International Journal of Contemporary Educational Research

Volume 11, Number 3, September 2024, Page 363-375 Article Type: Research Article

Teacher Observation Skills Scale: Validity and Reliability

Abdullah Çetin[1] | | abdcetin46@gmail.com
Kahramanmaras Sütcü İmam University, Faculty of Education, Department of Educational Sciences, Kahramanmaraş, Türkiye

Abstract
The present research aims to develop a valid and reliable instrument to measure teachers' observation skills. The research was conducted during the 2023-2024 academic year with 651 teachers. Data were collected through online survey methods. Expert opinion was consulted for the content and face validity of the scale in the research. Exploratory (EFA) and confirmatory factor (CFA) analyses were performed to assess the scale's construct validity. Based on the results, it was found that the scale had a structure consisting of two dimensions and 18 items. It was determined that the factor loadings of the items in the scale were between .70 and .93. Considering the relevant literature, it was decided to name these dimensions as "natural observation skill" (NOS) and "scientific observation skill" (SOS). The Cronbach Alpha for the scales were both .95. The research developed a valid and reliable scale that can determine the level of teachers' observation skills.

Keywords: Observation, Observation skill, Teacher, Scale development

Citation
Çetin, A. (2024). Teacher observation skills scale: validity and reliability. *International Journal of Contemporary Educational Research*, *11* (3), 363-375. https://doi.org/10.52380/ijcer.2024.11.3.692

Received	31.05.2024
Accepted	01.09.2024
Publication	29.09.2024
Peer-Review	Double-anonymized-double-blind
Plagiarism Checks	Yes - iThenticate
Conflicts of Interest	The author(s) has no conflict of interest to declare.
Complaints	ijceroffice@gmail.com
Grant Support	The author(s) acknowledge that they received no external funding in support of this research.

[1] Corresponding Author

Introduction

Humans are natural observers of their surroundings to understand the world and shape their future actions (Merriam, 2009). In other words, observation is a part of daily life (Erden, 2011). Observation is often used to learn, acquire knowledge, collect data, evaluate, and communicate. Observation is a direct technique for obtaining information (Bahçeci, 2018). Observation has an important place in knowledge formation (Maral et al., 2012). Systematic observations during the formation of scientific knowledge are indispensable in testing and developing theories together with scientific inquiry processes (Yürümezoğlu & Öztaş Cin, 2019). Most research on people generally involves observation (Büyüköztürk et al., 2013). Observation is one of the most important data sources in qualitative research because it allows for first-hand access to data (Yıldırım & Şimşek, 2011). Observation is also a learning technique in lessons (Tok, 2023). The value of an activity or a product is also decided through observation (Karakütük et al., 1994; Poulson et al., 1996). How students communicate and interact with each other and with their teachers, and which positive and negative behaviors occur in the classroom, are also determined through observation (Erden, 2011).

The literature review reveals that observation is classified as routine/natural observations and systematic/scientific observations (Erden, 2011; Merriam, 2009; Yürümezoğlu & Öztaş Cin, 2019). Most observations in daily life are routine observations. These observations are made unconsciously and unsystematically to make sense of environmental events (Merriam, 2009), and their purpose is unclear (Erden, 2011). Natural observations are important in identifying the problem at the beginning of scientific research (Erden, 2011). Systematic/scientific observation refers to a structure covering scientific research processes that start with a scientific inquiry and continue with multiple inquiries. The data collection process in scientific observation has a methodological character (Yürümezoğlu, 2015). All of these classifications, as well as the use of observation in various fields, result in different definitions of observation.

Basic science process skills include observation, communication, classification, assessment, and prediction skills (Karamustafaoğlu, 2011; Ramig et al., 1995). The MoNE states that observation skills are among the most important skills to be gained by students, and all courses contain learning outcomes, including observation skills (MoNE, 2018). Both the 2005 and 2018 curricula emphasize the development of observation skills (Taze Karaçalı, 2021). Routine naturalistic observation skills help people connect with their surroundings, but they are not enough (Johnson, 2014). Individuals must improve their observation skills by learning about the methodological aspects of observation. This is possible by using sequential observation activities to systematize and scientificize children's natural spontaneous observation skills (Yürümezolu, 2015).

Observation is defined by Lederman (2007) as making sense of the world through the senses or extensions of the senses. Observation is a holistic process of receiving and processing data from the external environment, using not only the eyes but all sensory organs (Yürümezoğlu & Öztaş Cin, 2019). Observation is a method used to thoroughly describe the behavior that occurs in any environment or institution (Bailey, 1994; Yıldırım & Şimşek, 2011). Observation is defined as the process of gathering data for research by focusing on specific targets such as humans, society, or nature, either naked or with a tool (Büyüköztürk et al., 2013). Observation in education and training means the planned and purposeful examination of assets and events in their natural environment (Tok, 2014). Monitoring and examining the activity to be evaluated according to predetermined criteria is what observation in evaluation is all about (Karakütük et al., 1994). Although there are various definitions of observation in the Turkish Language Association's (TDK) dictionary, it has been observed that it includes all of the above definitions (TDK, 2023). Observation is at the basis of understanding nature and at the first step of scientific process skills (Yürümezoğlu & Öztaş Cin, 2019).

People's observation interpretations may also differ according to their observation skills. In other words, the interpretations of different people who observe the same thing may differ. The individual's knowledge, experiences, beliefs, and expectations influence his/her observations and interpretations (Chalmers, 2010). That is, the equipment used by individuals influences their observation skills and demonstrates that observation skills can be improved. Observation skill is a skill that can be developed through different activities (Cuthrell et al., 2016; Maral et al., 2012; Schwartz & Lederman, 2008). People can also learn to be careful and systematic observers by improving their observation skills (Merriam, 2009). Acquiring scientific knowledge and conducting activities on observation contribute significantly to the development of observation skills (Yürümezoğlu & Öztaş Cin, 2019).

Classroom observations have long been the undisputed foundation of teacher education programs in the United States. Pre-service teachers (PSTs) at various stages of development visit teachers' classrooms to complete observation hours each semester (Cuthrell et al., 2016). Although observation has been used in educational research for a long time, it was heavily used as a data collection tool in various dimensions of the classroom environment in the 1950s and 1960s (Yıldırım & Şimşek, 2011). Similar research is being conducted in Türkiye as part of the Teaching Practice course. However, it was found that pre-service teachers in Türkiye have deficiencies in terms of observation (Cansız & Cansız, 2018; Karslı et al., 2010). Young and Bender Slack (2011) emphasize that pre-service teachers and teachers need to acquire observation skills and that these skills should be continuously improved.

The observation skill is the most important basic skill that affects the development of other science process skills as well as being the first step of these skills (Abruscato, 2000; Ramig et al., 1995; Yürümezoğlu & Öztaş Cin, 2019). Observation is not an innate skill (Young & Bender Slack, 2011). Observation skill is a skill that can be developed through different activities (Cuthrell et al., 2016; Maral et al., 2012; Schwartz & Lederman, 2008). This skill is used in many activities such as learning, acquiring information, studying, problem-solving, data collection, evaluation, etc. Furthermore, teachers' observation skills are critical for the positive realization of teaching (Young & Bender Slack, 2011). Using observation skills in learning and teaching environments at school and in daily life is extremely important (Ergin et al., 2005). Individuals with high levels of these skills can solve problems in their daily lives quickly and effectively (Smith & Scharman, 1999).

Individuals who do not use scientific process skills are unlikely to be successful in business life (Rillero, 2010). Teachers are responsible for designing learning experiences that allow students to gain the most from observation. Therefore, activities that improve students' observation skills must be included (Anagün & Yaşar, 2009). However, it is not known to what extent teachers have observation skills. The literature review revealed no measurement tool for teacher observation skills in international and national studies. The first study to be conducted in the development of observation skills is to reveal the observation skill level of individuals. Therefore, a valid and reliable instrument to measure teacher observation skills was developed in the present study. This scale is expected to contribute to eliminating the deficiency in the literature. Moreover, this scale is expected to mediate the use of observation skills in the social sciences. The present study is important in terms of showing that observation is more than just a data collection tool used in science.

Method

This research was conducted to develop a measurement tool to determine teachers' observation skill levels. The Teacher Observation Skills Scale was developed in this context. In this section, the study group and the procedure of the study were explained in detail.

Study Group

The study group consisted of 651 volunteer teachers working in different provinces of Türkiye, mostly (50%) in Kahramanmaraş province in the 2023-2024 academic year. Of the participants in the study, 347 (53%) were male and 304 (47%) were female. Participants work in kindergarten (5%), elementary school (30%), middle school (29%), high school (29%), science and art centers (5%) and special education institutions (2%). When the participants are analyzed in terms of years of service, it is observed that there are 84 (13%) participants with 0-5 years of seniority, 96 (15%) with 6-10 years of seniority, 220 (34%) with 11-20 years of seniority, and 250 (38%) with more than 20 years of seniority. The participants were 27 different branch teachers, including classroom teachers, religious culture and ethics teachers, Turkish teachers, mathematics teachers, and English teachers. Participants were randomly divided into two groups systematically in the study. The data of the first group (339 participants) were used for EFA, and the data of the second group (312 participants) were used for CFA and analyzed.

Procedure

The researcher who decides to develop a scale must first clearly define "what is to be measured." The theoretical structure of the variable to be measured and related variables must be revealed by reviewing the literature. The format of the measurement tool must be determined in the second stage, and an item pool must be created. This is followed by the submission of the items to the experts, piloting the designed scale, evaluating the items, and finalizing the scale (Şahin & Boztunç Öztürk, 2018).

If a new scale is to be developed, the first step is to conduct a literature review on the subject. This requires paying attention to whether there are similar scales in the literature and which questions/topics should be addressed for the scale (De Vellis, 2003). In this study, no valid and reliable scale to assess teachers' observation skills was found in either national or international literature. Accordingly, it was decided to develop a scale to determine teachers' observation skills. The literature focuses on studies involving observation skills and related skills. The theoretical structure of observation skills was revealed.

Upon the decision to develop a scale, the format of the scale must be determined and an item/question pool must be created accordingly (De Vellis, 2003). According to the mathematical characteristics of the data, this format can be Thurstone type, Likert type, or Osgood dimensional separation scale (Tavşancıl, 2005). The scale form was prepared as a five-point Likert scale ranging from "(1) strongly disagree" to "(5) strongly agree" in this study. Another type of scale that is most commonly used in social studies and is based on the principle that participants provide information about themselves is the Likert-type scale. Participants state what the items in the Likert-type scale mean to them (Ekici et al., 2012). A pool of 55 items, which were designed to measure teachers' observation skills, was created by examining the literature thoroughly (Karamustafaoğlu, 2011; Ramig et al., 1995; Young & Bender Slack, 2011; Yürümezoğlu, 2015; Yürümezoğlu & Öztaş Cin, 2019). While designing the items, attention was paid to the fact that the items must be simple and comprehensible and that an item should not have more than one expression of judgment and thought (Ekici et al., 2012).

Expert opinions are obtained on how well the scale and each item in the scale serve the purpose (Karakoç & Dönmez, 2014). Experts can ask for corrections or deletions of items in the draft scale. However, the researcher should be responsible for accepting the suggestions (De Vellis, 2003). The items designed in this study were also reviewed by field experts in terms of scope, appropriateness, language, form, and comprehensibility. An expert evaluation form was then used to obtain the opinions of experts working in different university faculties of education (four in curriculum and instruction, two in assessment and evaluation, and one in educational administration) and researchers (two in science education) who have published on observation. Experts expressed their opinions on the form as "appropriate" or "must be corrected or removed." In line with expert opinions, eight of the items were removed, new items were added in their place, and some of the items were reorganized. The edited items were examined by a linguist, and needed arrangements were made. Content and face validity were thus ensured (Taşkın and Akat, 2010).

Revising the scale, developed based on expert opinions, by implementing it on a small number of participants who share similar characteristics to the target group. Piloting is important in terms of the readability of the items, revealing misunderstandings, and the completion time of the scale (Crocker & Algina, 1986). The pilot implementation of the scale prepared in this study was conducted with 35 teachers. The pilot implementation was analyzed and evaluated, and the instructions for the scale were prepared. It is adequate to select between 30 and 50 participants representing the target group for pilot implementation (Şeker & Gençdoğan, 2014). The 55-item draft scale, which was plotted, was filled in via Google Forms. The number of teachers who completed the form was 651. SPSS 22 package program for EFA and Mplus 5.1 package program for CFA were used for data analysis. The Cronbach's alpha internal consistency coefficient and split-half method were calculated to determine the reliability of the scale.

Ethical Approval

Ethical permission (Date: 26.08.2023-Number: 237539) was obtained from the Kahramanmaraş Sütçü İmam University Ethics Committee for this research.

Results and Discussion

EFA and CFA must be conducted in different samples (Orcan, 2018; Worthington & Whittaker, 2006). The participants were divided into two groups by using *ranbetween(1,2)* EXCEL formula. Based on the EXCEL grouping results, the data of the first group (339 participants) were used for EFA, and the data of the second group (312 participants) were used for CFA and analyzed. In factor analysis studies, it is recommended that all items on the scale have at least five times the number of participants and that the number of participants be at least 100 (Ferguson & Cox, 1993; Gorsuch, 1983).

Results Related to Explanatory Factor Analysis

SPSS was used for exploratory factor analysis (EFA). The principal axis factoring estimation method and promax rotation were used to run EFAs. In order to analyze the suitability of the EFA data obtained from the participants (339), the Kaiser-Meyer-Olkin (KMO) coefficient value was examined. Table 1 presents the data on KMO and Barlett's test.

Table 1. KMO and Bartlett's test.

KMO Measure of Sampling Adequacy.		.98
Bartlett's Test of Sphericity	Approx. Chi-Square	7017.38
	df	153
	Sig.	.00

The KMO value of .98 in Table 1 shows that the sample is perfect. The data obtained from Barlett's Sphericity Test (7017.38; $p < .001$) showed a multivariate normal distribution, indicating that the other assumptions for the EFA analysis were met. The fulfillment of these two assumptions shows that factor analysis can be performed. In order to determine the factor structure of the scale, Kaiser's criterion and the scree slope graph were examined. Data on the total variance explained according to the Kaiser criterion are presented in Table 2.

Table 2. Results for factor eigenvalues and explanation variances.

	Starting Eigenvalues			Total After Rotation		
Factor	Total	% Vary	Cum %	Total	% Vary	Cum %
1	12.36	68.68	68.68	12.12	67.33	67.33
2	1.47	8.17	76.85	1.23	6.83	74.16
3	.58	3.25	80.10			
4	.57	3.20	83.31			

According to the Kaiser criterion, factors with a loading value of 1 or higher are evaluated in factorization (Pallant, 2016). As a result of the EFA conducted by choosing principal axis factoring, two factors with eigenvalues above 1 were proposed. When the variance value of the first factor is analyzed, it is observed that it contributes 68.68%. The variance for the second factor was 8.17%. A Scree Plot graph was analyzed to decide the number of factors. Figure 2 depicts the Scree Plot graph.

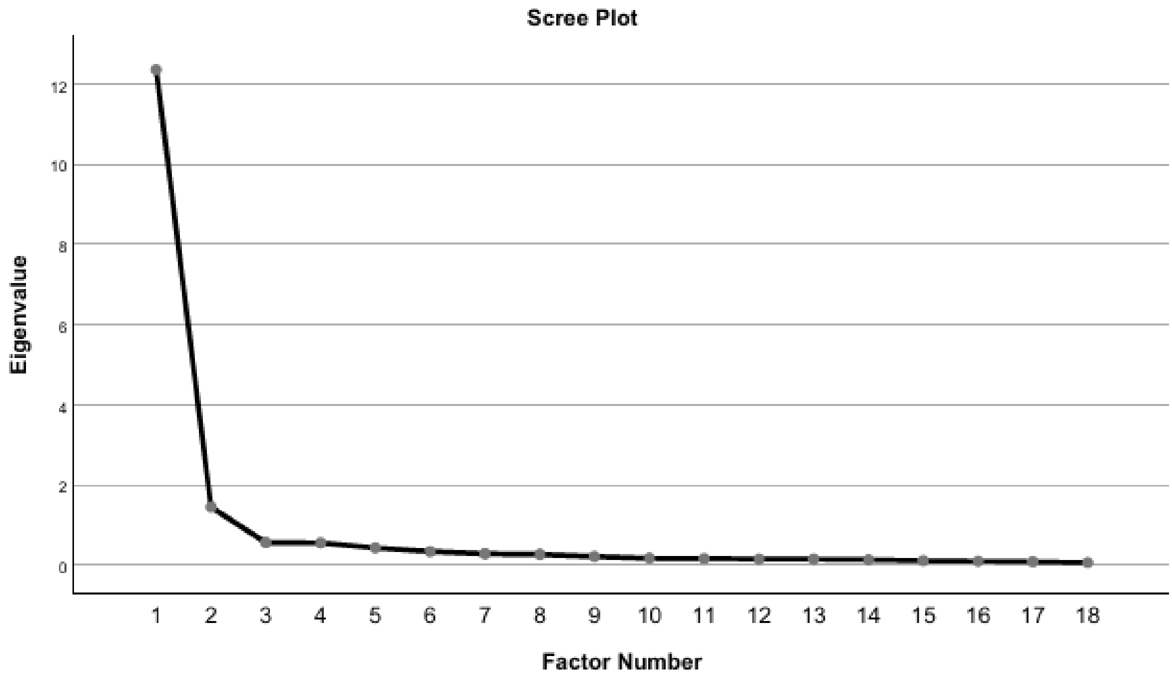

Figure 1. Scatter plot of teacher observation skill scale eigenvalues

When the graph in Figure 1 is examined, the slope plateaus at the third point, the contribution of the subsequent factors to the variance is quite small, and the values are numerically close to each other. When the total variance table and the slope graph are evaluated, the first two factors are accepted and the other factors are rejected.

Although there is no consensus among researchers in the literature, the acceptance points for factor loading values are .20 (Şencan, 2005). Based on the initial EFA results, some of the items (such as the items numbered 3, 19, and 40) were cross-loaded to multiple factors. After these items were removed, the EFA was conducted again. The factor loadings resulting from the final EFA are presented in Table 3.

Table 3. Results for factor loadings.

Item	Factor 1 (NOS)	Factor 2 (SOS)
S1	.83	
S2	.93	
S3	.89	
S4	.62	
S5	.87	
S6	.78	
S7	.80	
S8	.80	
S9	.69	
S10	.82	
S11	.70	
S12		.86
S13		.87
S14		.89
S15		.74
S16		.91
S17		.75
S18		.78

Table 3 shows that the Teacher Observation Skills Scale consists of two factors and 18 items. It is observed that the factor loadings of the scale items consisting of two dimensions are between .93 and .70. The EFA analysis revealed that the variance explained by the structure of the scale was 74.16%. Based on the EFA results, the correlation between extracted factors was .75.

Results related to Confirmatory Factor Structure

Mplus 5.1 program was used to run the CFA models. Confirmatory factor analysis was performed to confirm the construct validity of the 18-item two-factor scale formed by exploratory factor analysis. The first model tested showed some improvement gaps. Therefore, few error correlations were added to the model. Specifically, the correlations between item 2 and item 3, item 5 and item 11, and item 8 and item 9 were added. Figure 2 shows the modified model and standardized values of the model result.

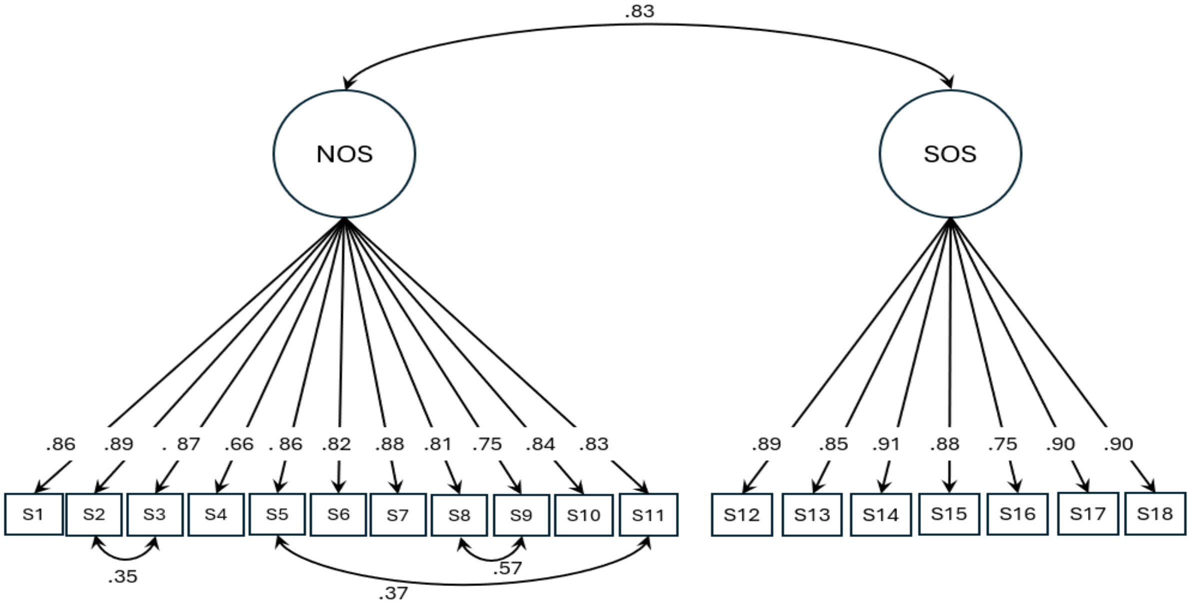

Figure 2. Standardized regression weights diagram

The boxes from S1 to S18 refer to the observed variables in the scale set, and NOS and SOS refer to the latent variables. The one-sided arrows in the figure indicate a one-way linear relationship. Standardized data for latent variables provide information about how well the latent variable is represented. The first latent factor, which was named Natural Observation Skill (NOS), was represented by 11 items. The values of standardized factor loadings ranged from .75 to .89. The second latent factor, which was named Scientific Observation Skill (SOS), was defined by 7 items. The values of standardized factor loadings for SOS ranged from .75 to .91. The standardized regression weight diagram shows that the observed variables that best represent NOS and SOS latent variables are S2 and S14, respectively. Results also show a correlation between NOS and SOS factors. The value of the correlations was .83. Table 4 provides information on the goodness of fit values.

Table 4. Teacher observation skills scale goodness of fit index.

Fit Indices	Acceptable Fit	Good Fit	Goodness of Fit Values	Conclusion
p*	$.05 < p \leq 1.00$	$.01 \leq p \leq .05$	.000	
x2/df	$2 \leq \chi 2/sd \leq 5$	$0 \leq \chi 2/sd < 2$	266.52/131=2.03	Acceptable
RMSEA	$.05 \leq RMSEA \leq .08$	$0 \leq RMSEA < .05$	.05	Acceptable
SRMR	$.05 \leq SRMR \leq .08$	$0 \leq SRMR < .05$	.03	Good Fit
CFI	$.95 \leq CFI < .97$	$.97 \leq CFI \leq 1.00$	.96	Acceptable

As it is not clearly defined which indices will be used in reporting the goodness of fit values, it is up to the researcher to decide which indices to use. When Table 4 is examined, the "p" value must not be significant. If this value is significant, other goodness of fit values are examined. While SRMR (.03) values showed good fit, x2/df (2.03), RMSEA (.05), and CFI (.96) values showed acceptable goodness of fit values. These results confirm the 18-item two-factor structure of the Teacher Observation Skills Scale.

Findings Related to Reliability

The reliability of the scale was evaluated with two different estimations. First, the Cronbach's alphas were estimated for both sub-dimensions. Later, the split-half reliabilities were calculated. The results were reported at table 5.

Table 5. The reliability results of the teacher observation skills scale.

	Number of Items	Cronbach's Alpha	Split-Half
Natural Observation Skill	11	.95	.94
Scientific Observation Skill	7	.95	.93

Table 5 shows that the Cronbach's alpha internal consistency coefficient of the scale is .94, which shows that it has an excellent reliability coefficient. Similarly, the split-half also shows high reliability values.

Conclusion

As a result of this study, a valid and reliable scale that can determine the level of teachers' observation skills was developed. While creating the TOSS, De Vellis' (2003) scale development stages and the practices in the literature were followed (İğde & Yakar, 2022; Şahin, & Boztunç Öztürk, 2018). A literature review was conducted, and the theories and concepts related to observation skills were examined. The format of the scale to be prepared was then decided, and an item pool was created. The draft scale items were submitted to expert opinions to ensure content and face validity. Adjustments were made to the item content, dimensions, and expressions in line with expert opinions. Following that, the pilot implementation of the devised scale was conducted, followed by an assessment of the item analysis, leading to the preparation of the ultimate version of the scale.

EFA and CFA were utilized to evaluate the construct validity of the TOSS. The EFA and CFA yielded a two-factor structure consisting of 18 items explaining 74.16% of the total variance. The fit indices of the single-factor structure of the TOSS were found to be sufficient. When 30% for the variance ratio explained in EFA and .30 lower limit for the factor loadings of the items in the scale are considered as criteria (Büyüköztürk, 2007; Costello & Osborne, 2005), as well as the fact that the fit indices calculated in CFA are within acceptable and good fit values, it can be said that the construct validity of the TOSS is ensured. Considering the relevant literature, it was decided to name these dimensions as "natural observation skill" (NOS) and "scientific observation skill" (SOS).

Reliability of the measurements obtained from the TOSS was evaluated using Cronbach's alpha method. Cronbach's alpha was found to be .95 for both dimensions in the scale. As a result of the research, a valid and reliable scale that can determine the level of teachers' observation skills was developed. Reliability coefficients of .70 and above are considered reliable in literature (Nunnaly & Bernstein, 1994; Tezbaşaran, 1997).

The items in the scale developed in this study are found to be compatible with the theories and concepts on observation skills as described in the literature. For instance, Aydoğdu's (2014) sentence "If you can notice a change around you on your way from home to school, it means that you are using your observation skills." is similar to the sentences in the scale developed. The characteristics related to observation skills in the studies in the literature (Arslan & Temiz, 2004; Harlen & Jelly, 1989) are as follows: "Determines the relevant details of the object and its surroundings." "Identifies significant differences/similarities between situations, events, and phenomena." "Groups events and objects by their common characteristics." "Determines the order of occurrence of events." is similar to the items in the scale developed in this study. The studies on observation skills in the literature also support the scale items obtained in this study (Abruscato, 2000; Aydoğdu, 2014; Chalmers, 2010; Cuthrell et al., 2016; Maral et al., 2012; Ramig et al., 1995; Young & Bender Slack, 2011; Yürümezoğlu, 2015; Yürümezoğlu & Öztaş Cin, 2019).

All people need to observe their surroundings to understand and explore the world around them (Ramig et al., 1995). Observation is one of the most important science process skills (Abruscato, 2000). Observation skills are the first step of scientific process skills and are the basis for other skills (Çepni & Çil, 2012). According to Akdeniz (2006), observation serves as the foundation for the acquisition of advanced skills such as prediction, communication, measurement, and classification. The essence of science is observation, and it is impossible to conduct scientific study without observation (Martin, 2003). Teachers should include activities that improve students' observation skills. Students' study skills are thus developed through observation (Anagün & Yaşar, 2009). Therefore, for teachers to develop students' observation skills, they first need to have observation skills themselves. A measurement tool is required to assess teachers' observation skills to determine whether they have this skill. The scale developed in this study serves this purpose.

Recommendations

The TOSS (Teachers' Observation Skills Scale) was created in this study to assess the levels of observation skills between teachers in kindergarten, primary, secondary, and high schools. Studies examining the relationship between teachers' observation skills and different skills (e.g., problem-solving, study, communication, categorization, etc.) can be conducted by using the TOSS.

The literature review revealed that no measurement tool could be used to measure teacher observation skills in international and national studies. This scale is expected to contribute to eliminating this deficiency in literature. The present research is also important in terms of showing that observation is more than just a data collection tool used in science. Observation is a measurable skill and is used in many areas such as learning, data collection, and evaluation in daily life. Therefore, it is required to determine the observation skills of individuals. The scale developed in the research can mediate the use and determination of observation skills in the social sciences. Researchers can use the developed scale to determine the level of observation skills in different studies. They can develop similar scales with reference to the scale.

Acknowledgements or Notes

The author received no financial support for the research and/or authorship of this article.

Author (s) Contribution Rate

Be sure to specify the Author (s) contribution rates (Author contribution rates are recommended to be stated in %)

Ethical Approval

Ethical permission (Date: 26.08.2023-Number: 237539) was obtained from the Kahramanmaraş Sütçü İmam University Ethics Committee for this research.

References

Abruscato, J. (2000). *Teaching children science: A discovery approach*, 5th Ed. Boston: Allyn and Bacon.

Akdeniz, A. R. (2006). Problem çözme, bilimsel süreç ve proje yönteminin fen eğitiminde kullanımı. S. Çepni (Ed.). *Kuramdan uygulamaya fen ve teknoloji öğretimi* (5. Baskı, s.107- 133) içinde. [The use of problem solving, scientific process and project method in science education. S. Çepni (Eds.). In Science and technology teaching from theory to practice (5th Edition, pp.107- 133). Ankara: Pegem Akademik Yayıncılık.

Anagün, Ş. S., & Yaşar, Ş. (2009). Developing scientific process skills at Science and Technology course in fifth grade students. *İlköğretim Online, 8*(3), 843 -865.

Arslan, A., & Tertemiz, N. (2004). İlköğretimde bilimsel süreç becerilerinin geliştirilmesi [Development of scientific process skills in primary education]. *Türk Eğitim Bilimleri Dergisi*, *2*(4), 479-492.

Aydoğdu, B. (2014). Bilimsel süreç becerileri. Ş. S. Anagün ve Duban, N. (Ed.), *Fen bilimleri öğretimi* (ss. 87-113) içinde. [Scientific process skills. In S. S. Anagün and Duban, N. (Eds.) Science teaching (pp. 87-113). Ankara: Anı Yayıncılık.

Bahçeci, F. (2018). Grupla öğretim teknikleri, Gömleksiz, N. (Ed.), *Öğretim ilke ve yöntemleri* (ss. 143-181) içinde. [Group teaching techniques, In Gömleksiz, N. (Eds.) Teaching principles and methods (pp. 143-181). Ankara: Asos Yayınları.

Bailey, K. D. (1994). *Methods of social research*. The Free Press, New York.

Büyüköztürk, Ş. (2007). *Sosyal bilimler için veri analizi el kitabı [Manual of data analysis for social sciences*] Ankara: Pegem Akademik Yayıncılık.

Büyüköztürk, Ş., Çakmak, E. K., Akgün, Ö. E., Karadeniz Ş., & Demirel F. (2013). *Bilimsel araştırma yöntemleri [Scientific research methods*]. Ankara: Pegem Akademik Yayıncılık.

Cansız, N., & Cansız, M. (2018). Enhancing preservice teachers' observation and inference skills. *Inonu University Journal of the Faculty of Education*, *19*(3), 362-373. https://doi.org/10.17679/inuefd.335762

Chalmers, A. F. (2010). *What is this thing called science?* Trans. Hüsamettin Arslan. (2nd Ed.). İstanbul: Paradigma.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, *10*(7), 1-9.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart and Winston, Inc.

Cuthrell, K., Steadman S. C., Stapleton, J., & Hodge E. (2016). Developing expertise: Using video to hone teacher candidates'classroom observation skills. *The New Educator 12*(1), 5–27 http://dx.doi.org/10.1080/1547688X.2015.1113349

Çepni, S., & Çil, E. (2012). Fen ve teknoloji programı ilköğretim 1. ve 2. kademe öğretmen el kitabı [Science and technology program primary school 1st and 2nd grade teacher handbook] Ankara: Pegem Akademik Yayıncılık.

De Vellis, R. F. (2003). *Scale development: Theory and applications* (2nd Ed.). Thousand Oaks, CA: Sage Publications.

Ekici, F., Ekici, E., Ekici, F. T., & Kara, İ. (2012). Öğretmenlere yönelik bilişim teknolojileri öz-yeterlik algısı ölçeğinin geçerlik ve güvenirlik çalışması [Validity and reliability study of ICT self-efficacy perception scale for teachers]. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi, 31*(31), 53-65.

Erden, M. (2011). *Eğitim bilimlerine giriş [Introduction to educational sciences*]. Ankara: Arkadaş Yayınevi.

Ergin, Ö., Şahin Pekmez, E., & Öngel Erdal, S. (2005). *Kuramdan uygulamaya deney yoluyla fen öğretimi [Teaching science through experimentation from theory to practice*]. İzmir: Dinazor Kitapevi.

Ferguson, E., & Cox, T. (1993). Exploratory factor analysis: A users' guide. *International Journal of Selection and Assessment, 1*, 84-94. http://dx.doi.org/10.1111/j.1468-2389.1993.tb00092.x

Gorsuch, R. (1983). *Factor analysis* (2nd Ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Harlen, W., & Jelly, S. (1989) *Devoloping science in the primary classroom*. London.

İğde, H., & Yakar, L. (2022). A study of reliability, validity and development of the teacher expectation scale. *International Journal of Assessment Tools in Education, 9*(3), 787-807. https://doi.org/10.21449/ijate.976585

Johnson, K. (2014). Creative connecting: Early childhood nature journaling sparks wonder and develops ecological literacy. *International Journal of Early Childhood Environmental Education, 2*(1), 126–39.

Karakoç F. Y., & Dönmez L. (2014). Basic principles of scale development. *Tıp Eğitim Dünyası, 13*(40), 39-49. https://doi.org/10.25282/ted.228738

Karakütük, K., Aksoy, H., & Akçay, R. C. (1994). Eğitim yönetiminde değerlendirme süreci [Evaluation process in education management]. *Ankara University Journal of Faculty of Educational Sciences (JFES), 27*(1), 361-383. https://doi.org/10.1501/egifak_0000000426

Karamustafaoğlu, S. (2011). Improving the science process skills ability of science student teachers using Idiagrams. *Eurasian Journal of Physics and Chemistry, 3*(1), 26-38. https://doi.org/10.51724/ijpce.v3i1.99

Karslı, F., Yaman, F., & Ayas, A. (2010). Prospective chemistry teachers' competency of evaluation of chemical experiments in terms of science process skills. *Procedia - Social and Behavioral Sciences, 2*(2), 778–781. https://doi.org/10.1016/j.sbspro.2010.03.101

Lederman, N. G. (2007). Nature of science: Past, present, and future. In S. K. Abell & N. G. Lederman (Eds.), Handbook of research on science education (pp. 831–879). Mahwah, NJ: Erlbaum.

Maral, S., Unver, A. O., & Yürümezoğlu, K. (2012). An activity-based study on providing basic knowledge and skills of measurement in teaching. *Educational Sciences: Theory and Practice 12*(1), 558–63.

Martin, D. J. (2003). *Elementary science methods: A constructivist aproach* (3rd Ed.). USA: Thomson Publishing Company.

Merriam, S. B. (2009). *Qualitative research a guide to design and implementation*. Jossey Bass.

Ministry of Education (2018). *Sosyal bilgiler dersi öğretim programı (İlkokul ve Ortaokul 4, 5, 6 ve 7. Sınıflar) [Social studies course curriculum (Primary and Secondary School 4th, 5th, 6th and 7th Grades)].* Ankara: Milli Eğitim Yayınları.

Nunnally, J., & Bernstein, I. (1994). *Psychometric theory*. New York: McGraw-Hill.

Orcan, F. (2018). Exploratory and confirmatory factor analysis: Which one to use first?. *Journal of Measurement and Evaluation in Education and Psychology, 9*(4), 414-421. https://doi.org/10.21031/epod.394323

Poulson, D. Ashby, M., & Richardson, S. (1996). *Userfit : A practical handbook on user-centred design for assistive technology*. TIDE EC-DG XIII, ECSC-ECEAEC, Brussels-Luxembourg.

Ramig, J. E., Bailer, J., & Ramsey, M. J. (1995). *Teaching science process skills*. Torrance, California: Good Apple.

Rillero, P. (2010). Process skills and content knowledge. *Science Activities, 35*(3), 3–4. https://doi.org/10.1080/00368129809600910

Schwartz, S. R., & Lederman, N. G. (2008). What scientists say: Scientists' views of nature of science and relation to science context. *International Journal of Science Education 30*(6). 727–771. https://doi.org/10.1080/09500690701225801

Smith, M. U., & Scharmann, L. C. (1999). Defining versus describing the nature of Science: A pragmatic analyses for classroom teachers and science educators. *Science Education, 83*(4), 493-509.

Şahin, M. G., & Boztunç Öztürk, N. (2018). Eğitim alanında ölçek geliştirme süreci: Bir içerik analizi çalışması [Scale development process in educational field: A content analysis research]. *Kastamonu Eğitim Dergisi, 26*(1), 191-199. https://doi.org/10.24106/kefdergi.375863

Şeker, H., & Gençdoğan, B. (2014). *Psikolojide ve eğitimde ölçme aracı geliştirme [Developing measurement tools in psychology and education*] (2nd Ed.). Ankara: Nobel Yayıncılık.

Şencan, H. (2005). *Sosyal ve davranışsal ölçümlerde güvenilirlik ve geçerlilik [Reliability and validity in social and behavioral measures*]. Ankara: Seçkin Yayıncılık.

Taşkın, Ç., & Akat, Ö. (2010). *Araştırma yöntemlerinde yapısal eşitlik modelleme [Structural equation modeling in research methods*]. Bursa: Ekin Yayınevi.

Tavşancıl, E. (2005). *Tutumların ölçülmesi ve SPSS ile veri analizi [Measurement of attitudes and data analysis with SPSS*]. Ankara: Nobel Yayınları.

Taze Karaçalı, H. (2021). Place of observation skill in 5. grade social studies textbooks. *Review of International Geographical Education (RIGEO), 11*(1), 100-117. https://doi.org/10.33403/rigeo.840610

Tezbaşaran, A. (1997). *Likert tipi ölçek hazırlama kılavuzu [Likert-type scale preparation guide*]. Ankara: Türk Psikologlar Derneği.

Tok, T. N. (2023). Etkili öğretim için yöntem ve teknikler. Doğanay, A. (Ed.) *Öğretim ilke ve yöntemleri* (ss. 176-252) içinde. [Methods and techniques for effective teaching. In Doğanay, A. (Ed.) Teaching principles and methods (pp. 176-252)] Ankara: Pegem Akademik Yayıncılık.

Türk Dil Kurumu (2023). *Türk Dil Kurumu sözlüğü [Turkish Language Association dictionary*]. Ankara: Türk Dil Kurumu Yayınları.

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist, 34*, 806-838. https://doi.org/10.1177/0011000006288127

Yıldırım, A., & Şimşek, H. (2011). *Qualitative research methods in the social sciences*. Ankara: Seçkin Yayıncılık.

Young, T., & Bender Slack, D. (2011). Scaffolding pre-service teachers' observations: Eye on the future. *Teaching Education, 22*(3), 325- 337. https://doi.org/10.1080/10476210.2011.593166

Yürümezoğlu, K., & Öztaş Cin, M. (2019) Developing children's observation skills using a fractal pattern from nature, *Science Activities*, *56*(2), 63-73. https://doi.org/10.1080/00368121.2019.1696734

Yürümezoğlu, K. (2015). Bilimin doğası ve gözlem, Yenice, N., Ed. *Bilimin doğası gelişimi ve öğretimi* içinde. [Nature of science and observation, In Yenice, N. (Ed). Nature of science development and teaching. Ankara: Anı Yayıncılık.

Appendix A: Turkish Version

Öğretmen Gözlem Becerileri Ölçeği

Madde No.	Boyutlar	Ölçek Maddeleri	Kesinlikle Katılmıyorum	Katılmıyorum	Kararsızım	Katılıyorum	Kesinlikle Katılıyorum
1	Doğal Gözlem Becerisi	Okula gidip gelirken çevreyi incelerim.					
2		Okulun görünümünde bir değişiklik olunca hemen fark ederim.					
3		Çevremde olan bitenler hemen dikkatimi çeker.					
4		Arkadaşlarımla buluşunca her zaman kılık-kıyafetlerine dikkat ederim.					
5		Okul/sınıf içerisinde düzensizliğe sebep olan durumu hemen fark ederim.					
6		Bulunduğum ortamdan ayrıldıktan sonra dahi önce bulunduğum ortamı net bir şekilde betimleyebilirim.					
7		Yolda yürürken olağan dışı bir şey olmuşsa onu hemen fark ederim.					
8		İdare-öğretmen, öğretmen-öğretmen arasında bir sorun varsa onu anında fark ederim.					
9		Okuldaki öğretmenin yaşadığı herhangi bir sorunu hemen fark ederim.					
10		Okul/sınıf kurallarına uymayan öğrencilerin kim/ler olduğunu hemen fark ederim.					
11		Okul dışında da sınıfımdaki öğrencilerimi tanırım.					
12	Bilimsel Gözlem Becerisi	Gözlem sürecini planlarım.					
13		Bir olayı/olguyu uzun süre gözlemleyebilirim.					
14		Gözlem yaparken hangi sırayla gözlem yapacağımı belirlerim.					
15		Gözlem sırasında uymam gereken etik kurallarına uyarım.					
16		Gözlem verilerimi ayrıntılı bir şekilde yazarım.					
17		Gözlemlerimde elde ettiğim tutarsız verileri ayıklayabilirim.					
18		Gözlemlerimin her aşamasını titizlikle yaparım.					