

Data Literacy Scale Development for High School Students

Gürbüz OCAK¹ Burak OLUR^{2*} Ayşegül KUTLU ÇAKIN³

1.Faculty of Education, Afyon Kocatepe University, PO box 03200, Afyonkarahisar, Turkey

2.Faculty of Science and Letters, Afyon Kocatepe University, 03200, Afyonkarahisar, Turkey

3.National Education Assembly, 03100, Afyonkarahisar, Turkey

Abstract

In this study, it was aimed to develop a scale to determine the data literacy level of high school students. The sample of the research consists of the students studying at high school in the province of Afyonkarahisar, Turkey in the 2020-2021 academic year. Exploratory sequential design, one of the mixed methods was employed in the study. The data were analyzed by Exploratory Factor Analysis (EFA) and it resulted in a structure with 21 items and four subdimensions named as data literacy awareness, self-awareness, data access, data literacy self-efficacy. Cronbach's alpha values for the subscales were found to be 0.910, 0.821, 0.722 and 0.648, respectively and 0.918 for the scale. It can be accepted that the data literacy scale for high school students, consisting of 21 items and 4 sub-dimensions, is a reliable scale.

Keywords: data literacy, high school students, scale development

DOI: 10.7176/RHSS/12-8-02

Publication date: April 30th 2022

1. Introduction

The development of humanity and societies has shown parallelism with the ability to produce and access information. People have developed various ways and methods to reach and disseminate it by aiming to use information as the ultimate goal. The most basic of these is spreading information with signs. The invention of writing is a turning point in this regard. Being able to read and understand the text was simply a state of being literate. Changes in the world have affected the definition and structure of information as well as the concept of literacy.

While the internet and technology were not as developed and widespread as they are today, people used to access information through the written media such as books, magazines, newspapers, articles. The sources of access to information and the information reached were also limited. However, as Polat and Odabaş (2008) stated, rapid technological changes in the world have also affected the information phenomenon. Changes in cultural, economic, political, technological and social areas have also changed the access to information and the quality of information. Technological developments such as smartphone, tablet, computer, and the increase in internet speed have provided easier access to the information that people are always after. Today, the sources of access to information are not limited like the period before the internet. Anyone with internet access anywhere in the world can access the information they need at that moment (Seker, 2005). Anyone can use the information s/he has reached at any point in the world and share it.

The fact that information resources and information are unlimited has also imposed some responsibilities on the modern world people. Some of these responsibilities are to understand that the information obtained is the information that the individual needs, to reach the information that s/he cannot access directly through the sources she has accessed, and to be able to use the information s/he has obtained for a specific purpose. All responsibilities have led to the emergence of new skills. The resulting skills revealed the concept of data literacy. Data literacy is a broader concept but superimposed on the general concept of literacy. Literacy is a concept related to the individual's perception, understanding and interpretation of the world s/he lives in and the objects, phenomena and events in this world, and to make sense of the relationships that make up her/his social life, or to make some sense out of something in a better way (AŞICI, 2009). Today, the concept of literacy is also discussed in a very broad way. Being literate does not only include reading and writing. According to Lankshear and Knobel (2006), "literacy" started to take on a different meaning from the 90's and gained its current meaning. It has come to a point where any information that is thought to be educationally valuable is considered literacy. Literacy is now used as a metaphor for "competence", "proficiency" and "functionality". Being computer literate or being technology literate has recently been used to indicate proficiency in those subjects. "Digital literacy" has the most important place among these concepts. Literacy has gained a broader meaning, such as understanding existing knowledge, out of the meaning of being literate.

Information literacy, or in other words data literacy, has gone far beyond the concept of literacy; As stated, competence has become a skill. It is necessary for today's people to be equipped with data literacy skills, to follow the developments in the world closely and to make sense of it.

Since the rapidly changing world, the limitlessness of knowledge and technological developments expect today's people to have data literacy skills, learners should also be equipped with this skill. The source of

information for students is not just the teacher as it used to be. Information is not just a phenomenon transmitted only by the teacher at school. Students can access the information they need from anywhere at any time. It is the duty of teachers to prepare suitable learning environments so that students can access the data they want in the way they want and work comfortably with the data they reach. Learning environments in 21st century should be arranged in a way that students can access data from different sources by using technology as required, and can use the data they reach in their daily lives and use them in problem solving (Kurbanoğlu & Akkoyunlu, 2001). If learning environments are organized in this way, it can contribute to students gaining data literacy skills.

Accessing information is easier than ever before. Especially with the development of technology, the information needed can be accessed instantly. Since it is easy to reach the information, but the sources of information are almost limitless, the problem arises whether the information is correct or not. Reaching the right information, making sense of the information, interpreting, analyzing and using it as a product are the skills expected in this century. What is expected from students is to know the data sources they will reach, to be able to use the data they have reached while solving problems and to transfer it to their daily life (Kurbanoğlu & Akkoyunlu, 2001). When the literature has been examined, it has been concluded that there are few scales on this subject and these scales focus on the competency of teacher candidate and high school and secondary school students (Adıgüzel, 2011; Hazar, 2018). Since a scale just measuring the data literacy level of high school students has not been found in the literature review, it is thought that the data literacy scale to be developed for just high school students in this study will contribute to the literature. Therefore, this study aims to develop a scale for determining the data literacy level of the high school students.

2. Method

In the scale development study, exploratory sequential design, one of the mixed method designs, was used. The first step in the exploratory sequential design is the process of collecting and analyzing qualitative data; then, it is followed by the process of collecting and analyzing quantitative data. Qualitative and quantitative data collected in the exploratory sequential design are not independent from each other, but are interrelated (Creswell & Plano Clark, 2017). During the scale development process, qualitative data were collected through interviews with people directly related to the subject of the scale. After the analysis of the data collected through the interview, the item pool was created. In the quantitative step of the scale development process, the general survey model was applied. The general survey model can be defined as a way of collecting information mostly through scales or interviews (Hutchinson, 2004) and aims to describe a situation as it is and to clearly reveal what the situation is (Karasar, 2020).

Different sampling methods were used in the scale development study. The data needed in the study and the way of data collection are shown in Table 1.

Table 1. Samples selected for the scale developed to measure the data literacy level of high school students

Data Collection Tools	Sampling Method	Sample
Interview form	Convenience Sampling	8
Pre-application for intelligibility of items	Convenience Sampling	21 High School Students (9 th , 10 th , 11 th , 12 th , Grades)
Application of the draft scale (Exploratory factor analysis)	Convenience Sampling	300 High School Students (9 th , 10 th , 11 th , 12 th , Grades)

When Table 1 is examined, written opinions by interview form were obtained from 5 11th grade high school students and 3 mathematics teachers working in high schools in Afyonkarahisar in the 2020-2021 academic year in order to create the categories related to the data literacy scale. Furthermore, the studies conducted in the relevant field were reviewed. As a result of these process, the keywords of data literacy have been determined. Based on these keywords and the that emerged, an item pool was created and the the items were sent to the expert opinion. After the expert opinion, the final version of the items was decided. For the intelligibility of the items, a preliminary application was made with 21 high school students studying at all grade levels in the 2020-2021 academic year in Afyonkarahisar. A pilot application was made after the necessary controls with the feedback received after the preliminary application. By using convenience sampling for the pilot application, 300 high school students studying at 9th, 10th, 11th and 12th grades in public schools in the 2020-2021 academic year in Afyonkarahisar were reached.

2.1. Participants

Since the reliability of the correlation coefficient of the sample is not reliable if it is obtained from a small sample, the sample size should be large enough to ensure the reliability of the correlation coefficient. A sample size of 50 was stated as "quite poor", 100 as "poor", 200 as "moderate", 300 as "good", 500 as "very good" and 1000 as "excellent" (Comrey & Lee, 1992, cited in Tabachnick & Fidell, 2006). The minimum number of participants should be more than 100 or 10 times the number of items in the scale (Floyd & Widaman, 1995 cited

in Hatcher & O'Rourke, 2013). Tekindal (2015) stated that a sample group of at least 300 people is required to conduct a scale development study. Based on these information on the literature, The data for the scale development study were collected from 300 high school students studying at high schools in Afyonkarahisar, Turkey during the 2020-2021 academic year by convenience sampling. The grade distribution of the sample is given in Table 2.

Table 2. The grade distribution of the sample

The grade level	N
9 th Grade	69
10 th Grade	98
11 th Grade	73
12 th Grade	60
Total	300

2.2. Scale Development Process

The scale was developed as a 5-point Likert scale in order to determine the data literacy level of high school students. The order which has been followed during the scale development process is as:

1. Reviewing the literature and creating interview questions
2. Conducting the interview and deciding on the key concepts according to the data
3. Creating an item pool and asking for expert opinion
4. Pre-application and necessary checks
5. Pilot application to high school students
6. Performing Exploratory Factor analysis
7. Calculating the Cronbach-Alpha reliability
8. Upper and lower group reliability analysis
9. Deciding on the final version of the scale items

2.2.1. Qualitative Data

The qualitative process which is employed during the scale development is given in Table 3.

Table 3. Qualitative stages and products used in the development of the data literacy scale for high school students

	Qualitative Data Collection	Qualitative Data Analysis	Developing the Measurement Tool
Process	<i>Convenience Sampling</i>	Coding	Creating sub-factors Writing the items and creating the item pool
	<i>Interview form for data literacy</i>	Creating categories	Checking item intelligibility check
Product	<i>Interview form</i>	Coded document Sub-factors	The form consisting 45 items for pilot application

At the first stage of the scale development study, a data literacy interview form was prepared in order to describe the data literacy levels of high school students. Expert opinion was sought while preparing the interview form. The open-ended questions in Table 5 were asked to the students to reveal their opinions. After the students' opinions were taken, they were subjected to content analysis and divided into codes and themes. Content analysis is the process of breaking down the raw data into key concepts or themes (Matthews & Ross, 2010). Coding of raw data is a method of reducing data through symbols or numbers, that is, collecting them under a title (Richards, 2015). The data of the study was coded by another researcher to ensure the consistency of the data. According to Miles and Huberman (1994), the checking of the codes does not only aim the certainty, but is also good for reliability. The agreement between intercoders was calculated according to the formula proposed by Miles and Huberman (1994) (Reliability = number of agreements/number of agreement + Disagreement X 100), and the reliability is 88%.

2.2.2. Quantitative Data

The quantitative process which is employed during the scale development is given in Table 4.

Table 4. Quantitative stages and products used in the development of the data literacy scale for high school students

	Quantitative Data Collection	Quantitative Data Analysis	Interpretation
Stages	<i>Pilot Application</i>	Exploratory factor analysis Item analyzes Reliability of the scale	Interpretation of the items under the factors Determining the level of validation of qualitative data
Products	<i>Scores</i>	Factor loads Proportional factor variance Item-total correlation Mean Standard deviation Cronbach -Alpha	Description of factors A valid and reliable measurement tool for high school students

After taking the opinions of high school students on data literacy, the scale items were started to be written by consulting the expert opinion. An item pool was created for the scale items and the scale was developed in a 5-point Likert type. Expert opinion was asked in order to find out whether the items in the item pool were intelligible and whether the scale served its purpose. In this direction, items that did not serve the purpose and were not understood were removed through the feedback from the expert; Items with poor intelligibility were corrected. Expert opinion was asked again for the remaining 45 items and applied to 21 high school students for pre-application. The scale items, which were rearranged according to the feedback from 21 high school students and experts, were made ready for the pilot application.

The scale, which was developed to determine the data literacy levels of high school students, was applied to 300 high school students (9th grade, 10th grade, 11th grade, 12th grade) from high schools in Afyonkarahisar, Turkey in 2020-2021 academic year. After applying the scale, the results of Kaiser-Mayer-Olkin (KMO) and Barlett's Test of Sphericity (BTS) tests were examined to decide the appropriateness of performing exploratory factor analysis (EFA) on the data. While the KMO coefficient indicates the suitability of the data for factor analysis, the Barlett (sphericity) test allows to examine whether there is a relationship between the variables as partial correlations (Büyüköztürk, 2020). After KMO and BTS results, exploratory factor analysis (EFA) was applied to the data.

2.2.3. Creating Item Pool

Before starting the scale development study, the literature was reviewed and the studies on data literacy were examined. There are not many studies on data literacy in the literature but information literacy scales. It has been concluded that there are few scales on this subject; one by Adıgüzel (2011) named as "Information literacy scale" and the other by Hazar (2018) as "Information, Media and Technology Skills Sufficiency Scale". Information literacy scale was used to measure pre-service teachers' level of accessing and structuring information; The scale developed by Hazar (2018) was developed for secondary school students and high school students. As a result of the research, data literacy can be defined as access to information, the ways followed in accessing information, the levels of making sense of information, evaluating, interpreting, reaching conclusions and using information. According to Koltay (2015), a data literate individual is one who can access data, interpret and evaluate data, and look at data critically. The data literate individual is also aware of the context in which data is created and used. Furthermore, the data literate individual thinks that s/he is responsible while using the data. Ethics and data use are part of how data is used and how privacy is protected (Mandinach, Parton, Gummer, & Anderson, 2015). After the literature review, key concepts such as ways of accessing key data, using-sharing data, interpreting data and ethical principles in accessing data were created. After the key concepts were created, 5 high school students and 3 high school mathematics teachers were asked 16 questions and their written opinions were obtained. Since there are few studies on data literacy in the literature, it was tried to provide data for key concepts by keeping the interview questions more. Since data literacy develops critical thinking in science and mathematics lessons and provides meaningful integration, the opinions of high school mathematics teachers were also asked in the study.

The open-ended questions which are used for creating the item pool is given in Table 5.

Table 5. Open-ended interview questions created with the key concepts determined after the literature review

Open-ended Questions on the Interview form
1-How do you get information about any subject you are researching? (Via internet - Via print media) Why?
2-How can you understand the content of the subject from the visuals you reach about the subject you are researching?
3- How can you understand the content of the subject from the graphics you reach about the subject you are researching?
4- How can you understand the content of the subject from the tables you reach about the subject you are researching?
5- How do you understand the content of written texts such as texts and articles that you have reached about the subject you are researching?
6-How do you create a product with the information you get? (Project, article, graphic, table, etc.)
7-Do you need to share the information you have obtained with others? Why?
8-Do you describe yourself as a data literate? In what ways would you define yourself as data literate?
9-What kind of behavior is expected from data literate people?
10-Do you pay attention to ethical principles and legal situations in accessing information? What do you do?
11-Can you distinguish the altered and distorted content from the correct content? How?
12-What do you do to follow the changes in the technology, method or resources used to access the data?
13-How do you combine (synthesize) information about a topic from different sources in a meaningful way?
14-What do you do to question the reliability of the information obtained on a subject?
15-How do you systematically bring together the information you have obtained from different sources to use in solving the problems you encounter?
16-How do you associate the new knowledge with the old knowledge?

After taking the opinions of 5 high school students and 3 math teachers working in high school about the questions, key concepts were created and item writing was started. After receiving expert opinion, 45 items remained in the item pool with the necessary editing. Sub-dimensions were determined by considering key concepts. The factors of the scale were named as data literacy awareness, self-awareness, access to data and literacy self-efficacy, respectively.

3. Findings

Findings related to the analysis of the research are given under this section.

3.1. Exploratory Factor Analysis

The purpose of factor analysis is to create a small number of factors by bringing together the interrelated variables and to complete the measurement process with the few factors created (Büyüköztürk, 2020; Ural & Kılıç, 2006). In this study, exploratory factor analysis was performed in order to reveal the structural features of the scale.

Before performing the exploratory factor analysis, the KMO (Kaiser-Meyer-Olkin) test was performed to check whether the sample size was suitable for EFA. The KMO value was calculated as 0.950 and the KMO value being 0.7 greater is accepted for sampling convenience (Şencan, 2005). In addition to the KMO value, the Barlett test should also be significant to decide whether the data are suitable for factor analysis (Field, 2005). It was significant as a result of Bartlett's test ($p=0.000$, $p<0.050$). The KMO value was recalculated after the items removed after the factor analysis. In Table.6, the results of the first and last analysis of the KMO and Barlett tests are given.

Table 6. First and final KMO and Barlett test results

First Test Results			Last Test Results		
KMO Value		0.950	KMO Value		0.927
Barlett Test	Approx. Chi.Square	8079,539	Barlett Test	Approx. Chi.Square	2800,471
	Df	.990		Df	210
	Sig.	.000		Sig.	.000

Maximum likelihood and vertical rotation were used during the scale development phase. In addition to being the most used method in exploratory factor analysis, the maximum probability factor is also the most preferred method for performing confirmatory factor analysis to the scale (Cokluk, Şekercioğlu, & Büyüköztürk, 2014). In scale development, axis rotation was performed in order to find the items with high correlation with the factors and to interpret them more easily. There are two methods of axis rotation. In the factor analysis, the varimax method, which is frequently used in scale development in social sciences, was used. This method aims to collect the load values close to 1.0 of the items of the scale under a factor (Büyüköztürk, 2020). As a result of the analysis, items that loaded on more than one factor and had low factor loads were determined. Items with less

than 0.10 difference between the highest and lowest item loads and items with factor load values less than 0.40 were determined. .40 and above is a good criterion for scale item selection (Nunnally & Bernstein, 1994). Items that did not meet these conditions were removed from the scale by performing factor analysis one by one and each time. However, the load of the 3rd item under the 4th factor after rotation was found to be .390. In line with the expert opinion, it was decided that the item was suitable for the relevant factor and should remain in the scale. After removing 24 items step by step from the scale, it was seen that there were 4 sub-dimensions in the factor analysis. After the Varimax rotation, the first sub-dimension of the data literacy scale for high school students was “Data literacy awareness” (23,420%), the second sub-dimension “Individual awareness” (10,355%), the third sub-dimension “Access to data” (8,928%) and the fourth sub-dimension “Data literacy self-efficacy” (7,650%). The ratio of explaining the total variance of the four factors is 50,354%. It is considered appropriate that the total variance in the scale is between 40% and 60% (Tavşancıl, 2018).

The eigen values and explained variance percentages of the factors are given in Table 7. below.

Table 7. The Eigen values and the explained total variance of data literacy scale for high school student

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	8,234	39,210	39,210	7,731	36,813	36,813	4,918	23,420	23,420
2	1,885	8,978	48,188	1,441	6,861	43,675	2,175	10,355	33,775
3	1,220	5,811	53,999	,795	3,788	47,462	1,875	8,928	42,703
4	1,125	5,359	59,358	,607	2,891	50,354	1,607	7,650	50,354

The factor loadings of the items are given below in Table 8.

Table 8. Rotated component matrix table of the data literacy scale for high school student

Items	Components			
	1	2	3	4
A1	,797			
A2	,694			
A3	,677			
A4	,637			
A5	,627			
A6	,616			
A7	,607			
A8	,556			
A9	,525			
A10	,520			
A11	,500			
A12	,494			
B1		,794		
B2		,692		
B3		,629		
C1			,769	
C2			,589	
C3			,538	
D1				,599
D2				,587
D3				,390

Varimax rotation method was used in order to ensure that the factors are independent and to bring together the items that are highly related to each other (Gürbüz & Şahin, 2014). After the rotation, there are 12 items in the 1st factor, 3 items in the 2nd factor, 3 items in the 3rd factor, and 3 items in the 4th factor. While an item load value of 0.40 and above is a good criterion for item selection (Nunnally & Bernstein, 1994); Büyüköztürk (2020) stated that load values between 0.30 and 0.60 are moderate and 0.60 and above are high. The load factors of 21 items in the scale are between 0.390-0.797.

The comparison of the answers of the students in the upper and lower groups, who were included in the research during the scale development process, with the independent t test is given in Table.9. It was seen that the difference between the lower and upper groups of all the items decided after the factor analysis was significant.

Item analysis of data literacy scale for high school students is given in Table 9.

Table 9. Item analysis of the scale

Factors	Items	Varimax Factor Load	Common Factor Variance	Item Total Correlation Coefficient	t	p
Data literacy awareness	A1	0,797	0,688	0,763	15,318	0
	A2	0,694	0,569	0,688	11,476	0
	A3	0,677	0,558	0,695	12,293	0
	A4	0,637	0,465	0,604	9,017	0
	A5	0,627	0,584	0,717	13,879	0
	A6	0,616	0,504	0,669	12,593	0
	A7	0,607	0,486	0,652	14,368	0
	A8	0,556	0,44	0,61	12,522	0
	A9	0,525	0,454	0,62	10,058	0
	A10	0,52	0,428	0,619	11,423	0
	A11	0,5	0,374	0,541	11,178	0
	A12	0,494	0,374	0,572	10,986	0
Individual awareness	B1	0,794	0,7	0,703	10,499	0
	B2	0,692	0,641	0,68	12,088	0
	B3	629	0,524	0,642	10,937	0
Access to data	C1	0,769	0,617	0,567	6,139	0
	C2	0,589	0,47	0,558	9,896	0
	C3	0,538	0,464	0,507	11,058	0
Data literacy self-efficacy	D1	0,599	0,437	0,47	7,916	0
	D2	0,587	0,473	0,524	7,344	0
	D3	0,39	0,326	0,39	7,635	0

Internal consistency coefficients of each factor and total were calculated. The scores obtained from the data literacy scale for high school students were sorted from highest to lowest and divided into two groups as the upper 27% (n=81) and the lower 27% (n=81). As a result of the independent sample t-test, it is seen that there is a significant difference in all items between the upper 27% and lower 27% scores of the items in Table 9 ($p < .05$). There are variance and Cronbach's alpha coefficients for each factor in Table 10.

Table 10. The number of items, variance rates and internal consistency coefficients of the data literacy scale and its factors

Factors	Items	Exp. Variance	α
1 st Factor	12	23,42%	0,91
2 nd Factor	3	10,36%	0,821
3 rd Factor	3	8,93%	0,722
4 th Factor	3	7,65%	0,648
Total	21	50,35%	0,918

As seen in Table 10, the variance rate of the 1st factor is 23,420%, the variance rate of the 2nd factor is 10,355%, the variance rate of the 3rd factor is 8,928%, and the variance rate of the 4th factor is 7,650%. The total variance rate of the scale is 50.354%. The alpha coefficients of the factors of the scale were calculated as 0.910, 0.821, 0.722 and 0.648, respectively. A reliability coefficient between 0.00 and 0.40 means that the scale is not reliable, between 0.30 and 0.60 means that the scale has low reliability, between 0.60 and 0.80 it is highly reliable, and between 0.80 and 1.00 it is "highly reliable" (Akgül & Çevik, 2003). While it is considered

sufficient that the reliability is around 0.70; Gürbüz and Şahin (2016) and Kılıç (2016) state that it can be accepted up to 0.60 in their scale studies. Therefore, the reliability coefficient of the 4th factor of the scale can be accepted as 0.648. Since the total alpha coefficient of the scale is 0.918, it can be said that the data literacy scale for high school students is highly reliable.

The mean and the correlation coefficients of the factors and the total scale is given in Table 11.

Table 11. Mean and correlation coefficients of the factors of the data literacy scale for high school students

Factors	N	\bar{x}	1 st Factor	2 nd Factor	3 rd Factor	4 th Factor
1 st Factor	300	47,52	-	-	-	-
2 nd Factor	300	9,47	0,525	-	-	-
3 rd Factor	300	10,2	0,46	0,472	-	-
4 th Factor	300	11,83	0,569	0,426	0,335	-

According to Table 11, the correlation coefficient of the scale for high school students indicates that there is a significant and moderate correlation between the factors.

4. Discussion, Conclusion and Recommendations

The data literacy scale developed for high school students is a valid and reliable scale that can be used to determine the data literacy levels of high school students. Since the scale that determines the data literacy levels of high school students was not found in the literature review, it can be said that the developed scale will fill the gap in this sense. The definition of duties and responsibilities expected from 21st century students has changed considerably. What is expected from students is to reach the data sources they need, to be able to make sense of and use the data they have obtained and be data literate. In today's world, accessing data, knowing how many ways to access data, interpreting the data they reach, turning the data into a product are the skills expected from students and individuals of the 21st century. It is important to determine data literacy levels to contribute to students' being data literate or to raise awareness of data literacy. The data literacy scale for high school students will be able to meet the need in this field and determine the level of the students. The total score obtained from the scale items can give reliable results regarding the data literacy levels of high school students. "Data literacy scale for high school students" is a scale that can make reliable decisions about high school students' data literacy levels.

In this scale development study, the "Data Literacy Scale for High School Students", which was developed to measure the data literacy levels of high school students, is in a 5-point likert format. There were 45 items in the scale before the analysis. Before the analysis, the scale was applied to 300 high school students. Exploratory factor analysis was performed and 24 items were removed from the scale one by one. The final version of the scale consists of 21 items and 4 dimensions. As a result of the analyzes made on the scale items, the item-total correlation results, the independent groups t-test based on the difference between the lower and upper group averages, and the results of the exploratory factor analysis indicate that there are appropriate factors and items in the scale that will determine the data literacy level. The calculated reliability coefficient of the scale, which consists of 21 items and 4 sub-dimensions, is 0.918. While the lowest score that can be taken from the scale is 21, the highest score that can be obtained is 105. The 1st sub-dimension of the scale was determined as "Data literacy awareness", the 2nd sub-dimension "Individual awareness", the 3rd sub-dimension "Data access paths" and the 4th sub-dimension "Data literacy self-efficacy".

References

- Adigüzel, A. (2011). Bilgi okuryazarlığı ölçeğinin geliştirilmesi. *Dicle Üniversitesi Ziya Gökalp Eğitim Fakültesi Dergisi*, 17, 15-28.
- Akgül, A. & Çevik, O. (2003). *İstatistiksel analiz teknikleri:SPSS'te işletme yönetimi uygulamaları*.Ankara:Emek Ofset.
- Aşıcı, M. (2009). Kişisel ve sosyal bir değer olarak okuryazarlık. *Değerler Eğitimi Dergisi*, 7(17), 9-26.
- Büyüköztürk, Ş. (2020). *Sosyal bilimler için veri analizi el kitabı* (28.Baskı). Ankara: Pegem Yayıncılık.
- Çokluk, Ö., Şekercioğlu, G. & Büyüköztürk, Ş. (2014). *Sosyal bilimler için çok değişkenli istatistik spss ve lisrel uygulamaları*. (3. Baskı). Ankara:Pegem Akademi.
- Creswell, J.W. & Plano Clark, V. L. (2017). *Designing and conducting mixed method research*. (3th Ed.). Los Angeles: Sage.
- Field, A. (2005). *Discovering statistics using spss*. London:Sage.
<https://www.futurelearn.com/>.
- Gould, R. (2017). Data literacy is statistical literacy. *Statistics Education Research Journal*, 16(1), 22-25.
- Gürbüz, S. & Şahin, F. (2016). *Sosyal bilimlerde araştırma yöntemleri*. Ankara:Seçkin Yayıncılık.
- Hatcher, L. & O'Rourke, N. (2013). *A step-by-step approach to using SAS for factor analysis and structural*

- equation modeling*. (2nd Ed.) Cary NC: Sas Institute.
- Hazar, E. (2018). Bilgi, Medya ve Teknoloji Becerileri Yeterlilik Ölçeği: Geçerlilik ve güvenilirlik çalışması. *Journal of Human Sciences*. 15(2), 1306-1316.
<https://hbr.org/2020/02/boost-your-teams-data-literacy>.
- Hutchinson, S.R. (2004). Survey Research. In Kathleen deMarrais and Stephen D. Lapan, (Ed.), *Foundations for Research Methods of Inquiry in Education and the Social Sciences* (283-301). Mahwah:Lawrence Erlbaum Associates, Publishers.
- Karasar, N.(2020).*Bilimsel araştırma yöntemi: Kavramlar ilkeler teknikler*.(35.Basım).Ankara:Nobel.
- Koltay, T. (2015). Data literacy: in search of a name and identity. *Journal of Documentation*, 71(2),401-415.
- Kılıç, S. (2016). Cronbach'ın alfa güvenilirlik katsayısı. *Journal of Mood Disorders*,6(1), 47-48.
- Kurbanoğlu, S. & Akkoyunlu, B. (2001). Öğrencilere bilgi okuryazarlığı becerilerinin kazandırılması üzerine bir çalışma. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 21(21), 81-88.
- Lankshear, C. & Knobel, M. (2006). *New literacies:Everyday practices& classroom learning*.(2nd Ed.). Berkshire: Open University Press.
- Mandinach, E. B., Parton, B. M., Gummer, E. S., & Anderson, R. (2015). Ethical and appropriate data use requires data literacy. *Phi Delta Kappan*, 96(5), 25-28.
- Matthews, B. & Ross, L. (2010).*Research methods. A practical guide for the social sciences*.Essex:Pearson.
- Miles, M.B. & Huberman, M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd Ed.) Thousand Oaks, CA:Sage
- Nunnally, J.C. & Bernstein.I.H.(1994).*Psychometric theory*.New York:McGraw Hill.INC.
- Öğüt, A. (2003). Bilgi çağında yönetim. (2. Baskı) Ankara: Nobel Yayıncılık.
- Polat, C. & Odabaş, H. (2008, March) *Bilgi toplumunda yaşam boyu öğrenmenin anahtarı: Bilgi okuryazarlığı*. Küreselleşme, Demokratikleşme ve Türkiye Uluslararası Sempozyumu, Antalya.
- Richards, L. (2015). *Handling qualitative data:A practical guide*. (3rd Ed.)London:Sage.
<https://schoolofdata.org/2016/01/08/>.
- Şeker, T. B. (2005). Bilgi teknolojilerindeki gelişmeler çerçevesinde bilgiye erişimin yeni boyutları. *Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 13, 377-391.
- Şencan,H. (2005). *Sosyal ve davranışsal ölçümlerde güvenilirlik ve geçerlilik*. Ankara:Seçkin Yayıncılık.
- Tabachnick, B. G. & Fidell, L.S. (2006). *Using Multivariate Statistics* (4th Ed.). Boston: Ally And Bacon.
<https://www.techlearning.com/how-to/how-to-develop-data-literacy-in-elementary-aged-student>.
<https://teknotower.com/veri-ve-bilginin-onemi-1/>.
- Tavşancıl, E. (2018). *Tutumların ölçülmesi ve spss ile veri analizi*. (6.Basım)Ankara:Nobel.
- Tekindal, S. (2015).*Duyuşsal özelliklerin ölçülmesi için araç oluşturma*.(3.Basım).Ankara:Pegem Akademi.
www.tdk.gov.tr.
- Ural, A. & Kılıç, İ. (2006).*Bilimsel araştırma süreci ve spss ile veri analizi*. Ankara:Detay Yayıncılık.
<https://www.veribilimiokulu.com/veriokuryazarligi>.
- Yeşilorman, M. & Koç, F. (2014). Bilgi toplumunun teknolojik temelleri üzerine eleştirel bir bakış. *Fırat Üniversitesi Sosyal Bilimler Dergisi*, 24(1), 117-133.