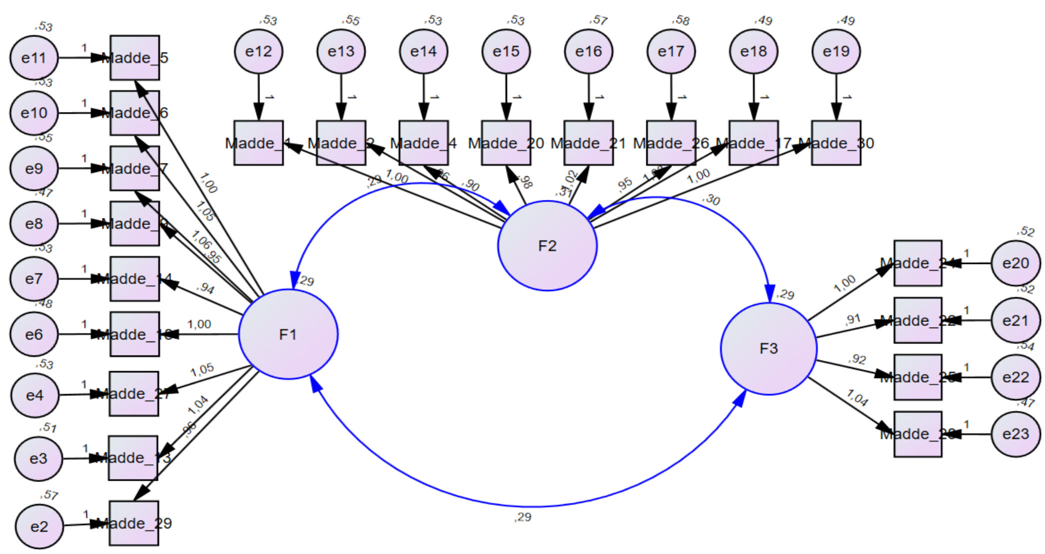

Figure 2. Path Diagram

3.4. Findings on the Reliability of the Scale

The reliability of the developed scale was evaluated using the Cronbach's Alpha (alpha) internal consistency coefficient at both the sub-dimension and whole-scale levels.

The internal consistency coefficient is a widely used statistical indicator for determining the extent to which items measure the same construct and the internal homogeneity of the scale. The Cronbach's alpha values for the three sub-dimensions of the scale were calculated as .836 for F1, .818 for F2, and .876 for F3, respectively. All of these values are above the conventional thresholds of .60 and .70, indicating a high level of internal consistency within each sub-dimension. In the literature, an alpha coefficient above .80 is considered to represent "high" reliability, while a value above .90 indicates "very high" reliability. In this context, the alpha values for all sub-dimensions, which are above .80, show that these sub-dimensions consistently and reliably represent the construct they measure. The Cronbach's alpha coefficient for the entire scale was found to be .919. This value demonstrates that the scale as a whole has a very high level of reliability and can be used as a dependable measurement tool for both individual and group-level assessments. These findings strongly support that the three-sub-dimensional structure of the developed Science Curiosity Scale is a statistically reliable and psychometrically consistent measurement instrument.

4. Discussion

The Science Curiosity Scale developed in this research measures middle school students' curiosity toward science through three distinct sub-dimensions: "Curiosity for Discovering Scientific Knowledge," "Curiosity for Using Science in Daily Life," and "Experience-Based Scientific Curiosity." With its high internal consistency coefficients and a strong factor structure, the scale provides a framework for the multi-dimensional evaluation of curiosity within the domain of science education.

The sub-dimension, "Curiosity for Discovering Scientific Knowledge," encompasses students' eagerness to acquire scientific information, their willingness to conduct research, and their habit of asking questions about science-related topics. This is conceptually similar to the "Epistemic Curiosity Scale (ECS)" developed by Litman and Spielberger (2003), which sought to evaluate individuals' intrinsic motivation to acquire scientific knowledge. That scale's "specific" curiosity sub-dimension, for example, directly overlaps with our scale’s items such as "I can't rest until I learn the answer to a science question" or "Learning important scientific facts satisfies me." Similarly, the first sub-dimension of our scale parallels the core purpose of the "Science Curiosity Scale" by Landrum et al. (2016), which focuses on "interest in scientific research or discovery" and "collecting information about scientific topics." Both scales contain items designed to measure individuals' specific motivation for scientific knowledge and their drive for discovery in this area. A clear conceptual similarity also exists between the first dimension of our scale and the "Children's Science Curiosity Scale" developed by Harty and Beall (1984). The "tendency to acquire basic knowledge" sub-dimension in Harty and Beall's scale measures children's fundamental pursuit of information about objects and events in their environment, driven by questions like "how" and "why." Our "Curiosity for Discovering Scientific Knowledge" sub-dimension targets this same basic pursuit with items such as "I try to learn new words and concepts from science class." This congruence reinforces the theoretical argument that the initial and foundational stage of scientific curiosity is rooted in the desire to discover and acquire information.

The scale's second sub-dimension, "Curiosity for Using Science in Daily Life," pertains to students' ability to connect scientific knowledge with everyday life, apply scientific thinking to daily events, and maintain awareness of scientific processes in their environment. This dimension moves beyond approaches that view curiosity solely as an intellectual pursuit by focusing on the practical application of theoretical knowledge. While many existing scales tend to treat scientific curiosity as a more theoretical endeavor, our second dimension distinguishes itself with items like "I want to learn how to use new technologies or scientific tools," "I look for solutions to daily problems using scientific knowledge," and "I love thinking about how I could invent something." By evaluating students' inclination to use new knowledge for practical purposes and their curiosity toward innovations in their environment, this dimension diverges from other scales. For example, while the "Epistemic Curiosity Scale" by Litman and Collins (2001) measures individuals' desire for knowledge on a specific topic, it focuses more on the cognitive aspect. Similarly, while the "flexibility" sub-dimension of the "Science Curiosity in Learning Environments (SCILE)" by Weible and Zimmerman (2016) addresses an individual's willingness to learn new information, it does not directly focus on concrete applications and problem-solving in daily life, unlike our developed scale. Consequently, this sub-dimension of our scale fills a gap in the existing literature by addressing scientific curiosity in a practical and applied context. The "curiosity for environmental events" and "curiosity for technology" dimensions of the scale developed by Sarışan Tungaç and Yaman (2023) for the preschool level also show a conceptual similarity with this sub-dimension. Both scales are designed to measure scientific curiosity by shifting it from an abstract, theoretical concept to one that is connected to the environment, technology, and practical situations encountered in daily life.

The third factor, "Experience-Based Scientific Curiosity," reflects students' tendency to satisfy their scientific curiosity through hands-on, applied processes such as observation, experimentation, and discovery. This dimension aligns with constructivist learning theories, which posit that learning is an active process of discovery and interaction, rather than passive information reception. Items in this sub-dimension, such as "I try mixing different substances to observe what happens," support the strong link between scientific process skills and curiosity. The "challenge-based curiosity" dimension in the "Science Curiosity Scale" developed by Landrum, Mills, and Johnston (2016) measures an individual's interest in complex, challenging scientific tasks that involve uncertainty. This concept directly overlaps with our scale's items, such as "I find it fun to design my own science project" and "Trying something new and unknown excites me." Similarly, the

"Experience-Based Scientific Curiosity" sub-dimension of the "Scientific Curiosity Scale" developed by Cındıl Kopan (2020) assesses students' inclinations to satisfy their curiosity by actively participating in projects, observations, or laboratory settings. This directly corresponds with this sub-dimension of our developed scale and reinforces the theoretical view that scientific curiosity is founded on concrete applications.

The scale's three sub-dimensions effectively capture the multi-faceted nature of scientific curiosity. This structure presents a robust alternative to existing measurement tools by integrating the epistemic, practical, and experiential aspects of curiosity. The newly developed "Science Curiosity Scale" is a valid and reliable instrument that can be used to measure the impact of various science education activities on students' curiosity levels. The scale's three sub-dimensions comprehensively evaluate students' drive to acquire scientific knowledge, their capacity to connect this knowledge with daily life, and their interest in hands-on learning. Unlike the single-dimensional structures prevalent in much of the international literature, this scale offers a unique and multi-dimensional structure appropriate for the developmental level of middle school students, thereby making a qualitative contribution to the literature.

5. Conclusion

In this research, a 21-item, three-sub-dimensional Science Curiosity Scale with a Cronbach's Alpha coefficient of .919 was developed to measure the scientific curiosity levels of middle school students in a valid and reliable manner. The scale, which uses a 5-point Likert format, provides a comprehensive evaluation of curiosity's different facets, including students' tendencies for active participation in scientific processes, their ability to establish connections between science and daily life, and their interest in experience-based learning. In this context, the developed scale has been introduced into the literature as a valid and original measurement tool that enables the quantitative monitoring of the effects of science education practices on students' curiosity.

References

Berlyne, D. E. (1954). A theory of human curiosity. British Journal of Psychology, 45(3), 180-191.

Bollen & J. S. Long (Eds.), Testing structural equation models (pp. 136–162). Sage.

Braghini, K. M. (2017). Scientific demonstration classes and the teaching of observation. *Revista Brasileira de História da Educação, 17*(2), 45–70.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A.

Büyüköztürk, Ş. (2012). Sosyal bilimler için veri analizi el kitabı. Pegem Akademi.

Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2021). Bilimsel araştırma yöntemleri (30th Edition). Pegem

Bybee, R. W. (2010). Advancing STEM education: A 2020 vision. *Technology and Engineering Teacher, 70*(1), 30–35.

Byrne, B. M. (2010). Structural equation modeling with AMOS: Basic concepts,

Cindil Kopan, T. (2020). Ortaokul öğrencilerinin bilimsel merak düzeylerinin çeşitli değişkenler açısından incelenmesi. (Unpublished Master's Thesis). Karadeniz Teknik University, Graduate School of Educational Sciences, Trabzon.

Costello, A. B., ve Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. Practical Assessment, Research & Evaluation, 10(7), 1–9. https://doi.org/10.7275/jyj1-4868

Çokluk, Ö., Şekercioğlu, G., ve Büyüköztürk, Ş. (2012). Sosyal bilimler için çok değişkenli istatistik: SPSS ve LISREL uygulamaları (Vol. 2). Ankara: Pegem

Dewey, J. (1938). Experience and education. The Macmillan Company.

Driver, R., Asoko, H., Leach, J., Mortimer, E., & Scott, P. (1994). Constructing scientific knowledge in the classroom. Educational Researcher, 23(7), 5–12. https://doi.org/10.3102/0013189X023007005

Engel, S. (2011). Children's need to know: Curiosity in schools. Harvard Educational Review, 81(4), 625–645. https://doi.org/10.17763/haer.81.4.h054131316473115

Field, A. (2009). Discovering Statistics Using SPSS (3rd ed.). Sage Publications.

George, D., & Mallery, P. (2010). SPSS for Windows step by step: A simple guide and reference, 17.0 update (10th ed.). Pearson.

Gottfried, A. E., Fleming, J. S., & Gottfried, A. W. (2016). Science motivation questionnaire II (SMQ-II). In S. L. A. E. G. (Ed.), The Handbook of Motivation at School (pp. 408–426). Routledge.

Guthrie, J. T., & Wigfield, A. (1997). Reading engagement: A framework for promoting motivation and comprehension. Reading & Writing Quarterly, 13(2), 167-177.

Harty, H., & Beall, D. (1984). The development and validation of the Children's Science Curiosity Scale. *Journal of Research in Science Teaching, 21*(6), 633-640.

Harty, H., & Beall, D. (1984). Toward the development of a children's science curiosity measure. Journal of Research in Science Teaching, 21(4), 425–436. https://doi.org/10.1002/tea.3660210408

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling, 6(1), 1–55. https://doi.org/10.1080/10705519909540118

Jirout, J. J., & Klahr, D. (2012). Children's scientific curiosity: In search of an operational definition of an elusive concept. *Developmental Review, 32*(2), 125–160. https://doi.org/10.1016/j.dr.2012.04.002

Kahan, D. M., Landrum, A. R., Carpenter, K., Helft, L., & Jamieson, K. H. (2017). Science curiosity and political information processing. *Political Psychology, 38*(S1), 179–199. https://doi.org/10.1111/pops.12396

Kashdan, T. B., Gallagher, M. W., Silvia, P. J., Winterstein, B. P., Breen, W. E., Terhar, D., & Steger, M. F. (2009). The curiosity and exploration inventory-II: Development, factor structure, and psychometrics. *Journal of Research in Personality, 43*(6), 987–998. https://doi.org/10.1016/j.jrp.2009.04.011

Kashdan, T. B., Rose, P., & Fincham, F. D. (2004). Curiosity and exploration: A new look at an old construct. Journal of Research in Personality, 38(3), 167-196. http://dx.doi.org/10.1207/s15327752jpa8203_05

Kashdan, T. B., Stiksma, M. C., Disabato, D. J., McKnight, P. E., Bekier, J., Kaji, J., & Lazarus, R. (2018). The five-dimensional curiosity scale: Capturing the bandwidth of curiosity and identifying four unique subgroups of curious people. Journal of Research in Personality, 73, 130–149. https://doi.org/10.1016/j.jrp.2017.11.011

Kline, R. B. (2011). Principles and practice of structural equation modeling (3rd ed.). Guilford Press.

Landrum, A. R., Hilgard, J., Akin, H., & Kahan, D. M. (2016). Measuring interest in science: The science curiosity scale. ResearchGate. 10.13140/RG.2.1.3352.8727/1

Landrum, A. R., Mills, C. M., & Johnston, A. M. (2016). Measuring interest in science: The Science Curiosity Scale. Science Education, 100(6), 849–873. https://doi.org/10.1002/sce.21230

Litman, J. (2005). Curiosity and the pleasures of learning: Wanting and liking new information. Cognition and Emotion, 19(6), 793–814. https://doi.org/10.1080/02699930541000101

Litman, J. A., & Spielberger, C. D. (2003). Measuring Epistemic Curiosity and Its Diversive and Specific Components. *Journal of Personality Assessment, 80*(1), 75–86. https://doi.org/10.1207/S15327752JPA8001_16

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. Psychological Bulletin, 116(1), 75–98. https://doi.org/10.1037/0033-2909.116.1.75

Marsh, H. W., Hau, K. & Wen, Z. (2004). In Search of Golden Rules: Comment on Hypothesis-Testing Approaches to Setting Cutoff Values for Fit Indexes and Dangers in Overgeneralizing Hu and Bentler's Findings. Structural Equation Modeling, 11, 320-341. http://dx.doi.org/10.1207/s15328007sem1103_2

McKee, A., Williamson, M., & Ruebush, L. (2007). Science Rocks! Engaging students through hands-on science demonstrations. *The Science Teacher, 74*(7), 38–42.

Millî Eğitim Bakanlığı [MEB]. (2024). Türkiye Yüzyılı Maarif Modeli – Fen Bilimleri Dersi Öğretim Programı. https://mufredat.meb.gov.tr

Minner, D. D., Levy, A. J., & Century, J. (2010). Inquiry in science education: The role of experience, reflection, and discourse. *Journal of Science Education and Technology, 19*(3), 263-281.

National Research Council. (2012). A framework for K–12 science education: Practices, crosscutting concepts, and core ideas. The National Academies Press. https://doi.org/10.17226/13165

OECD. (2019). PISA 2018 Results (Volume I): What Students Know and Can Do. OECD Publishing. https://doi.org/10.1787/5f07c754-en

Piaget, J. (1970). Science of education and the psychology of the child. Orion Press.

Post, T., Walma van der Molen, J.H. Development and validation of a questionnaire to measure primary school children's images of and attitudes towards curiosity (the CIAC questionnaire). *Motiv Emot, 43*, 159–178 (2019). https://doi.org/10.1007/s11031-018-9728-9

Renninger, K. A., & Hidi, S. (2015). The power of interest for motivation and engagement. Routledge.

Sackett, C., Krumm, S., & Hirschi, A. (2015). The role of curiosity in career adaptability. *Journal of Vocational Behavior, 88*, 148-155. https://doi.org/10.1016/j.jvb.2014.11.008

Sarışan Tungaç, A., & Yaman, S. (2023). Erken çocukluk döneminde fen bilimlerine yönelik merak ölçeği'nin geliştirilmesi: Geçerlik ve güvenirlik çalışmaları. Ondokuz Mayıs University Eğitim Fakültesi Dergisi, 42(2), 1035–1072. https://doi.org/10.7822/omuefd.1291463

Schumacker, R. E., & Lomax, R. G. (2016). A beginner's guide to structural equation modeling (4th ed.). Routledge.

Serin, G. (2010). İlköğretim 7. sınıf öğrencilerinin fene karşı meraklarının incelenmesi. Mustafa Kemal University Journal of Graduate School of Social Sciences, 7(13), 237–252.

Tabachnick, B. G., & Fidell, L. S. (2013). Using Multivariate Statistics (6th ed.). Boston, MA: Pearson.

Von Stumm, S., Hell, B., & Chamorro-Premuzic, T. (2011). The hungry mind: Intellectual curiosity is the third pillar of academic performance. *Perspectives on Psychological Science, 6*(6), 574–588. https://doi.org/10.1177/1745691611421204

Weible, J. L., & Zimmerman, H. T. (2016). Science curiosity in learning environments: Developing an assessment tool. *International Journal of Science Education, 38*(8), 1235–1255. https://doi.org/10.1080/09500693.2016.1186853

Appendix A. Science Curiosity Scale

Faktörler	Maddeler	Kesinlikle Katılıyorum	Katılıyorum	Kararsızım	Katılmıyorum	Kesinlikle Katılmıyorum
1. Bilimsel Bilgiyi Keşfetmeye Yönelik Merak	1.İnsan vücudu, uzay, canlılar gibi konular hakkında daha fazla bilgi edinmek isterim.					
	2.Yeni bilgiler öğrendikçe daha fazlasını merak ederim.					
	3.Bir fen sorusunun cevabını bilmesem bile öğrenmeden rahat edemem.					
	4.Fen dersinde geçen yeni kelimeleri ve kavramları öğrenmeye çalışırım.					
	5.İklim değişikliği gibi konuların nedenlerini ve sonuçlarını öğrenmek isterim.					
	6.Bir sorunun çözümünü bulmak için çeşitli kaynaklardan bilgi toplamaya çalışırım.					
	7.Doğadaki olayların nasıl gerçekleştiğini öğrenmek beni heyecanlandırır.					
	8.Kendi fen projemi tasarlamak bana eğlenceli gelir.					
	9.Yeni ve bilinmeyen bir şeyi denemek beni heyecanlandırır.					
2. Günlük Yaşamda Bilimi Kullanmaya Yönelik Merak	10.Yeni teknolojileri veya bilimsel araçları kullanmayı öğrenmek isterim.					
	11.Bilim insanlarının hangi konular üzerinde çalıştığını merak ederim.					
	12.Günlük hayatta karşılaştığım sorunlara fen bilgisiyle çözüm ararım.					
	13.Fen dersinde duyduğum bilimsel kavramları internette araştırırım.					
	14.Her zaman bildiğim şeylerden farklı olanı keşfetmeye çalışırım.					
	15.Farklı maddeleri karıştırıp ne olacağını gözlemlemeyi denerim.					
	16.Bir şeyi nasıl icat edebileceğimi düşünmeyi severim.					
	17.Daha önce duymadığım bir konuyu öğrenmek ilgimi çeker.					
3. Deneyim Temelli Bilimsel Merak	18.Fen dersinde öğrendiklerimi kullanarak yeni fikirler üretmek isterim.					
	19.Fen kitaplarındaki resimleri veya şemaları dikkatle incelerim.					
	20.Bilim insanlarının geçmişte neleri keşfettiğini öğrenmek ilgimi çeker.					
	21.Gözlemler yaparak yeni bir şeyler keşfetmek isterim.					