

Reading The World With Statistical Literacy: Results of An Empirical Study

Deniz Ünal¹, Begüm Çığsar*¹, Dibanur Alam¹, Şerife Ezgi Atalar¹,
Sait İsmail Gümüş¹, Ali Çakıtlı¹

Abstract: As in all kinds of literacy, this study has been motivated by considering the contribution of statistical literacy to the development reaching from the individual to the society and its ability to bring critical thinking. Statistical literacy scale indexed by Turkish Surveying Index was applied to 398 university students and the correlations between the literacy levels of these participants and some socio-demographic characteristics were examined. As a result of the analyses, it was seen that there was no difference among the statistical literacy scores of the students in terms of years spent at school, parental education status or universities. It was observed that there was a weak positive relationship with grade points averages. The most striking finding is that numerical sciences students got higher scores than social sciences.

Keywords: Critical Interpretation, Critical Thinking, Linear Regression, Logistic Regression, Statistical Literacy

¹**Address:** Çukurova University, Faculty of Arts and Sciences, Dept. of Statistics, Adana, Türkiye.

***Corresponding author:** begumcigsar@gmail.com

Citation: Ünal, D., Cıgsar, B., Alam, D., Atalar, Ş.E., Gümüş, S.İ., Cakitli, A. (2023). Reading The World with Statistical Literacy: Results of An Empirical Study. Bilge International Journal of Science and Technology Research, 7(2): 198-203.

Citation:

1. INTRODUCTION

According to Turkish Language Association (TDK), literacy is defined as the state of being literate, while literate means educated (TDK, 2022).

Literacy according to United Nations Educational, Scientific and Cultural Organization's (UNESCO) definition; It is the ability to identify, understand, interpret, create, communicate and calculate different types and varying types of printed and written materials (UNESCO, 2022). UNESCO also emphasizes that literacy is a continuum of learning that includes the integration and adaptation of individuals into their communities and wider societies by enabling them to develop their potential and knowledge.

Thus, it is emphasized that these people who fully participate in the society will contribute to the acquisition of the right to education and the improvement of their livelihoods. For these reasons, UNESCO (2022) includes global literacy among its targets. In order to emphasize and remind the

public of the importance of literacy, 8 September has been celebrated as the literacy day every year since 1966.

The contribution of literacy to social development and the adaptation of the individual to society over time has caused this concept to go beyond literacy skills. Because the advances and deepening in scientific fields, technology and art in the century we live in have brought with the need to understand, make sense of, and adapt these fields to life.

Thus, concepts such as digital literacy, mathematical literacy, financial literacy, cinema literacy, media literacy, and statistical literacy have emerged.

In this sense, literacy can be defined as having and understanding basic information about an information or problem encountered in daily life, being able to think, comment and criticize this information.

From this perspective, it is clear that literacy will be a driving force for development and will guide societies in issues such as deep poverty, access to health and food, and sustainability,

security, discrimination, media (digital media) which are among the biggest problems of today. Because, literate individuals will be able to comment on the problems they face in the light of their basic knowledge, create solution methods for these problems, and criticize the existing system. The increase of these individuals will reveal social awareness and the power to put the necessary pressure on the decision-makers in this regard.

This variety in the field of literacy and the social driving force it can create has revealed the necessity of including these types of literacy in formal education over time. As a matter of fact, in the report published by the Board of Education and Discipline in 2017 with the title of 21st century skills, has health, environment, financial, media etc. literacy types.

In the same report, the item “acquiring literacy skills” is also included under the heading of competences and skills aimed at gaining students. It has been noticed all over the world as well as in Turkey that literacy skills are criteria that show the development levels of societies.

With this motivation, financial literacy is addressed as a government policy in Australia and New Zealand (Taylor and Wagland, 2011). Similarly, media education is provided for the development of media literacy in Russia and academic studies are carried out to develop this education (Yoon, 2009). In this regard, literacy studies related to the subjects in school curricula have started to increase in the literature. Mathematics also stands out among these subjects. Academic research has also been done on the connections between mathematical literacy and scientific literacy, in addition to studies on the relevance of the mathematics we learn in school to our daily life (Ojose, 2011; Wilkins et.al.,2002). Schields (2005), on the other hand, focused on the relationship between information and data literacy and statistical literacy. While emphasizing the importance of information and data literacy by the relevant researchers in the same study, Schields stated that statistical literacy is necessary for students to make sense of these literacy types by bringing them to the level of critical thinking.

Gal (2004) emphasized that the way to be an informed, questioning citizen and employee is through statistical literacy. Thus, he added that people's ability to make effective decisions in solving their daily problems will develop. In his other work (Gal, 2002), he described statistical literacy as the ability to interpret, critically evaluate and communicate statistical outputs. Engel (2017), on the other hand, emphasized that it is indispensable in the construction of an active citizenship process. In the same study, Engel underlines that it is one of the most important duties of statistics educators to prevent even the false presentation of evidence of empirical data in public, political and social fields. It is stated in the same study that directing public debates from emotional or subjective areas to areas based on evidence and science and making political decisions on these facts will be possible with statistical literacy.

It is clear that learning all types of literacy is necessary for social progress, for moving discussions to a scientific level,

and for speaking objective facts in societies. Statistical literacy, on the other hand, is essential for dealing with, understanding, and considering and criticizing multivariate phenomena together.

In this study, statistical literacy has been specifically addressed with the perspective that all kinds of literacy are an important element of personal and social development. The aim is to introduce the concept of statistical literacy and to draw attention to this area. In the study, which examines the relationship between university students' statistical literacy levels and socio-demographic characteristics, their field of education and universities, it is also examined whether there is a correlation between the students' literacy level and their general grade points averages (GPA), parental education levels, and their knowledge and confidence levels in official statistics.

2. MATERIAL AND METHOD

With the above-mentioned purpose, a questionnaire, including the statistical literacy scale (Şahin, 2012) indexed by TOAD, was prepared. Necessary ethics committee approvals were obtained for the application of the questionnaire and the Google Forms application was used during the implementation of the survey.

Assuming 95% significance level and $p=0.05$, the sample size 398 was obtained by using the sample calculation equation developed by Yamane (1967) according to the population.

All of the individuals participating in the survey are university students. 398 students from different schools such as Çukurova, Mersin, METU, Çağ, Başkent and Boğaziçi universities participated in the survey.

The analysis in Table 2 was carried out by calculating the total score of the Statistical Literacy Scale. In all analyzes, it was concluded that the skewness and kurtosis values of the variables were between -1.5 and +1.5 and the normality assumption was provided by examining the histogram graphs (Tabachnick & Fidell, 2013). Due to the assumption of normality, ANOVA and independent groups t-test were used in the analysis, and the mean and standard deviations of the variables were given with “Mean±S.S.” notation. As a result of the reliability analysis performed for the Statistical Literacy Scale, the Cronbach Alpha value was found to be 0.71. The scale is reliable. (Carmines and Zeller, 1979).

3. EMPIRICAL ANALYSIS

When the participants are examined, it is seen that there are differences both in the universities and in the departments they study. This allows us to see the statistical literacy trends of different departments. For example, in addition to life sciences such as Computer Engineering, Environmental Engineering, Industrial Engineering, Statistics, Mathematics, Chemistry; students studying in social sciences such as Law, Journalism, Public Relations and Advertising, English Language Teaching participated in the survey. When examined percentages, students studying in life sciences are 65.3% of the sample, and students studying

in social sciences constitute 34.7% of the sample. 51% of the study consisted of females and 49% of males. When asked how many years the students have been studying, 16.1% said it was for one year, 17.8% for two years, 27.1% for three years, 24.9% for four years, 14.1% for five years or a longer period of time.

When the education of the parents of the students participating in the study was examined, it was determined that 23.6% of the parents had primary school or lower, 18.1% had secondary school, 33.7% had high school and 24.6% had undergraduate or higher education.

3.1 Literacy Scale

The answers given by the participants to the Statistical Literacy Scale are given in Table 1.

Table 1. Statistical Literacy Scale Questions' Frequency

Questions of Literacy Scale	Incorrect Answers Frequency (Percentage)	Correct Answers Frequency (Percentage)
Question 1	208 (52,3)	190 (47,7)
Question 2	162 (40,7)	236 (59,3)
Question 3	153 (38,4)	245 (61,6)
Question 4	101 (25,4)	297 (74,6)
Question 5	92 (23,1)	306 (76,9)
Question 6	200 (50,3)	198 (49,7)
Question 7	298 (74,9)	100 (25,1)
Question 8	294 (73,9)	104 (26,1)
Question 9	298 (74,9)	100 (25,1)
Question 10	240 (60,3)	158 (39,7)
Question 11	272 (68,3)	126 (31,7)

It was determined that more than half of the students gave incorrect answers to the 1st, 6th, 7th, 8th, 9th, 10th and 11th questions. The questions that most people answered incorrectly were the 7th and 9th questions. More than half of the students had correct answers to the 2nd, 3rd, 4th and 5th questions. The question that most people answered correctly is the 5th question.

In the literacy scale, it is not aimed to measure the level of knowledge, but to measure the cognitive interpretation capacities of individuals and whether they can make these cognitive interpretations critically (Şahin, 2012). In the study (Şahin, 2012), in which the scale was defined, the categories of interpretation or critical interpretation of each question were also defined. What is meant by critical interpretation here is to express the ability to question the data that provide the statistical results, the interpretations made regarding these results, and the arguments based on them.

In Question 7, one of the questions in which the most misinterpretation were recorded, the measurement results of the weight of an object were given by 9 students, and they were asked to comment on the actual weight of this object. The background of this question is related to population, sample means, and outlier. The question aims to measure

both the cognitive interpretation and the critical level of this interpretation that is, the connection between interpretations and statistical information (Şahin, 2012).

Similarly, the 9th question, in which the most misinterpretation were made by the participants, includes the concepts of measures of central tendency in general and concepts such as median, outlier, extreme value or maximum value in particular. In this question, "The median price of a meal in a student cafeteria is \$30. Therefore, she states that most of the students in this sample spend approximately \$30 on food each week." proposition is discussed. It is a noticeable point that this question is one of the most incorrectly answered question whose category is critical interpretation. Because the accompaniment of critical thinking with knowledge and thought is the most important parameter that makes literacy meaningful.

Again, question 8, which 73,9% of the students answered incorrectly, is a question that includes the concept of conditional probability in its background and requires critical interpretation (Şahin, 2012). In this question, the students were asked to comment on probability based on information about the musicians at a flute concert.

The question with the highest number of correct answers is as follows: A decision will be made to give up one of the equipment in a sports center. While making this decision, it is desired to be made by paying attention to the interests and experiences of the athletes. The learning outcome of this question was determined as interpretation in Şahin's (2012) study.

3.2 Relations

Independent groups t-test and ANOVA were used to determine whether there was a difference between the variable groups according to the scale mean. The results can be seen in Table 2.

Table 2. Relations to Statistical Literacy

Variables	Literacy (Mean±S.S.)	Statistics
Gender		t = -1,23 p = 0,22
Female	5,02 ± 2,49	
Male	5,34 ± 2,65	
Departments		t = 4,72 p<0,001
Life Sciences	5,61 ± 2,50	
Social Sciences	4,36 ± 2,51	
Educational Status of Parents		F = 0,55 p = 0,65
Primary school or lower education level	5,14 ± 2,43	
Middle School	5,49 ± 2,50	
High school	5,01 ± 2,67	
Bachelor or Higher	5,21 ± 2,63	
Trust in Official Statistics		F = 3,36 p = 0,04
Yes*	4,63 ± 2,75	
No	5,10 ± 2,53	
Occasionally*	5,48 ± 2,48	
GPA	Pearson Correlation = 0,10	p = 0,04

There was no significant difference between gender groups according to the Statistical Literacy scale ($p=0,22$).

There is a significant difference between individuals studying in life and social sciences in terms of statistical literacy ($p<0,001$). Statistical literacy averages of individuals studying in life sciences are higher than individuals studying in social sciences.

There is no significant difference between the educational status of individuals' parents in terms of statistical literacy averages ($p=0,65$). There is a weak positive correlation between the GPA of students and their statistical literacy scores ($p=0,04$).

There is a significant difference between people's confidence in official statistics ($p=0,04$). Accordingly, as a result of the Post-Hoc test, a significant difference was found between individuals who had confidence and those who sometimes ($p=0,04$). Statistical literacy averages of confident individuals are lower than individuals who say they have confidence occasionally. Those with the lowest average are the group that says having confidence on official statistics.

From this section onwards, models were created using regression analysis. First, a regression model was created by accepting the statistical literacy level as the dependent variable. Since it was noticed that the results obtained in the analysis on trust in official statistics were also worth examining, the second regression model, which accepted the official statistics as the dependent variable, was added to the study.

3.3 Regression Analysis for Statistical Literacy

A multiple linear regression model was created in which statistical literacy score was chosen as the dependent variable, and gender, department status in life/social sciences and GPA were defined as independent variables. All assumptions necessary to build the model are provided. Analysis results are given in Table 3.

Table 3. Statistical Literacy's Regression Analysis

Regression					
	Unstd Coeff		Std Coeff	t	Sig.
	B	Std. Error	Beta		
Const.	3,22	0,65		4,99	$p<0,001$
Female	-0,25	0,25	-0,05	-0,98	0,33
Life Sciences	1,19	0,26	0,22	4,50	$p<0,001$
GPA	0,47	0,22	0,11	2,15	0,03
R=0,25 R ² =0,06	F = 8,76 p<0,001				

Then,

$$\text{Statistical Literacy} = 3,22 - 0,25 * \text{Female} + 1,19 * \text{Life Sciences} + 0,47 * \text{GPA} \quad (1)$$

The model in which statistical literacy score is chosen as the dependent variable, and gender, reading status in life/social sciences and grade points average is taken as the independent variable is significant ($p<0,001$). Independent variables have a 6% contribution to statistical literacy.

In our model, reading in life/social sciences makes a significant contribution to the model ($p<0,001$). Statistical literacy scores of individuals studying in life sciences are 1,19 points higher than individuals studying in social sciences.

GPA predicts statistical literacy positively ($p=0,03$). A one-unit change in the GPA causes an increase of 0,47 points in the statistical literacy score.

3.4 Logistic Regression Analysis for Official Statistics

The analysis on trust in official statistics which accepted as the dependent variable, and gender, department status in life/social sciences and literacy were defined as independent variables. All assumptions necessary to build the model are provided (Table 4).

Table 4. Official Statistics’ Logistic Regression Analysis

Parameter Estimates							
Confidence in Official Statistics		B	Std. Error	Wald	df	Sig.	ExpB
Sometimes	Intercept	0,65	0,38	2,96	1	0,09	
	Literacy	0,15	0,05	8,03	1	0,005	1,17
	Life Sciences	-0,58	0,29	3,90	1	0,05	0,56
	Social Sciences	0			0		
	Male	-0,44	0,28	2,46	1	0,12	0,65
	Female	0			0		
No	Intercept	1,13	0,39	8,31	1	0,004	
	Literacy	,076	0,058	1,720	1	0,19	1,08
	Life Sciences	-0,49	0,32	2,33	1	0,13	0,61
	Social Sciences	0			0		
	Male	-1,57	0,30	26,85	1	p<0,001	0,21
	Female	0			0		
a. The reference category is: Yes.							
Nagelkerke = 0,12							
Pearson chi-square = 87,573 p = 0,22							

$$\text{Logit}(y_1) = 0.65 + 0.15 * \text{Statistical Literacy} - 0.58 * \text{Life Sciences} - 0.437 * \text{Male} \quad (2)$$

$$\text{Logit}(y_2) = 1.13 + 0.08 * \text{Statistical Literacy} - 0.49 * \text{Life Sciences} - 1.57 * \text{Male} \quad (3)$$

where y_1 denotes “sometimes” and y_2 denotes “no”.

Based on the Pearson chi-square statistic model fits the data well ($p=0.22>0.05$).

One unit increase in total score of scale would be expected to increase relative risk for preferring “Sometimes” to “Yes” by a factor of 1.17. It means that if all variables in the model are held constant, the subject with higher scale score is more likely to select “Sometimes” over “Yes” with the lower scale score.

In other words, as participants' scores on the statistical literacy scale increase, they are more likely to answer "sometimes" to trusting official statistics than "yes". Participants with a high literacy score do not unconditionally trust official statistics and answer the question about trusting the official statistics as "in some cases they trust".

Students in life sciences are more likely than students in social sciences to select “Yes” over “Sometimes”. In other words, individuals studying in life sciences have more confidence in official statistics than those studying in social sciences.

When examined in terms of gender, males are more likely than females to prefer “Yes” over “No”.

4. RESULTS AND DISCUSSION

The Statistical Literacy Scale developed by Şahin (2012) was used in this study to measure the level at which individuals understand and interpret statistical information. The ability of students to interpret and apply statistical information was evaluated by applying a questionnaire to 398 students studying at different universities and departments. In addition to the literacy scale questions, the participants were also asked to answer questions such as gender, family education level, reading status in the life/social sciences, GPA, how long have they been at university, knowledge and confidence in official statistics. It was determined that 51% of the individuals participating in the survey were female and 49% were male, and the majority of the participants' parents were high school graduates.

If we look at the answers given for the scale questions one by one, we see that most incorrectly answered questions are 7th and 9th ones. According to the scale averages, variables of gender, family education, GPA, being in life/social science, knowledge and confidence in official statistics were analyzed by independent groups t-test and ANOVA analysis. Accordingly, no difference was found between the groups in terms of genders and parental education levels, according to statistical literacy averages. However, it has been determined that the statistical literacy averages of the students studying in the life sciences are higher than the students studying in the social sciences. This reveals the necessity of taking measures to increase the statistical literacy levels of the students of the social sciences.

A positive correlation was observed in the correlation between the scale and the GPA. The higher the GPA, the higher the literacy level. When the knowledge and confidence groups on official statistics are compared according to the statistical literacy averages, it is seen that the literacy averages of the individuals who "sometimes" trust are higher than those of the individuals who "trust". This shows that the increase in literacy scores actually increases the ability of individuals to interpret, think and question critically.

In addition, regression model was constructed by gender, being in life/social science and GPA variables and their contribution to statistical literacy was examined. According to this analysis, it has been determined that the statistical literacy level of students studying in life sciences is 1,19 points higher than students studying in social sciences. Again, with the same analysis, it was determined that a one-point increase in GPA caused a 0,47-point increase in statistical literacy.

In addition to all these analyses, further analyzes regarding the trust of the participants in official statistics show that as their statistical interpretation skills increase, they give the answer "sometimes" rather than "yes" for trusting the official statistics. In other words, it has been observed that as the statistical literacy of individuals' increases, their confidence

in official statistics decreases. On the other hand, it is also reflected in the results that males have more confidence in official statistics than females.

As a result of the analysis, it was determined that the statistical literacy levels of the individuals were affected by the students' grade point average and the department they studied. This situation has an impact on the issue of trust in official statistics. With the study, it was aimed to reveal an awareness by associating the statistical literacy levels of the students from different universities and departments with various variables. Thus, regardless of the department or university, it is aimed to contribute to personal and social development by determining the statistical thinking and interpretation abilities of the students.

In order to increase statistical literacy, compulsory courses on reading data, making sense of histograms/graphs, questioning and interpreting statistical results should be added to the curriculum starting from kindergarten age. Currently, due to the belief in the driving force that social enlightenment will bring, the Ministry of National Education has published a report by showing the necessary sensitivity to the issue since 2017.

In addition, the development of the society in this area can be achieved through public service announcements or entertaining cartoons for children. In addition, it is essential that all non-governmental organizations that desire social development, that dream of a more enlightened society, where individuals can question what is presented to them, and which can create the necessary pressure on decision-makers, should take a role in out-of-school education on statistical literacy. In short, anyone who aims to divert public debate from emotional or subjective areas to evidence and science-based areas should take a role in the steps to be taken on statistical literacy. Because political decisions will be based on objective facts instead of subjective areas, and social progress will be possible with statistical literacy.

REFERENCES

- Board of Education (2017). Retrieved from: https://ttkb.meb.gov.tr/meb_iys_dosyalar/2017_07/18160003_basin_aciklamasi-program.pdf.
- Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment*. Sage publications.
- Engel, J. (2017). Statistical literacy for active citizenship: A call for data science education. *Statistics Education Research Journal*, 16(1), 44-49. <https://doi.org/10.52041/serj.v16i1.213>.
- Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International statistical review*, 70(1), 1-25. <https://doi.org/10.1111/j.1751-5823.2002.tb00336.x>.
- Gal, I. (2004). Statistical literacy. In *The challenge of developing statistical literacy, reasoning and thinking* (pp. 47-78). Springer, Dordrecht. https://doi.org/10.1007/1-4020-2278-6_3.
- Ministry of National Education (2017). Retrieved from: https://ttkb.meb.gov.tr/meb_iys_dosyalar/2017_07/18160003_basin_aciklamasi-program.pdf.
- Ojose, B. (2011). Mathematics literacy: Are we able to put the mathematics we learn into everyday use. *Journal of mathematics education*, 4(1), 89-100. <https://doi.org/10.12691/education-7-10-1>.
- Shields, M. (2005). Information literacy, statistical literacy, data literacy. *IASSIST quarterly*, 28(2-3), 6-6. Retrieved from: https://iassistquarterly.com/public/pdfs/iqvol28_2_3shields.pdf.
- Şahin, F. (2012). A study for development of statistical literacy scale for undergraduate students (Doctoral dissertation, Boğaziçi University. Retrieved from: <https://toad.halileksi.net/sites/default/files/pdf/lisansogrencileri-icin-istatistiksel-okuryazarlik-olcegi-toad.pdf>.
- Tabachnick, B.G., Fidell, L.S. (2013). *Using Multivariate Statistics* (sixth ed.) Pearson, Boston.
- Taylor, S., and Wagland, S. (2011). Financial literacy: A review of government policy and initiatives. *Australasian Accounting, Business and Finance Journal*, 5(2), 101-125. Retrieved from: <https://ro.uow.edu.au/aabfj/vol5/iss2/7/>.
- TDK, Turkish Language Association (2022). Retrieved from: <https://sozluk.gov.tr/>.
- Turkish Surveying Index (TOAD) (2022). Retrieved from: <https://toad.halileksi.net/>.
- UNESCO, United Nations Educational, Scientific and Cultural Organization (2022). Retrieved from: <http://uis.unesco.org/en/glossary-term/literacy>.
- Wilkins, J. L., Zembylas, M., and Travers, K. J. (2002). Investigating correlates of mathematics and science literacy in the final year of secondary school. In *Secondary analysis of the TIMSS data* (pp. 291-316). Springer, Dordrecht. https://doi.org/10.1007/0-306-47642-8_18.
- Yoon, J. (2009). The Development of Media Literacy in Russia: Efforts from Inside and Outside the Country. *Issues in Information and Media Literacy*, 1, 189-213.