



Testing data literacy: Numeracy, statistical literacy, and sociological knowledge among journalists in Estonia and Turkey

Ragne Kõuts-Klemm ^{1*}

 0000-0001-9277-6700

Ayşen Yalman ²

 0000-0003-2280-02824

Mehmet Arif Arık ²

 0000-0001-9118-3022

Liis Auväärt ¹

 0000-0002-9786-1696

¹ Institute of Social Studies, University of Tartu, Tartu, ESTONIA

² Department of Journalism, Akdeniz University, Antalya, TURKEY

* Corresponding author: ragne.kouts@ut.ee

Citation: Kõuts-Klemm, R., Yalman, A., Arif Arık, M., & Auväärt, L. (2026). Testing data literacy: Numeracy, statistical literacy, and sociological knowledge among journalists in Estonia and Turkey. *Online Journal of Communication and Media Technologies*, 16(2), Article e202631. <https://doi.org/10.30935/ojcm/18574>

ARTICLE INFO

Received: 27 Nov 2025

Accepted: 10 Apr 2026

ABSTRACT

International comparative studies show that adults' data literacy is considerably lower than that of younger generations. Many older adults do not engage in continuing education and instead depend on journalists, functioning as knowledge brokers, to interpret and contextualize data. The news media have traditionally carried both an informative and educational function for their audiences. This article examines journalists' own data literacy, given their potential to shape the public's capacity to navigate data-rich environments. The study innovates by distinguishing data literacy as comprising numeracy, statistical literacy, and sociological data knowledge, and tests data literacy across various journalistic contexts. Drawing on a data literacy test conducted among journalists in Estonia and Turkey (respectively 10 and 20 interviews), the study identifies notable gaps in these three data literacy components. Whereas general math skills of journalists align with the numeracy level of the general population, knowledge of statistics and sociological data depends more on the particular educational contexts. The study reveals the necessity of additional data training for journalists to respond to the challenges facing in datafying societies. The study argues for targeted educational initiatives to strengthen data literacy among journalists and the wider population.

Keywords: data literacy, journalism, Estonia, Turkey

INTRODUCTION

As society becomes increasingly datafied, people outside the formal education system also need opportunities to develop their data literacy. Availability of data for daily practical decisions challenges them to find alternative avenues to formal education to acquire such skills and act competently (Karger et al., 2024). Mass media has historically served not only an informative but also an educational function (McQuail, 2006; Oyebamiji, 2012). Yet studies suggest that journalists often rely on a narrow range of data practices, lack critical engagement, and frequently reproduce the narratives of data holders. Assessing journalists' own data literacy is thus essential to understanding their capacity to support the wider public. The objective of this study is to find extent of journalists' data literacy and to identify which aspects require the most development.

A recent global survey shows that data skills among journalists are mainly self-taught: seven in 10 acquired data journalism skills without formal training (Bisiani et al., 2025). These findings highlight the need to better understand journalists' current competencies and gaps and point to opportunities for academia to support the field through advanced training programs and targeted curricula. As reported in the aforementioned survey, only a quarter of the data journalists reported having obtained any formal education related to data, online or offline (DataJournalism.com, 2021).

The study at hand focuses on data literacy, defined as a combination of numerical, statistical, and technical capacities (Gray et al., 2018). We broaden the definition of data literacy, and stress also their "basic ability of knowledge construction and communication to collect, analyze, and use data and process it as information for problem-solving" (Kim & Kim, 2021, p. 700). For journalists, these skills are essential for producing factual and objective reporting in a data-driven society, where governance and social processes are increasingly quantified and publicly available data continues to grow. This process of "datafication"—the quantification of aspects of life previously experienced in qualitative, non-numeric forms (Engebretsen et al., 2018, p. 3)—positions journalists as "knowledge brokers" (Pentzold et al., 2021). Given that public data literacy levels remain uneven—although data journalists themselves strongly feel that the COVID-19 pandemic increased audience data literacy as suggested by the 2021 survey (DataJournalism.com, 2021)—the public relies on journalists to interpret and use large volumes of data produced by governments and other data owners (Frank et al., 2016). Journalists working with data can encounter issues similar to, for example, an enterprise data analyst (Kasica et al., 2020).

Research on the skills required for data-oriented journalism is growing (Bhaskaran et al., 2024; Burns & Matthews, 2018; Heravi, 2019; Lewis et al., 2020). Most studies examine self-reported skills, categorizing data competencies in various ways (Appelgren, 2017; Haim, 2024; Larrondo et al., 2016; Örnebring & Mellado, 2018). One example divides data-related work into seven categories: journalism, data analysis, data visualization, statistics, data wrangling, scraping, and machine learning (Bisiani et al., 2025, p. 509). Some approaches classify skills according to the workflow, while others examine sub-skills individually (Kim & Kim, 2021), or investigate the inherent structure of skills (Fischer, 1980; Sharma, 2017). Individuals can often assess their mathematical abilities with reasonable accuracy (Fagerlin et al., 2007), but only a few studies have attempted to measure journalists' data literacy beyond self-assessment (Köuts-Klemm, 2019).

We test data literacy as a combination of numeracy, statistical literacy, and sociological data knowledge across different journalistic contexts. Comparative research in this area is scarce (Bisiani et al., 2025). Our study addresses this gap by testing data literacy among journalists in Turkey and Estonia, countries familiar to the authors, allowing in-depth contextual interpretation of differences in skill components. The most notable difference between the two countries is their level of datafication (Open Data Maturity Report, 2023). Estonia has a global reputation as a "digital state," focusing on a national-level, open-data ecosystem, primarily due to improvements in the open government data portal usability and legislation amendments (Rajamäe & Nikiforova, 2024), as well as adult competencies in literacy, numeracy, and information processing (OECD, 2016a).

Participants from both countries completed a data literacy questionnaire developed by Köuts-Klemm (2019), which combined the questions with "think-aloud" commentaries from the respondents. This approach provides not only information on journalists' ability to use and interpret numbers, but also insight into their reasoning when working with different types of data. In terms of research methods used to study data journalism, this approach is an attempt to address the higher prevalence of qualitative methods and the lack of quantitative approaches, such as surveys and questionnaires (Heravi & Lorenz, 2020). Therefore, we propose the following research questions (RQs):

1. **RQ1.** Are the main components of journalistic data literacy (numeracy, statistical literacy, and sociological data knowledge) apparent among Estonian/Turkish journalists?
2. **RQ2.** Do any notable differences emerge concerning journalistic data literacy components in Estonia/Turkey? If so, what could be the reasoning behind these differences?

We start this article by providing a systematic overview of

- (1) what kind of data journalists are dealing with and need to deal with,

- (2) data literacy components, and
- (3) additional context concerning datafication and data journalism in Estonia and Turkey.

LITERATURE REVIEW

Data Commonly Used in Data Journalistic Projects

Data journalism can be substantially different according to the data journalists who are using it (Ramsälv et al., 2024) and the human resources available for the data-rich project at hand. De Maeyer et al. (2015, pp. 410-411) differentiate between ordinary data journalism that “is manageable by one individual [and] can be done on a daily basis” and thorough data journalism that is produced by teams with “a range of skills.” More complex forms of journalism, where diverse teams produce investigative data-intensive long-reads, can be viewed as an independent form of data journalism.

From an audience perspective, a simple usage of numbers and infographics might be understood as data journalism—like results of an online poll conducted by a media outlet, a summary of a sociological study on the population’s attitudes, or presentation of statistics in interactive info-graphics (Borges-Rey, 2016; Cushion et al., 2017; Steensberg, 2023). Research on data journalistic projects has revealed that such projects tend to cover the domains where data is routinely available. Common themes include politics, societal issues such as census results and crime reports, health and science issues, business, and the economy (Cushion et al., 2016; Loosen et al., 2020).

Very often, the data used for journalistic work is pre-processed and interpreted by third parties, that data being publicly available or available upon request (Auväärt, 2023; Beiler et al., 2020; Cengiz & Kaya, 2026; Loosen et al., 2015, 2020; Parasie & Dagiral, 2013; Stalph, 2018; Van Witsen, 2020). In this case, a journalist is not an autonomous interpreter of data. On the other hand, the availability of non-pre-interpreted data has also been growing as a result of digitalization and wider use of Big Data in economics, governance, and other areas (Baack, 2015). If a journalist is able to analyze data for him/herself, this offers the possibility to make one’s own conclusions about the data and elicit facts that might go missing or not be emphasized by another party. It gives the journalist more control over verifying facts (Appelgren & Nygren, 2014). In such instances, data literacy becomes increasingly relevant.

Few studies attempt to test or measure data skills. “Despite the extensive scholarship about how journalists handle statistics, little is known about what they understand of the measured facts they report on—e.g., gross domestic product, inflation, crime rates and so forth” (Van Witsen, 2020). To address this gap in research, the authors of this article propose that due to the variety of data (Kitchin, 2014), a more diverse understanding of skills needed for data journalism is due.

Data Literacy Components

The term *data* is most often viewed as self-explanatory in academic papers (Auväärt & Kõuts-Klemm, 2025), yet following the classifications in the social sciences one can describe data in numerous categories such as form (qualitative or quantitative), structure (structured, semi-structured or unstructured), source (captured, derived, exhaustive or transient), producer (primary, secondary, or tertiary) and type (indexical, attribute or metadata) (Kitchin, 2014, p. 4). For journalistic purposes, it is a far too detailed classification. For the journalist, the main value of data is in the information it represents. However, to work with different data, knowledge of differences in specific data types would be useful to understand their scope and limits for usage in news stories.

From a journalist’s viewpoint, data is used in at least three functions in reporting. There are stories that

- (1) are enriched by data,
- (2) use data to investigate, and
- (3) explain data (Rogers et al., 2017).

As a result, the term *data literacy* can be viewed as a combination of numerical, statistical, and technical capacities (Gray et al., 2018) or, to expand, the “basic ability of knowledge construction and communication to collect, analyze, and use data and process it as information for problem-solving” (Kim & Kim, 2021, p. 700).

To “be fluent” in data literacy requires the mastery of various competencies, such as mathematical knowledge and reasoning, but also more advanced skills such as critical thinking and problem solving (Cui et al., 2023).

We propose to combine data science, sociological, and journalism approaches to data for researching journalistic data literacy. The three components we distinguish in this study while researching journalistic data skills are numeracy, statistical literacy, and sociological data knowledge. For all three, more basic and more advanced appearances can be seen. In brief, if data is discussed specifically as numbers, the term *numeracy* can be used to describe the ability to interpret and communicate mathematical information and ideas, whereas the term *literacy* indicates the ability to evaluate and engage with written texts (OECD, 2016a).

As *basic numeracy* is being able to deal competently with numbers, tables, and graphs (Parsons & Bynner, 2005), it is taught in schools all over the world. Math skills are needed for every citizen to be able to make decisions in the modern society. However, research has shown that numeracy is complicated for a big share of the population, and even the knowledge of the simplest forms of numerical reasoning cannot always be taken for granted (Hill & Brase, 2012; Peters et al., 2006). On an advanced level, numeracy includes basic logic, performing multi-step operations, a fundamental understanding of chance and basic statistical principles, and comfort with proportions, fractions, probabilities, and risks (Dieckmann et al., 2009)—knowledge that is not accessible to most of the population. However, research also shows that with training, it is possible to improve numeracy abilities (Howard et al., 2016).

Statistical literacy is different from numeracy (Wolff et al., 2016), referring to the use of statistics as evidence in arguments, and the core element here is assembly—specifically, how the statistics are defined, selected, and presented, making context a crucial part of the term (Schild, 2005). Statistical literacy goes above numeracy while stressing the relations, and for calculations, detectable patterns between numbers and variables (Wolff et al., 2016). Basic statistics is a part of Western school programs, too (Frank et al., 2016). Everyone needs to understand at least basic risk estimations to be able to make decisions that impact on their well-being or health, e.g., to follow the treatment instructions by medical personnel, if needed. The advanced levels of statistical literacy, like risk estimations, cumulative probabilities, or Bayesian reasoning, are complicated even for highly educated groups (Lipkus et al., 2001). Journalists need to possess at least basic statistical literacy.

Furthermore, we claim that it is relevant to pay special attention to the *sociological data*, where data is extracted from and constructed in social reality, and according to the different theoretical views, can have different truth values. The positivist approach in sociology, which asserts the possibility of extracting data from social reality similarly to natural sciences with high truth value (see criticism of this viewpoint, e.g., Hindess, 1973), was thought to be history. But in the era of Big Data, sociology’s positivist approach has experienced a rebirth (Diaz-Bone et al., 2020). The constructivist approach is still prevalent today. It posits that the data used is always a theoretical choice. Sociological data, then, is always arguable. Diversity of social reality and social actions is hardly reducible to few high level generalizations, representing different groups and world views (Larsson, 2009). Data literacy necessarily includes recognition of what social context plays a relevant role in data interpretations.

As data literacy components, the different levels of numeracy, statistical literacy, and sociological data knowledge can be analytically distinguished, as we outline in the analysis scheme (**Figure 1**). For all three components, more basic and more advanced appearances can be seen. The basic numeracy and statistics are covered in Western school systems as “school-math” in the first steps of formal education, and the advanced numeracy or statistics are taught in high schools and universities. Sociological knowledge became a part of journalism undergraduate and master’s training programs during the previous century (Weaver & McCombs, 1980). For all three components, one can only understand them or move towards using them like journalists ought to do for data-driven stories.

In practice, it can be complicated to separate the levels of numeracy, statistical literacy, and sociological data knowledge an individual possesses because they may be intertwined. The three components must be analytically separated. **Figure 1** outlines the general scheme for empirical work.

Before moving on to testing these data literacy components among journalists in Estonia and Turkey, the following section gives a brief overview of country contexts and the differences in journalists’ working conditions.

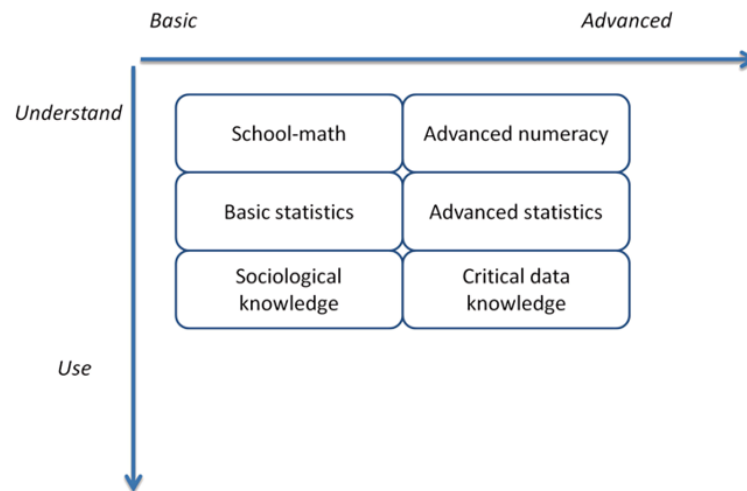


Figure 1. General scheme of data literacy components (the authors' own drawing)

Country Contexts: Datafication and Data Journalism in Estonia and Turkey

Estonia and Turkey represent two distinct types of media systems. Estonia can be considered the most Westernized media model among Central and Eastern European countries (Greskovits, 2015; Mancini, 2015). It features high media freedom, significant commercialization, and an almost oligopolistic media market (Loit & Harro-Loit, 2017). However, public service media remains strong (Jõesaar & Kõuts-Klemm, 2019), similar to the media systems found in Nordic countries.

In contrast, the Turkish media system has a media structure based on freedom of communication rather than media freedom, acting in line with the state-centered policy of the country and focusing on news on local government and business circles (Gencil Bek, 2014). As a result of the development of cross-media ownership in the Turkish media, the sector has come under the control of a small number of media holdings (Kuyucu, 2013). Although it is possible to find examples of qualified, critical, and research-based alternative journalism in Turkey, it can be observed that the mainstream media in general reproduces the existing political culture rather than being critical. For instance, Cengiz and Kaya (2026) argue that data journalism, which requires in-depth analysis, long-term research, and specialization, struggles to thrive in Turkey due to the fast-paced news cycle and political pressures. It is observed that the number of investigative data journalism practices in daily news in Turkey is limited. The Turkish news media is reported to lag behind in its use of data-driven reporting and visualization compared to newsrooms in the United States, Europe and Asia (Smith, 2018).

Digital systems and services are advancing in all societies. Estonia is widely recognized for its availability of digital services in the public and private sectors and its high level of digitalization in various aspects of life (Himma-Kadakas & Kõuts-Klemm, 2023). According to the Open Data Maturity Report (2023), Estonia is one of the trend-setters in open data availability, usage, and policy. Turkish context is different. The last international assessment of Turkey occurred in 2016 and 2017. The Global Open Data Index (2024) ranked it 34th of 94 total countries regarding access to data. The audiences in both countries are differently prepared to navigate the data-rich environment. When we compare the literacy and numeracy rates in each country, we see that Estonia performs above the average of OECD member states, while Turkey significantly underperforms (Kavuncu & Polat, 2019; OECD, 2016a, 2024).

We hypothesize that journalism in Estonia, while acting in more datafied context compared to Turkey, creates in its daily reporting more contacts with data for audiences. It makes investigating the differences in journalists' skills and their preparedness for the ongoing datafication of their societies a challenging task.

METHOD OF STUDY

To test the three components of data literacy, a standardized data literacy questionnaire developed by Kõuts-Klemm (2019) was used in both Estonia and Turkey. The original questionnaire was developed based on three main sources: tasks from math schoolbooks, tasks used in PIAAC adults' numeracy testing, and

framing analyses developed by psychologists Kahneman and Tversky (1984) (Kõuts-Klemm, 2019). The questionnaire includes only tasks that are used and validated in various learning and testing contexts. Whereas previous testing experiences indicate, a few tasks already provide us with insights into participants' numeracy and statistical skills (Cokely et al., 2012). Additionally, to grasp the specifics of the journalism profession, a few tasks related to journalism have been included in the questionnaire. The questionnaire's validation is based on two expert groups. Firstly, scholars in the field of mathematics and statistics, who have long experience in numeracy and statistical literacy teaching, supported the decision on the optimal number of tasks. Secondly, feedback from professional journalists helped to improve the wording of the questions related to journalism practice.

For the current study, participants answered the questionnaire in their respective native languages, the forward and backward translation was used to grant the reliability of the translations (Tsang et al., 2017). (see example of questions in English in [Appendix A](#)). The questionnaire consisted of 10 questions that assessed different components of data literacy. The numeracy category included questions related to "school-math" tasks, statistics, logical and probability with and without framing, sociological knowledge, and the presentation of research data (offered in journalistic and sociological formats). The questionnaire was multiple choice to minimize advantages and/or disadvantages to each participant that could be caused by stress and time-constraint pressures due to queries that required manual calculations. Only two questions were open-ended. The aim of adding them was to collect information about "reading" the data: Do participants use social-scientific or professional journalistic argumentation in explaining their choices, or make other claims about the presented information?

Participants were selected using purposive sampling. The inclusion criteria required journalists to have published at least one news article containing numerical data or statistics within the two months prior to the data collection. This criterion ensured relevance to data-related journalism practices. We collected data from 10 journalists in Estonia and 20 in Turkey to assess their data skills. All participants signed consent documents that informed them of the relevant aspects of the study. The data collection and processing were carried out in compliance with the general data protection regulation. The study followed the European code of conduct for research integrity (revised edition 2023).

Participants represent various newsrooms and cover different topics. The Estonian group comprised urban and rural journalists. All journalists in the Turkish group worked for media based in major cities. Most journalists in both groups held bachelor's or master's degrees. The gender balance is also worth noting. The Estonian group was generally equally divided between males and females. In Turkey, the majority of journalists were male. None of the participants identified themselves as data journalist. More precise characteristics of participants are provided in [Table 1](#).

After completing the questionnaire, we requested that participants do follow-up interviews. Journalists from Estonia gave comments through face-to-face interviews that spanned 28 to 50 minutes, while the Turkish journalists participated in e-mail interviews. Consequently, the Estonian data collection employed "think-aloud" principles, which help gain insight into the cognitive processes of participants (van Someren et al., 1994). Conversely, during the Turkish data collection, we expected participants to take a more active role in commenting on the questionnaire and providing their answers. This was evident in the longer personal explanations given by respondents after researchers described the purpose of the study.

After participants from both countries had completed the questionnaire, the data collected in both groups were merged into an Excel table to analyze the patterns in groups separately and comparatively. The standard deviation of answers was calculated for both country groups. Differences between groups for each of the data literacy components were calculated separately.

The answers to open questions were translated into English by the authors and analyzed qualitatively. Researchers from both countries read and discussed the comments and distinguished between social-scientific, professional journalistic, and other arguments journalists gave in explaining their answers in the questionnaire.

Since the number of participants is relatively small, the results of this study cannot be used as a basis to assess data skills in journalist populations in the respective countries. The most important point of the study is the composition of their data literacy and the identification of gaps in data literacy components.

Table 1. Overview of participant journalists in Estonia and Turkey*

Estonia (N = 10 participants)	Turkey (N = 20 participants)
E1–female, master’s degree, 18 years experience, local outlet, no special beat	T1–male, bachelor’s degree, age group 45-54, public broadcaster, no special beat
E2–male, bachelor’s degree, 10 years experience, national outlet, no special beat	T2–male, bachelor’s degree, age group 35-44, public broadcaster, no special beat
E3–female, secondary-school education, nine years experience, national outlet, no special beat	T3–female, bachelor’s degree, age group 25-35, public broadcaster, no special beat
E4–female, bachelor’s degree, 10 years experience, national outlet, no special beat	T4–male, secondary-school education, age group 45-54, freelancer, no special beat
E5–male, bachelor’s degree, eight years experience, public broadcaster, no special beat	T5–male, bachelor’s degree, age group 25-34, public broadcaster, no special beat
E6–female, master’s degree, 4 years experience, public broadcaster, science and technology beat(s)	T6–male, bachelor’s degree, age group 45-54, public broadcaster, no special beat
E7–female, bachelor’s degree, three years experience, national outlet, economy and finance beat(s)	T7–male, bachelor’s degree, age group 45-54, public broadcaster, no special beat
E8–male, bachelor’s degree, two years experience, national outlet, no special beat	T8–male, master’s degree, age group 35-44, public broadcaster, no special beat
E9–male, master’s degree, three years experience, national outlet, no special beat	T9–female, bachelor’s degree, age group 35-44, national outlet, no special beat
E10–male, bachelor’s degree, eight years experience, regional outlet, no special beat	T10–female, bachelor’s degree, age group 25-34, national outlet, no special beat
	T11–male, bachelor’s degree, age group 35-44, national outlet, no special beat
	T12–female, bachelor’s degree, age group 25-34, national outlet, no special beat
	T13–male, bachelor’s degree, age group 45-54, national outlet, no special beat
	T14–male, bachelor’s degree, age group 25-34, national outlet, no special beat
	T15–male, bachelor’s degree, age group 18-24, national outlet, no special beat
	T16–female, bachelor’s degree, age group 35-44, national outlet, no special beat
	T17–male, bachelor’s degree, age group 35-44, national outlet, no special beat
	T18–male, master’s degree, age group 25-34, national outlet, no special beat
	T19–male, bachelor’s degree, age group 35-44, public broadcaster, no special beat
	T20–male, bachelor’s degree, age group 35-44, public broadcaster, no special beat

Note. *To achieve diversity among participants, in Estonia, work experience was taken into consideration, and in Turkey, age was considered

RESULTS

After checking the completed questionnaires and summing up the correct answers, the resulting data were grouped to form **Table 2**. It shows that basic numeracy is the most-skilled category in both country groups, while statistics are more complicated to understand and use. There are visible differences between country groups in all components of data literacy.

We will comment on each of the data literacy components in the next three sections.

Numeracy Among Journalists: In Accordance With PISA and PIAAC Results

Basic numeracy questions in the questionnaire derive from the first level of formal education’s math study books and were presented using general mathematical symbols and verbal expressions, following the sample questions of the PIAAC study (see OECD, 2016a). Only three questions were selected for the questionnaire to keep the questionnaire as short as possible. The difficulty level of the questions varied.

The questions, “Which is the largest (1/2, 1/3, ¼, or 1/5)?” and percentage calculations (“Please calculate the percentage of the number 16 from the number 64”) yielded the most correct answers. Respondents chose among four multiple-choice answers. Both types of calculations have been “trained” in math classes from

Table 2. Results of testing of data literacy components among Estonian and Turkish participants

Data literacy component	Number of questions and answer variants in the questionnaire	Number of correct answers in the Estonian group (N = 10)	Number of correct answers in the Turkish group (N = 20)
Basic numeracy (“school-math”)*	Three questions and 10 answer variants	73%	66%
Statistics*	Four questions and 14 answer variants	55%	38%
Including logical and probability calculations presented without framing*	Two questions and eight answer variants	60%	30%
Including logical and probability calculations presented with framing	Two questions and six answer variants	50%	45%
Sociological knowledge	Three open-ended questions	Analyzed qualitatively	Analyzed qualitatively

Note. *Differences between Estonian and Turkish groups are statistically significant at the level of $p < 0.001$

early school years, and participants expressed high self-confidence by commenting on their answers in face-to-face or e-mail interviews later on. Studies show that it is possible to acquire a basic numeracy level even for individuals who have lower abilities for numerical thinking (Lipkus et al., 2001; Peters et al., 2006).

Answering the math question presented in textual form (originally presented in Estonian and in Turkish, respectively) proved more complicated: “There are flowers on a lake. The area with flowers will grow twice as large every day. If the flowers cover the whole lake in 48 days, then in how many days will one-half of the lake have flowers?”

The question resulted in 55% of correct answers among the Estonian group and in 30% of correct answers among the Turkish group. This kind of information processing also depends on language and literacy, while mathematics discourse in context integrates linguistic and nonlinguistic modalities of communication (Martiniello, 2009). Research shows that mathematics and literacy are strongly connected (Purpura & Napoli, 2015).

Patterns visible in numeracy among Estonian and Turkish participants are in accordance with the study results among countries’ general populations. PISA tests among school children show that in mathematics, Estonian students are above the OECD average, and 85% of students attained at least Level 2 proficiency. That is significantly more than the OECD average of 69%. Meanwhile, Turkish students scored below the OECD average in mathematics (OECD, 2023). Additionally, adults results from the most recent PIAAC study show that the average proficiency in numeracy is very high in Estonia—ranking 6th overall among the 31 countries (OECD, 2024). No new comparable data exists for Turkey’s adult population. In an earlier, 2014-2015 study wave, the OECD found that adults in Turkey showed below-average proficiency in all three assessed domains: literacy, numeracy, and problem solving in technology-rich environments. The score among Turkish adults was the third-lowest among the researched countries in numeracy (OECD, 2016b). The OECD study also revealed the strongest gender gap among the participating countries, where females significantly underperformed in numeracy (OECD, 2016b). Another clear determinant of numeracy level was education: Higher educated individuals performed significantly better (Atasoy & Güçlü, 2020).

The primary reasons for the low levels of mathematical knowledge and data literacy among journalists in Turkey may include the insufficient emphasis on statistics and data analysis in journalism education, the dominance of traditional text-based reporting, biases against mathematics, and the limited interest of media organizations in data journalism (Dincer et al., 2021). Journalism educators see the weakness of math education at the first school levels as the foundational problem (Guzel & Gondel, 2022). Additionally, the fast-paced news cycle and heavy workload make it difficult for journalists to allocate time for data analysis, while the prevalence of short and rapid content in digital media further restricts in-depth data utilization (Cengiz & Kaya, 2026).

Statistical Literacy: Notable Differences

Statistical literacy questions were given with and without framing; last ones prescribing the interpretation following examples by Entman (1993) and Kahneman and Tversky (1984). Questions without framing derived from math quizzes for students in the 10th class, e.g., “Which probability is higher:

- (1) For the flip of a coin to be “tails” following a result of “heads” in the previous flip, or
- (2) Today will be a rainy day if yesterday was rainy?”, or questions about the meaning of statistical correlation.

Statistics presented without framing indicated the highest gap between Estonian and Turkish participants. Fewer than one-third of questions were answered correctly by the Turkish group. Estonian participants got the correct answer 60% of the time.

Questions with framing derived from Kahneman and Tversky (1984, p. 343), translated by the authors into Estonian and Turkish. While one significant part of numeracy is related to the understanding and ability to use numbers in texts (Kirsch et al., 2002), the framing allows us to see how heuristics may influence decision-making based on the numbers. Additionally, for interpreting statistics, usually there are many textual elements that are necessary to consider. This can be complicated. Research shows that the more elements that need to be considered to interpret the numerical information, the higher the probability of misinterpretations (Cokely et al., 2012; Lowrie & Diezmann, 2009).

Statistical data presented with framing resulted in fewer differences between Estonian and Turkish participants, both giving correct answers almost half the time. If critical thinking and statistical analysis skills are generally not sufficiently developed, individuals tend to rely on heuristics and prior knowledge rather than engaging in analytical evaluation when confronted with numerical data.

However, both Estonian and Turkish participants show the ability to self-reflect and are themselves critical of their statistical literacy. Several of them indicated less self-confidence in answering the statistics questions, labeling it, “this awful math,” or revealing a lack of time for this type of thinking. These kinds of expressions in follow-up interviews were meant to excuse in advance the mistakes the participants might make, indicating their uncertainty in interpreting information provided in these formats.

Sociological Knowledge: Journalists With Academic Journalism Education Perform Better

Sociology is taught at the university level, but unevenly in different study programs. Sociological data knowledge includes knowledge about the sociological data collection methods and the contextual determination of data. To test this component of data literacy, the questionnaire consisted of one data table presenting daily ratings of the main TV channels in both countries, one short human interest news story, and one question about a concept of data in sociology. The news story originally appeared in an international magazine that was also published in Estonian and Turkish and asserts a correlation between makeup and income for females (see link to the text in [Appendix A](#)).

The follow-up interviews show that participants feel more comfortable commenting on the data from the media branch with which they are familiar. They expressed themselves as more confident, interpreted data based on their expert knowledge from the field (e.g., what kinds of shows TV channels have or who are the most popular TV journalists working at a station), and less in sociological terms (e.g., how the ratings are measured, attention to the seasonal effects, etc.). The data from an unfamiliar branch (makeup and income) raised more questions and hesitations about data and their interpretations. Journalists were asking to get more information about the sampling and data collection methods, the reliability of statistical calculations, and the accuracy of causal interpretations. They showed a general critical attitude toward the information provided and a professional readiness for fact-checking.

For the Estonian participants, answering all questions containing sociological information appeared easiest. They expressed feelings of being on solid ground when reading the information enriched by data and getting a picture of the contexts. They commented on the information presented in this form in three ways: using

- (1) professional journalistic arguments (most of them),
- (2) knowledge about the methodologies of social sciences (a few respondents), or
- (3) common sense.

Referring to the methodologies, they paid attention to the following aspects: who published the study, whether the results were convincing, and how big the sample was. So, some journalists used sociological vocabulary to express their hesitation towards a piece of news about research data.

Participants from Turkey addressed the following aspects of news pieces presenting research data:

- (1) quality of data collection and presentation,
- (2) causalities in data, and
- (3) journalistic professional criteria to present the data.

The first two employed sociological arguments, such as the idea that the data is insufficient, that the data is manipulated, or that the data cannot show causality in real life. Only a few respondents paid attention to the problems in following the “Five W’s” rule (What? When? Where? Why? How? Who cares?), which states that a lead should contain all the essential points of the story. Furthermore, individual approaches to data-driven news production in Turkey are closely linked to working conditions and institutional structures. For instance, participant T5 emphasized that the pressure of producing rapid news content leaves little to no time for data analysis, leading editorial teams to rely heavily on external sources. Similarly, T11 expressed skepticism toward the reliability and completeness of data provided by public institutions, which, in turn, led to a deliberate avoidance of data-based content. On the other hand, female journalists such as T9 and T16 adopted a more critical and attentive stance particularly regarding the presentation and interpretation of gender-related data. Participant T3 argued that data literacy is not merely a technical skill, but also an ethical responsibility in journalism—underscoring the necessity of contextualizing data within news narratives. T18 acknowledged personal shortcomings in statistical literacy but reported actively trying to address these gaps through online resources and open-access courses.

These insights indicate that Turkish journalists’ engagement with data is largely shaped by individual efforts, editorial guidance, and the level of institutional support available. Moreover, the perception among some journalists that data serves merely as a decorative element in news stories suggests that data-driven thinking has not yet been fully integrated into professional journalistic routines.

Generally, sociological concepts have not been widely used as arguments in both groups; instead, the participants used their professional knowledge if available.

DISCUSSION AND CONCLUSIONS

In this article, we have demonstrated the use of a data literacy questionnaire to gather information about the three main components of journalistic data literacy: numeracy, statistical literacy, and sociological data knowledge. This method was put to test in two countries: Estonia and Turkey.

Results show (**RQ1**) that these three main components were indeed clearly identifiable among both Turkish and Estonian respondents. Notable differences were visible in each of the three cases (**RQ2**): In case of numeracy, the results mirrored international PISA and PIAAC tests, with Estonian journalists giving more correct answers in basic math problems than their Turkish colleagues. But there were also significant differences while looking at the questions dealing with statistical literacy: an accuracy of 50%-to-60% in the case of Estonian journalists, and 30%-to-45% by the Turkish. While addressing sociological knowledge, the Turkish journalists interpreted the data, focusing on their expert knowledge from journalism, whereas their Estonian counterparts were keener to use sociological terms. Our research revealed that journalists identify the need for better data skills, but their everyday working context enables them to work meaningfully without advanced data skills, while journalism is still a profession where the main focus is working with people rather than numbers.

Interviews with journalists in Turkey revealed that data-driven journalism is shaped not only by individual skill levels, but also by structural limitations of the media environment and institutional culture. Participant T5’s remark, “We don’t have time for data; editors prefer to rely on external graphics,” illustrates how the pressure for fast news production discourages in-depth engagement with data and instead promotes superficial visual supplementation. T11’s comment, “Using such data is risky; it is often incomplete or contradictory,” highlights the pervasive mistrust in official data sources and reflects a broader culture of self-censorship, which ultimately constrains the development of data literacy skills. T3’s statement, “It’s easy to play with numbers, but the real responsibility lies in establishing the correct context,” underscores an ethical understanding of journalism that positions data not merely as a decorative element, but as central to the process of meaning-making. However, such critical awareness tends to remain at the level of individual

initiative and does not translate into institutional practice. In the Turkish media landscape marked by political pressures, economic dependency, and the challenges of digital transformation, systematic investment in data journalism remains limited. As a result, the integration of data-driven practices into journalistic routines cannot rely solely on personal motivation; rather, it requires broader structural reforms, including strengthened independent media institutions, access to reliable data sources, and the expansion of high-quality professional training.

By using the same data literacy questionnaire in both countries, the authors of this paper conclude that the basic structure of journalistic data skills remains the same (**RQ2**). But it can appear differently in different contexts—and countries. Specific skills show significant differences depending on the amount of data encountered in everyday work and preparedness for data journalism. While in both countries, numeracy will be trained in the first stages of schooling consistently up to 7 years old, statistics literacy and sociological data knowledge will be obtained differently. In Estonia, the basic statistics are a part of secondary education. At universities that provide academic journalism education, advanced statistics are also part of the curricula. In Estonia, journalism higher education is widely spread and will be taught in strong relation to social sciences, it gives a good basis for data skills among journalists. In the case of vocational journalism education or when the media organizations themselves need to prepare data journalism skills, the outcome is less effective. A practice-centered approach is not enough to support abstract thinking and interpreting skills necessary for data literacy. Yet previous research has also shown that data literacy is trainable (Howard et al., 2016) and the impact of meaningful practice on competencies is inevitable (Lave, 1988). The results of a recent worldwide survey (Bisiani et al., 2025) show that most journalists currently gain knowledge about data journalism by self-learning. This refers to the uneven distribution of skills in the journalist population.

Journalists, who serve as intermediaries of data literacy for the broader public and whose work increasingly relies on open-source data, must themselves possess adequate levels of data literacy. The data literacy of the general population can only improve if these “intermediaries”/“data brokers” are sufficiently skilled. In a datafying society, journalists need to constantly improve their abilities to produce factual and truthful stories that rely on data. While our research has shown that data literacy is an outcome of individual motivations and structural and contextual conditions, both need to be approached more systematically, e.g., by providing adult education programs and academic upskilling training courses depending on the gaps in individual data literacy. Universities could provide more follow-up courses or micro-credentials to the professional community. Dedicated programs should be designed to strengthen the data literacy of practicing journalists.

The data has inherent diversity, while it can be derived from physical, digital, and social worlds. Each type of data requires specific knowledge in assessing its reliability and usability for reporting. Our research indicated that a lack of skills in advanced statistics and sociological knowledge can be a major determinant of the scarcity of critical data stories in journalism.

For future studies, new questions will arise. What kind of options will provide AI tools for journalism for data-driven reporting is a question for future research. Robots and algorithms are already in newsrooms used for numerous tasks across all stages of news production: news gathering, writing and presentation, editing and promoting (Wu, 2026). There is a definite option to use prompting as an AI tool to acquire data. However, the need to be able to interpret the data and to be knowledgeable about the limits and scope of interpretation requires journalism data literacy—an understanding of how data is obtained from social reality.

Author contributions: **RK-K:** Conceptualization; Investigation; Methodology; Supervision; Visualization; Writing - original draft; **AY:** Data curation; Formal Analysis; Investigation; Writing - original draft; **MAA:** Data curation; Formal Analysis; Investigation; Writing - review & editing; **LA:** Conceptualization; Investigation; Validation; Writing - review & editing. All authors approved the final version of the article.

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

Ethics declaration: Data collection and processing comply with the General Data Protection Regulation and the European Code of Conduct for Research Integrity. Permission 101004811 was obtained by the Tartu University ethics committee.

AI statement: Grammarly was used for language editing. Authors assume full responsibility for the final content, including any editing assistance provided by generative AI.

Declaration of interest: The authors declared no competing interests.

Data availability: Data generated or analyzed during this study are available from the authors on request.

REFERENCES

- Appelgren, E. (2017). An illusion of interactivity: The paternalistic side of data journalism. *Journalism Practice*, 12(3), 308-325. <https://doi.org/10.1080/17512786.2017.1299032>
- Appelgren, E., & Nygren, G. (2014). Data journalism in Sweden: Introducing new methods and genres of journalism into "old" organizations. *Digital Journalism*, 2(3), 394-405. <https://doi.org/10.1080/21670811.2014.884344>
- Atasoy, R., & Güçlü, N. (2020). Evaluation of numeracy skills of adults according to the results of PIAAC 2015 in Turkey. *World Journal of Education*, 10(2), 27-49. <https://doi.org/10.5430/wje.v10n2p27>
- Auväärt, L. (2023). Fighting COVID-19 with data: An analysis of data journalism projects submitted to Sigma Awards 2021. *Central European Journal of Communication*, 15(3), 379-395. [https://doi.org/10.51480/1899-5101.15.3\(3\).3](https://doi.org/10.51480/1899-5101.15.3(3).3)
- Auväärt, L., & Köuts-Klemm, R. (2025). Turning data into news: Five different skill sets of journalists working with data. *Mediální Studia/ Media Studies*, 19(1), 6-30. https://medialnistudia.fsv.cuni.cz/wp-content/uploads/sites/6/2025/07/medialni_studia_volume2025_01.pdf
- Baack, S. (2015). Datafication and empowerment: How the open data movement re-articulates notions of democracy, participation, and journalism. *Big Data & Society*, 2(2). <https://doi.org/10.1177/2053951715594634>
- Beiler, M., Irmer, F., & Breda, A. (2020). Data journalism at German newspapers and public broadcasters: A quantitative survey of structures, contents and perceptions. *Journalism Studies*, 21(11), 1571-1589. <https://doi.org/10.1080/1461670X.2020.1772855>
- Bhaskaran, H., Kashyap, G., & Mishra, H. (2024). Teaching data journalism: A systematic review. *Journalism Practice*, 18(3), 722-743. <https://doi.org/10.1080/17512786.2022.2044888>
- Bisiani, S., Abellan, A., Arias Robles, F., & García-Avilés, J. A. (2025). The data journalism workforce: Demographics, skills, work practices, and challenges in the aftermath of the COVID-19 pandemic. *Journalism Practice*, 19(3), 502-522. <https://doi.org/10.1080/17512786.2023.2191866>
- Borges-Rey, E. (2016). Unravelling data journalism: A study of data journalism practice in British newsrooms. *Journalism Practice*, 10(7), 833-843. <https://doi.org/10.1080/17512786.2016.1159921>
- Burns, L. S., & Matthews, B. J. (2018). First things first: Teaching data journalism as a core skill. *Asia Pacific Media Educator*, 28(1), 91-105. <https://doi.org/10.1177/1326365X18765530>
- Cengiz, V., & Kaya, S. (2026). Data journalism practices in Turkey during COVID-19: Methods, narratives, news production, and future directions. *Journalism Practice*, 20(2), 621-642. <https://doi.org/10.1080/17512786.2024.2369106>
- Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin numeracy test. *Judgment and Decision Making*, 7(1), 25-47. <https://doi.org/10.1017/S1930297500001819>
- Cui, Y., Chen, F., Lutsyk, A., Leighton, J. P., & Cutumisu, M. (2023). Data literacy assessments: A systematic literature review. *Assessment in Education: Principles, Policy & Practice*, 30(1), 76-96. <https://doi.org/10.1080/0969594X.2023.2182737>
- Cushion, S., Lewis, J., & Callaghan, R. (2017). Data journalism, impartiality and statistical claims. *Journalism Practice*, 11(10), 1198-1215. <https://doi.org/10.1080/17512786.2016.1256789>
- Cushion, S., Lewis, J., Sambrook, R., & Callaghan, R. (2016). Impartiality review of BBC reporting of statistics: A content analysis. *BBC*. http://downloads.bbc.co.uk/bbctrust/assets/files/pdf/our_work/stats_impartiality/content_analysis.pdf
- DataJournalism.com (2021). State of Data Journalism Survey 2021: 11 surprising findings. Retrieved 8 May 2026, from <https://datajournalism.com/read/blog/data-journalism-survey-2021>
- De Maeyer, J., Libert, M., Domingo, D., Heinderyckx, F., & Le Cam, F. (2015). Waiting for data journalism: A qualitative assessment of the anecdotal take-up of data journalism in French-speaking Belgium. *Digital Journalism*, 3(3), 432-446. <https://doi.org/10.1080/21670811.2014.976415>

- Diaz-Bone, R., Horvath, K., & Cappel, V. (2020). Social research in times of big data. The challenges of new data worlds and the need for a sociology of social research. *Historical Social Research/Historische Sozialforschung*, 45(3), 314-341. <https://www.jstor.org/stable/26918415>
- Dieckmann, N. F., Slovic, P., & Peters, E. M. (2009). The use of narrative evidence and explicit likelihood by decisionmakers varying in numeracy. *Risk Analysis*, 29(10), 1473-1488. <https://doi.org/10.1111/j.1539-6924.2009.01279.x>
- Dincer, O., Kaderoglu Bulut, C., & Comu, T. (2021). Türkiye'de yükseköğretim politikaları bağlamında gazetecilik bölümleri: Karşılaştırmalı bir inceleme [Journalism departments in the context of higher education policies in Turkey: A comparative study]. *Türkiye İletişim Araştırmaları Dergisi*, 37, 193-214. <https://doi.org/10.17829/turcom.861477>
- Engelbrechtsen, M., Kennedy, H., & Weber, W. (2018). Data visualisation in Scandinavian newsrooms: Emerging trends in journalistic visualisation practices. *Nordicom Review*, 39(2), 3-18. <https://doi.org/10.2478/nor-2018-0007>
- Entman, R. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, 43(4), 51-58. <https://doi.org/10.1111/j.1460-2466.1993.tb01304.x>
- Fagerlin, A., Zikmund-Fisher, B. J., Ubel, P. A., Jankovic, A., Derry, H. A., & Smith, D. M. (2007). Measuring numeracy without a math test: Development of the subjective numeracy scale. *Medical Decision Making*, 27(5), 672-680. <https://doi.org/10.1177/0272989X07304449>
- Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review*, 87(6), 477-531. <https://doi.org/10.1037/0033-295X.87.6.477>
- Frank, M., Walker, J., Attard, J., & Tygel, A. (2016). Data literacy—What is it and how can we make it happen? *The Journal of Community Informatics*, 12(3), 4-8. <https://doi.org/10.15353/joci.v12i3.3274>
- Gencil Bek, M. (2014). Karşılaştırmalı perspektiften Türkiye'de medya sistemi [The media system in Turkey from a comparative perspective]. *Mülkiye Dergisi*, 34(269), 101-125. <https://dergipark.org.tr/tr/pub/mulkiye/article/631>
- Global Open Data Index. (2024). Global open data index. OKFN. <http://index.okfn.org/place.html>
- Gray, J., Gerlitz, C., & Bounegru, L. (2018). Data infrastructure literacy. *Big Data & Society*, 5(2), Article 205395171878631. <https://doi.org/10.1177/2053951718786316>
- Greskovits, B. (2015). Capitalist diversity and the media. In J. Zielonka (Ed.), *Media and politics in new democracies* (pp. 58-70). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198747536.003.0004>
- Guzel, M., & Gonden, M. (2022). Journalism education in Turkey in the spiral of neoliberal policies and technological transformation: Problems, pursuits, and suggestions. *Journal of Qualitative Research in Education*, 31, 223-245. <https://doi.org/10.14689/enad.31.1610>
- Haim, M. (2024). The German data journalist in 2021. *Journalism Practice*, 18(6), 1378-1397. <https://doi.org/10.1080/17512786.2022.2098523>
- Heravi, B. R. (2019). 3WS of data journalism education: What, where and who? *Journalism Practice*, 13(3), 349-366. <https://doi.org/10.1080/17512786.2018.1463167>
- Heravi, B. R., & Lorenz, M. (2020). Data journalism practices globally: Skills, education, opportunities, and values. *Journalism and Media*, 1(1), 26-40. <https://doi.org/10.3390/journalmedia1010003>
- Hill, W. T., & Brase, G. L. (2012). When and for whom do frequencies facilitate performance? On the role of numerical literacy. *Quarterly Journal of Experimental Psychology*, 65(12), 2343-2368. <https://doi.org/10.1080/17470218.2012.687004>
- Himma-Kadakas, M., & Kõuts-Klemm, R. (2023). Developing an advanced digital society: An Estonian case study. In S. Davydov (Ed.), *Internet in the post-soviet area: Technological, economic and political aspects* (pp. 109-133), Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-32507-6_6
- Hindess, B. (1973). *The use of official statistics in sociology: A critique of positivism and ethnomethodology*. Springer. <https://doi.org/10.1007/978-1-349-01579-5>
- Howard, S. J., Woodcock, S., Ehrich, J., & Bokosmaty, S. (2016). What are standardized literacy and numeracy tests testing? Evidence of the domain-general contributions to students' standardized educational test performance. *British Journal of Educational Psychology*, 87(1), 108-122. <https://doi.org/10.1111/bjep.12138>

- Jõesaar, A., & Kõuts-Klemm, R. (2019). Is the role of public service media in Estonia changing? *Baltic Screen Media Review*, 7(1), 96-116. <https://doi.org/10.2478/bsmr-2019-0006>
- Kahneman, D., & Tversky, A. (1984). Choice, values, and frames. *American Psychologist*, 39(4), 341-350. <https://doi.org/10.1037/0003-066X.39.4.341>
- Karger, T., Kalenda, J., Vaculíková, J., & Kočvarová, I. (2024). Online learning platforms and resources in adult education and training: New findings from four European countries. *International Journal of Lifelong Education*, 43(4), 417-431. <https://doi.org/10.1080/02601370.2024.2358896>
- Kasica, S., Berret, C., & Munzner, T. (2020). *Table scraps: An actionable framework for multi-table data wrangling from an artifact study of computational journalism*. arXiv. <https://doi.org/10.48550/arXiv.2009.02373>
- Kavuncu, F., & Polat, S. (2019). A brief assessment of adults' skills in Turkey: Results from survey of adults skills (PIAAC). *TÜSİAD-Sabancı Üniversitesi Rekabet Forumu*. https://ref.sabanciuniv.edu/sites/ref.sabanciuniv.edu/files/2021-03/a_brief_assessment_of_adult_skills_in_turkey.pdf
- Kim, S., & Kim, T. (2021). A study of the definition and components of data literacy for K-12 AI education. *Journal of the Korean Association of Information Education*, 25(5), 691-704. <https://doi.org/10.14352/jkaie.2021.25.5.691>
- Kirsch, I. S., Jungeblut, A., Jenkins, L., & Kolstad, A. (2002). *Adult literacy in America: A first look at the findings of the national adult literacy survey*. National Center for Education, U.S. Department of Education. <https://nces.ed.gov/pubs93/93275.pdf>
- Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures & their consequences*. SAGE. <https://doi.org/10.4135/9781473909472>
- Kõuts-Klemm, R. (2019). Data literacy among journalists: A skills-assessment based approach. *Central European Journal of Communication*, 12(24), 299–315. [https://doi.org/10.19195/1899-5101.12.3\(24\).2](https://doi.org/10.19195/1899-5101.12.3(24).2)
- Kuyucu, M. (2013). Türkiye'de çapraz medya sahipliği: Medya ekonomisine olumsuz etkileri ve bu etkilerin önlenmesine yönelik öneriler [Cross-media ownership in Turkey: Negative impacts on the media economy and suggestions for preventing these impacts]. *Selçuk İletişim*, 8(1), 144-163. <https://dergipark.org.tr/tr/pub/josc/article/200536>
- Larrondo, A., Domingo, D., Erdal, I. J., Masip, P., & Van den Bulck, H. (2016). Opportunities and limitations of newsroom convergence: A comparative study on European public service broadcasting organisations. *Journalism Studies*, 17(3), 277-300. <https://doi.org/10.1080/1461670X.2014.977611>
- Larsson, S. (2009). A pluralist view of generalization in qualitative research. *International Journal of Research & Method in Education*, 32(1), 25-38. <https://doi.org/10.1080/17437270902759931>
- Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511609268>
- Lewis, N. P., McAdams, M., & Stalph, F. (2020). Data journalism. *Journalism & Mass Communication Educator*, 75(1), 16-21. <https://doi.org/10.1177/1077695820904971>
- Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making*, 21(1), 37-44. <https://doi.org/10.1177/0272989X0102100105>
- Loit, U., & Harro-Loit, H. (2017). *Media pluralism monitor 2016: Monitoring risks for media pluralism in EU and beyond: Country report: Estonia*. European University Institute.
- Loosen, W., Reimer, J., & Schmidt, F. (2015). When data become news—A content analysis of data journalism pieces. *Academia*. https://www.academia.edu/16029530/When_data_become_news_A_content_analysis_of_data_journalism_pieces
- Loosen, W., Reimer, J., & De Silva-Schmidt, F. (2020). Data-driven reporting: An on-going (r)evolution? An analysis of projects nominated for the data journalism awards 2013-2016. *Journalism*, 21(9), 1246-1263. <https://doi.org/10.1177/1464884917735691>
- Lowrie, T., & Diezmann, C. M. (2009). National numeracy tests: A graphic tells a thousand words. *Australian Journal of Education*, 53(2), 141-158. <https://doi.org/10.1177/000494410905300204>
- Mancini, P. (2015). The news media between volatility and hybridization. In J. Zielonka (Ed.), *Media and politics in new democracies* (pp. 25-37). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198747536.003.0002>
- Martiniello, M. (2009). Linguistic complexity, schematic representations, and differential item functioning for English language learners in math tests. *Educational Assessment*, 14(3-4), 160-179. <https://doi.org/10.1080/10627190903422906>

- McQuail, D. (2006). Media roles in society. In K. Nordenstreng (Ed.), *Researching media, democracy and participation*.
- OECD. (2016a). *Skills matter: Further results from the survey of adult skills*. OECD Publishing. <https://doi.org/10.1787/888933366458>
- OECD. (2016b). Turkey–Country note–Skills matter: Further results from the survey of adult skills. *OECD*. <https://www.oecd.org/content/dam/oecd/en/about/programmes/edu/piaac/country-specific-material/cycle-1/Turkiye-Country-Note-ENG.pdf>
- OECD. (2023). *PISA 2022 results (volume I): The state of learning and equity in education*. OECD Publishing. <https://doi.org/10.1787/53f23881-en>
- OECD. (2024). *Do adults have the skills they need to thrive in a changing world?: Survey of adult skills 2023*. OECD Publishing. <https://doi.org/10.1787/b263dc5d-en>
- Open Data Maturity Report. (2023). Open data maturity report. *European Union*. <https://data.europa.eu/en/publications/open-data-maturity/2023>
- Örnebring, H., & Mellado, C. (2018). Valued skills among journalists: An exploratory comparison of six European nations. *Journalism*, 19(4), 445-463. <https://doi.org/10.1177/1464884916657514>
- Oyebamiji, M. A. (2012). The role of mass media in the delivery of adult education for national development. *Global Journal of Educational Research*, 11(1), 57-63. <https://scispace.com/pdf/the-role-of-mass-media-in-the-delivery-of-adult-education-38lsosnoov.pdf>
- Parasie, S., & Dagiral, E. (2013). Data-driven journalism and the public good: “Computer-assisted-reporters” and “programmer-journalists” in Chicago. *New Media & Society*, 15(6), 853-871. <https://doi.org/10.1177/1461444812463345>
- Parsons, S., & Bynner, J. (2005). *Does numeracy matter more?* NRDC.
- Pentzold, C., Fechner, D. J., & Zuber, C. (2021). “Flatten the curve”: Data-driven projections and the journalistic brokering of knowledge during the COVID-19 crisis. *Digital Journalism*, 9(9), 1367-1390. <https://doi.org/10.1080/21670811.2021.1950018>
- Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological Science*, 17(5), 407-413. <https://doi.org/10.1111/j.1467-9280.2006.01720.x>
- Purpura, D. J., & Napoli, A. R. (2015). Early numeracy and literacy: Untangling the relation between specific components. *Mathematical Thinking and Learning*, 17(2-3), 197-218. <https://doi.org/10.1080/10986065.2015.1016817>
- Rajamäe, K., & Nikiforova, A. (2024). Exploring Estonia’s open government data development as a journey towards excellence: Unveiling the progress of local governments in open data provision. *arXiv*. <https://doi.org/10.48550/arXiv.2403.11952>
- Ramsälv, A., Ekström, M., & Westlund, O. (2024). The epistemologies of data journalism. *New Media & Society*, 26(11), 6307-6324. <https://doi.org/10.1177/14614448221150439>
- Rogers, S., Schwabish, J., & Bowers, D. (2017). *Data journalism in 2017: The current state and challenges facing the field today*. Google News Lab. <https://newslab.withgoogle.com/assets/docs/data-journalism-in-2017.pdf>
- Schild, M. (2005). Information literacy, statistical literacy, data literacy. *IASSIST Quarterly*, 28(2-3), 6-11. <https://doi.org/10.29173/iq790>
- Sharma, S. (2017). Definitions and models of statistical literacy: A literature review. *Open Review of Educational Research*, 4(1), 118-133. <https://doi.org/10.1080/23265507.2017.1354313>
- Smith, J. (2018). *Data journalism in global newsrooms: Trends and challenges*. Oxford University Press.
- Stalph, F. (2018). Classifying data journalism. A content analysis of data-driven stories. *Journalism Practice*, 12(10), 1332-1350. <https://doi.org/10.1080/17512786.2017.1386583>
- Steensberg, M. T. (2023). In numbers we trust–The use of statistics in political news. *Journalism Practice*, 17(7), 1500-1518. <https://doi.org/10.1080/17512786.2021.1987298>
- Tsang, S., Royse, C. F., & Terkawi, A. S. (2017). Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. *Saudi Journal of Anesthesia*, 11(Suppl 1), S80-S89. https://doi.org/10.4103/sja.SJA_203_17
- van Someren, M., Barnard, Y., & Sandberg, J. (1994). *The think aloud method—a practical guide to modelling cognitive processes*. Academic Press.

- Van Witsen, A. (2020). How daily journalists use Numbers and statistics: The case of global average temperature. *Journalism Practice*, 14(9), 1047-1065. <https://doi.org/10.1080/17512786.2019.1682944>
- Weaver, D. H., & McCombs, M. E. (1980). Journalism and social science: A new relationship? *Public Opinion Quarterly*, 44(4), 477-494. <https://doi.org/10.1086/268618>
- Wolff, A., Gooch, D., Montaner, J. J. C., Rashid, U., & Kortuem, G. (2016). Creating an understanding of data literacy for a data-driven society. *The Journal of Community Informatics*, 12(3), 9-26. <https://doi.org/10.15353/joci.v12i3.3275>
- Wu, S. (2026). Journalists as individual users of artificial intelligence: Examining journalists' "value-motivated use" of ChatGPT and other AI tools within and without the newsroom. *Journalism*, 27(2), 388-406. <https://doi.org/10.1177/14648849241303047>

APPENDIX A

Table A1. Example of questions in the data literacy questionnaire

Data literacy component	Example of a question in the questionnaire	Researcher's comments
Basic numeracy (the "school-math")	Question: Please calculate the percentage of the number 16 from the number 64. Variants: 10%, 25%, 40%, 45%	Questions derived from the math study books for 4 th and 8 th class; based on the example of PIAAC, for numeracy see sample questions).
Statistics presented without framing	Question: Which probability is higher: A. To have after the throw the obverse of the coin up; or B. That today is a rainy day if yesterday was rainy? Variants: A, B, probabilities are the same	Questions derived from the math quizzes for the students in the 10 th class.
Statistics presented with framing	Question: Imagine that the country is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the programs are as follows: If Program A is adopted, 200 people will be saved. If program B is adopted, there is a one-third probability that 600 people will be saved and a two-thirds probability that no people will be saved. Which of the two programs would you favor? Variants: I will favor A; I will favor B; It makes no difference, because the probabilities are the same	Exact wording derives from Kahneman and Tversky (1984, p. 343), translated by the authors into Estonian and Turkish.
Presentation of sociological data	A short news story about correlation between make up and income for females (see: English version in https://www.glamourmagazine.co.uk/article/women-who-wear-makeup-get-paid-more).	The text has been published in both countries' news media, translators are not known.

