International Journal of Educational Researchers
14(3): 43-56
ISSN: 1308-9501, Copyright © 2023
DOI: 10.29329/ijer.2023.595.4

Educational Research Association Turkiye • 2008
www.eab.org.tr

ORIGINAL ARTICLE

Development of the Scientific Literacy Scale: A Study of Validity And Reliability

Ceren Mutluer*

Asst. Prof. Dr., Measurement and Evaluation in Education, Bolu Abant Izzet Baysal University, Bolu, Turkiye.

Ethical Statement
The study was approved by the institutional board of Bolu Abant Izzet Baysal University Ethics Committee (Protocol No: 2022/122-05.05.2022) and informed consent was taken from all individual participants.

Funding Information
No funding was received for the study.

Conflict of Interest
No conflict of interest is present in the conduction or the reporting of this study.

ABSTRACT

In this study, it was aimed to develop a scale that measures the concept of scientific literacy, which is often emphasized as one of the 21st century skills. Literature was reviewed and qualitative data were collected, and 48 items were initially created in the item pool. Data were collected from 976 university students for the study. When testing the validity of the scope, 2 items were removed. Construct validity was first examined by EFA. After exploratory factor analysis (EFA), 2 items were removed. AVE and CR values were calculated for convergent and divergent validity. Confirmatory factor analysis (CFA) was performed to test construct validity. In this process, 3 items were removed. Reliability analysis was performed with 41 items. In the process of collecting data for the measurement tool, it was expected for measurement invariance to prove whether the measured feature for men and women changed and whether it proved the validity of the structure. While formal and metric invariance did not change according to gender variable, scalar invariance could not be achieved.

Keywords: the scientific literacy scale, scale development, validity and reliability.

Received: 18/08/2023 **Accepted**: 10/09/2023

***Corresponding Author**
Asst. Prof. Dr. Jackie Chan, English Language Teaching, Çanakkale Onsekiz Mart University, Turkiye
Email: jackiechan@email.com

INTRODUCTION

From an objective point of view, the process of explaining nature, and the universe more generally, is one of the main issues that individuals are curious about. Observing the events in the universe and examining the events in detail constitutes the subject area of science. Subject areas in science vary according to the situation examined. According to their subject areas, they are divided into positive sciences (physics, chemistry, biology, psychology, sociology, anthropology), formal sciences (mathematics and logic), art and humanities (literature, painting, music) and metaphysics (religion). Although these sciences have different focuses, a scientific study basically involves the same systematic steps. According to Erkuş (2021), these steps are as follows: I) identification of the problem; II) preparation of hypotheses or research questions related to the problem; III) deciding the means of research to test hypotheses; IV) collection of data; V) analyzing the data to make it meaningful; VI) making interpretations and generalizations; VII) writing the report; VIII) orientation to new problems depending on the results of the research. One of the qualities that researchers should have for this stage-based process is scientific literacy.

Scientific literacy was first mentioned by Hurd (1958) in science and technology education and became the subject of scientific research (Turgut, 2005). The concept of science literacy has taken its place as a concept that is more considered and emphasized thanks to the progress in science and technology. This concept is explained by establishing the nature of science, scientific content knowledge, and the science-technology-society relationship (Miller, 1983). Scientific literacy requires the ability to use scientific knowledge, understand the world by identifying problems, producing evidence-based conclusions, and make decisions about changes caused by humans (Bybee, 1997). Individuals who are scientifically literate are able to know the relationship and interaction between science and society, understand the nature of science, know the ethical values that the scientist adheres to, understand the basic ideas in science, and understand the difference between science and the humanities (Pella et al., 1966).

The basic literacy rates that show the development level of countries seem to have left their place to the concept of science literacy in the 21st century (Ulukan, 2021). Scientific literacy has now become an internationally recognized educational motto and contemporary educational goal (Laughsch, 2000), and in a world where science and technology are increasingly being shaped, science and technology literacy is now a universal necessity.

The concept of scientific literacy has become the subject of many studies considering its emphasis on 21st century skills. On the level of scientific literacy (Şahin and Say, 2010; Şahin and Ateş, 2018; Turgut, 2007), there are some studies which searched for the relationship with critical thinking (Bakırcı et al., 2020; Tekin et al., 2016), and its relationship with problem-solving skills (Ertek et al., 2013; Mertoğlu and Öztuna, 2004; Tezel and Tezgören, 2019).

Purpose

When the measurement tools used in the research studies were examined, either the informational and gradual process related to the scientific research process or the measurement tools which emphasize only scientist characteristics were encountered. When the measurement tools for cognitive and affective awareness about scientific research processes including the concept of scientific ethics and the recognition of the individual's own competencies in the study process were examined, it was determined that there were measurement tools that emphasized a different part in each study. For this reason, in order to close this gap in the field, it was aimed to develop a measurement tool that addresses the concept of 'scientific literacy' in detail.

METHOD

This study is a descriptive study in which the validity and reliability analyses of Scientific literacy scale were conducted, and the psychometric properties of Scientific literacy scale were determined.

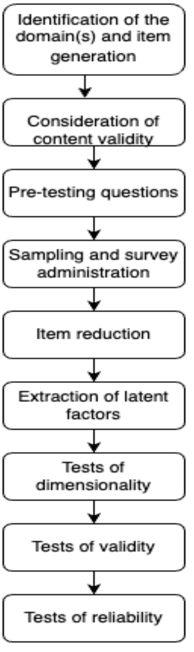
Participants

In this study, data were collected from 976 (524 female and 452 male) university students. Bolu Abant İzzet Baysal University, Faculty of Education teacher candidate students participated in the study voluntarily. Random sampling method and GPower 3.1 program were applied to decide the number of people to achieve a major impact. The program was run on the basis of a two-way hypothesis, 2 intercept (slope, intersection value), Type I error (α) value 0.05 (significance level), Type II error (β) value 0.95. When the calculated values were examined, a sample of at least 567 people with a 95% confidence interval showed that they would give appropriate results for this research. The expected power value with this sample number was determined as 0.9502 (α=0.05; β=0.95; $t_{critical}$=1.9642). When the power analysis results and the current sample number are compared, it is found out that more people are reached than the minimum number of people who were supposed to be reached. In the study, 436 people took part in EFA and the remaining 538 people took part in CFA.

Developing the Scale

The stages proposed by Boehmer et al. (2018) were followed to improve the measuring instrument. These stages are indicated as follows.

Figure 1. Scale development process



Initially, the item pool has been created. In order to create a pool of substances, the literature in which the concepts of 'science literacy' and 'scientific literacy' mentioned were reviewed. While creating the item pool, university students who took 'research methods in education' or 'scientific research methods' courses were also asked to write essays on the subject. 6 participants wrote the essays voluntarily and 48 items have been prepared for the substance pool. Then, expert opinions (2 measurement and evaluation experts, 1 science and technology education specialist and 1 Turkish education specialist) were taken for the items, and 2 items were removed from the scale. By excluding the substances issued in line with the expert opinions, a total of 46 items were created. In terms of readability and comprehensibility of the articles, scale items were given to 12 university students independent from 6 students who had written essays. These 12 university students were randomly selected and were volunteers. Together with the feedback received from the teacher candidates, the final version of the 46-item pool was provided.

Procedure

Data was collected between January 2023 and March 2023. In the study process, 976 university students were reached. 350 of participants' answers were collected through face-to-face and 626 of them were collected online (Google forms). There is no missing data in the study. Participants first read the consent form for their voluntary participation, and they were informed of their random participation in the study. They were then informed about the purpose of the survey. It took approximately 20 minutes for the participants to answer.

Data Analysis

In the study, exploratory factor analysis was performed as the first stage of construct validity to determine whether the items were good indicators of the measured properties. SPSS V 25.0 for EFA (IBM; Armonk, NY, USA). Kaise-Meyer Olkin (KMO) test and Barlett test results were reported to determine the factorization status of the substances. With the Varimax transformation, it was examined under which dimensions the factors were grouped. In the analysis process, a value of 0,40 and above was accepted as the critical value for the factor load value. A factor loading above 0,40 is a good indicator/representative for the size studied (Stevens, 1992). The results of convergent and divergent validity were also reported in the construct validity. The Average Variance Extracted (AVE) value was required to be greater than 0,50 and the Composite Reliability (CR) value to be greater than 0,70. The results were interpreted according to these criteria. Confirmatory factor analysis was used as another stage of construct validity. LISREL V 8.80 was used for this analysis. When looking at the multiple normality assumption, the Robust Maximum Probability (R-ML) estimation was used. Chi-square, chi-square/df ratio, Comparative Fit Index (CFI), Non-Normed Fit Index (NNFI), Tucker Lewis Index (TLI), Root-Mean-Square Error of Approximation (RMSEA) and Standardized Root Mean Square Residual (SRMR) values were reported for model fit indices. Multi-group CFA was performed to check whether the developed measurement tool has measurement invariance according to gender variable. The difference in model fit values (ΔCFI, ΔRMSEA and ΔSRMR) in the measurement invariance examined gradually was examined by considering criterion 0,01 (Chung et al., 2016).

RESULTS AND DISCUSSION

Testing Validity

The Results of Explanatory Factor Analysis (EFA)

Factor analysis was first performed to ensure the construct validity of the measurement tools used in the research. Factor analysis is carried out in order to collect a large number of observed variables under a smaller number of dimensions, to make definitions using the observed variables, to determine the sub-dimensions of the developed scale, to determine which sub-dimension the items are in, and to determine the suitability of the dimensions of the scale for different cultures and groups in adaptation studies (Tabachnick and Fidell, 1996; Thompson, 2004). In order to determine the validity of the scope of the scale, EFA was made with the item pool consisting of 46 items remaining after the expert opinion. Items 6, 11, 18 and 22 on the scale were reversed because they indicated semantically inverse expressions. Then, the results of the KMO test and the Barlett test were reported to determine whether the substances showed factorization or not. Table 1 shows the KMO and Barlett test results.

Table 1. The Results of KMO and Barlett's Tests

Kaiser-Meyer-Olkin	0,923
Barlett's test	11234,74
df	1035
p	0,00

The Kaiser-Meyer-Olkin value should be above 0.50 (Field, 2000). A five-factor structure was obtained as a result of EFA. The eigenvalue, variance and cumulative variance values indicating the number of factorizations according to the items clustered in the study are given in Table 2.

Table 2. Eigenvalues, Variance and Cumulative Variance Values

Factors	Eigenvalues	(%) Variance	Total Variance Explanied
F 1	8,861	25,784	48,436
F 2	4,283	9,311	
F 3	2,544	5,536	
F 4	1,947	4,232	
F 5	1,044	3,573	

F1 = Scientist characteristics; F2=Method and reporting competence. F3=Interest in scientific research F4=Marrow of the scientific research method course. F5 = Scientific ethics

The eigenvalue results were examined to determine the number of factors in the measurement tool. The number of factors were decided by considering criterion 1 for eigenvalue. The number of factors of the eigenvalue greater than 1 give the factorization number of the items of the scale (Pedhazur & Pedhazur Schelkin, 1991). In this study, when the eigenvalues were examined, it was observed that the eigenvalues of 5 factors were greater than 1 criterion and ranged from 8.861 to 1.044. When the variance values are examined, the "scientist characteristics" dimension explains the variance explained by the factor 1 at most. The total explained variance was calculated as 48.436%. Between 40% and 60% is the ideal level for the described variance (Scherer et al., 1988). After EFA, the factor loads were examined to what size they were loaded by varimax rotation. Factor loads by factors and dimensions are presented in the table below.

Table 3. Results of Item Factors and Factor Loadings

		Factor Loadings (λ)				
	Items	F1	F2	F3	F4	F5
m32*	I enjoy doing scientific research.	0,823				
m31	Doing research has always intrigued me.	0,791				
m33*	I like to scan literature.	0,779				
m36	I follow scientific developments closely.	0,698				
m35	I like to help those who do scientific research.	0,645				
m29	I know how to analyze data in my research.	0,606				
m30	I know which pattern to use in the research.	0,552				
m28	I know the databases I will search the literature with	0,540				
m34	I care about scientific research.	0,538				
m27	I have no difficulty in forming a hypothesis for my study by scanning the literature in detail from the detailed literature.	0,510				
m37	I carefully examine the literature to produce original products and choose topics that will close the gap in the literature.	0,503				
m45	I report my study by paying attention to the literature review (introductory part), method, findings, discussion, conclusion and recommendations sections.		0,623			
m26	I try to access the data by getting permission for the data collection tools.		0,579			
m47	I cite my citations in the text in full in the bibliography.		0,573			
m2	I take care to cite in a scientific study.		0,567			
m23	I plan my research in accordance with scientific research processes.		0,546			
m48	In a scientific research, I pay attention to the title order (spelling and level).		0,535			
m46	I format tables and figures appropriately according to spelling guidelines, not as they are from analysis programs. I mention it in the research.		0,513			
m5	In the process of scientific research, I pay attention to the confidentiality of personal data.		0,504			
m21	I thoroughly review the relevant literature on the problem situation.		0,497			
m4*	When planning a scientific study, I try to encourage the voluntary participation of participants.		0,477			
m24	When I search the literature, I pay attention to the topics mentioned in the proposals of the previously written theses.		0,476			
m3	I am afraid of falling into repetition in a scientific research.		0,471			
m1	I know the importance of hiding the names of participants when planning a scientific study.		0,460			
m44	I know according to which guideline (APA, etc.) the research will be reported.		0,431			
m10	In scientific research, I consider the possibility of physically and psychologically harming the participants.		0,411			
m41	A scientist is a person who thinks universally, is objective, enlightened and has a high moral responsibility.			0,796		
m39	A scientist is a productive person who is productive on behalf of humanity.			0,796		
m43	A scientist is a person who is open to continuous improvement and self-improvement.			0,766		
m40	A scientist is an honest, knowledgeable, productive, outspoken person who respects other scientists.			0,751		
m42	A scientist is a person with strong foresight.			0,680		
m38	The scientist is the individual who transforms the feeling of curiosity into research.			0,663		
m8**	I can directly import the introductory chapters of similar studies into my own work.				0,769	
m9**	I can manipulate with the data to support the hypothesis I established in my study.				0,711	
m6**	I see no problem in importing the parts of other research that I have read and inspired before into my work.				0,703	
m11**	I think that research that is not concluded quickly is unnecessary research.				0,636	
m25**	I choose the subjects I like in the studies as my own study location.				0,563	
m7	I can share my research data with everyone.				0,463	

m16	I love the scientific research methods course.	0,733
m15	One of the courses that the course process passes quickly is the scientific research methods course.	0,658
m14	I strive not to be absent from the course of scientific research methods	0,641
m19	I look forward to the arrival of the scientific research course.	0,601
m17**	I look forward to the end of the scientific research methods course.	0,547
m18**	I get bored of extracurricular repetition and reading in scientific research class.	0,458

F1 = Interest in scientific research; F2=Competence in method and reporting; F3= Scientist characteristics; F4= Scientific ethics; F5= Interest in scientific research course

Note: * After CFA, these items were removed from the scale

Note: ** Marked items are reverse coded

When table above is examined, it is seen that the factor load values are between 0,411 (item 10) and 0,823. In the analysis process, item 20 and item 22 were removed from the scale because their factor loads were lower than 0,40. The naming was done for the measuring tool by examining the properties measured by the post-EFA dimensions. In the first dimension, the dimension measured by 11 items was called "Interest in scientific research", the dimension measured by 15 items was called "Competence in method and reporting", the dimension measured by 6 items was called "Scientist characteristics", the dimension measured by 6 items was called "Scientific ethics" and the dimension measured by 6 items was called "Interest in scientific research course". After EFA, 44 substances remained.

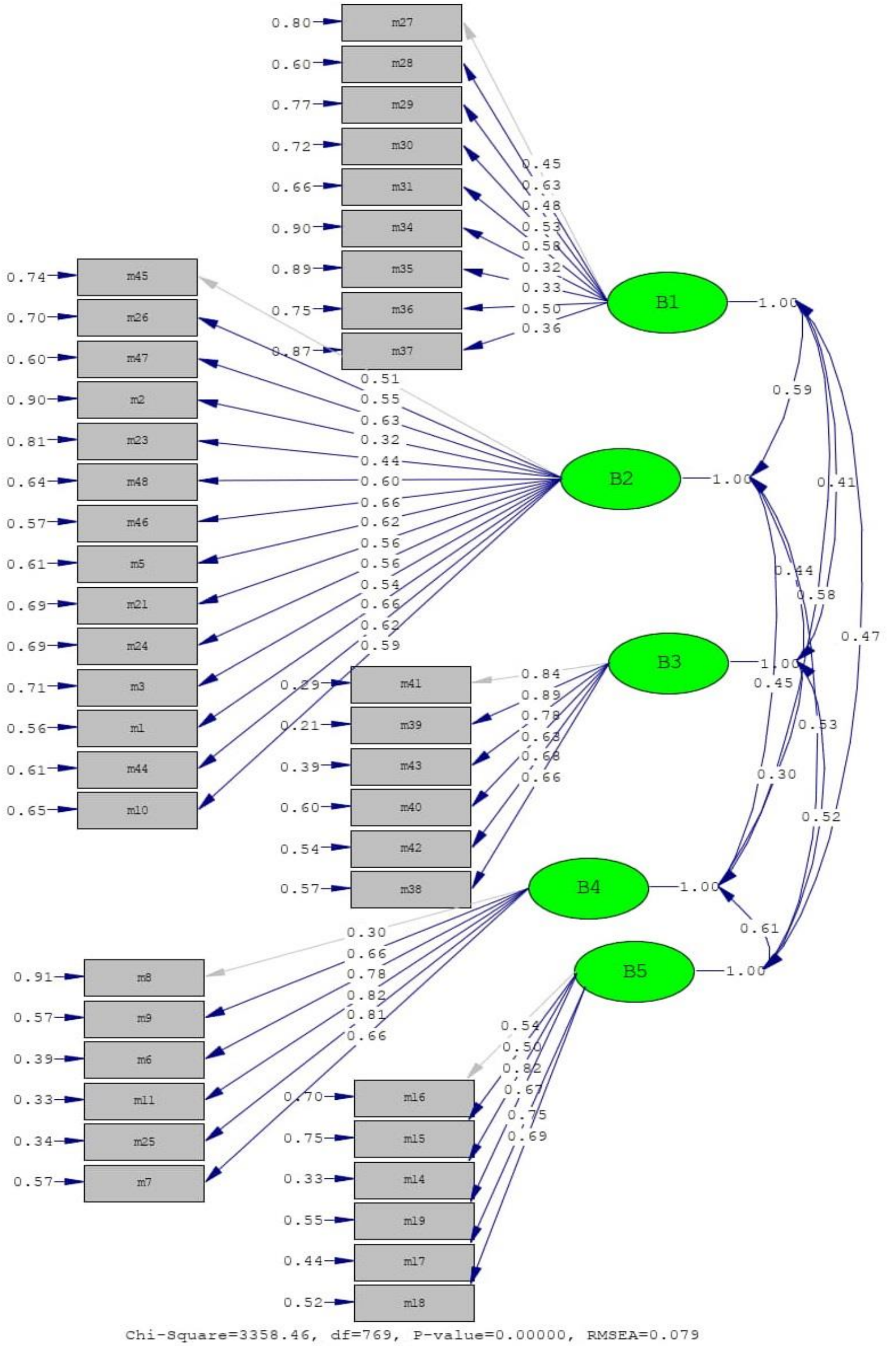
Convergent and Discriminant Validity

Convergent validity evaluates the extent to which items considered under the same factor are related to each other (Hair et al., 2013). The AVE value was examined to determine the convergent validity. To ensure convergent validity, the AVE value is expected to be equal to or above 0,50 (Fornell and Lacker, 1981). AVE values have been calculated for all individual dimensions. These values were AVE_1 = 0,42; AVE_2 = 0,27; AVE_3 = 0,55; AVE_4 = 0,42; AVE_5 = 0,38. For the third dimension, convergent divergence is within the desired range of values, while in other dimensions this is not achieved. For divergent validity, CR values are evaluated according to the criterion of being equal to 0,70 and greater (Fornell and Larcker, 1981). The calculated CR values are as follows: CR_1 = 0,88; CR_2 = 0,84; CR_3 = 0,88; CR_4 = 0,81; CR_5 = 0,78. Compared to the critical value, divergent validity is achieved.

The Results of Confirmatory Factor Analysis (CFA)

Confirmatory factor analysis (CFA) is the process of testing the hypothesis specifically for the relationship between measurement models, i.e. observed measurements or indicators (e.g., test items, test scores, behavioral observation ratings) and hidden variables or factors. A key feature of CFA is its hypothesis-driven nature. It differs from its counterpart, EFA, in that the researcher must determine in advance all aspects of the CFA model (Brown, 2006). In CFA, the current model is retested with research data and applied to test the suitability of the items for the target audience. In this study, CFA analysis was performed with 46 items and the obtained CFA path diagram is given as follows. In this figure the symbol of "B" means to factors.

Figure 2. *Path Diagram of the Scientific Literacy Scale*



The model fit indices obtained for the tested CFA analysis are given in Table 4.

Table 4. Result of Model Fit Indices

Model Fit Indexies	Values	Comment of Values
χ^2	3358,46	Bad
χ^2/df	4,37	Bad
RMSEA	0,079	Good
NNFI	0,92	Good
CFI	0,92	Good
GFI	0,87	Acceptable
AGFI	0,85	Acceptable
λ	0,30-0,89	Good
Error Covariance (ε)	0,21-0,91	Good

The scale, which had 44 remaining items after the EFA process on the scale was removed from the analysis because the item 32 and item 33 from the first dimension, and the item 4 factor load value from the second dimension were lower after the CFA. In the process of analysis 1, analysis was performed by R-ML method for level 5-factor structure. Chi-square statistics were examined as the first model data fit index. The chi-square value was calculated to be quite large. Since chi-square (3358,46) is a fit index that is affected by the sample size, it is greater than 3 even when divided by degrees of freedom (4,37). The fact that $\chi 2$ /df is greater than 3 is an evidence of a significant difference between the established model and the existing model. When we look at the chi-square fit index, it is seen that the model data harmony cannot be achieved. The RMSEA value provides a robust statistic among model data fit indices. Since the calculated RMSEA value (0,079) is 0,08 and below, the result given in the table shows that the model established for the sample in the study provides model data compliance. GFI, AGFI, CFI and NNFI values are required to be 0,90 and above (Büyüköztürk et al., 2010). Model data compliance for CFI and NNFI values (CFI=0,92; NNFI=0,92), but GFI and AGFI values are close to 0,90, so they are acceptable. According to standardized solutions, factor load values and error covariance values were reported. The factor load values of the scale range from 0,30 to 0,89, while the error covariance values range from 0,21 to 0,91. These values indicate that substances are good representatives of dimensions.

Investigation of Measurement Invariance

Item responses may vary according to a specific audience (such as gender, educational status, cultural characteristics) in a measurement tool. For this reason, it is expected to retest the structural and measurement models of the target group according to various demographic characteristics and to investigate their representativeness according to a group. The statistical process used in this process is measurement invariance. Measurement invariance is the independence of the group membership that includes demographic qualities in the scores expected to be obtained by individuals at a certain level in terms of the measured psychological structure (Önen, 2009). Considering the testing measurement invariance, there are types of measurement invariance in which the stage-by-stage relationship must be maintained. These types are "formal invariance", "metric invariance", "scale invariance", "strict invariance" (Knight and Hill, 1998; Vandenberg and Lance, 1998).

In the study, measurement invariance was examined according to gender variables. The first level 5-factor structure was tested by formal, metric, scale, and strict invariance analyses, respectively by gender variable. The results obtained are given below.

Table 5. *Examination of Measurement Invariance with Multigroup-CFA Results*

Measurement Invariance	χ^2	*df*	*RMSEA*	*p*	*GFI*	*CFI*	*SRMR*
Configural invariance	9786,45	1254	0.08	$p < 0.001$	0.90	0.92	0.098
Metric invariance	9895,63	1314	0.08	$p < 0.001$	0.90	0.92	0.098
Scalar invariance	9912,02	1310	0.08	$p < 0.001$	0.89	0.86	0.097
Metric - Configural	109,18	60	0.08	$p < 0.001$	0	0	0.000
Scaler - Configural	125,57	56	0.08	$p < 0.001$	-0.01	0.06	0.001
Scaler - Metric	16,39	4	0.08	$p < 0.001$	-0.01	0.06	0.001

Measurement invariance includes four stages in itself. Formal invariance results, which are the first of these invariance types, are presented in the table above. As a result of the analysis, χ^2/df value does not support model fit, while RMSEA, GFI, CFI and SRMR values show model fit. In line with this result, shape invariance was achieved, and metric invariance stage was initiated. For metric invariance, model fit indices such as shape invariance are interpreted that only χ^2/df f index does not provide model fit, but other fit indices prove that metric invariance is provided. For scalar invariance, deviations from model fit indices GFI and CFI values were observed. When the differences of model fit indices are compared with 0.01 in the types of invariance, it is seen that there is a difference. Scalar invariance was not achieved.

Testing Reliability

After testing the validity during the development phase of the measurement tool, reliability should be tested as well because the reliability of a measurement tool means that there are no errors in the answers given. Therefore, in this study, reliability analysis was performed using the relationship between the item, the total score and the Cronbach alpha coefficient for reliability. Reliability analysis results are given in the table below.

Table 6. *Result of Item-Total Score Correlations*

	Item-Total Score Correlation				
Items	Interest in scientific research	Competence in method and reporting	Scientist characteristics	Scientific ethics	Interest in scientific research course
m32	0,512				
m31	0,506				
m33	0,501				
m36	0,514				
m35	0,502				
m29	0,492				
m30	0,436				
m28	0,378				
m34	0,432				
m27	0,462				
m37	0,345				
m45		0,521			
m26		0,514			
m47		0,456			
m2		0,478			
m23		0,526			

m48		0,479			
m46		0,453			
m5		0,482			
m21		0,412			
m4		0,432			
m24		0,405			
m3		0,502			
m1		0,469			
m44		0,443			
m10		0,423			
m41			0,526		
m39			0,512		
m43			0,507		
m40			0,467		
m42			0,452		
m38			0,378		
m8				0,503	
m9				0,476	
m6				0,498	
m11				0,456	
m25				0,453	
m7				0,471	
m16					0,465
m15					0,437
m14					0,416
m19					0,452
m17					0,348
m18					0,378
Cronbach Alpha (α)	**0,900**	**0,864**	**0,883**	**0,713**	**0,805**
Total α					**0,905**

In the study, the consistency of the items among themselves was reported with the correlation results between the item and the total score. When the correlation coefficients between the item and the total score are examined, they vary between 0,521 and 0,348. After the items' total score correlation was examined, the reliability coefficient of the factors separately was calculated with the Cronbach alpha coefficient. Cronbach alpha values for each dimension and the whole scale were examined for the calculation of reliability. The alpha coefficients examined were above 0,70 (α_{F1}=0,900; α_{F2}=0,864; α_{F3}=0,883; α_{F4}=0,713; α_{F5}=0,805; α_{Scale}=0,905). In this case, it can be said that the measuring instrument measures with little error and the reliability is high. It is seen with the reliability coefficients above that the substances have a high level of reliability in themselves. A reliability coefficient that can be considered as sufficient in a Likert-type scale should be as close to 1 as possible (Tezbaşaran, 1997). According to these results, it can be said that the reliability of the measurement tool used in the research is at a high level.

CONCLUSION

In this study, it was aimed to develop a scientific literacy scale. For the use of a measurement tool, it must first be proved that it is valid and then reliable. To determine the validity and reliability of the scale to be developed in this research, a number of stages were followed. First of all, when the literature is reviewed, the absence of another measurement tool that covers the dimensions in the measurement tool is the reason for this research. In this study, an item pool was created by using the results of the research studies in which the concept of scientific literacy was studied with an emphasis on 21st century skills and by asking university students to write compositions.

A 5-point Likert-type scale was used for the answers in the developed scale. In the scale, the expressions as “1=strongly disagree”, “2=disagree”, “3=partially agree”, “4=agree” and “5=strongly agree” were used. After the expert opinion, the number of items were decreased to 46.

EFA was performed as the first stage of construct validity with 46 items. Items with factor loadings less than 0.40 were excluded from the scientific literacy scale. After EFA, 44 items remained. To prove validity, divergent and convergent validity values (AVE, CR) were examined. While a dimension fit AVE, the CR values were in the desired critical value band. CFA was performed to test the reconstruct validity of the design. Although the $\chi2$ /df value did not support the model fit, the model fit was achieved because the RMSEA value was calculated as 0.078, the NNFI value as 0.92, and the CFI value as 0,92. Since the GFI and AGFI values were very close to 0.90, it can be said that the model fit was achieved as a whole. In the CFA process, when item 32 and item 33 in the first dimension and item 4 in the second dimension were examined in terms of factor load and error covariance, it was deemed appropriately to remove them. In the scientific literacy scale, items 27, 28, 29, 30, 31, 34, 35, 36 and 37 were in the " Interest in scientific research " sub-dimension; items 45, 26, 47, 2, 23, 48, 46, 5, 21, 24, 3.1, 44 and 10 were in the “Competence in method and reporting” sub-dimension; items 38, 39, 40, 41, 42, 43 were in the " Scientist characteristics " sub-dimension; items 6, 7, 8, 9, 11 and 25 were in the sub-dimension " Scientific ethics"; items 14, 15, 16, 17, 18 and 19 were included in the " Interest in scientific research course” sub-dimension.

As a result, the validity of the structure and scope of the scientific literacy scale were tested. With this result, this measuring instrument consisted of substances representing the measured properties without mixing properties other than the intended structure. Reliability analysis was performed to check whether the measurement instrument whose validity was tested contained errors or not, and it was concluded that the reliability coefficient was high.

REFERENCES

Bakirci, H., Kahraman, F., & Artun, H. (2020). Ortak Bilgi Yapılandırma Modelinin Biyoçeşitlilik Konusunda Beşinci Sınıf Öğrencilerinin Bilimsel Süreç Becerilerine ve Eleştirel Düşünme Becerilerine Etkisi. *Fen Matematik Girişimcilik ve Teknoloji Eğitimi Dergisi*, *3*(1), 51-64.

Boateng, G.O., Neilands, T.B., Frongillo, E.A., Melgar-Quiñonez, H.R. & Young, S.L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Front Public Health*, *6*.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. The Guilford Press.

Büyüköztürk, Ş., Şekercioğlu, G., & Çokluk, Ö. (2010*). Sosyal bilimler için çok değişkenli istatistik: SPSS ve LISREL uygulamaları*. Pegem Akademi.

Bybee, R. W. (1997). *Achieving scientific literacy: From purposes to practices*. Portsmouth, NH: Heinemann.

Chung, H., Kim, J., Park, R., Bamer, A. M., Bocell, F.D. & Amtmann, D. (2016). *Testing the measurement invariance of the University of Washington Self-Efficacy Scale short form across four diagnostic subgroups*. Qual Life Res.

Erkuş, A. (2013*) Davranış bilimleri için araştırma süreci*. Ankara: Nobel.

Ertek, Y., Ertek, E., & Güneş, B. (2013). Bilimsel süreç becerileri ile fizik öğretim programında yer verilen problem çözme becerileri arasındaki ilişkinin incelenmesi. *Fen Bilimleri Öğretimi Dergisi*, *1*(2), 110-121.

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

Hair, J. F., Black, W. C., Babin, B. J., ve Anderson, R. E. (2013). *Multivariate data analysis: Pearson new international edition PDF eBook*. Pearson Higher Ed.

Hurd, P. D. (1998). Scientific Literacy: New Minds for a Changing World. *Science Education*, *82*, 407-416.

Knight, G. P., ve Hill, N. E. (1998). Measurement equivalence in research involving minority adolescents. In *Studying minority adolescents* (pp. 197-224). Psychology Press.

Laugksch, R. C. (2000). Scientific literacy: A conceptual overview. *Science Education*, *84*(1), 71–94.

Mertoğlu, H., & Öztuna, A. (2004). Bireylerin Teknoloji Kullanimi Problem Çözme Yetenekleri ile Iliskili Midir?. *TOJET: The Turkish Online Journal of Educational Technology*, *3*(1).

Miller, J. (1983). Scientific literacy: A conceptual and empirical review, *Daedalus*, *112*(2), 29–48.

Pella, M. O., O'Hearn, G. T., & Gale, C. G. (1966). Referents to scientific literacy. *Journal of Research in Science Teaching*, *4*, 199-208.

Pedhazur, E.J. & Pedhazur, Schmelkin, L. 81991). *Measurement, design, and analysis: An integrated approach.* Erlbaum Associates, Inc.

Önen, E. (2009). *Ölçme değişmezliğinin Yapısal Eşitlik Modelleme Teknikleri ile incelenmesi* (Yayımlanmamış Doktora Tezi). Ankara Üniversitesi/Eğitim Bilimleri Enstitüsü: Ankara.

Stevens, J. P. (1992). *Applied multivariate statistics fort the social sciences* (2nd ed.). Erlbaum.

Şahin, C. T., & Say, Ö. (2010). İlköğretim öğrencilerinin bilimsel okuryazarlik düzeylerinin incelenmesi. *Uluslararası Yönetim İktisat ve İşletme Dergisi, 6*(11), 223-240.

Şahin, F., & Ateş, S. (2018). Ortaokul öğrencilerine yönelik bilimsel okuryazarlık ölçeği adaptasyon çalışması. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi, 38*(3), 1173-1205.

Savaş, M. A. (2019). *Zekâ oyunları eğitiminin fen bilimleri öğretmen adaylarının eleştirel düşünme becerileri üzerine etkisi* (Master's thesis, Bartın Üniversitesi, Eğitim Bilimleri Enstitüsü).

Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). Harper Collins.

Tekin, N., Aslan, O., & Yağız, D. (2016). Fen bilimleri öğretmen adaylarının bilimsel okuryazarlık düzeyleri ve eleştirel düşünme eğilimlerinin incelenmesi. *Amasya Üniversitesi Eğitim Fakültesi Dergisi, 5*(1), 23-50.

Tezbaşaran, A. A. (1997). Validity issues of a Likert type scale (a case study). *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 13*, 41-45.

Tezel, Ö., & Tezgören, I. (2019). Sekizinci sınıf öğrencilerinin bilimsel okuryazarlık düzeyleri ile problem çözme becerileri arasındaki ilişkinin incelenmesi. *Eskişehir Osmangazi Üniversitesi Türk Dünyası Uygulama ve Araştırma Merkezi Eğitim Dergisi, 4*(2), 68-84.

Thompson, B. (2004). Exploratory and confirmatory factor analysis: Understanding concepts and applications. *Washington, DC, 10694*(000), 3.

Turgut, H. (2007). Herkes için bilimsel okuryazarlık. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi, 40*(2), 233-256.

Turgut, H. (2005). *Yapılandırmacı Tasarım Uygulamasının Fen Bilgisi Öğretmen Adaylarının Bilimsel Okuryazarlık Yeterliliklerinden"Bilimin Doğası" ve "Bilim-Teknoloji-Toplum ilişkisi" Boyutlarının Gelişimine Etkisi*. Doktora Tezi, Yıldız Teknik Üniversitesi, İstanbul.

Ulukan, G. (2019). *PISA 2015 verileri ve ülkelerin gelişmişlik düzeyleri ile öğrencilerin fen okuryazarlığının incelenmesi.* Yüksek Lisans Tezi. Aksaray Üniversitesi, Fen Bilimleri Enstitüsü, Aksaray.

Vandenberg R. J., Lance C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3(1), 4-69.