Article History
Received: 27.02.2022
Received in revised form: 27.08.2022
Accepted: 29.08.2022
Article Type: Research Article

International Journal of Contemporary Educational Research (IJCER)

www.ijcer.net

Validity and Reliability Study on The Development of Data Literacy Scale for Educators

Serap Öz[1], Ali Özdemir[2]

[1]Republic of Turkey Ministry of National Education, 0000-0002-7291-59342
[2]Marmara University, 0000-0001-6089-1966

To cite this article:

Öz, S., & Özdemir, A. (2022). Validity And Reliability Study on The Development of Data Literacy Scale For Educators. *International Journal of Contemporary Educational Research, International Journal of Contemporary Educational Research*, *9*(3), 649-661. https://doi.org/10.33200/ijcer.1079774

This article may be used for research, teaching, and private study purposes.

Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing,
systematic supply, or distribution in any form to anyone is expressly forbidden.

Authors alone are responsible for the contents of their articles. The journal owns the copyright of the articles.

The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of the research material.

International Journal of Contemporary Educational Research
Volume 9, Number 3, September 2022, 649-661 ISSN: 2148-3868

Validity and Reliability Study on the Development of Data Literacy Scale for Educators

Serap Öz[1*], Ali Özdemir[2]
[1]Republic of Turkey Ministry of National Education
[2]Marmara University

Abstract

The purpose of this study is to develop a valid and reliable Likert-type scale that can be used to measure the data literacy skills of educators. In the development process of the scale, after reviewing the relevant literature, a pool of 130 items was designed and presented to the experts for their view. After the evaluation of experts, the content validity rate and content validity indexes of the items were calculated by using the Lawshe method, and 39 items were formed. The draft scale of 39 items was applied to 820 teachers and administrators working in public schools in the 2021-2022 academic year. In order to determine the construct validity of the scale, principal component analysis and factor analysis were performed. As a result of these analyses, a scale consisting of 3 factors and a total of 30 items was developed. For reliability, Cronbach α (.953) a coefficient was calculated for the overall dimensions of the scale and varimax rotation. The results show that the scale is valid and reliable.

Keywords: Data literacy, Scale development, Educational data mining, Data literacy scale

Introduction

The term "data" constitutes the basis of information, which is considered as the most important power and capital of human beings today. The amount of data that is not meaningful on its own and turns into "information" only when interpreted has increased enormously through the development of information and communication technologies. Since the amount of this data is so great and as it is constantly changing, the concepts of big data and big data analysis have emerged. Big data is formed by patterns and templates that cannot be reached with classical mathematical methods, and big data analysis means the extraction of these patterns and templates meaningfully as an important decision-making process today **(**Özen, Kartal & Emre, 2017). When the "data" is mentioned in an educational context, it means systematically collected information that reflects education in school in many ways (Schildkamp & Lai, 2013). Many components, such as test results of students or observations made by teachers (Ikemoto & Marsh, 2007), the frequency of their visits to some platforms or even their facial expressions may be (Özen, Kartal & Emre, 2017) accepted as educational data.

Through the 21st century, the role expected of teachers has also changed significantly. In an age where information increases twice every four years, data becomes more important day by day, it is possible for countries to become information societies with individuals who can use and control this information so that they do not fall behind the world (Drucker,1996). Therefore, educators are expected to develop their 21st century skills. "Data literacy" is the skill required to interpret, evaluate, and make decisions based on data analysis today, where huge amounts of data are collected in almost all institutions thanks to the rapidly developing information technologies in the last 20 years (Bollier, 2010).

Data literacy is defined as "the ability to understand and use data effectively to make decisions." (Mandinach & Gummer, 2013). Mandinach (2012) states that educators should go through stages such as collecting, examining, analyzing, and interpreting data systematically while making data-based decisions. Green, Wilson, Versland, Gibson & Nollmeyer (2015) argue that very few educators, including administrators, have received training on that subject and suggest that professional development policies should be adopted to develop these skills for both teachers and school administrators. Bocala and Boudett (2015) state that the training for developing data literacy skills in education faculties is limited. It is stated here that current teachers must improve their data literacy skills in order to develop teaching strategies (Schildkamp, Karbautzki, & Vanhoof, 2014). Ingram, Louis, and

* Corresponding author, *Serap Öz*, *serapozyigit@hotmail.com*

Schroeder (2004) indicate that in order to improve the quality of education, educators should develop data-based decision-making skills.

The use of data in education has frequently been emphasized by politicians in the last decade to ensure sustainable development in many areas. Moreover, many governments aiming to take greater responsibility on this issue provide professional learning opportunities for educators to learn how to use data responsibly that is, how to protect students' privacy"to facilitate data-based instructional decision-making" (Mandinach & Gummer, 2016). Carlson, Borman, and Robinson (2011) argue that improving data literacy skills in teachers will have educational outcomes such as developing more effective classroom and teaching practices and ultimately increasing student performance. It is predicted that making decisions based on data in education can help students determine their strengths and weaknesses in learning and determine the appropriate instructional actions for what they need (Hoogland et al. 2016).

Although there are attempts to develop data literacy competencies as part of 21st century skills for both teachers and students in Turkey, the relevant literature does not suggest any instrument to assess and evaluate these competencies. Through this study, it aims to develop a data literacy scale in educators and to contribute to the literature.

21st Century Teacher Skills

The changing role of the teacher throughout the 21st century has changed from being the source of information to being a mediator between knowledge and student. Today, where access to all kinds of resources and information is so common, students have become individuals who can instantly check the accuracy and reliability of the information provided by their teachers. In this context, it is important for teachers to access and gain correct data and transfer it to students by associating it with daily life (Driscoll, 2019). 21st century skills are defined by many international organizations such as P21 (Partnership for 21st Century Learning), OECD (Organization for Economic Co-operation and Development), ASIA Society (Asia Society Partnership for Global Learning), ISTE (International Society for Technology in Education), NCREL (North Central Regional Educational Laboratory), and EU (European Union). According to the definitions put forward by these organizations, the competencies that 21st century teachers need to develop are data literacy and using digital tools to analyze and evaluate data (Anagün, Atalay, Kılıç, & Yaşar, 2016).

The research in the USA and Australia shows that there have been many studies on the use and interpreation of educational data, such as statistical analysis of standard assessment and evaluation results and their display on data walls in teachers' rooms and corridors (Sahlberg & Hasak, 2016). Romero & Ventura (2008) state in their research around the world that educational data analysis applications are more frequently used in regions such as North America, Western Europe, Austria, and New Zealand compared to other countries in the world. According to Long and Siemens (2001), the use of big data in education has a very significant role as a guide in determining the necessary innovations in the curriculum and pedagogical strategies. Among the studies aiming to specify the 21st century skills, Trilling and Fadel (2009) stated that under the term digital literacy, data literacy, media literacy, and information and technology literacy are very fundamental for educators. Wagner (2008) also emphasizes the importance of skills for accessing and analyzing data among the seven skills of the 21st century.

The 21st Century Skills Assessment and Teaching Organization (ATC21S) within Cisco, Intel, and Microsoft has categorized 21st century skills and emphasizes data literacy skills for people as a very significant instrument for the success of institutions (Kyllonen, 2012). Moreover, Hoogland (2016), states that data literacy is a critical skill for the 21st century. He defines the data literacy skill as asking and answering questions about the available data and following the data exploration, the visualization of it to understand things better. Data literacy differs from the data analysis stage at the point that it encompasses the ability to clean and prepare data for analysis by casting a critical eye on the outcomes.

In Turkey, the Ministry of National Education (MEB) stated that teachers should have the competence to analyze and interpret the data after determining the assessment and evaluation methods and techniques within the scope of student development monitoring and evaluation (MEB, 2008). The general qualifications of teachers, which were updated in the year 2017, underlie the teachers' competence to prepare and use measurement and evaluation tools suitable for their developmental characteristics among the professional knowledge and skills related to the 21st century.

Turkey has made many attempts to ensure the integration of changing world conditions into educational institutions with some great projects such as, Movement for Increasing Opportunities and Technology

Improvement (FATİH), Maine, Apple Future Classrooms (Göksun, 2016). The concept of "data literacy" is emphasized under the theme of "Data Based Management" in the 2023 Education Vision Report published in December 2018. The term is quite new for our country and needs to be investigated in detail. "A Competent Data Control Unit" and "Learning Analytics Platform" where the academic data of the students is evaluated together with the data regarding their interests, abilities, and temperaments are among the goals of the Ministry of National Education (2018). These studies show that our country will need more educators specializing in data analysis in the near future. As a result, establishing a data literacy culture and measuring it in educational institutions has become a very fundamental objective for Turkey, as well as in other countries of the world (2023 Education Vision Document, 2018).

Educators' Data Literacy Ability

The term "literacy" is defined by UNESCO (2004) as "the ability to define, understand, interpret, create, and calculate through relevant written and printed sources". Data literacy is defined as "the desire and ability to attract the attention of society through data" (Dağ, 2019). According to Morrow (2018), data literacy is the ability to read, study, analyze, and discuss data. As one of the fundamental competencies of the 21st century, data literacy is defined as knowing how to access data in different ways, asking questions, and making basic statistical analysis (Dağ, 2019).

In recent years, the terms as "data collection" and "data analysis" has been a significant necessity, especially in governmental institutions where a very huge amount of data is gathered. Consequently, a need to benefit from the power of statistics by analyzing the collected data in a swift and meaningful way has arisen. In this context, it has become more important than ever to provide the next generation with statistical analysis knowledge and thinking skills by adapting them to various fields (Aydeniz, 2017). Although many schools, not only in our country but around the world, collect a lot of data about students, they do not have a clear vision of how the collected data can be used to increase student success (Creighton, 2006). Kekahio and Baker (2013) point out that using data to improve educational policies and practices has become a high priority in the context of increasing accountability practices in recent years, but educators do not have sufficient infrastructure in this regard. Accordingly, educators need to gain data literacy skills and raise future generations who are also able to use these skills. In a study by the U.S. Department of Education (2011), teachers responded to scenarios involving hypothetical student data designed to investigate their understanding of the types of data available to support teaching decisions. It was observed that teachers had difficulties to finding data, recognizing different types of data, making appropriate calculations, interpreting data tables and graphs, and applying the findings to improve students' learning.

Green et al. (2016) aimed to improve teachers' three skills in their studies to increase their data literacy level in order to improve their data-based decision-making skills. These skills are, recognizing multiple data sources and identifying how they can be used to improve teaching, discovering how to analyze and interpret data in a way that helps improve teaching; and creating a school team that supports teachers' use of data to improve teaching. The criteria determined for data literacy in educators within the scope of the data-based decision-making applications training program developed by the University of Twente consist of eight steps. These are as follows: defining problems, developing hypotheses or questions, gathering data, data quality control, data analysis, interpretation and conclusion, implementation of improvement measures, and evaluation (Kippers et al.2018).

Mandinach, Friedman, and Gummer (2015) emphasize the importance of data literacy for teachers as a key ability to transform information into actionable teaching knowledge and practices by collecting, analyzing, and interpreting all kinds of data and determining appropriate teaching steps. Accordingly, they also define a teacher with data literacy skills as someone who is able to make his data literacy skills meaningful by relating discipline knowledge and practices, curriculum knowledge, pedagogical content knowledge, and how children learn. Within the educational context where education has turned into an evidence-based profession, data literacy has become a basic skill for teachers and all teachers to apply effective teaching practices (Mandinach & Gummer, 2016). Until today, an effective education policy has not been followed regarding effective data use, which is one of the main determinants of teaching quality in institutions. As a result, data use has been perceived by many teachers as a burden to deal with rather than an instrument to be used to improve the teaching process (Almy, Chong & Dorrington, 2019).

It is known that in our education system (MEB, 2019), which has more than 18 million students and nearly one million teachers as of 2018, educational data has increased very rapidly depending on this number. In particular, digital platforms such as e-school, EBA, DYS, MEBBİS, TEFBİS are known to be very important data sources for the education system (Akgün, 2019). When the data collected on these platforms is not processed

meaningfully, it turns into data stacks that we call "big data". The studies conducted in our country show that the educators in fact have a sufficient amount of data that can be implemented in the decision-making process (Demir, 2009). However, in our country, the competence in "data literacy", which means analyzing and extracting data masses among educators in a meaningful way and attracting the attention of society through these data, is very limited (MEB, 2018). In a study conducted by Özdemir Saylam and Bilen (2018) among educators in 2018, the rate of participants who had an idea about educational data mining was found to be 17%. Considering that this rate is quite low, it is important to develop a policy to empower teachers with data and to ensure that they follow career paths that support their ability to use data to improve teaching. Without the necessary data handling skills, teachers do not have a powerful tool to make the best decisions to improve student achievement (Almy, Chong, & Dorrington, 2019). The current policies implemented to date have mainly focused on the field of measurement and evaluation (MEB, 2017). The importance of data use in educational decision making and in improving the quality of education has not been emphasized or addressed by educational policies yet. Today, the knowledge that educators know about students and education is far beyond their current test scores. They need to know and make sense of how to use a variety of student data in order to improve teaching and student achievement (Almy et al.2019). One of the main steps that Turkey should have taken to compete in the international arena in the field of science and technology in line with the objectives of the 2023 vision is to provide all citizens, not only educators, with data literacy competence. It is of great importance to increase the data literacy level among educators in order to achieve the goal of "making data-based decisions in education" set forth within the scope of the 2023 Education Vision Report.

The first attempt to develop data literacy skills in educators should be to determine their competencies in that subject. However, in our country, a measurement system to determine the teacher's competencies has not yet been available. Issues such as performance evaluation that come up from time to time are also postponed for certain reasons. Accordingly, there is no information about the competencies of teachers in Turkey except for the KPSS scores, in which only general cultural and pedagogical knowledge scores are measured. Considering the systems of countries with high success in exams such as PISA and TIMMS, it is seen that serious studies are carried out in these countries on subjects such as the training, selection, and professional development of teachers. In the study conducted by Aydeniz (2017), it is stated that curriculum reform is one of the main things to be done to adapt to the requirements of the age. Nevertheless, to what extent teachers are competent in statistical thinking and data analysis skills, which are among the subjects that should be added to the curriculum in order to establish the 21st century competencies that students should acquire, is another subject that needs to be investigated. Although many in-service training programs have been planned for the professional development of teachers in recent years, there has been no study on measuring data literacy skills and improving their competencies in this area.

Method

This research was carried out on the screening model. The aim of the study is to develop a valid and reliable scale to measure educators' data literacy levels. The three-stage (the creation of scale items and applying for expert opinion, pilot study, validity and reliability stages (Karasar, 2007), scale development process was followed during the study.

Study Group

The final study group consists of 820 educators who work as teachers or educational administrators in different regions of İstanbul in the 2021-2022 education year. The demographic information of the study group is given in Table 1. The "maximum diversity" sampling method, which is one of the purposive sampling methods, was used to determine the study group. By creating a relatively small sample in the maximum diversity sampling method, it is to reflect the diversity of individuals who may be a party to the problem studied in this sample at the maximum level (Yıldırım and Şimşek, 2016). For this purpose, the questions were sent to a group working in different regions and levels of Istanbul, working in different branches and having different experiences. At the time of data collection, the scale was sent to 1200 targeted people online due to the pandemic, but 820 people responded.

Table 1. The Demographic Information of the Study Group

Variable	Group	f	%
Gender	Woman	593	72,3
	Man	227	27,7
Age	Under 25	20	2,4
	26-35	306	37,3
	36-45	297	36,2
	46-55	168	20,5
	55+	29	3,5
Experience	1-3 years	68	8,3
	3-6 years	91	11,1
	6-9 years	150	18,3
	10+ years	511	62.3
Education	Bachelor Degree	603	73,5
	MA Student	22	5,1
	MA	168	20,5
	PHD Student	7	0,9
Title	Teacher	758	92,4
	Assistant Director	39	4,8
	Director	23	2,8
Institution Type	Official	544	66,3
	Private	276	33,7
Total		820	100

Table 2 shows the branches of the 820 participants. The table shows that maximum diversity among the participants was achieved by including teachers from all fields in the study.

Table 2. The Demographic Information of the Study Gruop According to Their Branches

Branch	f	%
Social Studies	139	17,0
Science Studies	92	28,2
Mathemathics	73	37,1
Foreign Language	117	51,3
Preschoool Teaching	30	55,0
Primary School Teaching	164	75,0
Vocational Courses	71	83,7
Pschological Guidance	62	91,2
Visual Arts / Music / Physical Education	72	17,0
Total	820	100

Scale Development Steps

The Forming of an Item Pool

In the process of scale development, the literature on "data literacy in education" was first reviewed by scanning the local and international studies in the last ten years on the subject. After the expert opinions, new items were added to the items prepared based on the data obtained according to the literature review. As a result, 130 candidate items thought to cover data literacy skills for educators were created.

Expert Opinions

Candidate items were submitted to 11 expert (referee) opinions. Seven of these people are experts in education management; four of them are educators who have completed their expertise in information technologies and data analysis. These experts are currently working as researchers or have their PHD degree at Marmara University, Yıldız Technical University and Hasan Kalyoncu University

In the scoring process of the "Data Literacy Scale for Educators" (DLSFE), reviewers were asked to scale the items to 3 points as (Remove, Review, Remain). In the items that need to be reviewed, changes were made through the opinions of experts. The Lawshe technique (Yurdugül & Bayrak, 2012) was used to determine the content validity rates (CGO) of the items in line with the expert opinions received. For this analysis, the formula [KGO = Nu/(N/2)-1) was used. The CGO indicated in the formula = Content validity rate, Nu = the number of experts who agree with the item and N = the total number of experts. The minimum value of the content validity criterion (CGO) for 11 experts is 0.59 (Ayre & Scally, 2014). As a result of the analysis, some of the items that were deemed irrelevant by the referees were removed, and some of the others were revised. The final 39-item scale was obtained.

The Creation of the Draft Form and the Pilot Study

The draft scale form was prepared by taking into account the cumulative rating scale (Likert-type scale) that allows multiple item use aiming to measure a single structure. The draft form was prepared in the form of (1) I strongly disagree, (2) I do not agree, (3) I am undecided, (4) I agree, (5) I completely agree, in a 5-point likert type. In addition, an "instruction" indicating the purpose of the scale and a "personal information form" were added to the draft form. An ethical consent was also attached to the participants, detailing their rights, including the right to withdraw at any time during the research process. The 39-item draft form was applied to 60 people as a pilot study. The Cronbach alpha reliability value was calculated as.953 which is considered a good value. As a result of the analysis, no item that reduces the reliability of the scale was found.

Study on Validity and Reliability

The draft scale was sent online due to the pandemic to the study group, education administrators and teachers working in public and private schools in different regions of İstanbul in the years 2021-2022. 820 people filled the scales completely. The data obtained from DLSFE were analyzed statistically and first, the construct validity of the scale was examined by performing exploratory factor analysis. Whether the DLSFE is suitable for factor analysis was determined by looking at the results of the Kaiser-Meyer-Olkin (KMO) Coefficient and Bartlett's Sphericity Test. As a result of the analysis, it was seen that the draft scale form was suitable for factoring, and an exploratory factor analysis was performed. Then, the structure of the scale was revealed by rotating the maximum variability (varimax) using the principal component analysis method. The 5 items that were not suitable in terms of factor loadings were eliminated, and the total item correlations of the scale were calculated over the remaining items and tested. For the reliability of the scale, the internal consistency coefficient was calculated with the Cronbach alpha formula, and the Cronbach alpha value was found to be.953.

Findings

An Exploratory Factor Analysis (EFA) was conducted to determine the construct validity of the scale. To decide whether the data set is suitable for factoring, Kaiser-Meyer-Olkin (KMO) and Bartlett's tests were conducted. The findings of the analysis are shown in Table 3.

Table 3: KMO and Bartlett's Test Analysis

Kaiser-Mayer-Olkin (KMO) Measure of Sampling Adequacy		0.923
Bartlett Test of Sphericity	X^2	19105.475
	df	465
	p	0.000*

According to Table 2, the Kaiser-Meyer-Olkin (KMO) test value is.923. This means that scale items are suitable for factoring. Büyüköztürk (2002) states that a KMO value of over.70 means that the scale is suitable for factoring. According to Bartlett's test, the data came from a multidimensional universe ($p<.001$) (Büyüköztürk, 2002).

The type of analysis in which the researcher does not have any information about the number of factors measured by the measuring tool and tries to obtain information about the nature of the factors measured with the measuring tool instead of testing a certain hypothesis is called exploratory factor analysis (Tavşancıl, 2006). In scale development studies, it is aimed to reveal the factor structure with principal component analysis, which is frequently used to test construct validity (Çokluk, Şekercioğlu & Büyüköztürk, 2018).

Confirmatory Factor Analysis

Then, Confirmatory Factor Analysis (CFA) was applied in order to make the confimation of the three-factor structure of the scale (Büyüköztürk, 2012). The model for the three-factor structure, which was determined by EFA as a result, was tested with CFA. The results can be seen in Table 4 below.

Table 4: Goodness-of-fit Values for CFA

Fit Indices	Statistics
CMIN/DF	4.67
RMSEA	0.067
SRMR	0.059
χ^2/d.f	1.784
GFI	0.869
AGFI	0.841
CFI	0.91
NFI	0.97
RFI	0.97

The table 4 indicates that the factor loads, and goodness-of-fit values are within the acceptable levels. The results of the tests show that the developed scale was valid. Also, the variables fit with the dimensions (Yaşlıoğlu, 2017).

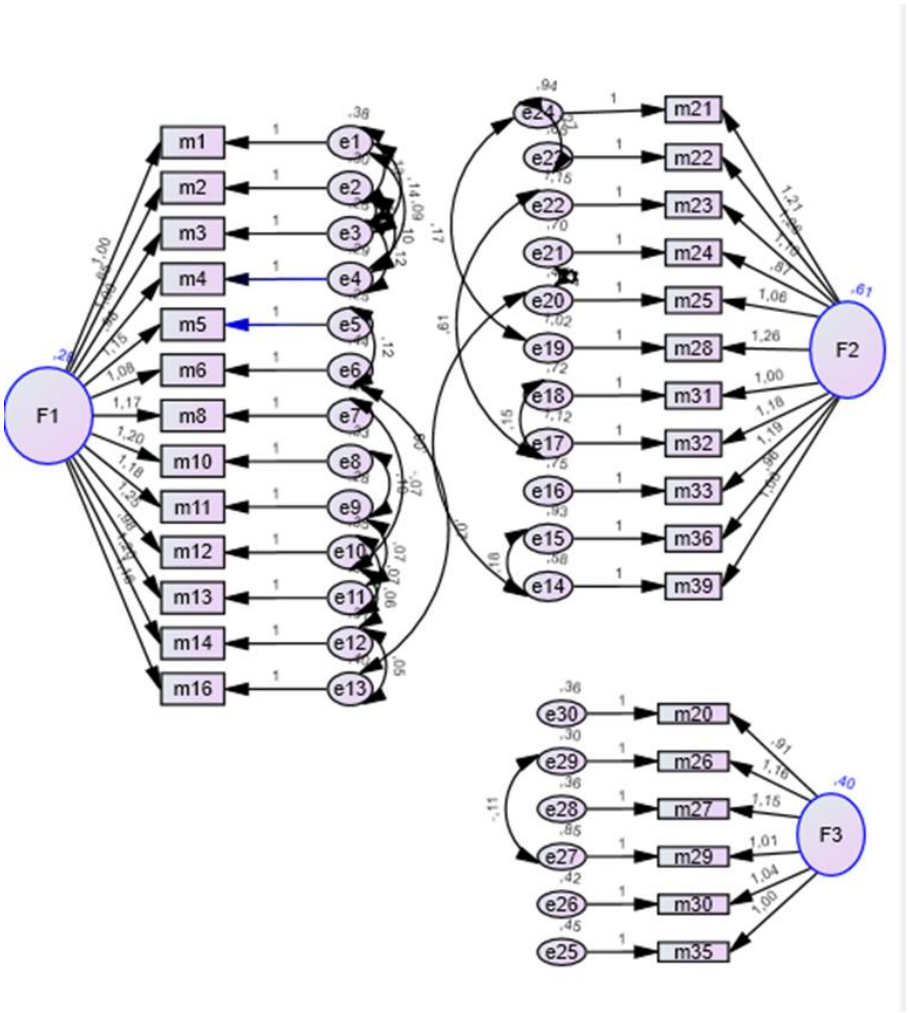

Figure 1. CFA Results

The varimax rotation method is based on the idea that the factors are interrelated and the oblique rotation methods (Büyüköztürk, 2002). As a result of the exploratory factor analysis, it is seen that the scale consists of

three factors. When we look at the percentage of factors explaining the total variance, the first factor explains 43,795, the second factor 10,180, and the third factor, 5,357. Explain why the total variance of factors in the scale is 59.332% in the total rate. The scattering diagram showing the breaking points of the scale is shown in Figure 2.

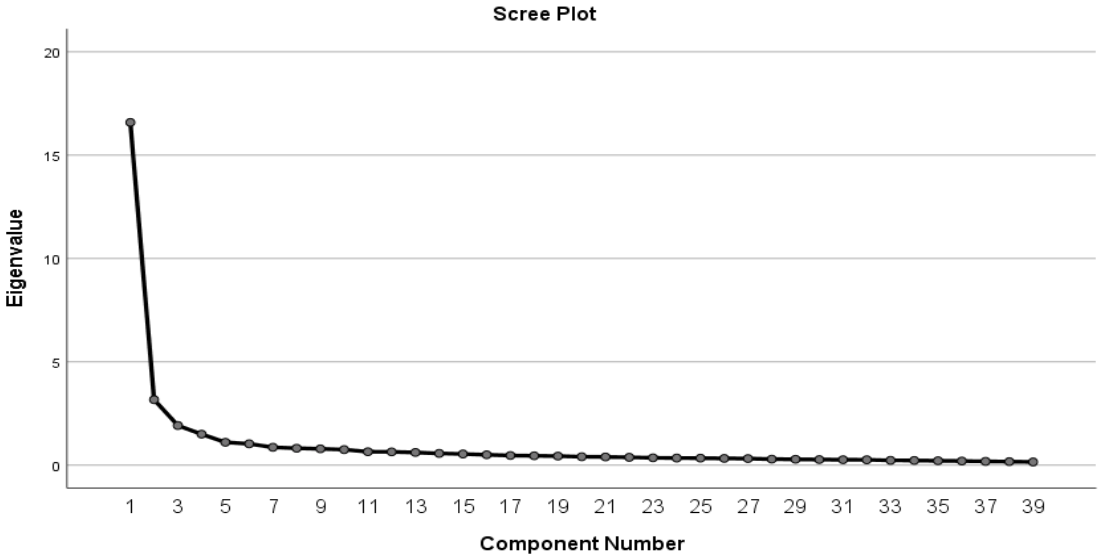

Figure 2. Screeplot Values

When the table is examined, it is seen that there are 3 different breaking points greater than one. High acceleration and rapid declines in the factor line chart are important in deciding the number of important factors (Büyüköztürk, 2002). According to the scatter diagram, there are 3 factors with an eigenvalue greater than one. Accordingly, the scale has 3 factors. The factor values of the loads are shown in Table 5.

Table 6. Factor Loads in Dimensions as a Result of Rotation

	Self-Efficacy	Experience	Attitude
item 1	0,795		
item 2	0,826		
item 3	0,922		
item 4	0,808		
item 5	0,715		
item 6	0,711		
item 8	0,704		
item 10	0,663		
item 11	0,707		
item 12	0,662		
item 13	0,642		
item 14	0,700		
item 16	0,617		
item 21	0,795		
item 22		0,717	
item 23		0,810	
item 24		0,777	
item 25		0,586	
item 28		0,627	
item 31		0,707	
item 32		0,618	
item 33		0,782	
item 36		0,609	
item 39		0,564	
item 20		0,587	
item 26			0,572
item 27			0,683

item 29	0,725
item 30	0,510
item 35	0,552

In scale development studies, factor loadings should be more than.30 (Büyüköztürk, 2002). For this reason, items with a factor load of.30 and above were evaluated. When Table 6 is examined, it is seen that items 1,2,3,4,5,6,8,10,11,12,13,14, and 16 had the highest factor loadings in the first factor. The items 21,22,23,24,25,28,31,32,33,36, and 39 had the highest factor loadings in the second factor. Lastly The items 20,35,26,27,29, and 30 had the highest loadings for the third factor.

Table 6. The Items Take Loadings in Sub Dimensions

Factor	Number of Items	Item Numbers
1	13	1,2,3,4,5,6,8,10,11,12,13,14,16
2	11	21,22,23,24,25,28,31,32,33,36,39
3	6	20,26,27,29,30,35

It has been observed that the scale has a three-factor structure with an eigenvalue greater than 1. The eigenvalues and cumulative variance percentages of the three factors found are shown in Table 7.

Table 7. Sub-Factors and Total Variance Explained

Factor	Eigenvalue	Variance Explained	Cumulative%
1	13,577	43,795	43,795
2	3,577	10,180	53,975
3	1,661	5,357	59,332

When Table 7 is examined, it is seen that the total variance explained under two factors is 53%. It is seen that the first factor has 43.7%, the second factor has 10%, and the third factor has 5% of the explained variance. A common meaning was sought for each item gathered under these three factors, and naming was done according to the literature. In this context, the first factor has been named "self-efficacy", the second factor as " experience", and the third factor was named "attitude". Factor load values, common factor variance, means and standard deviation of the items for factor 1 in the scale are shown in Table 8.

Table 8. Factor Loading Values, Common Factor Variance, Means and Standard Deviation of Items for the Factor 1

Self-Efficacy	Factor1	h^2	X	sd
1. I can distinguish the data I encounter in my daily life (numerical data, character string data, logical data, etc.).	,795	,549	4,489	,818
2. I can tell the difference between quantitative and qualitative data.	,826	,576	4,632	,709
3. I am able to link data together.	,922	,656	4,528	,729
4. I can understand whether the two data points are directly or inversely proportional to each other.	,808	,576	4,543	,748
5. When I need to access some data, I know where to find it.	,715	,654	4,372	,796
6. I can access any kind of data I need.	,711	,548	4,093	,879
8. Data gives me an idea of why some goals are not achieved.	,704	,559	4,193	,821
10. I can extract the data I have collected and achieve meaningful results.	,663	,582	4,198	,857
11. The data not only gives me an idea but also helps me see the gaps and fill them in.	,707	,640	4,364	,823
12. I can interpret tables and graphs containing statistical information.	,662	,576	4,254	,894
13. I think interpreting data is as basic a skill as reading and writing.	,642	,559	4,442	,798
14. I think I am competent in evaluating the data I encounter in my profession.	,700	,615	4,176	,846
16. When I see data about education, I can confirm the accuracy of this information from various sources.	,617	,578	4,030	,888

Factor loading values, common factor variance, means and standard deviation of the items for factor 2 in the scale are shown in Table 9.

Table 9. Factor Loading Values, Common Factor Variance, Means and Standard Deviation of Items for the Factor 2

Experience	Factor2	h^2	X	Sd
21. I am familiar with at least one data analysis program (such as Excel, SPSS, Maxquda, Nvivo, and so on).	,717	,538	3,403	1,361
22. I think I am competent at visualizing data with graphics.	,810	,652	3,231	1,273
23. I know about the concept of data mining.	,777	,587	2,395	1,377
24. I check the course achievement graphics of my students over the digital platforms (e.g., school, eba, zoom, etc.) they use.	,586	,422	4,052	1,077
25. I can make discussions and comments on data analysis.	,627	,639	3,792	1,074
28. I have created graphics and tables using a computer program before.	,707	,513	3,391	1,408
31. I follow the data about education in the world.	,618	,509	3,501	1,150
32. I know the concept of a data set.	,782	,617	2,481	1,404
33.I examine statistically the change in my students' exam scores.	,609	,531	3,489	1,269
36. I think I am familiar with basic statistical concepts (mode, median, standard deviation, arithmetic mean, etc.).	,564	,451	3,563	1,223
39. I use data effectively to achieve a professional set of goals.	,587	,620	3,708	1,090

Factor loading values, common factor variance, means and standard deviation of the items for factor 3 in the scale are shown in Table 10.

Table 10. Factor Loading Values, Common Factor Variance, Means and Standard Deviation of Items for the Factor 3

Attitude	Factor3	h^2	X	Sd
20. I think the data is necessary for the structuring, recording, and easy analysis of the information.	0,572	,536	4,291	,827
26. I find it important that decisions are made based on data in the education system.	0,683	,571	4,214	,911
27. Data analysis skills make me a better educator.	0,725	,523	4,184	,938
29. I think in-service training is necessary in data analysis.	0,510	,329	4,046	1,118
30. When I come across statistical data about education, I examine it carefully.	0,552	,608	4,082	,923
35. Visualizing data is important for seeing meaningful relationships	0,610	,642	4,282	,920

Şencan (2005) states that the model matrix will be sufficient for easier interpretation of the factors. The reliability of the scale was calculated with the Cronbach Alpha internal consistency coefficient. Cronbach Alpha values are shown in Table 11.

Tablo 11. The Cronbach-Alpha Values for Sub-Dimensions of the Data Literacy Scale for Educators

Sub-Dimensions	Cronbach-Alpha
Self-Efficacy	.932
Experience	.913
Attitude	.846
Total Scale	.953

According to Table 11, the reliability coefficent value of the scale sub-dimensions is.932 for the “Self-Efficacy”, .913 for the “Experience” and.846 for the “Attitude“. The value for the whole scale is.953 after eliminating 5 factors according to the factor analysis results.

Discussion, Conclusion and Suggestions

Data literacy skills have become a basic need in today's world, where teachers have to work with increasing

amounts of data every day. Accordingly, data literacy, which must be included in the program in the teacher training process (Mandinach and Gummer, 2016), is a competency that needs to be researched, measured, and discussed in our country as well. Salmacia (2017), found strong agreement that a critical aspect of being a successful teacher is being data-literate. However, how to measure and evaluate this skill is still a matter of debate among researchers.

The aim of this study is to develop a scale to measure the data literacy level of educators and to analyze its validity and reliability. The Data Literacy Scale for Educators is a Likert-type scale with a total of 30 items under three factors. These three factors explain 59% of the total variance. The fact that the scale explains 40% of the general variance is considered sufficient in terms of social sciences (Kline, 2014). The factor load values of the items of attitude towards statistics related to factor analysis are given in Table 7. Items with an item-total correlation of 0.30 and higher indicate that they can measure the feature to be measured (Çokluk, Şekercioğlu & Büyüköztürk, 2012).

The KMO and Barlett tests were used to determine the suitability of the data for factor analysis. The data are suitable for factor analysis if the KMO is greater than 0.60 and the Barlett test is significant (Çokluk, ekerciolu, & Büyüköztürk, 2012). The Kaiser-Meyer-Olkin sample suitability value of 0.960 and the significance level of Bartlett's test of sphericity of 0.000 (for $p \leq 0.05$) indicate that the data are suitable for factor analysis. When the three factors are considered together, they explain 55,554% of the variation in the total variance. This value depends on the load values of the items in each factor and is aimed at being increased (Çokluk, Şekercioğlu & Büyüköztürk 2012). The fact that the explained variance exceeds 50% of the total variance is an important criterion of factor analysis. At this rate, it is an acceptable value within the expected disclosure percentage rates (Tavşancıl, 2014).

The scree plot made to support the eigenvalue assumption is given in Figure 1. When the scree plot, which includes the eigenvalues on the vertical axis and the factors on the horizontal axis, is examined, it is seen that the high-accelerated decline decreases after the third point. From the first point, the downward trend seen from the beginning is indicated by the dots in the degree of contribution to the variance, and each interval between two points represents a factor (Çokluk, Şekercioğlu & Büyüköztürk 2012). As a result of the exploratory factor analysis, three factors with eigenvalues above 1 are named as self-efficacy, experience and attitude. Fifteen items in the self-efficacy dimension reveal the knowledge level of educators in applying data analysis. Thirteen items in the experience dimension reveal the past experiences of educators in data analysis and data literacy in their education or business life. Finally, six items in the attitude dimension reveal the feelings and thoughts of educators about the necessity of data literacy skills in their profession.

The Cronbach Alpha Coefficient was used for the reliability of the scale. Accordingly, the alpha value of the self-efficacy dimension was found to be.953, the alpha value of the experience sub-dimension as.913 and the alpha value of the attitude dimension as.846. The Cronbach Alpha reliability coefficient of the total scale is.952. Cronbach's alpha value of the scale is between $0.81 < \alpha < 1.00$ and this points out that it is highly reliable (Tavşancıl, 2014).

As a result, it can be said that the Data Literacy Scale for Educators is a convenient assessment instrument in terms of validity and reliability. The scale is expected to contribute to the field of data collection by revealing the data literacy level of educators and raising awareness about this issue.

Limitations and Recommendations

The basic limitation of this study is the number of participants. It is recommended that the reliability and validity of the scale be determined in future studies with larger samples, different age groups, and random samples. It is recommended that the survey be conducted in other areas of the country as the current research is limited to İstanbul.

Author (s) Contribution Rate

The authors contributed equally to the article.

Conflicts of Interest

The authors declare that they have no conflict of interest.

Ethical Approval

Ethical permission (12.01.2022-215807) was obtained from the Marmara University Research and Publication Ethics Committee for this research.

References

Akgün, E. (2019). *Educational data mining in 2023 Education Vision*. *SETA*.

Almy, S., Chong, H., & Dorrington, A. (2019, October 11). *Data Literacy*. Retrieved from, https://dataqualitycampaign.org/wp- content/uploads/2016/03/DQC-Data-Literacy-Brief.pdf

Anagün, Ş. S., Atalay, N., Kılıç, Z., & Yaşar, S. (2016). Developing the 21st century skills efficiency perceptions scale for teacher candidates: A validity and reliability study. *Journal of PAU Faculty of Education, 40*(40), 160 - 175. doi: 10.9779/PUJE768

Aydeniz, M. (2017). Our education system and our 21st century dream: a stem-oriented economic roadmap for Turkey as we move towards 2045 goals. *Trace: Tennessee Research and Creative Exchange*.

Ayre, C., & Scally, A. J. (2014). Critical values for Lawshe's content validity ratio: Revisiting the original methods of calculation. *Measurement and Evaluation in Counseling and Development, 47*(1), 79-86.

Bocala, C. and Boudett, K.P. (2015, May 8). *Teaching educators habits of mind for using data wisely, teachers college record*. Retrieved from https://www.tcrecord.org/Content.asp?ContentId=17853

Bollier, D. (2010*). The promise and peril of big data*. Washington: Aspen Institute, Communications and Society Program.*T*he Aspen Institue Publishing Washington DC.

Büyüköztürk, Ş. (2002). Faktör analizi: Temel kavramlar ve ölçek geliştirmede kullanımı. *Kuram ve Uygulamada Eğitim Yönetimi, 32*(32), 470-483.

Carlson, D., Borman, G.D., & Robinson, M. (2011). A multistate district-level cluster randomized trial of the impact of data-driven reform on reading and mathematics achievement. *Educational Evaluation and Policy Analysis, 33*(3), 378-398.

Creighton, T. (2006). *Schools and data: The educator's guide for using data to improve decision making*. Corwin Press.

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

Dağ, P. (2019 November, 11). *Veri okuryazarlığını anlamak – I.* Retrieved from https://www.egitimreformugirisimi.org

Demir, K. (2009). İlköğretim okullarında verilere dayalı karar verme. *Theory and Practice Education Management, 59*(59), 367 - 397.

Driscoll, M. (2019, November 19). *Future-focused strategic planning for schools: the 'what' and the 'why' need a 'how'*. Retrieved from https://thinkstrategicforschools.com/education-21st-century/

Drucker, P. (1996). *21. yüzyıl için yönetim tartışmaları.* Epsilon Publishing.

Göksün, D. O. (2016*). Öğretmen adaylarının 21. Yy. öğrenen becerileri ve 21. Yy. öğreten becerileri arasındaki ilişki* [Unpublished doctoral dissertation]. Anadolu Üniversitesi, Eskişehir.

Green, J. L., Wilson, S, Versland, T., Gibson L. & Nollmeyer G. (2016). Teachers and data literacy: A blueprint for professional development to foster data driven decision making. *Journal of Continuing Education and Professional Development, 3*(1),14-32. doi:10.7726/jcepd.2016.1002

Eğitim Reformu Girişimi (2019 May 5). *Veri okuryazarlığını anlamak* Retrieved from https://www.egitimreformugirisimi.org/veri-okuryazarligini-anlamak-i/

Hoogland, I., Schildkamp, K., Van der Kleij, F., Heitink, M., Kippers, W., Veldkamp, B., & Dijkstra, A. (2016). Prerequisites for data-based decision making in the classroom: *Research evidence and practical illustrations. Teaching and Teacher Education, 60*, 377–386. http://dx.doi.org/10.1016/j.tate.2016.07.012

Ikemoto, G. S., & Marsh, J. A. (2007). Cutting through the data-driven mantra: Different conceptions of data-driven decision making. In P. A. Moss (Ed.). *Evidence and decision making* (pp. 105–131). Wiley Blackwell.

Ingram, D., Louis, K. R. S., & Schroeder, R. (2004). Accountability policies and teacherdecision making: Barriers to the use of data to improve practice. *Teachers College Record, 106*(6), 1258–1287.

Karasar, N. (2010). *Bilimsel araştırma yöntemi: kavramlar, ilkeler, teknikler*. Nobel.

Kekahio, W., & Baker, M. (2013*). Five steps for structuring data-informed conversations and action in education* (REL 2013–001). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Pacific.

Kippers, W. B, Poortman, C.L., Schildkamp, K. and Visscher, A. J. (2018). Data literacy: What do educators learn and struggle with during a data use intervention? *Studies in Educational Evaluation*. *56*, 21-31 doi: 10.1016/j.stueduc.2017.11.001.

Kline, P. (2014*). An easy guide to factor analysis.* Routledge.
Kyllonen, P. C. (2012). Measurement of 21st century skills within the common core state standards. In *Invitational Research Symposium on Technology Enhanced Assessments*, 7-8.
Long and Siemens, G. (2001). Penetrating the analytics in learning and education. *Educause,15,* 30 - 32.
Mandinach, E. B. (2012). A perfect time for data use: Using data-driven decision making to inform practice. *Educational Psychologist, 47*(2), 71–85.
Mandinach, E. B., & Gummer, E. S. (2013). A systemic view of implementing data literacy in educator preparation. *Educational Researcher, 42*(1), 30-37.
Mandinach, E., & Gummer, E. (2016). *Data literacy for educators: Making ıt count in teacher preparation and practice*. Teachers College Press
Mandinach, E., & Gummer, E. (2016). Every teacher should succeed with data literacy. *Sage Journals 97*(8) 43-46 https://doi.org/10.1177/0031721716647018.
Mandinach, E. (2014 November 20). *Data literacy for teachers: What is it, what do we know about how to make it happen and why is it important?* Retrieved from https://studylib.net/doc/5348687/data-literacy-for-teachers
Mandinach, E., Friedman, J. M., & Gummer, E. (2015). How can schools of education help to build educators' capacity to use data? A systemic view of the ıssue. *Teachers College Record, 117*(4), 1-50. https://eric.ed.gov/?id=EJ1056728.
MEB . (2008). *Öğretmen yeterlikleri: Öğretmenlik mesleği.* Milli Eğitim Bakanlığı Yayınları.
MEB. (2017). *Öğretmenlerin genel ve mesleki yeterlilikleri*. Milli Eğitim Bakanlığı Yayınları.
MEB. (2018 November 11). *2023 Eğitim Vizyon Belgesi*. Retrieved from http://2023vizyonu.meb.gov.tr/
MEB. (2019, May 19). *Örgün eğitim istatislikleri*. Retrieved from http://sgb.meb.gov.tr/meb_iys_dosyalar/2019_09/30102730_meb_istatistikleri_orgun_egitim_2018_2019.pdf a
Morrow, J. (2018 April, 28). *Developing a data literate workforce*. Retrieved from https://www.qlik.com/us/-/media/files/training/global-us/qlik-education-data-literacy-program-strategy-and-framework.pdf
Özdemir, A., Saylam, R., & Bilen, B.. (2018) Eğitim sisteminde veri madenciliği uygulamaları ve farkındalık üzerine bir durum çalışması. *Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*. *22*(2), 2159 – 2172.
Özen, Z., Kartal, E., & Emre, İ. (2017). Eğitimde büyük veri. *Turkish Online Journal of Educational Technology, 183*
Romero, C., & Ventura, S. (2008). Educational data mining: A Survey from 1995 to 2005. *Expert Systems with Applications*, 125 - 146.
Sahlberg, P., & Hasak, J. (2016, June,18). *Büyük veri ne kadar büyük?* Retrieved from https://tedmem.org/mem-notlari/degerlendirme/buyuk-veri-ne-kadar-buyuk
Salmacia, K. A. (2017). *Developing outcome-driven, data-literate teachers* [Unpublished doctoral dissertation]. University of Pennsylvania. ProQuest Dissertations & Theses Global.
Schildkamp, K., & Lai, M. K. (2013). Introduction. In K. Schildkamp, M. K. Lai, & L. EarSchildkamp, K., Karbautzki, L., & Vanhoof, J. (2014). Exploring data use practices around Europe: Identifying enablers and barriers. *Studies in Educational Evaluation, 42*(1),15–24. http://dx.doi.org/10.1016/j.stueduc.2013.10.007.
Şencan, H. (2005). *Sosyal ve davranışsal ölçümlerde güvenilirlik ve geçerlilik*. Seçkin.
Tavşancıl, E. (2014). *Tutumların ölçülmesi ve SPSS ile veri analizi*. Nobel.
Trilling, B., & Fadel, C. (2009). *21st century skills: Learning for life in our times.* John Wiley & Sons.
UNESCO Education Position Paper (2004, February,5). *The plurality of literacy and the implications of its policies and programs,* Retrieved from http://unesdoc.unesco.org images/0013/001362/136246e.pdf
Wagner, T. (2008). *The global achievement gap: Why even our best schools don't.* Basiz Books.
Yıldırım, A. & Şimşek, H. (2016). *Sosyal bilimde nitel araştırma yöntemleri.* Seçkin.
Yurdugül, H. ve Bayrak, F. (2012). Ölçek geliştirme çalışmalarında kapsam geçerlik ölçüleri: Kapsam geçerlik indeksi ve Kappa istatistiğinin karşılaştırılması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, Özel sayı 2*, 264-271.