Bartın University Journal of Faculty of Education
BUEFAD Volume 15, Issue 1

16-28
dergipark.org.tr/buefad
DOI: 10.14686/buefad.1763773

Development of Geographical Inquiry Skill Scale for Secondary School Students*

Samet Karakuş [a]* **& *Yavuz Akbaş*** [b]

a Phd Student, Trabzon University, https://orcid.org/0000-0003-4588-0163 *samet_karakus20@trabzon.edu.tr

b Prof. Dr., Trabzon University, https://orcid.org/0000-0002-3500-4701 yavuzakbas52@gmail.com

Research Article
Received: 12.08.2025
Revised: 02.12.2025
Accepted: 02.12.2025

Abstract

Although various measurement tools exist to quantitatively assess geographical inquiry skill, no scale has been specifically developed for secondary school students. The aim of this study is to develop a Geographical Inquiry Skill Scale for secondary school students. The scale consists of 16 items across three subscales: 'asking and collecting', 'organizing, analyzing and concluding', and 'sharing'. The research was conducted within the scope of a survey model, and data obtained from 421 secondary school students were analyzed. Data analyses were performed using SPSS and LISREL software. Expert opinions were sought to ensure content validity; exploratory and confirmatory factor analyses were conducted to establish construct validity. The scale's reliability was assessed using Cronbach's alpha, item-total correlations, and the 27% upper-lower group discrimination method. Confirmatory factor analysis results showed that the model had acceptable to excellent fit indices (CMIN/DF = 1.38; RMSEA = 0.047; SRMR = 0.05; GFI = 0.91; AGFI = 0.88; IFI = 0.95; CFI = 0.95). The Cronbach's alpha reliability coefficient for the scale was 0.85, indicating high reliability. Furthermore, all sub-dimensions of the scale demonstrated acceptable convergent validity and composite reliability. The relationships between the sub-dimensions were found statistically significant and moderate. In conclusion, it is suggested that the developed scale can be used to quantitatively determine the geographical inquiry skill levels of secondary school students.

Keywords: Geographical inquiry, secondary school students, validity, reliability

Ortaokul Öğrencileri İçin Coğrafi Sorgulama Becerisi Ölçeğinin Geliştirilmesi

Öz

Coğrafi sorgulama becerisini nicel olarak değerlendirmek için çeşitli ölçüm araçları mevcut olsa da, ortaokul öğrencileri için özel olarak geliştirilmiş bir ölçek bulunmamaktadır. Bu çalışmanın amacı, ortaokul öğrencileri için bir Coğrafi Sorgulama Becerisi Ölçeği geliştirmektir. Ölçek, "sorma ve toplama", "düzenleme, analiz etme ve sonuç çıkarma" ve "paylaşma" olmak üzere üç alt ölçekte toplam 16 maddeden oluşmaktadır. Araştırma, tarama modeli kapsamında yürütülmüş ve 421 ortaokul öğrencisinden elde edilen veriler analiz edilmiştir. Veri analizleri SPSS ve LISREL yazılımları kullanılarak gerçekleştirilmiştir. İçerik geçerliliğini sağlamak için uzman görüşleri alınmış; yapı geçerliliğini belirlemek için keşifsel ve doğrulayıcı faktör analizleri yapılmıştır. Ölçeğin güvenilirliği Cronbach alfa, madde-toplam korelasyonları ve %27 üst-alt grup ayrım yöntemi kullanılarak değerlendirilmiştir. Doğrulayıcı faktör analizi sonuçları, modelin kabul edilebilir ile mükemmel uyum indekslerine sahip olduğunu göstermiştir (CMIN/DF = 1,38; RMSEA = 0,047; SRMR = 0,05; GFI = 0,91; AGFI = 0,88; IFI = 0,95; CFI = 0,95). Ölçeğin Cronbach alfa güvenilirlik katsayısı 0,85 olup yüksek güvenilirliği göstermektedir. Ayrıca, ölçeğin tüm alt boyutları kabul edilebilir yakınsak geçerlilik ve bileşik güvenilirlik göstermiştir. Alt boyutlar arasındaki ilişkiler istatistiksel olarak anlamlı ve orta düzeyde bulunmuştur. Sonuç olarak, geliştirilen ölçeğin ortaokul öğrencilerinin coğrafi sorgulama beceri düzeylerini nicel olarak belirlemede kullanılabileceği önerilmektedir.

Anahtar kelimeler: Coğrafi sorgulama, ortaokul öğrencileri, geçerlik, güvenirlik

**This study was presented as an oral paper at the 11th International Eurasian Educational Research Congress (EJER), held in Kocaeli between May 21–24, 2024.*

To cite this article in APA Style:

Karakuş, S., & Akbaş, Y. (2026). Development of geographical inquiry skill scale for secondary school students. *Bartın University Journal of Faculty of Education, 15*(1), 16-28. https://doi.org/10.14686/buefad.1763773

© 2026 Bartın University Journal of Faculty of Education. This is an open-access article under the Creative Commons Attribution NonCommercial 4.0 license (https://creativecommons.org/licenses/by-nc/4.0/).

INTRODUCTION

Humanity in the 21st century is facing global problems on an unprecedented scale, surpassing any challenges in history. Traditional approaches to solving global issues are insufficient. Complex problems, such as climate change, economic inequality, and social injustice, require innovative solutions and qualified individuals capable of devising them. To this end, educational institutions aim to nurture individuals who are equipped with the ability to think globally, implement creative solutions, possess critical thinking skills, are sensitive to local and global issues, and have a high capacity for inquiry. To achieve these goals, skill-based education models are being promoted. Along with skills-based education, it is becoming increasingly important to nurture individuals equipped with 21st-century skills who are inquisitive, analytical, and capable of higher-level thinking, enabling them to produce solutions to global issues. In this context, inquiry-based learning approaches are being given more prominence in educational settings (Yiğit-Özüdoğru & Demiralp, 2022).

The inquiry-based learning method is defined as a student-centred process; it is a teaching approach in which students ask questions, research the answers to these questions, and complete the process based on the answers they obtain (Kaçar & Çakmak, 2020; Kidman & Casinader, 2017). This approach, which differs from traditional teaching methods, aims to enable students to examine the data they collect in detail, relate it to their prior knowledge, identify connections between data, and develop original perspectives (Roberts, 2010). Geographic inquiry is considered a process, skill, and method that reflects inquiry-based learning (Oberle et al., 2019; Yiğit-Özüdoğru, 2021). Geographic inquiry provides students with the opportunity to ask questions about real-world issues, seek answers using various knowledge and skills, construct their own knowledge, and understand that knowledge is dynamic (Davidson, 2006). Through geographic inquiry, students adopt the perspectives and methods of a geographer and become aware of the relationships between humans and the environment. Thus, by observing the patterns and trends that emerge in these relationships, they begin to establish social and environmental connections for solving problems (National Geographic Society, 2017). Teaching practices based on the geographic inquiry method have the potential to equip students with 21st-century skills such as time perception, problem solving, use of geographic information systems and other information and communication technologies, collaboration and communication, interpersonal relationships, self-confidence, alternative thinking, effective presentation, desire to learn, decision making, creative thinking, photography and interpretation, use of evidence, map reading and preparation, critical thinking, entrepreneurship, independent work, and flexibility. In addition, it is stated that this approach contributes to the development of students' social responsibility, cultural awareness, and citizenship skills regarding local, regional and global issues (Demirci, 2006; Roberts, 2013).

Geographical inquiry skill is defined in the literature through different stages. Edelson et al. (2013) proposed a three-stage model: (1) formulating geographic questions, (2) acquiring, organizing and analysing geographic information, and (3) explaining and communicating geographical patterns and processes. Roberts (2003) conceptualises this process in four stages: *creating a need to know, using data, making sense,* and *reflecting on learning*. The Environmental Systems Research Institute (ESRI, 2025) proposes a five-stage model consisting of *Introduction, Exploration, Explanation, Deepening,* and *Evaluation*. Similarly, the National Geographic Society (2017) defines the geographic inquiry process in five stages: *ask, collect, visualize, create,* and *act*. In the 2024 Turkey Century Education Model (TCEM), geographical inquiry skill is included among social science skills and are structured in five stages: *posing geographical questions, collecting geographical information, organising geographical information, analysing geographical information,* and *reaching and sharing geographical conclusions*. Overall, the literature indicates that, despite differences in the number of stages, geographical inquiry skill exhibits a holistic and sequential structure.

Examining these stages in detail, the first stage involves the formulation of geographical questions. At this stage, the student first identifies a problem and asks questions such as ‘What?’, ‘Where?’, ‘When?’, ‘Why there?’ and ‘Why is it important?’ (Yiğit-Özüdoğru, 2021). For example, a student who notices air pollution in their neighbourhood can develop geographical questions that ask where air pollution is most prevalent in Istanbul, the reasons for this situation, and the measures that can be taken to reduce air pollution.

The second stage involves collecting geographical data. In this stage, appropriate data types are determined to answer geographical questions (e.g., photographs, satellite images, statistics, measurements), and data collection tools (surveys, observations, interview forms, fieldwork plans, etc.) are prepared to conduct fieldwork, interviews,

and surveys (National Geographic Society, 2017). Measurements can be taken to observe changes in air pollution over time, surveys can be conducted with citizens, and expert opinions can be sought.

The third stage involves organising and analysing the collected data. At this stage, the data is analyzed by converting it into visual representation tools, such as graphs, tables, charts, and maps; comparisons are made, interpretations developed, and predictions are based on the results (Yiğit-Özüdoğru, 2021). For example, air pollution data measured in different neighborhoods and at various times can be converted into tables and graphs for analysis, and this data can also be visualized on maps.

In the fourth stage, various conclusions are drawn from the analysed data, and geographical questions are answered. These conclusions are supported by generalisations and inferences made from the findings (Demirci, 2006). For example, it may be concluded that areas with high levels of air pollution are generally neighbourhoods with high population and industrial activity. At this stage, solution proposals such as installing filters on factory chimneys, promoting the use of public transportation, and imposing penalties on industrial establishments that cause air pollution can be developed.

In the final stage, the target audience that should play an active role in solving the problem is identified, and the geographical information obtained is communicated to the relevant stakeholders using appropriate methods (National Geographic Society, 2017). For example, data on air pollution and the proposed solutions can be communicated to the relevant governor or local authorities in a slide presentation.

In Turkey, the curriculum was updated in 2024, implementing the Turkey Century Education Model and further emphasizing the importance of skills in the learning process. Within this scope, geographical inquiry skill has been positioned as one of the 17 core skills in the social sciences. This skill must be taught to students in both secondary school social studies and high school geography classes under the new curriculum. Strengthening the skills-based approach in teaching programmes necessitates the reliable measurement of secondary school students' competence in geographical inquiry processes. In this regard, it is anticipated that the assessment tools to be developed will make significant contributions to teachers by helping them understand their students' readiness levels for geographical inquiry skill and by enabling them to better organise learning environments for the acquisition of these skills.

A review of the literature reveals two scales that can be used to quantify geographical inquiry skill. One of these scales was developed for high school students (Yiğit-Özüdoğru & Demiralp, 2022), while the other was an adaptation of the same scale into the Kazakh language (Sadykova et al., 2023). No scale was found that measures the geographical inquiry skill of secondary school students. This study aims to fill the gap in the field. The purpose of the study is to develop a scale of geographical inquiry skill for secondary school students. In line with this purpose, the following sub-questions were addressed:

- How valid is the geographical inquiry skill scale for secondary school students?
- How reliable is the geographical inquiry skill scale for secondary school students?

METHOD

Research Design

The survey model aims to describe the current state of a situation as it is. In this model, the situation or event under investigation is defined within its own terms, without any influence (Karasar, 2022). In this study, this method was used because the aim was to describe the current state of secondary school students' geographical inquiry skill as they are.

Study Group

The study was conducted using convenience sampling, a method selected for its ability to facilitate the data collection process (Gürbüz & Şahin, 2018). The study was conducted during the second semester of the 2023-2024 academic year in five different districts (Ortahisar, Akçaabat, Beşikdüzü, Araklı, and Çaykara) of Trabzon Province, located in the northeastern part of Turkey. The study was applied to a total of 555 secondary school students in grades 6, 7, and 8. However, 134 students who did not complete the scale according to the prescribed protocol were excluded from the study. The analysis was conducted using data obtained from a total of 421 students. The 421 participants were randomly divided into two groups for exploratory factor analysis (EFA) and confirmatory factor analysis (CFA), and the necessary analyses were performed on each group. The demographic information of the study group is presented in Table 1.

Table 1. Demographic Information About Participants

Demographic Information		EFA (n)	CFA (n)
Gender	Female	125	83
	Male	125	88
Class level	6th grade	74	56
	7th grade	86	54
	8th grade	90	61
Total		250	171

Procedure

The research followed the steps suggested by DeVellis and Thorpe (2022) for developing a scale for secondary school students' geographical inquiry skill. The following steps were taken to develop the Secondary School Students' Geographical Inquiry Skill Scale (SSSGISS):

Literature Review and Item Pool Creation

To ensure the content validity of the scale, a comprehensive literature review was conducted, existing scales related to the topic were examined, and an item pool of 34 items was subsequently developed. The item pool was evaluated by nine experts with at least a master's degree and professional expertise in geography and social studies education, including faculty members and in-service geography teachers. During the expert evaluation process, in accordance with the Lawshe (1975) technique, experts were asked to classify each item as "essential," "useful but not essential," or "not necessary." The content validity ratio (CVR) for each item was calculated as CVR = $(N_e - N/2) / (N/2)$, where N represents the total number of experts and N_e represents the number of experts who rated the item as essential. The calculated CVR values were compared with the revised critical CVR of 0.778 reported by Ayre and Scally (2014) for nine experts. Since all items had CVR values equal to or greater than the critical value, no items were removed from the item pool. In line with Lawshe's approach, the content validity index (CVI) was calculated as the mean of the CVR values for the retained items and was 0.947. This value exceeds the minimum acceptable CVI threshold of 0.889 proposed by Ayre and Scally (2014) for nine experts. In this context, it can be concluded that the 34-item draft scale demonstrates strong content validity.

Pilot Study

As part of the pilot study, the draft scale was presented to two students from each of the sixth, seventh, and eighth grades. Researchers evaluated whether each item was understandable to students by interviewing them individually (average 40 minutes). As a result of the interviews and analyses, some items in the draft scale were rephrased to make the scale more understandable in terms of expression, and examples were provided to improve comprehensibility. The linguistic appropriateness of the scale was checked by a Turkish language expert. Subsequently, the pilot application of the scale was conducted with 26 seventh-grade students. Subsequently, the scale data were entered into the SPSS 29 programme, and reliability analyses were conducted. The Cronbach's alpha reliability analysis revealed that the scale had a reliability of 0.927. Item-total correlation calculations were then performed, and no items with a correlation below 0.20 were found. Therefore, no items were removed from the draft scale (Şencan, 2005). Thus, the draft scale consisting of 34 items was made ready for application.

Application of the Measurement Tool

A preparatory text was created to inform participants about the SSSGISS prior to the pilot study and main application. In this text, the definition of geographic inquiry is provided, along with explanations of the concepts of geographic questions and geographic data, accompanied by examples. This ensured that participants had prior knowledge of the skill in question. Participants were provided with a short instruction manual on how to complete the scale. The scale was designed as a 5-point Likert-type scale, with response levels (e.g., 'strongly agree' to 'strongly disagree') indicated on both the front and back pages. This application was carried out to facilitate the response process for participants, taking into account the requirement to complete the back page of the scale. Of the 555 participants who completed the scale, 134 students were excluded from the analysis due to incomplete or incorrect data. As a result, 421 valid scale forms were evaluated in the study.

Data Analysis

In this study, the validity and reliability of the scale were comprehensively analyzed. To determine content validity, Lawshe's (1975) expert opinion technique was applied. EFA and CFA were used together to examine construct validity. Within the scope of convergent validity, average variance extracted (AVE) and composite reliability (CR) values were calculated using the AVECR programme (Aydoğdu, 2023). To provide further evidence for the construct validity of the scale, Pearson's Moment Correlation coefficients were calculated to examine the relationships among the sub-dimensions. Cronbach's alpha coefficient was used to evaluate the

reliability of the scale. Item-test correlations were examined in item analyses, and the discriminative levels of the items were evaluated using an independent groups t-test. The relationships between the subscales were analyzed using Pearson's correlation coefficient. All data analyses were performed using SPSS 29.0 and LISREL 12 software. Prior to the analysis, the normality assumption was checked based on skewness and kurtosis values. The significance level was set at 0.05 for all statistical tests.

Validity and Reliability

The validity of the study was examined through content, factor, and convergent validity analyses (AVE and CR values) and correlation analysis between subscales (Pearson Moment coefficient). For reliability, Cronbach's alpha coefficient, item-total correlations, and independent sample t-test results for upper and lower group comparisons were presented as evidence to reveal the relationship between the scale's subscales.

FINDINGS

This section of the study presents findings regarding the validity and reliability of the scale developed.

Findings Related to EFA

Before EFA, the data distribution was evaluated for normality. The distribution characteristics of the data set were examined using skewness and kurtosis. According to Ramos et al. (2018), skewness and kurtosis values between -1.00 and +1.00 are acceptable. The skewness and kurtosis values obtained in this study ranged from −0.559 to 0.332, indicating that the data approximated a normal distribution and were appropriate for factor analysis.

During EFA, the data were analyzed using the principal component method and oblique rotation technique (Promax). The reason for using principal component analysis is that it is a commonly used method for reducing many variables to a smaller set. Oblique rotation methods, on the other hand, provide more accurate solutions when factors are related (Tabachnick & Fidell, 2014).

Prior to interpreting the factor structure, sampling adequacy and the suitability of the correlation matrix for factor analysis were evaluated. The Kaiser–Meyer–Olkin (KMO) measure was 0.916, indicating a high level of sampling adequacy, and Bartlett's Test of Sphericity was statistically significant ($\chi^2(231) = 1923.139$, $p < .001$), confirming that the data were appropriate for exploratory factor analysis (Büyüköztürk, 2024).

During the EFA process, items with factor loadings below 0.32 (Tabachnick & Fidell, 2014) and items loading on more than one factor with a difference in factor loadings of less than 0.10 (Büyüköztürk, 2024) were removed from the scale. As a result of this elimination process, the scale was reduced to a three-factor structure consisting of 16 items. The resulting three-factor structure accounts for 48.76% of the scale's total variance. Three factors with eigenvalues above 1 were identified. Factor 1 (eigenvalue = 5.179) explained 32.368% of the total variance, Factor 2 (eigenvalue = 1.531) explained 9.570%, and Factor 3 (eigenvalue= 1.092) explained 6.824%. These three factors account for 48.76% of the total variance. The three factors obtained on the scale have been named as follows in accordance with the literature:

1. Asking and Collecting
2. Organising, Analysing, and Drawing Conclusions
3. Sharing

The three-factor structure of the scale was confirmed by examining the scree plot (Figure 1).

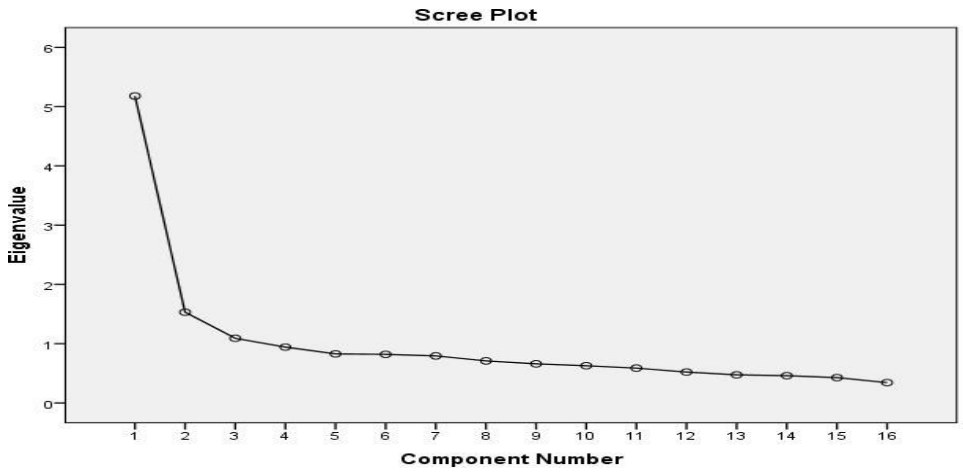


Figure 1. Scree plot view of the three-factor structure

Upon examining Figure 1, a significant decline is observed after points 1, 2, and 3, and the curve then follows a horizontal course from point 4 onwards. This finding provides evidence that the scale has a three-factor structure (Kılıç, 2022). Table 2 presents the factor loadings of the final three-factor solution obtained through EFA for the SSSGISS.

Table 2. Scale Item Loadings by Factor

	Factor loads		
Items	F1	F2	F3
2. I can explain in detail a problem I have noticed in my environment by researching various sources.		,544	
4. I can explain the basic characteristics of a geographical question	,357	,513	
7. I can ask geographical questions about a problem I have noticed in my neighborhood.		,540	
8. I can determine what data is necessary to answer a geographical question.		,700	
9. I can determine how to collect geographical data.		,597	
11. I can prepare data collection tools that can obtain geographical data.	- ,386	,872	
18. I can display the collected data using tools such as maps, tables, and graphs.	,786		
21. I can compare the information contained in tables, graphs, maps, and photographs.	,755		
22. I can interpret maps, graphs, tables, and photographs containing data.	,597		
24. Based on the analysed data, I can conclude the solution to the problem in my environment.	,338		
25. I can support the results achieved with the necessary evidence.	,405		
26. Based on the results obtained, I can make predictions about the future of the problem.	,464		
28. I can decide how the result I have achieved could affect the lives of living beings.	,545		
29. I can determine how I will share the result I have achieved (website, poster, billboard, oral presentation, etc.).			,701
31. I can use technological devices to share the answer to my geographical question in appropriate ways.		-,330	,712
32. I can correctly identify the best person or institution to share the answer to my geographical question with.			,741

Table 2 presents the factor loadings of the 16 items retained across three factors following EFA. All retained items exhibit factor loading above the acceptable threshold, ranging from .338 to .872.

Factor 1 consists of seven items with factor loadings ranging from .338 to .786, with item 18 having the highest loading (.786). Factor 2 consists of six items with loadings ranging from .386 to .700, with item 8 showing the strongest relationship with this factor (.700). Factor 3 consists of three items with relatively high and stable loadings ranging from .701 to .741, with the highest loading observed in item 32 (.741).

Although some items (items 4, 11, and 31) load on more than one factor in the three-factor structure, the differences between factor loadings exceed 0.10, indicating that these items should be retained in the scale and that there are no problematic cross-loadings.

Findings Related to CFA

To validate the EFA-identified structure, a confirmatory factor analysis (CFA) was conducted on a sample of 171 participants. According to Anderson and Gerbing (1984), the sample size for CFA should exceed 100. Therefore, it can be concluded that an adequate sample size was achieved for the CFA. The skewness and kurtosis values of the dataset were -0.402 and -0.642, respectively. These values indicate normal distribution, consistent with the criteria proposed by Tabachnick and Fidell (2023). The path diagram generated by the CFA is presented in Figure 2.

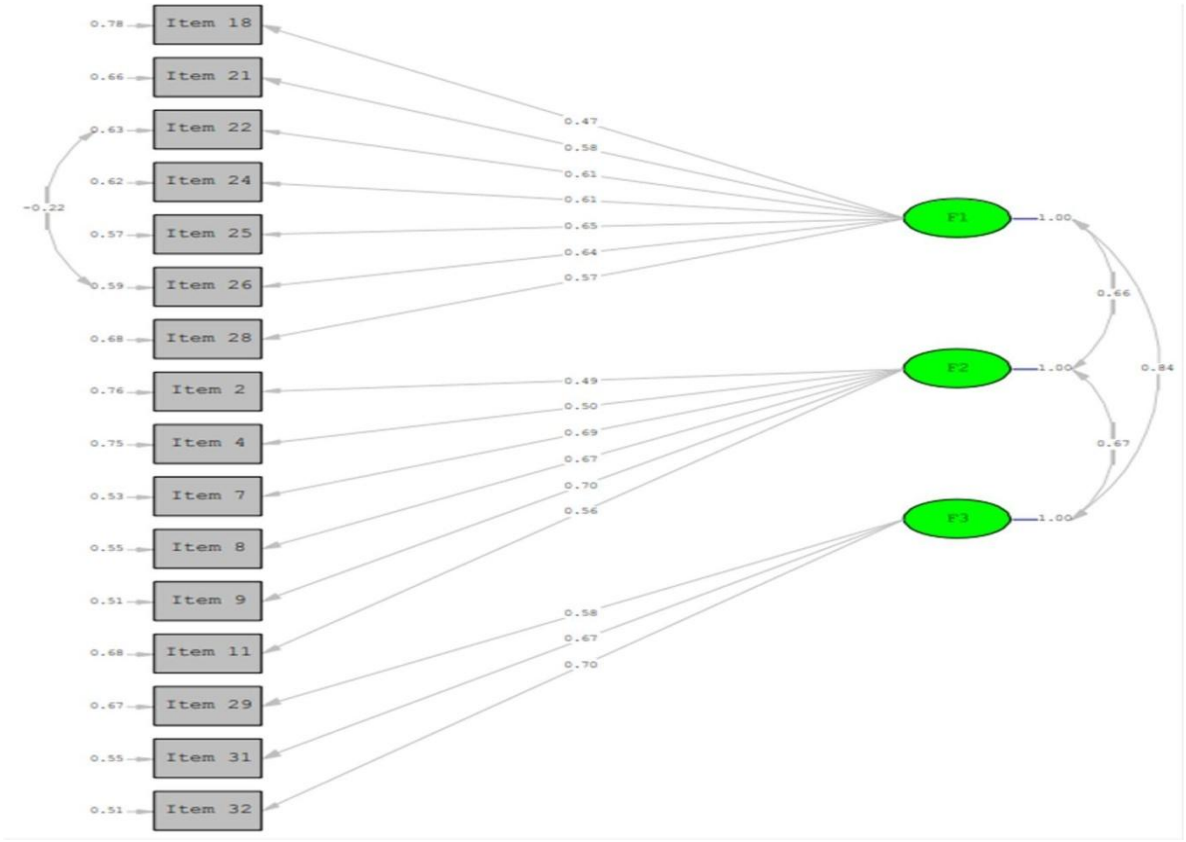


Figure 2. CFA results based on three-factor structure

When examining Figure 2, it can be seen that the standardized factor loadings of the items range between 0.47 and 0.70. According to the first-level CFA results, factor loadings must be at least 0.30 (Büyüköztürk, 2024). The values obtained in this regard meet the specified criteria. When the factor loadings of the items are examined, the highest factor loading (0.70) was found for item 9 ("I can determine how to collect geographic data") in the "asking and collecting" sub-dimension and item 32 ("I can correctly identify the best person or institution to share the answer to my geographic question") in the "sharing" sub-dimension. Additionally, the correlation coefficients between scale factors should not exceed 0.85. Kline (2023) states that in such cases, the correlation coefficients between factors indicate that both factors measure the same structure and that one of the factors should be removed from the model. Looking at the path diagram, the correlation coefficients between the factors are 0.84-0.67 and 0.66. These values indicate that the three factors measure different structures.

Model fit; 'Root Mean Square Error of Approximation' (RMSEA), 'Root Mean Square Error of Approximation' (RMSEA), 'Root Mean Square of Standardized Residuals' (SRMR), 'Incremental Fit Index' (IFI), 'Comparative Fit Index' (CFI), 'Non-Normalised Fit Index' (NNFI), Goodness of Fit Index (GFI), Incremental Fit Index (IFI), Comparative Fit Index (CFI), Adjusted Goodness of Fit Index (AGFI), and x2/df (chi-square/degrees of freedom) criteria. Table 3 presents the fit indices obtained in this study, along with the acceptable and excellent fit values from the literature.

Table 3. Goodness-of-Fit Indices

Indices	Excellent Goodness-of-Fit Criteria	Acceptable Goodness-of-Fit Criteria	Measured values	Conclusion
χ2/ df	$0 \leq \chi 2/df \leq 2$	$2 < \chi 2/ df \leq 3$	138.08/100=1,38	Excellent fit
GFI	.90≤ GFI ≤ 1.00	.85≤ GFI ≤ .90	0.91	Excellent fit
AGFI	.90≤ AGFI ≤ 1.00	.85≤ AGFI ≤ .90	0.88	Acceptable fit
CFI	.95≤ CFI ≤ 1.00	.90≤ CFI ≤ .95	0.95	Excellent fit
IFI	.95≤ IFI ≤ 1.00	.90≤ IFI ≤ .95	0.95	Excellent fit
NNFI	.95≤ NNFI ≤ 1.00	.90≤ NNFI ≤ .95	0.94	Acceptable fit
RMSEA	0≤ RMSEA ≤ .05	.05≤ RMSEA ≤ .08	0.047	Excellent fit
SRMR	0≤ SRMR ≤ .05	.05≤ SRMR ≤ .10	0.054	Acceptable fit

Reference: (Hu & Bentler, 1999; Kline, 2023)

The goodness-of-fit values and reference ranges according to the model's goodness-of-fit indices are shown in Table 3. The model was analyzed, and the X2/df value was found to be 1.38. If the χ2/df value is between 0 and 2, the model is generally considered to have excellent goodness of fit. The RMSEA value (90% CI = 0.016-0.050) was found to be 0.047, indicating an excellent goodness of fit. The same conclusion is reached when examining the CFI, IFI, and GFI values; for an excellent fit, the CFI and IFI values must be 0.95 or higher, and the GFI value must be 0.90 or higher. The model's AGFI value was found to be 0.88, SRMR value 0.054, and NNFI value 0.94, indicating that these values are acceptable for fit within this framework. It can be said that the majority of the analyzed values indicate an excellent fit.

Findings Related to Convergent Validity

Convergent validity demonstrates both the interrelationships among items loading on the same factor and their shared association with the underlying latent variable (Yaşlıoğlu, 2017). In determining convergent validity, the factor loadings, CR and AVE values of the items are calculated (Hair et al., 2014). Table 4 presents the AVE and CR values of the factors.

Table 4. AVE and CR values of factors

Factors	Number of items	AVE Value	CR Value
Asking and collecting	7	0.35	0.78
Organizing, analysing, and drawing conclusions	6	0.36	0.77
Sharing	3	0.42	0.68

When Table 4 is examined, it can be seen that the AVE values of the sub-dimensions are below 0.50, the CR values of the 'asking and collecting' and 'organizing, analyzing and drawing conclusions' sub-dimensions are above 0.70, and the 'sharing' sub-dimension is close to 0.70. Based on these values, it can be interpreted that the CR values meet the desired values. Fornell and Larcker (1981) state that if the CR value is above 0.60, the structure's convergent validity is sufficient even if the AVE is below 0.5. In this regard, it is possible to say that the convergent validity of the sub-dimensions is also sufficient.

Findings Regarding the Relationship Between the Subdimensions of the Scale

To provide further evidence for the construct validity of the scale, Pearson's Moment Correlation coefficients were calculated to examine the relationships among the sub-dimensions. The results are presented in Table 5.

Table 5. Relationship Between Sub-Dimensions

Factors		Asking and collecting	Organising, analysing, and drawing conclusions	Sharing
Asking and collecting	Pearson	1	.644	.250
	Sig.		.000*	.000*
	N	250	250	250
Organising, analysing, and drawing conclusions	Pearson	.644	1	.396
	Sig.	.000*		.000*
	N	250	250	250
Sharing	Pearson	.250	.396	1
	Sig.	.000*	.000*	
	N	250	250	250

*p<0.01

An examination of Table 5 revealed positive and statistically significant correlations among the subscales of the scale. While the strongest relationship was observed at a moderate level (r = .644), the remaining correlations ranged from low to moderate (r = .250–.396), suggesting that the subscales are related yet distinct constructs (Ural & Kılıç, 2013).

Findings Regarding the Reliability of the Scale

The reliability of the scale was assessed using Cronbach's alpha (α) coefficient. In addition, item–tota correlations were examined. By standard psychometric practices (Büyüköztürk et al., 2016), the mean scores of the upper group (top 27%) and lower group (bottom 27%) were compared using an independent samples t-test. Table 6 presents the reliability coefficients of the scale and its subscales, item–total correlation values, and t-test results for the 27% upper–lower group comparison.

Table 6. Reliability Statistics and Item Discrimination Results for Scale Items and Subscales

Items	**Item Total Correlation**	**t (Lower 27% -Upper 27%)** p<0.01*
	SSSGISS (α = 0.852)	
	Asking and collecting (α = 0.779)	
Item 2	,552	9.33
Item 4	,564	9.54
Item 7	,543	7.24
Item 8	,489	9.78
Item 9	,509	9.19
Item 11	,415	9.85
	Organising, analysing and drawing conclusions (α = .0.770)	
Item 18	,511	14.44
Item 21	,486	14.87
Item 22	,597	16.00
Item 24	,522	11.44
Item 25	,554	15.39
Item 26	,530	13.23
Item 28	,374	14.05
	Sharing (α = 0.608)	
Item 29	,356	5.34
Item 31	,224	10.54
Item 32	,405	9.04

Based on table 6 analysis, Cronbach's α (0.852) was calculated for the entire scale, indicating that the scale is reliable. Regarding the subscales, it can be concluded that the coefficients obtained for the 'asking and collecting' subscale and the 'organizing, analyzing and drawing conclusions' subscale are reliable, while the coefficient for the 'sharing' subscale is acceptable (Yurdugül, 2006; Karagöz, 2014).

As shown in the literature, the item-total correlations range from .224 to .597, indicating that the scale items fall within acceptable limits (Şencan, 2005). As a result, it can be concluded that all 16 items meet the necessary standards. To further assess the reliability of the geographical inquiry skill scale, a comparison was conducted between the lower and upper groups (27%) to examine the items' discriminative power. Upon examining the table, it was determined that the t-values were significant (p<0.01). Based on this finding, the scale items have sufficient discriminative power.

DISCUSSION & CONCLUSION

This study aims to develop a geographical inquiry skill scale for secondary school students. In line with this purpose, the study focuses on the process of developing a scale to measure geographical inquiry skill at the secondary school level. Considering the limited number of measurement and assessment instruments available for evaluating geographical inquiry skills (Sadykova et al., 2023; Yiğit-Özüdoğru & Demiralp, 2022), this study is expected to fill an important gap in the field.

During the scale development process, the relevant literature (Akça, 2014; Bednarz et al., 2013; Edelson et al., 2013; Demirci, 2006; Roberts, 2003, 2010; Sadykova et al., 2023; National Geographic Society, 2017; Yiğit-Özüdoğru & Demiralp, 2022) was examined in detail. Based on the analyses, it was determined that the developed scale consists of three sub-dimensions. This structure was supported by the findings of EFA and CFA. The three factors included in the scale were named, in accordance with the literature, as "asking and collecting," "organizing, analyzing, and drawing conclusions," and "sharing" (Demirci, 2006; Roberts, 2003, 2010; Yiğit-Özüdoğru & Demiralp, 2022).

Although the exact number of stages in the geographical inquiry process has not been clearly defined in the literature, it is generally stated that this process consists of three to seven stages (Yiğit-Özüdoğru, 2021). In this context, it can be concluded that the three-dimensional structure of the developed scale and the naming of its sub-dimensions are consistent with the literature. Specifically, the questioning and collecting sub-dimension includes three items for the questioning stage and three for the data collection stage. In the organizing, analyzing, and drawing conclusions sub-dimension, there is one item related to the organizing stage, two items related to the analyzing stage, and four items related to the conclusion-drawing stage. Finally, the sharing stage comprises three items. The representation of the organizing stage with a single item within the "organizing, analyzing, and drawing conclusions" sub-dimension is consistent with the study by Yiğit-Özüdoğru and Demiralp (2022). The item "I can display the collected data using tools such as maps, tables, and graphs" addresses the requirements in this

dimension for geographical inquiry skill. According to the literature, this item can be considered sufficient to represent the organizing dimension (Bednarz et al., 2013; Sadykova et al., 2023; Yiğit-Özüdoğru & Demiralp, 2022). Similarly, in the study by Yiğit-Özüdoğru and Demiralp (2022), the organizing stage was represented by a single item (Item 17: Being able to convert geographical information into maps, graphs, and tables). On the other hand, Bednarz et al. (2013) addressed the organizing stage with a single item, "organizing data and creating representations of data to help solve a problem or answer a question," which corresponds to the item in the present study.

The scale explains 48.76% of the total variance. In the literature, it is stated that an explained variance of over 40% is considered sufficient (Alpar, 2022; Tavşancıl, 2019; Tezci, 2016). The minimum score on the scale is 16, and the maximum is 80. Based on the analysis, the Cronbach's alpha internal consistency coefficient for the entire scale was 0.85. The alpha coefficients for the sub-dimensions were 0.779, 0.770, and 0.608, respectively.

The fit indices obtained from the confirmatory factor analysis (RMSEA = 0.047; GFI = 0.91; AGFI = 0.88; SRMR = 0.054; IFI = 0.94; NFI = 0.95; CFI = 0.95) were examined, and the model was found to have acceptable fit values. In addition, the item–total correlations of the scale items were found to be at a sufficient level, and the upper–lower group comparisons were statistically significant (Şencan, 2005). These findings indicate that the scale items effectively measure the intended construction and are capable of distinguishing between students who possess geographical inquiry skill and those who do not. Based on these results, it can be concluded that the scale is valid and reliable.

Furthermore, significant, positive, and moderate correlations were found among the sub-dimensions of the scale. This finding can be attributed to the interrelated and complementary nature of geographical inquiry skill, as noted by Bruce and Casey (2012), who emphasized that each stage of inquiry supports the next. Accordingly, the scale developed within the scope of this study can be considered a valid and reliable measurement tool for assessing the geographical inquiry skill of secondary school students in Grades 6–8.

The findings of this study support the results of geographical inquiry skills scales developed for high school students by Yiğit-Özüdoğru (2022) and adapted by Sadykova et al. (2023). In both studies, the scales were structured according to the five stages of geographical inquiry. In terms of reliability, Cronbach's alpha coefficients of 0.92 and 0.94 were reported, indicating very high internal consistency. The Cronbach's alpha coefficient of 0.85 obtained in the present study indicates a high and acceptable level of internal consistency within the recommended range for research scales (Hair et al., 2009). While alpha values exceeding 0.90 may indicate item redundancy and the need for scale shortening (Kanbay et al., 2021), the reliability coefficient obtained in this study suggests a balanced level of internal consistency without excessive item overlap. However, the relatively lower Cronbach's alpha value observed for the sharing sub-dimension compared to the other sub-dimensions may be considered a limitation of the study.

Finally, convergent validity was examined using Average Variance Extracted (AVE) and Composite Reliability (CR) values. Although AVE values for the sub-dimensions were below 0.50, CR values exceeded the recommended threshold of 0.60, indicating acceptable convergent validity (Fornell & Larcker, 1981). Previous research suggests that in exploratory or newly developed measurement models, convergent validity may be assessed using more flexible criteria when AVE values exceed 0.25 (Hair et al., 2006; Hsu & Wu, 2013; Mayerl, 2016). Similar findings have been reported in other studies, where some sub-dimensions exhibited AVE values below 0.50 (Akbağ et al., 2025; Gudekli et al., 2022; Kanchanapibul et al., 2014; Mumcu et al., 2022). Within this framework, the convergent validity of the present scale can be considered adequate. However, it should be noted that convergent validity analyses were not reported in some previous geographical inquiry skills studies (Yiğit-Özüdoğru, 2022; Sadykova et al., 2023), which limits direct comparisons. As a result of this study, several suggestions can be made regarding the developed scale:

- The scale can be used to examine the level of geographical inquiry skill at the secondary school level on different samples according to various variables.
- The scale developed in this study can be used to quantitatively assess the level of geographical inquiry skill at the secondary school level and to examine changes over time in longitudinal studies.
- Since the scale is the first scale developed to determine the level of geographical inquiry skill at the secondary school level, its validity and reliability can be retested on different student groups.
- The scale can be adapted to different cultural contexts through cross-cultural adaptation studies.
- The 'Sharing' dimension of the scale can be strengthened by adding new items in future studies.

Limitations

The study is limited to a sample of secondary school students enrolled in five districts of Trabzon province during the 2023-2024 academic year within Türkiye. Furthermore, the study focuses solely on students' geographic inquiry skills. Although the reliability coefficient for the 'Sharing' sub-dimension of the scale (0.60) is considered acceptable in the literature, this is a limitation of the study because it has a lower internal consistency coefficient than other sub-dimensions.

Statements of Publication Ethics

The study was approved by the Scientific Research and Publication Ethics Committee of the Social and Human Sciences at Trabzon University, dated December 4, 2023.

Researchers' Contribution Rate

Authors	Literature review	Method	Data Collection	Data Analysis	Results	Conclusion
Author 1	☒	☐X	☒	☐X	☒	☒
Author 2	☐X	☒	☐X	☒	☐X	☒

Conflict of Interest

There is no conflict of interest among the authors.

REFERENCES

Akbağ, M., Aydoğdu, F., & Rizzo, A. (2025). Developing a novel parental phubbing scale of mother and father forms for adolescents in Türkiye: A validity and reliability study. *Personality and Individual Differences*.

Akça, D. (2014). *Coğrafya öğretiminde coğrafi sorgulama becerisinin öğrenci başarısına, tutum ve kalıcılığa etkisi* [Yüksek lisans tezi]. Necmettin Erbakan Üniversitesi.

Alpar, R. (2022). *Applied statistics and validity and reliability with examples from sports, health and education sciences with analysis steps in SPSS*. Detay Publishing.

Anderson, J. C., & Gerbing, D. W. (1984). The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika, 49*(2), 155–173. https://doi.org/10.1007/bf02294170

Aydoğdu, F. (2023). AVE ve CR calculation. Personal webpage. https://www.fuataydogdu.com/avecr

Ayre, C., & Scally, A. J. (2014). Critical values for Lawshe's content validity ratio: Revisiting the original methods of calculation. *Measurement and Evaluation in Counseling and Development, 47*, 79–86. https://doi.org/10.1177/0748175613513808

Bednarz, S. W., Heffron, S., & Huynh, N. T. (Eds.). (2013). *A road map for 21st century geography education: Geography education research* (A report from the Geography Education Research Committee of the Road Map for 21st Century Geography Education Project). Association of American Geographers.

Bruce, C. B., & Casey, L. (2012). The practice of inquiry: A pedagogical 'sweet spot' for digital literacy? *Computers in the Schools: Interdisciplinary Journal of Practice, Theory, and Applied Research, 29*(1–2), 191–206. https://doi.org/10.1080/07380569.2012.657994

Büyüköztürk, Ş. (2024). *Sosyal bilimler için veri analizi el kitabı* (31. baskı). Pegem Akademi.

Büyüköztürk, Ş., Çakmak, E. K., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2016). *Scientific research methods*. Pegem.

Çokluk, Ö., Şekercioğlu, G., & Büyüköztürk, Ş. (2023). *Sosyal bilimler için çok değişkenli istatistik: SPSS ve Lisrel uygulamaları*. Pegem Akademi.

Davidson, G. (2006). Start at the beginning. *Teaching Geography, 31*(3), 105-108.

Demirci, A. (2006). Coğrafya öğretiminde coğrafi sorgulama becerisinin geliştirilmesi ve kullanımı. *Marmara Coğrafya Dergisi, 0*(14), 61-80.

DeVellis, R. F., & Thorpe, C. T. (2022). *Scale development: Theory and applications* (5th ed.). Thousand Oaks: SAGE Publications.

Edelson, D. C., Shavelson, R. J., & Wertheim, J. A. (Eds.). (2013). *A road map for 21st century geography education: Assessment* (A report from the Assessment Committee of the Road Map for 21st Century Geography Education Project). Washington, DC: National Geographic Society.

ESRI (2025). Getting to Know GeoInquiries an intro to teaching with maps https://storymaps.arcgis.com/stories/939b0102d1bb4af380edc6d9de2e5cc4

Field, A. (2024). *Discovering statistics using SPSS*. SAGE.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50.

Gudekli, A., Deren Van Het Hof, S., Eser, M. T., & Oruc, M. S. (2022). News media literacy scale: Validity and reliability study. *Journal of Economy Culture and Society, 65*, 61–80. https://doi.org/10.26650/JECS2021-837512

Gürbüz, S, Şahin, F. (2018). Sosyal Bilimlerde Araştırma Yöntemleri. Seçkin Yayınları. Ankara.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). *Multivariate data analysis* (7th ed.). Pearson.

Hair, J., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2014). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Sage Publications.

Hsu, L., & Wu, P. (2013). Electronic-tablet-based menu in a full-service restaurant and customer satisfaction — a structural equation model. *International Journal of Business, Humanities and Technology, 3*(2), 61–71.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional

Kaçar, T. ve Çakmak, Z. (2020). Sosyal bilgiler dersinde sorgulamaya dayalı öğretimin öğrencilerin ders başarısına, eleştirel düşünme becerilerine ve öğrenmenin kalıcılığına etkisi. *Elektronik Sosyal Bilimler Dergisi, 19*(76), 1651-1672.

Kanbay, Y., Firat, M., Akcam, A., Cinar, S., & Ozbay, Z.(2021). Development of Firat netlessphobia scaleand investigation of its psychometric properties.*Perspectives in Psychiatric Care, 58*(4), 1258-1266.https://doi.org/10.1111/ppc.12924

Kanchanapibul, M., Lacka, E., Wang, X., & Chan, H. K. (2014). An empirical investigation of green purchase behaviour among the young generation. *Journal of Cleaner Production, 66*, 528–536.

Karagöz, Y. (2014). SPSS ve AMOS uygulamalı istatistiksel analizler. Nobel Akademik Yayıncılık.

Karasar, N. (2022). Bilimsel araştırma yöntemi: Kavramlar, ilkeler, teknikler (37. Baskı). Nobel Yayıncılık.

Kidman, N., & Casinader, G. (2017). *Inquiry-based teaching and learning across disciplines comparative theory and practice in schools.* Basingstoke, England: Palgrave Macmillan.

Kılıç, A. F. (2022). "Açımlayıcı Faktör Analizinde Boyut Sayısına Karar Verme: Yöntemlere Kısa Bir Bakış", *Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 2022 Sayı 51*: Özel sayı 1, Denizli, ss. Ö305-Ö318.

Kline, R. B. (2023). *Principles and practice of structural equation modeling*. Guilford Press.

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology, 28*(4), 563–575. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x

Mayerl, J. (2016). Environmental concern in cross-national comparison: Methodological threats and measurement equivalence. In Green European. Routledge, 182-204.

Mumcu, A., Çakarel T. Y, & Demirbağ, O. (2022). Girişimcinin sosyal sermayesi ölçeği: Bir ölçek uyarlama çalışması. *Journal of Research in Business, 7*(2), 563-581. DOI: 10.54452/jrb.1148869

National Geographic Society (2017), *The geo-inquiry process: Educator guide*, National Geographic Society, Washington,DC.https://media.nationalgeographic.org/assets/file/Educator_Guide_Geo_Inquiry_Final_2.pdf

Oberle, A. Joy Bess, Kelsey Ehmke, Shelley Rath & Amber Robbins (2019) Geo-Inquiry: Informed Action to Make Our World a Better Place, *The Geography Teacher, 16*:4, 170-178, DOI: 10.1080/19338341.2019.1662467

Ramos, C., Costa, P.A., Rudnicki, T., et al. (2018). The effectiveness of a group intervention to facilitate posttraumatic growth among women with breast cancer. *Psycho-Oncology, (27),*258–264. https://dx.doi.org/10.1002/pon.4501

Roberts, M. (2003). *Learning through enquiry: Making sense of geography in the key stage 3*. Geographical Association.

Roberts, M. (2010). Geographical enquiry. *Teaching Geography, 35*(1), 6-9.

Roberts, M. (2013). *Geography through enquiry.* Sheffield: Geographical Association

Sadykova, B., Yesnazarova, U., & Tokbergenova, A. (2023). The geographic inquiry process skills scale: A validation study. *Eurasia Journal of Mathematics, Science and Technology Education, 19*(9), em2319. https://doi.org/10.29333/ejmste/13471

Şencan, N. (2005).Sosyal ve davranışsal ölçümlerde güvenilirlik ve geçerlilik, Ankara: Türkiye, Seçkin Yayıncılık, ss:105-403

Tabachnick, B. G., & Fidell, L. S. (2014). *Using multivariate statistics* (6th ed.). Pearson Education Limited.

Tavşancıl, E. (2019). *Tutumların ölçülmesi ve SPSS ile veri analizi* (6th ed.). Nobel Akademik Yayıncılık.

Tezci, E. (2016). *Eğitimde ölçme ve değerlendirme* (1st ed.). Detay Yayıncılık.

Türkiye Century Education Model (2024). Geographic inquiry skill. (https://tymm.meb.gov.tr/)

Ural, A., & Kılıç, İ. (2013). *Bilimsel araştırma süreci ve SPSS ile veri analizi*. Detay.

Yaşlıoğlu, M. M. (2017). Sosyal bilimlerde faktör analizi ve geçerlilik: Keşfedici ve doğrulayıcı faktör analizlerinin kullanılması. *İstanbul Üniversitesi İşletme Fakültesi Dergisi, 46*, 74–85.

Yiğit-Özüdoğru, H. (2021). *Coğrafya eğitiminde coğrafi sorgulama becerilerinin değerlendirilmesi* [Yayımlanmamış doktora tezi]. Gazi Üniversitesi.

Yiğit-Özudoğru, Y. H., & Demiralp, N. (2022). Developing a geographic inquiry process skills scale. *Education Inquiry, 13*(3), 374-394. https://doi.org/10.1080/20004508.2020.1864883

Yurdugül, H. (2006). Paralel, eşdeğer ve konjenerik ölçmelerde güvenirlik katsayılarının karşılaştırılması. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi, 39*(1), 15–37.