Friedrich *et al.*
International Journal of STEM Education (2024) 11:58
https://doi.org/10.1186/s40594-024-00517-z

International Journal of STEM Education

REVIEW

Open Access

What shapes statistical and data literacy research in K-12 STEM education? A systematic review of metrics and instructional strategies

Anja Friedrich[1*], Saskia Schreiter[2], Markus Vogel[3], Sebastian Becker-Genschow[4], Roland Brünken[1], Jochen Kuhn[5], Jessica Lehmann[1] and Sarah Malone[1]

Abstract

The pervasive digitization of society underscores the crucial role of data and its significant impact on decision-making across various domains. As a result, it is essential for individuals to acquire competencies in handling data. This need is particularly pertinent in K-12 education, where early engagement with data and statistics can lay a foundational understanding for future academic and professional endeavors. Additionally, K-12 education should provide students with critical skills necessary for navigating the complexities of daily life and making informed decisions in a data-rich society. This systematic review examines the state of research on statistical and data literacy in K-12 STEM (Science, Technology, Engineering, and Mathematics) education. It focuses specifically on cognitive, affective, and behavioral metrics and pedagogical approaches empirically investigated in this context. Using a rigorous selection process, we identified and synthesized 83 original empirical papers. Additionally, we invited the authors of these studies to share their perspectives on future strategies for addressing statistical and data literacy. The results indicate that the included studies primarily focus on the construct of statistical literacy, which is operationalized through a diverse array of metrics, predominantly within the context of mathematics education. We identified effective pedagogical approaches, such as authentic problem-solving and the integration of real-world data. The researchers surveyed emphasized the importance of interdisciplinary teaching, adapted curricula, and improved professional development for pre- and in-service teachers. Our findings underscore the growing relevance of this field, but suggest that integrated perspectives on statistical and data literacy within STEM subjects are limited.

Keywords Statistical literacy, Data literacy, STEM education, K-12 education

*Correspondence:
Anja Friedrich
a.friedrich@mx.uni-saarland.de
[1] Department of Education, Saarland University, Campus A4 2, 66123 Saarbrücken, Germany
[2] Institute of Mathematics and Computer Science, University of Education, Schwäbisch Gmünd, Germany
[3] Institute of Mathematics and Computer Science, Heidelberg University of Education, Heidelberg, Germany
[4] Faculty of Mathematics and Natural Sciences, University of Cologne, Cologne, Germany
[5] Faculty of Physics, Ludwig-Maximilians-Universität München, Munich, Deutschland

Increasing digitization has led to data and data representations playing a significant societal role and being used for decision-making in various areas of life. The developments in the field of Artificial Intelligence (AI) have further increased the relevance of these competencies, as AI systems—trained on ever-increasing amounts of data—exert more power over our behavior and decisions (Ng et al., 2021). In this context, competent handling of data—called statistical and data literacy—has become an essential skill in the twenty-first century (OECD, 2021). It is noteworthy that the perceived relevance of statistical and data literacy among the public and in research has increased (OECD, 2021). On the one hand, the constant advance of technical progress makes it easier to collect and process large amounts of data and access it from anywhere. On the other hand, stricter guidelines for the protection of personal data have been introduced, which have led to more frequent reminders that one's data are being collected and for what purposes.

As many as 30 years ago, Wallman (1993) called for the promotion of statistical literacy in school education to produce citizens who can handle data and data representations. Since then, the definitions of the term statistical literacy and the sub-competencies included have changed and essentially expanded.

While the term initially focused solely on the recipients of data and, according to Gal's (2002) prominent definition, involved drawing correct inferences from, critically appraising, and communicating statistical data, Gould's definition (2017) expanded this scope. He emphasized the importance of skills essential not only for recipients, but also for producers of data, such as data collection and the generation of data representations, thereby broadening the concept from statistical literacy to data literacy.

Later, concepts of data ethics and data protection were partially added (e.g., Wolff et al., 2017).

Concerning the promotion of statistical literacy in school education, K-12 education (i.e., from the end of kindergarten to grade 12) plays a critical role in equipping learners with the necessary skills and knowledge. STEM subjects (Science, Technology, Engineering, and Mathematics) have traditionally been the primary setting for the acquisition of fundamental concepts related to statistical and data literacy, including measurement, experimentation, data collection, analysis, and visualization, which have increasingly found their way into STEM curricula internationally (e.g., National Council of Teachers of Mathematics [NCTM], 2000; New Zealand Ministry of Education, 2014; NGSS Lead States, 2013). As a result, STEM subjects play a crucial role in promoting and developing these skills in students, which is also visible in emerging research on statistical and data literacy in the context of K-12 STEM education (e.g., Swan et al., 2013; Watson & Callingham, 2003).

Regarding the development of statistical and data literacy, as well as their implementation in STEM school education, the question arises to what extent and in what ways these areas are reflected in international research. Here, we present a systematic review that delves into the current research landscape surrounding statistical and data literacy in K-12 STEM education. Our objective is to explore the empirical investigation of metrics and instructional approaches in international research. Through this review, we aim to synthesize the state-of-the-art in this domain and, in doing so, identify existing research gaps that could guide future advancements in the field.

Theoretical framework

Statistical and data literacy

When mentioning *statistical literacy* or *data literacy*, researchers often do not adhere to a single, unified concept for either term. This variability can be attributed to the development of diverse definitions and conceptualizations of these competencies over time. In the following section, we aim to elucidate the most significant definitions and conceptual frameworks employed by researchers in their empirical work on statistical and data literacy.

A fundamental definition of statistical literacy comes from Wallman (1993, p. 1): "the ability to understand and critically evaluate statistical results that permeate our daily lives—coupled with the ability to appreciate the contributions that statistical thinking can make in public and private, professional and personal decisions." A first prominent framework of statistical literacy has been suggested by Watson (1997) who describes the acquisition of statistical literacy as a cyclic process consisting of three hierarchical tiers: (i) understanding basic statistical terminology; (ii) understanding terminology when it appears in social contexts; and (iii) questioning claims that are made in context without proper statistical justification. Watson and Callingham (2003) expanded and refined the knowledge components of the hierarchy by empirically confirming six different levels of statistical literacy:

1. Idiosyncratic (idiosyncratic engagement with context, tautological use of terminology, basic mathematical skills).
2. Informal (informal engagement with context, single elements of complex terminology and settings, understanding of basic table, graph, and chance calculations).
3. Inconsistent (selective engagement with context, appropriate recognition of conclusions, qualitative use of statistical ideas).
4. Consistent non-critical (appropriate but non-critical engagement with context, multiple aspects of terminology usage, appreciation of variation in chance settings, statistical skills, such as mean, simple probabilities, and graph characteristics).
5. Critical (questioning engagement in contexts, involving appropriate use of terminology, qualitative interpretation of chance, and appreciation of variation).
6. Critical mathematical (using proportional reasoning in media or chance contexts, showing appreciation for uncertainty in making predictions, and interpreting subtle aspects of language).

Gal (2002, 2004) introduced the ability to communicate through statistics as a further important element of statistical literacy. Furthermore, according to Gal (2004), statistical literacy consists of both a knowledge component (statistical knowledge, mathematical knowledge, context knowledge, critical questions, literacy skills) and a dispositional component (critical stance, beliefs, and attitudes). Watson (2006) also acknowledged the importance of dispositions in the development of statistical literacy. She used the following definition for her research with a focus on the school context: “Statistical literacy is the meeting point of the data and chance curriculum and the everyday world, where encounters involve unrehearsed contexts and spontaneous decision-making based on the ability to apply statistical tools, general contextual knowledge, and critical literacy skills” (Watson, 2006, p. 11).

A significant further conceptual expansion occurred when Gould (2017) overcame the limitation that definitions of statistical literacy focused solely on the recipients of data (also see Gal, 2004). He positioned statistical literacy as a subset within the broader framework of *data literacy*, which according to the author, encompasses skills that are essential both as a recipient and a producer of data. The concept of data literacy has been refined further by several research groups to include aspects of data protection and data ethics (Calzada Prado & Marzal, 2013; Ridsdale et al., 2015; Schüller et al., 2019; Wolff et al., 2017).

Frameworks that conceptualize data literacy as skills, competencies, tasks, and/or knowledge are presented by different authors. Calzada Prado and Marzal (2013) distinguished five different components of data literacy: (i) understanding data; (ii) finding and/or obtaining data; (iii) reading, interpreting, and evaluating data; (iv) managing data; and (v) using data and translated them into instructional topics or units for direct implementation within education. Ridsdale et al. (2015) derived a competencies matrix with five core aspects of data literacy (data, collection, management, evaluation, and application), operationalized into 23 competencies and 64 associated tasks or skills classified into either conceptual competencies (e.g., data tools, data ethics), core competencies (e.g., data organization and manipulation, data visualization, data sharing), or advanced competencies (e.g., data conversation, evaluating decisions based on data). Based on this, Data to the People (2018) developed a competency framework (called *databilities*) aiming to create an instrument to assess data literacy. *Data to the People* is a globally recognized organization specializing in building, fostering and nurturing data literacy. Wolff et al. (2017) conceptualize data literacy within their framework in a way that closely resembles existing models, yet they distinctly allocate the key aspects across the various stages of the PPDAC (Problem, Plan, Data, Analysis, and Conclusion) inquiry cycle (Wild & Pfannkuch, 1999). Wolff et al. (2017) proposed a definition of data literacy that encompasses core practical and creative skills, specialized data-handling abilities, and the capability to effectively and ethically use data to address real-world questions. Thus, according to them, key aspects of data literacy include data selection, cleaning, analysis, visualization, critique, interpretation, storytelling from data, and incorporating data into design processes. The most recent comprehensive framework developed by Schüller et al. (2019) is noteworthy; however, it is important to recognize that its applicability is limited to higher education, which represents a significant limitation for our research question. The framework adopts a cyclical process model which categorizes the various steps of the process and their associated competencies into productive and receptive components. The framework features main categories (establish data culture, provide data, evaluate data, interpret data products, interpret data, derive actions), subcategories (competencies), and levels of competencies.

The current literature on statistical and data literacy contains inconsistencies in terms of both the terminology used and the sub-competencies included. Especially in the research fields of social science and education, data literacy is often even considered synonymous with statistical literacy (Schield, 2004). Additionally, it can be observed that the term data literacy has gained prominence in discussions related to data analytics and data science, specifically referring to the facilitation of data-driven decision-making processes (Khan et al., 2018).

In our analysis of the various theoretical definitions and frameworks of statistical and data literacy, we can discern that these literacies are understood to encompass a range of competencies and skills that are crucial for effective data handling and interpretation. Based on our understanding, these competencies can be understood using a framework of three distinct components: cognitive, affective, and behavioral.

The *Cognitive Component* primarily addresses the intellectual skills and knowledge required to understand and process statistical information. This includes the ability to interpret data, understand probabilistic and statistical reasoning, and apply these insights to make informed decisions (see definition of Gal, 2002; Wallman, 1993; Watson, 1997; Watson & Callingham, 2003). The cognitive component is foundational, as it pertains to the theoretical knowledge and intellectual abilities that form the basis of statistical and data literacy.

The *Behavioral Component* reflects the practical application of statistical knowledge and affective dispositions.

It encompasses actions such as collecting data, analyzing it, drawing conclusions, and communicating findings effectively (e.g., see definitions and frameworks of Calzado Prado & Marzell, 2013; Gould, 2017; Schüller et al., 2019). This component is about translating cognitive understanding and affective dispositions into (real-world) practices, embodying the practical skills that enable individuals to perform data-related tasks proficiently.

The *Affective Component* involves the attitudes, beliefs, and values related to one's engagement with data (e.g., see definitions of Gal, 2002, 2004; Watson, 2006). This component acknowledges that a person's feelings toward data—such as their comfort level with numbers, their perception of data trustworthiness, and their motivation and interest in engaging with statistical information—can play a crucial role in shaping their statistical and data literacy, as well as in determining how effectively they develop and utilize these skills. Together, these components form a comprehensive framework that captures the multifaceted nature of statistical and data literacy. In the following figures (Figs. 1 and 2), we have assigned the most important aspects from the various theoretical frameworks and definitions on statistical and data literacy that have been taken up in the theoretical background to the three components (cognitive, affective, behavioral) and located the various authors with their theoretical frameworks within them. Even though we agree that the concepts overlap (e.g., Gould, 2017), we have chosen two visualizations here—one for statistical literacy (Fig. 1) and one for data literacy (Fig. 2)—to show how the respective theoretical conceptualizations position themselves among the three components. In the visualization of statistical literacy (Fig. 1) it becomes apparent that the cognitive and affective components are more prominently emphasized than the behavioral component, since more authors incorporate aspects of these components in their definitions and conceptualizations of statistical literacy. While Wallman (1993), Watson (1997), and Watson and Callingham (2003) mention exclusively cognitive aspects in their definitions, affective and cognitive aspects are included by Watson (2006) within her framework. Only Gal's framework Gal (2002, 2004) includes all three components (see positioning of the authors in Fig. 1).

In the representation of data literacy (Fig. 2), we have included those subcomponents (in bold) that are unique for the conceptualizations of data literacy. Regarding this concept, we find a strong theoretical emphasis on the behavioral component while the affective component seems to take a back seat. Figure 2 shows that three frameworks (Calzado Prado & Marzall, 2013; Gould, 2017; Wolff et al., 2017) focus on cognitive and behavioral aspects without incorporating any affective

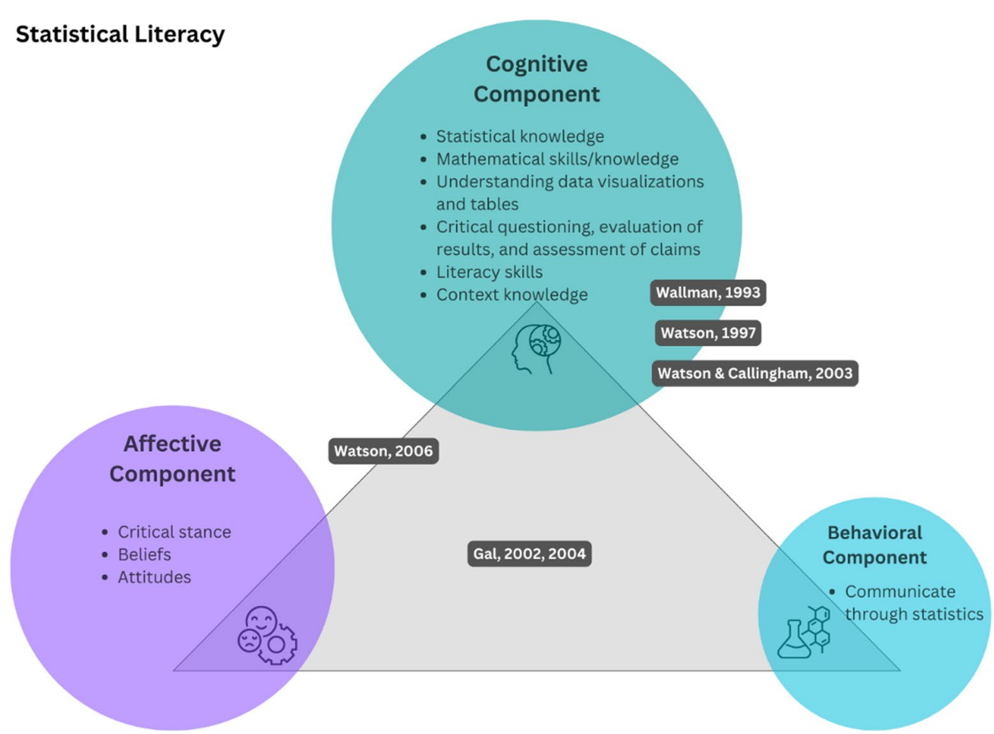

Fig. 1 Three-component framework of statistical literacy including positioning of the authors

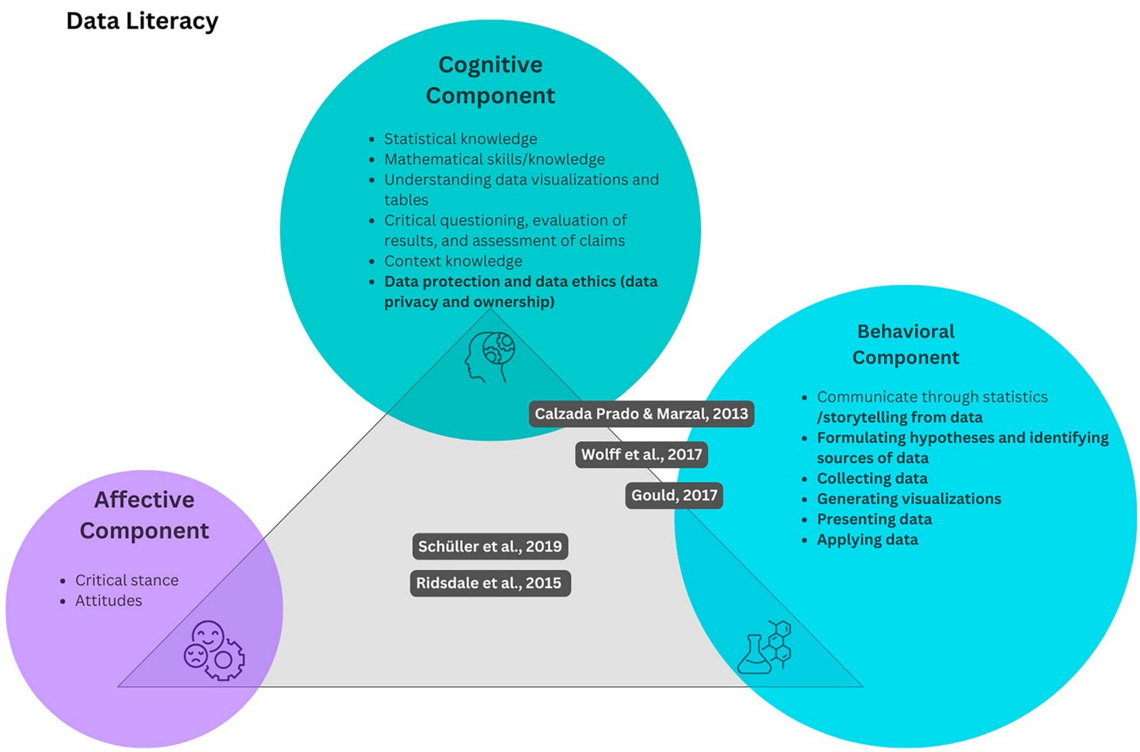

Fig. 2 Three-component framework of data literacy including positioning of the authors

subcomponents. Ridsdale et al. (2015) and Schüller et al. (2019) include all three components in their conceptualizations of data literacy.

Beyond the inconsistencies in the current definitions of statistical and data literacy, it is important to note that other related concepts are rising in prominence. *Information literacy*, for example, can be described as the capacity to locate, organize, and utilize information across diverse formats (Carlson et al., 2015). Statistical and data literacy differ from this definition as they specifically focus on information derived from data (Heidrich et al., 2018). Furthermore, parallels can be drawn to the concept of *scientific literacy*, which also exhibits overlaps with statistical and data literacy, but pertains to the knowledge and understanding of wider scientific concepts and processes (National Research Council, 1996). Other terms have been coined, such as *digital literacy*, *AI literacy*, and *data science literacy*, that increasingly integrate the use of computers and new technologies (such as AI) and emphasize competencies in domains of information technology. In this review, we focus only on the concepts and terms of statistical and data literacy, and do explicitly not include other concepts and terms, although we understand that they may be incorporated in some of the reviewed articles.

Statistical and data literacy in K-12 STEM education

As noted earlier and advocated by Wallman (1993), K-12 education plays a critical role in the formation of data-literate citizens. The teaching of statistical and data literacy has become an expected element of schooling (Chalkiadaki, 2018; Franklin et al., 2007). There are increasing calls to adapt international curricula to include statistical and data literacy, with a particular focus on K-12 education (Franklin et al., 2007; Redecker & Punie, 2017). Primarily, foundational concepts pertaining to statistical and data literacy have traditionally resided within the domain of statistics and have been integrated into the mathematics curriculum. Numerous efforts to enhance statistics education began in the 2000s, evidenced by reforms in statistics curricula across multiple countries and more governmental awareness (e.g., Kultusministerkonferenz, 2004a, b; Ministério da Educação, 2006; Ministry of Education, 2015; NCTM, 2000). These reforms advocate transitioning from a mere emphasis on descriptive statistics to actively engaging students in data exploration and probability modeling, as noted by Pfannkuch (2018). Burril et al. (2023) point out that today, secondary level statistical education globally varies between basic data analysis with a focus on calculations (e.g., in Brazil, Colombia, South Africa, and Turkey), and a more mathematically oriented approach emphasizing probability (e.g., in Finland or in the United Kingdom). Some countries prioritize a data-driven, simulation-based curriculum to introduce concepts like hypothesis testing for practical application (e.g., in New Zealand or Japan; Burril et al., 2023). Furthermore, the Next Generation Science Standards (NGSS Lead States, 2013) incorporate essential skills of statistical and data literacy by expecting students to analyze and interpret data, use mathematical and computational thinking, and explain scientific phenomena using models. In the course of these developments within the curricula, projects and guidelines also emerged with ideas for implementing them in the classroom, such as the Pre-K-12 Guidelines for Assessment and Instruction in Statistics Education (GAISE; Franklin et al., 2007, updated in 2020: Bargagliotti et al., 2020) or recently the Data Literacy Charta (Schüller et al., 2021) or the iNZight project (Wild & Inzight Team, 2023).

Fostering statistical and data literacy is not a simple concern, as the theoretical models and definitions above consider them as complex constructs that require not only a set of basic skills but also higher-order cognitive skills (e.g., interpretation, critical thinking), advanced behavioral skills, and dispositional demands. It is therefore assumed that the development of statistical and data literacy requires time and must begin in the early years of schooling (English, 2013; Franklin et al., 2007; Gal, 2004; Sharma, 2017). Approaches that are considered suitable for promoting statistical and data literacy in schools are the orientation towards elements of the statistical inquiry process (PPDAC, Wild & Pfannkuch, 1999) or statistical problem-solving process (Bargagliotti et al., 2020; Franklin et al., 2007), which should be actively practiced by the students (Bargagliotti et al., 2020; Garfield & Ben-Zvi, 2007; Petocz et al., 2018), preferably in real-world scenarios (Pfannkuch, 2018). Furthermore, research on teaching and learning statistics indicates that methods of cooperative learning, use of technological tools, project-based learning, model-based learning, and use of context-based real data may all be particularly suitable for the promotion of these competencies (Aziz & Rosli, 2021; Bargagliotti et al., 2020; Ben-Zvi & Makar, 2016; Garfield & Ben-Zvi, 2007).

The current review

As described in the theoretical background, not only the definition of the construct of statistical and data literacy has expanded in recent years, but also the curricular development, particularly in STEM subjects, as well as the technological progress. This has led to an increased focus on fostering statistical and data literacy within school education and thus also to increased research in this area. Considering the broad and varied definitions of these concepts and the diverse curricular approaches proposed, the open question here is how statistical and data literacy is operationalized and which components of statistical and data literacy have been investigated and/or promoted in international empirical research in this area to date. Existing literature reviews in this field have predominantly focused on adults and higher education, explored single aspects in isolation, conducted theoretical examinations of statistical literacy, or concentrated on other related concepts (Aziz & Rosli, 2021; Braun & Huwer, 2021; Chew & Dillon, 2014; Francois et al., 2020; Schüller & Busch, 2019; Sharma, 2017; Shaughnessy, 2007). However, there is a noticeable gap in comprehensive reviews that focus specifically on K-12 STEM education, considering the full range of metrics and instructional strategies for statistical and data literacy. To meet this gap, we formulated two research questions (RQ1 and RQ 2) with a specific focus on integrating empirical research on statistical and data literacy within the context of K-12 STEM education.

Understanding the metrics used to measure students' statistical and data literacy is crucial for identifying how these competencies are assessed and what aspects are prioritized in educational research. By examining the metrics, we can gain insights into the cognitive, affective, and behavioral components emphasized in current studies.

RQ1. Which student metrics are empirically investigated in studies related to statistical and data literacy in K-12 STEM education?

Identifying effective instructional strategies is essential for informing curriculum design and teaching practices. RQ2 therefore seeks to uncover which pedagogical approaches have been tested and proven successful in fostering statistical and data literacy among K-12 students, providing evidence-based guidance for educators.

RQ2. What instructional strategies to promote statistical and data literacy in STEM school education have been empirically investigated and which have been found to be effective?

Literature syntheses, particularly systematic reviews, provide a comprehensive overview of the current state of research and moreover, can help identify key authors in the field. The expertise of these key authors can be leveraged to further explore the future trajectory of statistical and data literacy in K-12 STEM education. Therefore, we have incorporated a third research question into our study. We will address this question through an explorative survey of the authors of the papers included in our review, leveraging their insights to gain a more profound perspective.

RQ3. How do key authors in the field perceive the future of statistical and data literacy in K-12 STEM education with respect to their theoretical conceptualization and strategies to best promote its teaching in the classroom?

Methods

To address the three research questions, we initially conducted a comprehensive systematic literature review (RQ1 and RQ2), followed by a written survey sent to the first and last authors of the included studies (RQ3).

Systematic review

For the systematic review, we closely followed the PRISMA 2020 guidelines (Page et al., 2021) in terms of planning, conducting, and scientific reporting of our work. A protocol describing the RQs and planned methodological procedures was pre-registered on the Open Science Framework (Friedrich et al., 2021). If it became necessary to deviate slightly from our original planning, it is transparently described in the method and discussion sections. Results of the systematic search and data reduction can be found in the PRISMA flowchart (Fig. 3).

Systematic literature search

We conducted a systematic literature search in September 2021, using equivalent search strings on PsychInfo, ERIC, and Web of Science. The search string was compiled in joint meetings involving the entire interdisciplinary review team, which consists of experts in mathematics and science education as well as educational psychology. As scoping searches had revealed that the range of terms used to describe the constructs of statistical and data literacy is very broad, we chose to use a wide search term for the search in titles, abstracts, and keywords (Table 1). First, we formed two categories of terms: (i) terms closely related to "data" and "statistical", and (ii) terms closely related to "literacy". In the search string, each term from category (i) was paired with every term from category (ii) (in both orders), and all possible combinations were connected using the Boolean operator OR. Four additional terms that are frequently used in curricula were included (for the complete search term, please see supplementary materials 1). Depending on the database, we used filters to reduce the amount of research that did not fit within our scope. A literature management program was used to merge all records and remove duplicates.

Eligibility criteria, article screening, and data extraction

To include very recent, high-quality studies that had not yet been published in journals, we additionally searched for relevant papers in proceedings of scientific conferences. We set the following eligibility criteria for conference inclusion: (i) the conference must be international (involving at least two countries) and (ii) have taken place between 2019 and 2021, (iii) the abstracts must be

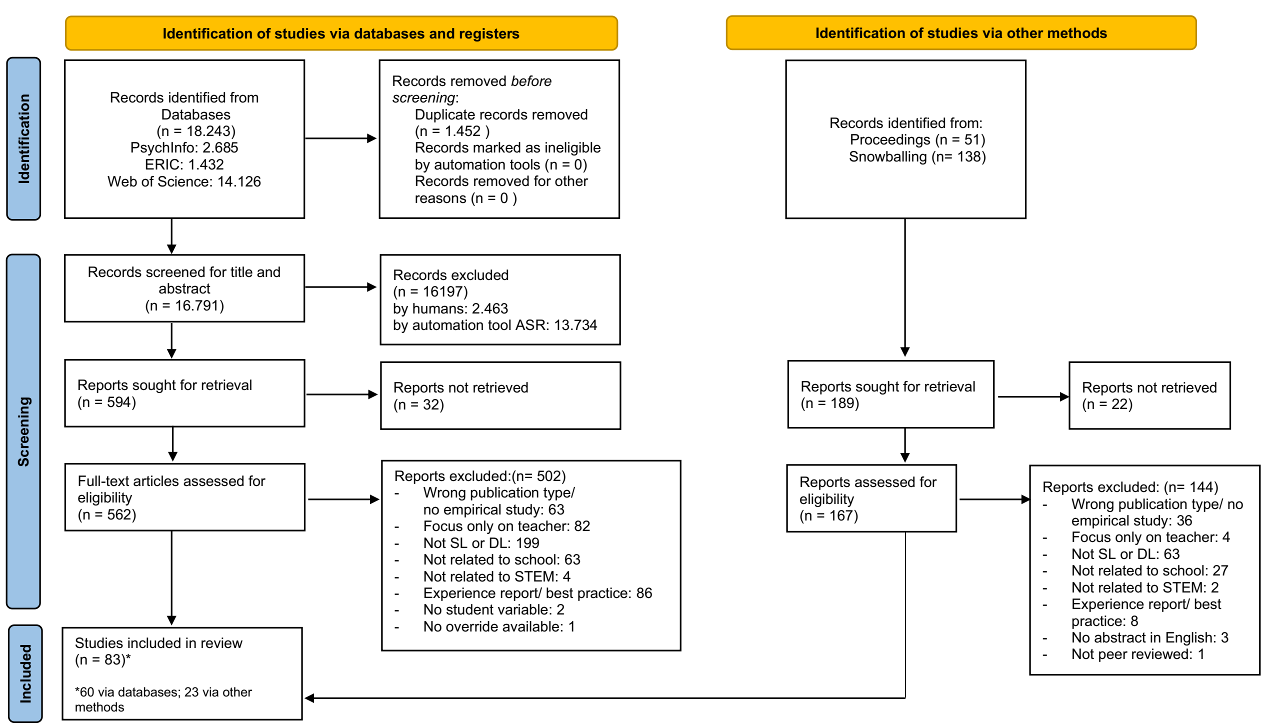

Fig. 3 Flowchart of the selection process (adapted from Page et al., 2021). Note. Systematic literature search conducted in September 2021 with no restrictions regarding time frame

Table 1 Composition of search string

Category 1		Category 2
Data Statistic* Quantitative Probabili* Data oriented Data based	+	Literacy Literate Reason* Competen* Think* Understand* Comprehen* Argu* Mastery
OR		
Data and chance Chance and data Uncertainty and data Data and uncertainty		

The asterisk (*) is used as a truncation symbol. It allows for the search of different word forms and endings of a term. For example: statistic*: This search includes terms like "statistic," "statistics," "statistical," and "statistically." By using the asterisk, we could ensure that all relevant studies using various forms of a term were included in our search

published in English, (iv) the proceedings must include full papers, (v) that were peer-reviewed, and (vi) that are related to education or STEM education. A list of conferences selected in this way can be found in supplementary materials 1.

Eligibility criteria were established to select articles that would be relevant to our research questions. To be included, articles must: (i) have an abstract in English, (ii) present original research, (iii) have a scope of investigation related to K-12 education, such that student metrics were assessed, (iv) include work related to STEM domains, (v) be peer-reviewed, and (vi) authors attribute their empirical research to either "statistical" or "data literacy" in the full texts. We chose this last eligibility criterion to avoid having to decide ourselves what falls under statistical and data literacy. Instead, we aimed to obtain a selection of studies that represents what empirically working researchers in the field cumulatively understand by these terms. If the main text was written in a language other than English, the eligibility criteria were assessed with the help of researchers who were proficient in the respective language. This concerned six articles: two in Portuguese, two in Turkish, one in Chinese, and one in Slovenian. There were no restrictions regarding the time frame for including articles.

Research syntheses, theory papers, and secondary analyses were not included in this systematic review. Articles dealing with learning in the context of kindergarten, vocational training, higher education, or informal education were excluded. The same holds true for articles related to education in non-STEM school subjects. Since STEM is sometimes defined in slightly different ways, we choose to include the following subject areas: Computer Science, Biological and Biomedical Sciences, Engineering, Mathematics and Statistics, and Physical Sciences (Chemistry, Physics, Geology, and Astronomy). Articles that were not peer-reviewed were excluded. To maintain a comprehensive scope in our search, we refrained from imposing the constraint that the terms *statistical* and *data literacy* must be explicitly present in the title, abstract, or keywords. However, it was a prerequisite for final inclusion that at least one of these terms be addressed within the entirety of the text.

In the first step, the titles, abstracts, and keywords of the articles found by the systematic search were screened to identify those articles that were most likely to meet the inclusion criteria. To increase the efficiency of the screening process and reduce human error, we used the open-source software ASReview (Utrecht University, 2021), which applies active learning to prioritize a set of uploaded articles (van de Schoot et al.,2021). After a short training phase with a few selected articles, the program presents one abstract after the other from the uploaded set to a rater, who must decide whether the article should be included. With each decision, the algorithm learns and the order in which the articles are presented to the rater is adjusted so that the most relevant articles are pushed to the top. We uploaded all article information into ASReview and two independent raters used the program individually to rate the articles. Both raters pre-trained the algorithm using the same pre-selected articles (10 eligible and 10 ineligible articles) in identical order. The termination criterion was reached as soon as fewer than five in 100 presented articles were deemed eligible. Rater 1 screened 2478 abstracts and Rater 2 2893 abstracts before reaching the termination criterion. It turned out that 723 abstracts had been screened by only one of the two raters because ASReview had not presented completely identical article sets to both. Therefore, each rater finally evaluated all abstracts from this non-overlapping set of articles that had not been presented to them by the program. Interrater reliability for all screened articles was moderate (Cohen's $\kappa = 0.568$, $p < 0.001$). The abstracts for which the raters had made different decisions were presented to a third person from the review team, who ultimately decided on inclusion or exclusion. Subsequently, the full texts for all included articles were uploaded into the software MAXQDA (VERBI Software, 2021); the full texts for the selected conference proceedings were also included. Screening of the full texts, which closely examined whether the articles met the inclusion criteria, was done in duplicate (interrater agreement $\kappa = 0.888$, $p < 0.001$). Disagreements were resolved by the review team. The reasons for the exclusion of articles during full-text screening were noted for each paper (see Fig. 3).

All 83 papers (60 studies identified via databases and registers, 23 via other methods) included in the final selection were further processed in MAXQDA. The papers were divided between three independent researchers, who highlighted, annotated, and extracted information relevant to the research questions. To support this, a framework was used to organize and code the extracted information (Table 2). Finally, backward, and forward snowball searches were performed for each study finally included during data extraction.

Survey

Systematic reviews compile and report the current state of research; however, they are unable to address questions not already explored in the existing literature. To go beyond a mere description of the current state and provide a forward-looking perspective, we aimed to enhance our review with a future vision. To achieve this, we proactively contacted the first and last authors of each article included in this review and send them a link to a survey implemented by means of the software Unipark (Tivian, 2021). We regarded these individuals as experts in statistical and data literacy within K-12 STEM education. While alternative approaches were possible, this method was deemed the most equitable for assembling an international cohort of experts believed to possess profound knowledge in the areas pertinent to our research questions. Their insights were considered valuable for projecting future developments in this field. In the questionnaire, we presented them with the pivotal findings of our review and solicited their insights through five open-ended questions, focusing on their vision regarding the role of K-12 education in fostering statistical and data literacy (see Survey, supplementary material 2). The questionnaire commenced with a brief introduction to the project, outlining the roles of the project leaders and the purpose of the survey. This was followed by a consent form. Upon granting their consent to participate, respondents were directed to a concise written introduction to the study. Thereafter, the five open-ended questions followed. Subsequently, additional data on teaching experience (K-12 and university level) and research experience in the field of statistical and data literacy were collected. Participation in the questionnaire was voluntary and anonymous. However, to encourage the sharing of personal insights and to acknowledge their contributions, participants were given the option at the conclusion of the questionnaire to forego anonymity, allowing for the possibility of being explicitly cited in our article.

The responses to the open-ended questions underwent exploratory qualitative analysis, with particular attention given to identifying areas of significant consensus and highlighting any innovative ideas shared. The responses provided by the participants were brief in most cases, typically consisting of one or two statements per question. Given the concise nature of these answers, we employed a straightforward approach for analysis. First, we extracted all the statements related to each question. We then grouped similar statements together. Subsequently, we counted the frequency with which each theme was mentioned to assess the level of consensus among respondents for each survey question. To ensure clarity in the review, we summarized each group of statements. In cases where a particularly noteworthy or representative statement was made, we included it as a direct quote in our results section.

Results

The results are presented in the following sub-sections. Firstly, an overview of the included studies is provided, detailing the year of publication, the STEM field, and the theoretical conceptualization of statistical and data literacy. Secondly, the findings related to the three research questions are presented.

Overview of the included studies

Table A1 (supplementary material 3) provides an overview of all 83 articles that focused on statistical or data literacy of students in K-12 STEM education. The

Table 2 Data extracted from eligible studies

	Information to be extracted
General	STEM domain (s) (Multiple; Nature Science; Mathematics & Statistics; Computer Science; Geography)
	School career (Pre-school; Primary school; Lower secondary education; Upper secondary education)
	Used Term, Definitions and Theoretical Framework of Statistical and Data Literacy
	Publication Year and Country
Research questions	
RQ 1: Metrics	Measured concepts (Cognitive Components; Behavioral Components; Affective Components)
RQ2: Instructional Strategies	Applied instructional Strategies

analysis of publication years showed that statistical and data literacy have been studied empirically since 2000, with a significant increase in published studies towards 2021. While papers focused on statistical literacy were found over the entire period, papers addressing data literacy were found once in 2006 and then steadily increasing from 2012. This reflects the development and refinement of terms discussed in the theoretical background.

Most studies focused exclusively on mathematics (75%). The other studies focused on multiple STEM domains (16%), natural sciences (physics, chemistry, and biology; 6%), or computer science (3%). They were conducted with students in all three school levels: mostly in lower secondary education (62%), followed by primary education (21%) and upper secondary education (17%). The studies collected data from students all over the world, including countries in Africa, North and South America, Asia, Australia, and Europe. The majority were in mostly English-speaking regions, especially Australia. Sample sizes varied widely across studies, ranging from $N=3$ (Bergner et al., 2020) to $N=7377$ (Mills & Holloway, 2013).

Most studies employ the term statistical literacy to frame their investigations. Of the 83 studies, 67 used the term *statistical literacy*, 14 the term *data literacy,* and two studies used both terms describing their research. Remarkably, despite the presence of theoretical conceptualizations and definitions regarding statistical and data literacy as highlighted in the theoretical background, empirical studies only partially align their research with or incorporate these frameworks (see Table A1, supplementary material 3). Thirty-two of the reviewed articles only refer to the terms statistical or data literacy—typically in the introduction, conclusion, or implications sections—without offering a definition or situating their work within the existing literature on the topic. These papers provide a theoretical framework based on the concept of statistical or quantitative reasoning, or adopt a different focus, such as data exploration, data skills, computational thinking, storytelling with data, statistical thinking, understanding of statistical concepts, statistical modeling, informal inference, understanding of variation, probabilistic thinking or (critical) mathematical literacy. Some papers also link their research to the PPDAC cycle by Wild and Pfannkuch (1999) or school curricula, e.g., the Guidelines for Assessment and Instruction in Statistics Education (GAISE) report (Franklin et al., 2007), NCTM Principles and Standards for School Mathematics (NCTM, 2000) or the disciplinary practices of Next Generation Science Standards (NGSS Lead States, 2013). Quite a large number of articles (43) that can be assigned to the concept of statistical literacy provide a definition or refer to one from a small group of influential authors: Wallman (1993), Gal (2002, 2004), Watson (1997), Watson (2006), or Watson and Callingham (2003). The majority of these articles relate to a hierarchical concept of statistical literacy and use Watson's three-tiered hierarchy Watson (1997) or Watson and Callingham's (2003) hierarchical construct of statistical literacy to frame their work. Also, assessments of the global construct statistical literacy have primarily utilized the instrument developed by Watson and Callingham (2003) or adaptations of it. In addition, five papers describe statistical literacy within the Statistical Literacy Competency Construct (SLCC) of Kuntze et al. (2008), who developed a hierarchical competency model for using models and representations in statistical contexts as a subcomponent of statistical literacy, which was in turn based on definitions of statistical literacy by Gal (2004), Wallman (1993), Watson (1997), and Watson and Callingham (2003). Beyond that, a notable amount of 11 papers mainly focusses on the definition of Gal (2002, 2004), who defined statistical literacy through two components—a knowledge and a dispositional component.

Data literacy is addressed less frequently in the reviewed literature and working definitions are drawn from a wider range of publications. Ten papers either provide their own definition of data literacy or refer to definitions provided by other authors. Swan et al. (2013), Vahey et al. (2006) and Vahey et al. (2012), for example, use a definition of data literacy that includes three different aspects: formulating and answering data-based questions, using appropriate data, tools, and representations, and developing and evaluating data-based inferences and explanations. Other authors (e.g., Gebre, 2018; Tedre et al., 2020; Wilkerson et al., 2021) go beyond this definition of a competence-oriented perspective of data literacy, by adding an empowerment-oriented perspective. They incorporate aspects of data privacy, security, ownership, community involvement, social justice, and power dynamics related to data access and usage, respectively, emphasize the ethical use of data literacy (also denoted as data agency). Furthermore, Gebre (2018) concurs with Gould (2017) that data literacy encompasses comprehension as both producers and consumers of data. Furthermore, it is noticeable that six of the 16 studies focus on research on the handling of complex data(sets) (*big data*).

RQ1. Metrics

For RQ1, we examined the included studies to identify which metrics researchers employed to investigate statistical and data literacy among students. The results are differentiated according to cognitive, affective, and behavioral components, in line with the distinction described in the theoretical background (see Figs. 1 and 2). It should be noted that some studies included metrics for a single one of the three components, while in others multiple components were addressed (Table A1, supplementary material 3).

Cognitive component

This section reports the findings of 63 studies examining cognitive student metrics associated with statistical and data literacy in four sub-sections that emerged as part of the analysis. On the one hand, we found articles that indicate to assess statistical or data literacy as global and overarching constructs, not especially emphasizing the theoretical grounding of the metrics used. The second sub-section presents the results of the studies focused on students' ability to understand data representations in statistical contexts. The third sub-section includes studies that addressed students' understanding of probability and probabilistic reasoning and, the fourth that measured other specific individual cognitive components that the respective authors associated with statistical and data literacy.

Statistical and data literacy as global constructs

In 17 of the identified papers, various methods were used to assess the overarching (cognitive) constructs of statistical or data literacy. Most of them drew upon the theoretical models of Watson (1997) and Watson and Callingham (2003). Seven of these papers are directly affiliated with Watson and Callingham's research group, focusing on the description and empirical testing of the hierarchical structure of the six-level model; using Rasch scaling (Callingham & Watson, 2005; Watson & Callingham, 2003), its progression during education (Callingham & Watson, 2017; Watson et al., 2005, 2006), and the relationship between teachers' knowledge of pedagogical content in statistics and students' statistical understanding (Callingham et al., 2016). An in-depth analysis of students' understanding of basic statistical concepts—such as probability and variation—is presented by Watson et al. (2007) using a combination of quantitative metrics and qualitative data. Watson et al. (2006) report longitudinal analyses on the development of statistical literacy over time, disentangling grade effects from cohort effects. While the latter may present challenges in interpretation, grade effects demonstrate a notable enhancement in performance, particularly between grades 3 and 5. Comparable effects are documented by Callingham and Watson (2017), who additionally identified a plateau phase in grade 9. Concerning the attained proficiency standard within the context of the six-level model, they further observed a tendency to plateau at level four. Beyond this core of research from the Australian research group led by Callingham and Watson, their model is also mentioned in a set of papers from other groups and countries that focus on different themes, such as student and teacher metrics and instructional formats. Based on metrics derived from these theoretical approaches, some studies investigated instructional approaches and assessed their effectiveness by evaluating learners' statistical literacy as a global construct (Aksoy & Bostan, 2021; Cakiroglu & Güler, 2021; Koparan & Güven, 2014b; Merriman, 2006; Yolcu, 2014). Merriman (2006) used a self-developed test and classified the results according to Biggs and Collis' (1982) SOLO-Taxonomy. The results support the importance of literacy skills (operationalized as English ability from the exams at the end of year 9) (Merriman, 2006) and mathematical knowledge as subcomponents of statistical literacy like mentioned by Gal (2002, 2004). Also using the SOLO taxonomy, Koparan and Güven (2014a) investigated the statistical thinking levels of 90 middle-school students and found a significant relationship between the grades and levels of statistical thinking. The study of Mills and Holloway (2013) focused—under the umbrella of statistical literacy—on evaluating student performance related to the data and chance content domain by using data from the TIMSS 2007.

Three studies explored—within the *Thinking with data* project—data literacy as a global construct by applying self-developed context-based tests in cross-disciplinary contexts (Swan et al., 2013; Vahey et al., 2006, 2012). In these tests, middle-grade students were asked, for example, to perform proportional reasoning and summarize and critically evaluate given arguments and positions with the help of data (Swan et al., 2013; Vahey et al., 2006, 2012).

Representations in statistical contexts

Another group of 8 studies particularly focused on assessing student's ability to read, understand, and critically interpret data representations. Aoyama and Stephens (2003) referred in their study to a Japanese model—introduced by Kimura (1999)—and investigated the ability of 17 Japanese grade 5and 38 Japanese grade 8 students to interpret graphs quantitatively and qualitatively. Classifying the results according to Biggs and Collis' (1982) SOLO-Taxonomy, they found that no student was able to argue on the intended highest level of the SOLO-Taxonomy. In a small qualitative study analyzing the structure of statistical literacy in six grade 8 students, Utomo (2021) referred to a heuristic model by Schield (2004) that focuses on the ability to read and interpret tables and graphs. Utomo (2021) revealed differences in data processing and communication, based on the level of mathematical abilities. Also, the results of Sharma's (2005) study showed that students had problems reading tables, which could be due to linguistic and contextual problems. Table and graph interpretation was at the forefront of the studies of Oslington et al. (2020) and Frischemeier (2019, 2020). Frischemeier (2019, 2020) investigated the progress of grade 4 students through between-group comparisons. His studies revealed that it was possible to foster statistical reasoning of even young students with the help of suitable digital representations, as well as positive attitudes towards statistics. Similarly, in a quasi-experimental pre-test versus post-test study, involving 72 students in grade 4, Ganesan and Leong (2018) showed that the students were able to examine multiple representations of the same data. The results of an interview study by Wilkerson et al. (2021) with older students—aged 15–18 years old—led to comparable results.

A corpus of five papers from a German research group dealt with the use of models and representations in statistical contexts, which they introduced as a subcomponent of statistical literacy (Kuntze et al., 2008, 2015; Lindemeier et al., 2007; Sproesser et al., 2014, 2018). They describe the development from a three-step (Lindmeier et al., 2007) to a five-step model (Sproesser et al., 2018). Moreover, Sproesser et al. (2014) report multilevel analyses on two data sets of 503 grade 8 and 535 grade 9 students, where they observed significant correlations between statistical literacy and cognitive abilities (verbal and non-verbal), as well as domain-specific knowledge and (on a class level) students' socioeconomic status. Notably, neither this particular model nor that proposed by Kimura (1999) has yet received significant recognition within literature.

Understanding probability and probabilistic reasoning

Seven studies addressed students' understanding of probability as a specific cognitive disposition of statistical or data literacy. These studies focused on assessing students' difficulties in understanding probabilities. English and Watson (2016) addressed the development of probability reasoning in 91 grade 4 students, while Watson and Kelly (2007) did the same with 69 students in grades 3–13. In these studies, the students investigated the relationship between the theoretical probabilities and relative frequencies of outcomes for random processes across different sample sizes. The researchers found that students developed a deeper understanding of probability when they realized that the relative frequencies stabilize around the theoretical probability as the sample size increases. Nacarato and Grando (2014) researched the development of probabilistic language and thinking in 12 students aged 10–12 years old. They found that the meaning of words from a probabilistic vocabulary needs to be clarified since many of them were known and used by students in daily but not probabilistic contexts. Almeida (2018) investigated the understanding of probability in 376 students in grades 6–9 of elementary school in Brazil by applying two items of a statistical literacy test instrument (from Watson & Callingham, 2003) and compared the results with Australian students as a reference group. The results showed that both groups had the same difficulties in understanding content that involves probability. Begolli et al. (2021), Primi et al. (2017), and Saidi and Siew (2022) investigated the role of proportional reasoning in attaining probabilistic knowledge. The main finding of their studies is that proportional, respectively, probabilistic reasoning is crucial for understanding probabilities.

Individual cognitive metrics

Another cognitive metric—to which four studies were devoted—is understanding variation (English & Watson, 2018; Fielding-Wells, 2018; Watson & Kelly, 2008; Watson et al., 2019). Garfield and Ben-Zvi (2007) consider understanding variation as a key to statistical reasoning and advocate integrating a focus on it from the earliest grades. English and Watson (2018) propose a framework with four components: working in shared problem spaces between mathematics and statistics, interpreting and reinterpreting problem contexts and questions, drawing informal inferences, and interpreting, organizing, and operating on data in model construction. Fielding-Wells (2018) demonstrated that through statistical modeling tasks using a combination of self-generated dot plots and dot plots generated by software (TinkerPlots), students aged 10–11 years old ($N=26$) could develop a conceptual understanding of variation. Watson et al. (2019) developed an approach by having students in grade 3 learn the principle of variation in a "STEM-related context where variation occurs in an easily measured and realistic fashion" (p. 1). The developed teaching sequence was successfully implemented in three schools and well over three-quarters of the 70 students achieved a passing grade. Watson and Kelly (2008) examine the development of understanding of three basic concepts of statistical literacy: sample, chance, and variation among students in grades 3, 5, 7 and 9. The results of the study show an increase from grade 3 to 5 that may represent a natural increase in language proficiency and emphasize the relevance of literacy skills within the promotion of statistical literacy.

Guler et al. (2016) investigated the critical views of 8 grade students regarding statistical data presented in newspaper articles. They found that students showed deficits in critically analyzing these articles and attributed this to a lack of relevant prior experience. Also, Budgett and Rose (2017) found that many students lacked the broad contextual knowledge required to critically evaluate statistically based reports. This aspect was also studied by Carvalho and Solomon (2012) with 533 Portuguese elementary school students and by Lee et al. (2021) in two classes of 5 and 6 grade students. They found that the outcomes of students' discussions were largely dependent on earlier experiences. Höper et al. (2021) measured students' awareness of where, how, and why data are collected and processed. In another four articles, the statistical modeling abilities of students were assessed (English & Watson, 2018; Makar & Allmond, 2018; Mendonca & Lopes, 2011).

Three studies investigated students' abilities concerning the concepts of sampling and bias (2000b; Watson & Kelly, 2005; Watson & Moritz, 2000a), and two more assessed the ability to use and understand statistical terminology (Watson & Moritz, 2003; Watson et al., 2014). Another two studies (Liu & Lin, 2011; Liu et al., 2010) were concerned with assessing students' misconceptions regarding correlation, by using a self-developed two-tier diagnostic instrument. Two studies develop an assessment for students' statistical reasoning in descriptive statistics (Chan et al., 2016; Saidi & Siew, 2019) and empirically demonstrated its consistency with respect to the selected constructs (describing data, organizing, and reducing data, representing data, and analyzing and interpreting data).

Finally, several concepts were investigated in only one study. These included data wrangling (Jiang & Kahn, 2020), data exploration (Irish et al., 2019), statistical problem-solving (Cotic, 2009) and, understanding certain coding concepts (Mosquera et al., 2020). Tedre's et al. (2020) study focuses on the component of data protection and data ethics. They used a mixed methods approach to explore children's attitudes and knowledge about machine learning (ML) concepts in everyday life and their use of personal data. They found that students generated data with little understanding, knowledge, or awareness about how and why ML services collect and use data, despite a high level of ML service use.

Behavioral component

Metrics, related to the behavioral component are investigated within 17 studies. Most of them focus on the PPDAC inquiry cycle or individual steps of it. Some studies investigated students' competent use of statistical data (e.g., Andre et al., 2020; English & Watson, 2015a; Watson & English, 2017, 2018). English and Watson (2015a) showed in their study, that participation in two consecutive lessons, in which 115 students in grade 4 (independently) collected and analyzed measurement data, enabled them to develop an understanding of the meaning of variation and to transfer it to other contexts. Birk and Frischemeier (2022) examined fourth grade students conducting statistical mini-projects on group comparisons. Results showed that these students not only take account of local perspectives when interpreting the data, but also use a more global perspective on the represented data. In a study of students' understanding of statistical data and its relevance, the entire process of finding, presenting, and visualizing statistical data was integrated into a biology inquiry project (Gebre, 2018). Findings revealed a difference between the narrow scope of students' written description of data and their broader use of data in their projects, and this was attributed to the fact that students' experience with data is limited to laboratory experiments and worksheets with structured data. Kochevar et al. (2015) demonstrated that students encountered challenges when making quantitative comparisons and presenting multiple measures and data sources. In contrast, they found that qualitative descriptions of data, data visualizations, and formulating claims related to the research question, all came naturally to them.

Furthermore, there are studies accomplishing several steps of an investigative activity such as problem posing (English & Watson, 2015b), pursuing meaningful statistical questions (Watson & English, 2015), identifying relevant sources of data beyond the given (Swan et al., 2013), creating new measures by norming data sets (Vahey et al., 2006, 2012), complex data handling (Wolff et al., 2019), critically evaluation of the statistical process (Pfannkuch, 2005) or, skills in creating data visualizations (Chin et al., 2016; Hourigan & Leavy, 2020; Kahn & Jiang, 2020). The study of Hourigan and Leavy (2020), about statistical association, indicates problems with creating graphs that adequately communicated covariation.

Affective component

About 16% of the papers reviewed in this study measured affective learner metrics such as attitudes, interest, anxiety, self-efficacy, motivation, and engagement in the context of statistical and data literacy. In this context, these aspects were not explicitly classified as subcomponents but regarded as relevant within the broader framework of statistical and data literacy. Affective learner metrics were measured in addition to cognitive outcomes in some studies (e.g., Saidi & Siew, 2022), whereas others focused exclusively on promoting affective outcomes (e.g., Bergner et al., 2020) or on developing instruments to measure them (e.g., Carmichael et al., 2010a).

Five studies report global affective measures, referred to as students' *attitudes* towards statistics or statistics education (see for example, Malaspina and Malaspina's (2020) semi-structured interview study). Carvalho and Solomon (2012) interviewed middle-school students about their attitudes toward statistics before the students participated in tasks in dyads to discuss statistical problems. The researchers found that high-performing dyads with positive attitudes towards statistics, confidence in their abilities, and relatively high understanding of the task context, engaged in the richest discussions. Some studies employed self-constructed questionnaires to measure students' attitudes toward statistical and data literacy. For instance, Irish et al. (2019) used such questionnaires to assess students' appreciation for learning data processing practices. Frischemeier (2020) utilized a three-point attitude scale to evaluate students' attitudes toward various aspects of a recently completed statistics learning unit. Saidi and Siew (2022) used a standardized self-rating instrument—related to perceived cognitive competence, value, difficulty, affect, effort, and interest (SATS; Saidi & Siew, 2019)—to assess the attitudes of middle-school students toward statistics. The results indicate that most students held a favorable view of statistics, and there was a small but significant positive correlation between positive attitudes and performance. Additionally, Saidi and Siew (2022) found—by administering the Statistics Anxiety Scale (SAS; Vigil-Colet et al., 2008)—that the students experienced moderate levels of statistics anxiety which were significantly negatively correlated with statistical reasoning and positive attitudes. A multiple linear regression analysis revealed that only a small portion of the variability in overall statistical reasoning could be explained, with important predictors being the subscales Interest (SATS), Interpretation Anxiety (SAS), and Value (SATS).

Eight of the reviewed papers included the investigation of learners' interest in statistics and data (). Mendonca and Lopes (2011) explored learners' interest and enthusiasm qualitatively, by observing and interviewing learners at different stages of an inquiry-based approach to learning statistics. Bergner et al. (2020) also conducted semi-structured interviews during an instructional approach that brought together embodiment and the use of technology to promote statistical literacy. The study revealed a high level of interest in the embodied context of the approach, which was found to be positively associated with interest in the statistical data and analysis. The remaining studies used quantitative approaches to measure interest. Dierker et al. (2017), for example, asked learners how interesting they had found a completed statistics course to be compared to other courses in a program, and how interested they would be in using statistics in the future. Irish et al. (2019) asked more specifically about students' interest in different practices with data (e.g., collecting or graphing data) in their rating items. They found that participants were especially interested in working with their own data. To define interest in statistical literacy as a multidimensional construct, Carmichael et al. (2010a) developed a theoretical model in the form of a taxonomy grid, with an interest axis (importance interest, reflective interest, curiosity interest) and a content axis differentiating between specific topics within statistics education (e.g., chance, inference) and activities and contexts within statistics education (e.g., use of technology). On this basis, they developed a multi-item, self-assessment scale for middle-school students: the Statistical Literacy Interest Measure (SLIM). The revised and final version of the SLIM was validated on two samples of students. A modest correlation was found between the SLIM scores of middle-school students and their general interest in mathematics. Moreover, as in an early version of the interest scale (Carmichael & Hay, 2009b), the authors found slight gender differences (e.g., boys expressed significantly more interest in problem-solving than girls). Furthermore, Carmichael and Hay (2009b) found a slight indication that younger students were more interested in statistical contexts related to games.

In their study, Carmichael et al. (2010b) found a further affective metric—students' self-efficacy—to be important in the development of interest in statistical literacy. Carmichael and Hay (2009a) developed a 9-item self-assessment instrument for measuring students' self-efficacy in the area of statistical literacy (the Self-efficacy for Statistical Literacy Scale, SESL) and showed that it was reliable using a sample of middle-school students. The items assess how competent students feel when they perform certain tasks related to averages, chance, graphs, inference, and sampling, within the context of school and media. The authors showed that self-efficacy in statistics exhibits a moderate positive correlation with self-efficacy in mathematics in general. The SESL was also used in the study by Carmichael et al. (2010a) where the authors found that interest and students' SESL scores were moderately and positively correlated. Carmichael et al. (2010b) replicated this finding and described a quadratic relationship between SLIM and SESL scores. Moreover, they revealed that the positive effect of teacher-estimated student achievement in mathematics on students' interest was entirely mediated by their perceived self-efficacy. In a similar vein, Dierker et al. (2017) assessed the extent to which learners felt more confident in different data-related tasks having completed a statistics course. To assess their engagement in the course, the authors asked the students to self-assess how much effort and work they had put into it, both in absolute terms (rating) and compared to other courses (dichotomic). As a further affective metric, students' motivation was estimated in two of the reviewed studies (Bergner et al., 2020; Cakiroglu & Güler, 2021) by qualitatively analyzing semi-structured interviews that were conducted with the students after an intervention.

RQ2. Instructional strategies to promote statistical and data literacy

Thirty-two studies reported empirical results regarding instructional strategies to promote statistical or data literacy in STEM education (Table A2, supplementary material 4). The majority of these focused exclusively on mathematics (59%) or addressed multiple STEM subjects (25%). Three studies (9%) focused exclusively on computer science and two (6%) on science. Sample sizes varied widely, ranging from only three participants in a case study with a design-based research approach (Bergner et al., 2020) to 606 participants in a quasi-experimental study (Vahey et al., 2012). The following section presents the results in a structured way, highlighting the key features that emerged from the analysis of the instructional strategies.

Exploiting meaningful data and authentic problems

A common feature of most approaches was the integration of authentic, open-ended problems and real-world data that students find meaningful. These features were frequently listed as typical of problem-based (Andre et al., 2020; Carvalho & Solomon, 2012; Swan et al., 2013; Vahey et al., 2012) or project-based learning approaches (Dierker et al., 2017; Koparan & Güven, 2014b). A large number of studies used complex, real-world data sets from a variety of sources. These included, for example, public global socioeconomic data (Jiang & Kahn, 2020), marine science data (Kochevar et al., 2015), water availability and use data (Swan et al., 2013; Vahey et al., 2006, 2012), public health data (Wilkerson et al., 2021), smart city datasets (Wolff et al., 2019), data on sustainable development of countries (Andre et al., 2020), data on a person's location (Höper et al., 2021), or data in media reports (Budgett & Rose, 2017; Merriman, 2006). Integrating authentic problems and real-world data has been shown to support the development not only of knowledge and (critical) understanding of statistical concepts (Andre et al., 2020; Carvalho & Salomon, 2012; Koparan & Güven, 2014b), but also relevant skills related to statistical and data literacy, such as students' ability to ask and answer data-based questions, create data visualizations, interpret data, and develop and evaluate data-based inferences (Kochevar et al., 2015; Merriman, 2006; Swan et al., 2013; Vahey et al., 2012). Furthermore, increases in students' confidence in data analysis and statistical skills, as well as positive effects on students' interest in data, have been found (Dierker et al., 2017; Kochevar et al., 2015).

Using digital tools

Many of the instructional strategies integrated digital tools. The use of specific data visualization and analysis software is effective in helping students deal with large and complex real-world data sets. Such software includes *Gapminder* (Andre et al., 2020; Jiang & Kahn, 2020), *Fathom* (Ganesan & Leong, 2018), *TinkerPlots* (Birk & Frischemeier, 2022; English & Watson, 2015a; Frischemeier, 2019), self-developed interactive app created using the *Shiny package in R* (Bergner et al., 2020), and *Excel* (Mendonca & Lopes, 2011), as well as novice-friendly *web-based interfaces* and *digital online platforms* (Cakiroglu & Güler, 2021; Kochevar et al., 2015; Wilkerson et al., 2021; Wolff et al., 2019). Instruction with digital tools for data visualization and analysis enabled learners—even at the primary and lower secondary school levels—to perform sophisticated statistical activities, such as filtering relevant data (Jiang & Kahn, 2020), analyzing and comparing data distributions concerning relevant distributional features like center, spread and shape (Birk & Frischemeier, 2022; English & Watson, 2015a; Frischemeier, 2019; Lee et al., 2021), observing variation in data (Watson & English, 2015), and examining patterns in data to support hypotheses (Kochevar et al., 2015) or make data-based decisions (Watson & English, 2017). In addition, teaching with digital tools has been shown to lead to higher levels of statistical reasoning compared to instruction without digital tools (Ganesan & Leong, 2018).

Three studies examined pedagogical approaches that integrated programming environments and software, such as *SAS Studio* (Dierker et al., 2017), *Python* (Höper et al., 2021; Mosquera et al., 2020), and *machine learning applications* (Tedre et al., 2020) to promote understanding of how technologies and services work, especially those that students use in their daily lives, even including data-driven machine learning and rule-based programming. Findings show that students evaluated their programming experiences as useful and rewarding and showed an increased affinity to further develop their programming skills (Dierker et al., 2017; Mosquera et al., 2020).

Applying methods of inquiry learning

Several studies used a statistical inquiry cycle to conduct data analysis projects and pursue meaningful statistical questions (Andre et al., 2020; Birk & Frischemeier, 2022; Frischemeier, 2019). Most studies referred to Wild and Pfannkuch's (1999) PPDAC cycle (or an adapted version of it), which consists of five phases: problem, plan, data, analysis, and conclusion.

Several studies examined the effectiveness of an instructional strategy in which students collected their own data. These included, for example, students recording their own physical activity (Bergner et al., 2020; Lee et al., 2021; Makar & Allmond, 2018; Watson & English, 2017), measures of their arm span or shoe size (Birk & Frischemeier, 2022; English & Watson, 2015a; Watson & English, 2015), or conducting surveys among the students' classes or families (Cakiroglu & Güler, 2021; Frischemeier, 2020). The overarching findings of these studies indicate that the data collection experience promotes critical thinking about data and supports students' abilities to interpret and understand data. Furthermore, positive effects on students' motivation were found, due to the perceived relevance and usefulness of data science (Bergner et al., 2020). Wolff et al.'s (2019) findings demonstrate that activities that provide the experience of collecting data can enhance students' critical thinking, even when the data are not collected in person.

Engaging students in games

A game-based approach to promote students' statistical and data literacy was investigated in four studies (Cakiroglu & Güler, 2021; Chin et al., 2016; Malaspina & Malaspina, 2020; Mosquera et al., 2020). These studies showed positive effects on students' enjoyment of learning, attitudes, and increased self-esteem regarding statistical or data literacy (Malaspina & Malaspina, 2020; Mosquera et al., 2020). Furthermore, positive effects on students' motivation, and partly on their statistical literacy achievement also, were reported (Cakiroglu & Güler, 2021).

Leveraging cross-disciplinarity

Eight studies used cross-disciplinary instructional strategies (i.e., addressing more than one STEM subject) to promote statistical and data literacy. One of these investigated the effectiveness of a cross-disciplinary approach (including mathematics, social studies, science, and English language arts) to promote data literacy compared to a non-disciplinary approach in a quasi-experimental design with independent groups (Vahey et al., 2012). Findings show that a cross-disciplinary approach led to increased levels of data literacy skills (e.g., creating and analyzing data-based arguments), indicating a general mechanism of transfer across the disciplines (Vahey et al., 2012).

RQ 3. A glance into the future: key insights from leading researchers

The articles included in the systematic review featured a total of 100 distinct first and last authors. We reached out to these individuals via email, inviting them to participate in our survey on the future of statistical and data literacy in K-12 STEM education. The survey yielded responses from 15 researchers, including internationally recognized experts in the field who provided their names: Ayse Aysin Bilgin, Rosemary Callingham, Joachim Engel, Rabia Karatoprak Ersen, Shian Jiang, Kristina Reiss, and Annie Savard. Among the 15 participants, 14 reported that they are still actively engaged in research related to statistical and data literacy. Seven participants have teaching experience in one or more STEM subjects within K-12 education, while 11 are or were in the past involved in STEM teaching at the university level.

The survey included brief summaries of our findings related to RQ1 and RQ2 of this review, which served as the context for the experts to answer the open-ended questions (see Survey, supplementary material 2). The first open-ended survey question pertained to the specific subcomponents of statistical literacy that ought to be targeted in future K-12 education: *Which subcomponents of statistical literacy should be addressed in future school education, and why? What minimum standards should be achieved from your point of view?* Overall, the authors concurred that a comprehensive set of subcomponents should be addressed; these included, for example, visual data representation, statistical concepts and analysis, critical evaluation, decision-making, critical thinking, integration of context knowledge, and statistical modeling. However, it was also suggested that in future statistical literacy education, there should be an emphasis on conveying statistics as an integrated concept. Associate Professor Rosemary Callingham, expert in mathematics and statistics education research and curriculum design at the University of Tasmania, succinctly articulated the main point: "One of the issues in school curricula with respect to statistics is that there is an emphasis on statistical skills, especially mathematical statistics, rather than on the meaning of the statistics" and "The pedagogical focus needs to shift from developing splinter skills […] to setting up and undertaking statistical investigations."

In response to our second question—*Which subcomponents of data literacy should be addressed in future school education, and why? What minimum standards should be achieved from your point of view?*—the respondents listed several skills that they considered important. In particular, the following terms were mentioned: data collection, communicating about data, (big) data management, data protection and storage, and ethical and legal considerations. One term that was frequently mentioned was *critical data literacy*, which according to Lynn English, a professor in mathematics education research and curriculum design at the Queensland University of Technology, can be defined as "knowing how to question data presented in various media, rather than just accepting the data as given" and should be integrated into the curriculum starting from elementary school.

In the third question, we explored the perceived relationship between statistical and data literacy*: In your opinion, how are statistical literacy and data literacy related, and how do they differ from further related competencies?* The respondents unanimously acknowledged a strong relationship between these concepts. Some of them perceived statistical literacy as the overarching concept. Others regarded data literacy as a fundamental prerequisite for statistical literacy, or as Associate Professor Ayse Aysin Bilgin at Macquarie University and current president of the International Association for Statistical Education (IASE) put it "without data literacy, statistical literacy is not sound."

Next, we sought opinions on how best to design future K-12 lessons to promote statistical and data literacy: *In your opinion, how should future K12-lessons be designed to promote statistical and data literacy? Should an interdisciplinary approach be taken, and which subjects should be included?* First, an interdisciplinary teaching and curriculum approach (across all STEM subjects and beyond) was deemed essential by all participants. Furthermore, the insights from the experts underscored the significance of project-based learning, relevance to real-life contexts, conceptual understanding, the integration of technology, and reflection on contemporary issues.

Finally, we inquired about strategies to enable educators to empower their students toward becoming proficient in statistics and data literacy: *In your opinion, how can we ensure that teachers in schools are appropriately trained and further educated in statistical and data literacy to prepare students for the demands of an increasingly data-driven society?* The experts proposed several key strategies to empower teachers, including the provision of professional development opportunities, offering improved statistic courses and workshops for pre-service and in-service teachers, fostering a positive attitude towards statistics, utilizing technology, focusing on deep understanding, forming partnerships and collaborations, creating high-quality instructional materials, establishing a better reward and recognition system, and integrating and reviewing curricula at various levels.

The experts conveyed that the implementation of these comprehensive adaptations, including the integration of new and continuously updated content into the curriculum, the utilization of interdisciplinary and innovative teaching methods, and the adjustment of teacher training and professional development, will necessitate substantial dedication, commitment, and support from their respective governments. Nonetheless, they were unanimous in their opinion that these efforts are crucial and will be rewarding for society.

Discussion

This systematic review demonstrates that statistical and data literacy in K-12 STEM education are receiving increasing research attention. A prevailing focus on statistical literacy is notably evident in the studies, while the number of studies addressing data literacy has been rising steadily since 2012. This trend aligns with the development of definitions and frameworks concerning statistical and data literacy, as described in the theoretical background.

Metrics

The empirical studies addressing statistical literacy predominantly draw upon definitions and conceptualizations by established authors such as Gal (2004), Wallman (1993), Watson (1997), and Watson and Callingham (2003), framing statistical literacy mostly as a hierarchical construct with a broad range of subcompetencies. Compared to statistical literacy, data literacy is less frequently discussed and defined in the existing literature, lacking a definitive and uniform definition. Consequently, a coherent understanding of both statistical and data literacy within K-12 STEM education is difficult to achieve. This inconsistency is evident in the varied usage of these terms in the reviewed research. We developed a three-component framework to unify the theoretical definitions proposed by various authors within a common structure of cognitive, behavioral, and affective components. This framework proved to be suitable for categorizing the metrics used to measure statistical and data literacy in the reviewed studies. The studies reviewed indicate that cognitive metrics of statistical and data literacy were predominantly measured., with only a few studies including behavioral and affective metrics, and even fewer exploring the relationships between them. Findings with regard to cognitive metrics underscore that a significant proportion of the studies focused on assessing single components of the constructs, including representations in statistical contexts, probability, and variation. They point to the difficulties and lack of experience of students in understanding and critically interpreting statistical representations and results. Studies that measured behavioral components, on the other hand, show that conducting their own experiments (developing research questions; data collection and analysis, creating visualizations) increases understanding and transfer and, also expands basic skills. Concerning affective student metrics examined in the context of statistical and data literacy, a clear focus on positive affect has become apparent. This emphasis on positive affect contrasts with the only study that explored how negative affect (statistics anxiety; Saidi & Siew, 2022) influences the development of statistical and data literacy, and how it can potentially be reduced or avoided through interventions. Research indicates that student's negative affect toward statistics is not a negligible factor when it comes to its impact on work behavior and achievement (Macher et al., 2015), even well before university. In terms of the affective component, the experts in the survey identified the development of a critical stance towards presented data and data-based decisions as an important learning objective for students. This indicates a potential future research focus that highlights this skill as a component of data literacy, referred to as critical data literacy.

The relative importance distribution among the three components, which emerged from the theoretical conceptualizations of statistical and data literacy (see Figs. 1 and 2), was exactly reflected in the focuses of the empirical studies. The cognitive and affective components were more pronounced in studies addressing statistical literacy, whereas studies pertaining to data literacy more frequently measured behavioral aspects (see Table A1 in supplementary material 3). The survey results also revealed that experts tend to associate the term statistical literacy with cognitive-mathematical skills, which they believe should be emphasized in future education, while they view data literacy as addressing behavioral components, related to the process of scientific inquiry. However, the experts do not consider it beneficial to explicitly separate statistical and data literacy or their subcomponents in instruction. Instead, they favor integrative teaching concepts for the future that also relate to real-world contexts.

Instructional approaches

Concerning our second research question, several instructional strategies were identified that are effective in promoting statistical and data literacy in K-12 STEM education. A common feature of most approaches is the integration of authentic, open-ended problems and complex real-world data that is meaningful for students. The data were often taken from complex real-world datasets from a variety of sources and domains, including public health data, sustainable development data, science data, and global socioeconomic data. The use of technological tools suitable for classroom use to analyze and visualize data (e.g., *Gapminder*, *TinkerPlots*, *Fathom*) has been demonstrated to help students deal with large and complex data sets and supported them to perform sophisticated statistical activities such as filtering relevant data, comparing data distributions, or observing variation in data, even at primary and lower secondary school level. The experience of collecting data seems to be effective to promote critical thinking about data and supports students' abilities to make sense of data. A few studies applied a game-based approach and reported positive effects primarily regarding students' affective variables such as students' enjoyment of learning, attitudes, and increased self-esteem regarding statistical or data literacy. However, further research is necessary to investigate the effectiveness of these approaches in controlled experimental designs and large upscaled implementation studies.

The instructional approaches identified in the systematic review as frequently studied and effective in promoting statistical and data literacy largely coincide with those that experts in our survey deemed promising for future instruction. Notably, the survey results highlighted a strong preference for interdisciplinary teaching methods. However, this focus was not evident in the reviewed empirical literature. While studies falling within the concept of data literacy are distributed across various STEM subjects, those focusing on the concept of statistical literacy are primarily found in the field of mathematics. Also concerning the reviewed instructional strategies, only a few of them used interdisciplinary approaches to promote statistical and data literacy (e.g., Vahey et al., 2012). Most studies in this review focused on a single STEM subject, primarily mathematics. This may be due to the current placement of fundamental aspects in mathematics curricula, such as reasoning, probability, variation, and handling data distributions. In the OECD Program for International Student Assessment (PISA) study of the year 2022, the topic area of uncertainty and data is also assigned to the field of mathematics (OECD, 2018). Although some studies can be found in multiple STEM and science fields (as a combination of biology, physics, and chemistry), there remain few studies that advocate for an integrated perspective on statistical and data literacy within STEM subjects, as recommended by policymakers (National Research Council [NRC], 2011, 2012) or researchers (e.g., Kelley & Knowles, 2016).

Limitations

When pre-registering this systematic review, we had intended to explore the potential relationships between teacher metrics and the development of students' statistical and data literacy through an additional research question. However, as the search process yielded only one paper (Callingham et al., 2016) that addressed this question, we decided to exclude this question from the current review. It is noteworthy that there is a scarcity of studies that explore relationships between teacher and student metrics in this context. At least we were able to ask the experts in our survey how teachers can be empowered to promote students' statistical and data literacy in a future-oriented manner. The experts evidently assumed that teachers would need explicit training and professional development in this area to meet future demands. They suggested fostering the cognitive aspects of teachers' statistical and data literacy and also identified the need to improve teachers' attitudes towards statistics and data on an affective level. To enable interdisciplinary teaching, the experts also recommended measures to facilitate networking among teachers. To provide a comprehensive overview of the current research landscape in this area, a separate review focusing on the statistical and data literacy of teachers in K-12 STEM education was conducted (Schreiter et al., 2024). Our review aimed to analyze the current body of empirical research on K-12 STEM education that explicitly establishes a conceptual link to statistical or data literacy. Although the individual contributions were treated equally in the review and presented accordingly, some focused more heavily on the concepts of statistical or data literacy while others only briefly mentioned them. Furthermore, it is likely that papers exist that address very similar topics but were not included in the review because they did not explicitly mention either of the two key terms. Thus, while our review provides insights into the research field, it is important to acknowledge that it may not be representative of the entire field, which may be more extensive than what is captured by our data. A sophisticated bibliometric review (e.g., Donthu et al., 2021) could provide additional insights by incorporating more quantitative aspects of the dissemination of the terms and identifying trends, developments, and gaps in the existing literature. Additionally, this approach might also capture the conceptual proximity of statistical and data literacy to related concepts and sub-concepts, perhaps leading to a possibility to establish a conceptual framework for it.

Conclusion

The current state of research on statistical and data literacy in STEM education, as presented in this review, reveals that despite the growing number of studies in this area, a comprehensive theoretical framework that integrates statistical and data literacy from an interdisciplinary perspective is currently lacking. Existing studies primarily focus on individual cognitive components and often overlook the interplay between various components. Additionally, research examining the influence of affective variables on statistical and data literacy remains scarce and is not explicitly categorized or examined as subcomponents. There is also a notable gap in understanding the influence and correlation of other variables related to statistical and data literacy.

The three-component framework developed in this study has proven useful for systematically organizing current studies and findings based on the three components (cognitive, behavioral and affective). This framework effectively maps the conceptual differences in theories regarding statistical and data literacy constructs. By combining diverse aspects of statistical and data literacy, the framework resolves the theoretical separation of these literacies, which is inconsistently reflected in existing studies. However, the review highlights the need for a more precise and comprehensive framework that examines and maps the relationships and overlap areas between the three components, such as critical data literacy, which intersects cognitive and affective components.

Despite the inherent promise of interdisciplinarity in this field, its current representation in K-12 education and the breadth of the existing research are insufficient. This indicates a significant discrepancy between what experts suggest and what is empirically researched to date, underscoring the urgency for expanded interdisciplinary in research and practice. Advancing this approach is crucial to more effectively realize the potential of fostering statistical and data literacy across all disciplines.

Based on this systematic review, three main conclusions can be drawn: first, a comprehensive, integrative framework for statistical and data literacy is desirable. Second, the relationships and overlaps among the three components, especially the affective component, are insufficiently differentiated and investigated. Third, interdisciplinarity should be strengthened in both research and practice.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40594-024-00517-z.

Supplementary Material 1.

Supplementary Material 2.

Supplementary Material 3.

Supplementary Material 4.

Acknowledgements

We would like to express our gratitude to Stephanie Haaß, Tobias Mosetter, and Hannah Fuhr for their invaluable support as student assistants during the initial screening of the search results. We also extend our thanks to the journal's reviewers and editors for their constructive feedback, which significantly enhanced the quality of the manuscript.

Author contributions

AF: conceptualization, methodology, investigation, data curation, formal analysis, writing—original draft. SS: conceptualization, methodology, investigation, data curation, formal analysis, writing—original draft. MV: conceptualization, writing—review and editing, funding acquisition. SB: conceptualization, writing—review and editing. RB: conceptualization, writing—review and editing, funding acquisition. JK: conceptualization, writing—review and editing, funding acquisition. JL: investigation, data curation, visualization, writing—review and editing. SM: conceptualization, methodology, writing—original draft, review and editing, supervision, project administration, funding acquisition.

Funding
Open Access funding enabled and organized by Projekt DEAL. This research was funded by the Federal Ministry of Education and Research, Germany (BMBF; Grant: 16MF1006) and supported by LMUexcellent, funded by the Federal Ministry of Education and Research (BMBF) and the Free State of Bavaria under the Excellence Strategy of the Federal Government and the Länder. The responsibility for the content of this publication lies with the authors.

Availability of data and materials
All data generated or analyzed during this study are included in this published article and its supplementary information files.

Declarations

Competing interests
The authors declare that they have no competing interests.

Received: 12 March 2024 Accepted: 18 October 2024
Published online: 13 November 2024

References

Aksoy, E. C., & Bostan, M. I. (2021). Seventh graders' statistical literacy: An investigation on bar and line graphs. *International Journal of Science and Mathematics Education, 19*(2), 397–418. https://doi.org/10.1007/s10763-020-10052-2

Almeida, C. C. (2018). The comparison of probability issues of a statistical literacy. *Cadernos Educacao Tecnologia E Sociedade, 11*(4), 624–633. https://doi.org/10.14571/brajets.v11.n4.624-633

Andre, M., Lavicza, Z., & Prodromou, T. (2020). Integrating 'education for sustainable development' in statistics classes: Visual analysis of social and economic data with gapminder. In: P. Arnold (Ed.), *New Skills in the Changing World of Statistics Education.* Proceedings of the Roundtable conference of the International Association for Statistical Education (IASE) (pp. 1–6). ISI/IASE.

Aoyama, K., & Stephens, M. (2003). Graph interpretation aspects of statistical literacy: A Japanese perspective. *Mathematics Education Research Journal, 15*(3), 207–225. https://doi.org/10.1007/BF03217380

Aziz, A. M., & Rosli, R. (2021). A systematic literature review on developing students' statistical literacy skills. *Journal of Physics Conference Series*. https://doi.org/10.1088/1742-6596/1806/1/012102

Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L., & Spangler, D.A. (2020). *Pre-K-12 guidelines for assessment and instruction in statistics education (GAISE) report II.* American Statistical Association and National Council of Teachers of Mathematics

Begolli, K. N., Dai, T., McGinn, K. M., & Booth, J. L. (2021). Could probability be out of proportion? Self-explanation and example-based practice help students with lower proportional reasoning skills learn probability. *Instructional Science, 49*(4), 441–473. https://doi.org/10.1007/s11251-021-09550-9

Ben-Zvi, D., & Makar, K. (2016). The teaching and learning of statistics. *Springer Cham*. https://doi.org/10.1007/978-3-319-23470-0

Bergner, Y., Mund, S., Chen, O., & Payne, W. (2020). Leveraging interest-driven embodied practices to build quantitative literacies: A case study using motion and audio capture from dance. *Educational Technology Research and Development, 69*(4), 2013–2036. https://doi.org/10.1007/s11423-020-09804-2

Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy*. Academic Press.

Birk, L., & Frischemeier, D. (2022, February 2–5). *Communicating data exploration insights through posters—A preliminary analysis of primary students´ learning outcomes* [Conference paper]. Twelfth Congress of the European Society for Research in Mathematics Education (CERME12), Bozen-Bolzano, Italy.

Braun, D., & Huwer, J. (2022). Computational literacy in science education–A systematic review. *Frontiers in Education*. https://doi.org/10.3389/feduc.2022.937048

Burrill, G. F., De Oliveria Souza, L., & Reston, E. (Hrsg.). (2023). *Research on reasoning with data and statistical thinking: International perspectives*. Springer International Publishing. https://doi.org/10.1007/978-3-031-29459-4

Budgett, S., & Rose, D. (2017). Developing statistical literacy in the final school year. *Statistics Education Research Journal, 16*(1), 139–162. https://doi.org/10.52041/serj.v16i1.221

Cakiroglu, Ü., & Güler, M. (2021). Enhancing statistical literacy skills through real life activities enriched with gamification elements: An experimental study. *E-Learning and Digital Media, 18*(5), 441–459. https://doi.org/10.1177/2042753020987016

Callingham, R., Carmichael, C., & Watson, J. (2016). Explaining student achievement: The influence of teachers' pedagogical content knowledge in statistics. *International Journal of Science and Mathematics Education, 14*(7), 1339–1357.

Callingham, R., & Watson, J. M. (2005). Measuring statistical literacy. *Journal of Applied Measurement, 6*(1), 19–47.

Callingham, R., & Watson, J. M. (2017). The development of statistical literacy at school. *Statistics Education Research Journal, 16*(1), 181–201. https://doi.org/10.52041/serj.v16i1.223

Calzada Prado, J., & Marzal, M. (2013). Incorporating data literacy into information literacy programs: Core competencies and contents. *Libri, 63*(2), 123–134. https://doi.org/10.1515/libri-2013-0010

Carlson, J., & Johnston, L. R. (Eds.). (2015). *Data information literacy: Librarians, data, and the education of a new generation of researchers*. Purdue University Press.

Carmichael, C., Callingham, R., Hay, I., & Watson, J. (2010a). Measuring middle school students' interest in statistical literacy. *Mathematics Education Research Journal, 22*(3), 9–39. https://doi.org/10.1007/BF03219776

Carmichael, C., Callingham, R., Hay, I., & Watson, J. (2010b). Statistical literacy in the middle school: The relationship between interest, self-efficacy and prior mathematics achievement. *Australian Journal of Educational & Developmental Psychology, 10*, 83–93.

Carmichael, C. S., & Hay, I. (2009a). The development and validation of the Students' Self Efficacy for Statistical Literacy Scale. In R. Hunter, B. Bicknell & T. Burgess (Eds.), *Proceedings of the 32nd Annual Conference of the Mathematics Education Research Group of Australasia* (Vol. 1, pp. 97–104). MERGA Inc.

Carmichael, C. S., & Hay, I. (2009b). Gender differences in middle school students' interests in a statistical literacy context. In R. Hunter, B. Bicknell & T. Burgess (Eds.), *Proceedings of the 32nd Annual Conference of the Mathematics Education Research Group of Australasia* (Vol. 1, pp. 89–96). MERGA Inc.

Carvalho, C., & Solomon, Y. (2012). Supporting statistical literacy: What do culturally relevant/realistic tasks show us about the nature of pupil engagement with statistics? *International Journal of Educational Research, 55*, 57–65. https://doi.org/10.1016/j.ijer.2012.06.006

Chalkiadaki, A. (2018). A systematic literature review of 21st Century skills and competencies in primary education. *International Journal of Instruction, 11*(3), 1–16. https://doi.org/10.12973/iji.2018.1131a

Chan, S. W., Ismail, Z., & Sumintono, B. (2016). A framework for assessing high school students' statistical reasoning. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0163846

Chew, P. K., & Dillon, D. B. (2014). Statistics anxiety update: Refining the construct and recommendations for a new research agenda. *Perspectives on Psychological Science, 9*(2), 196–208. https://doi.org/10.1177/1745691613518077

Chin, D. B., Blair, K. P., & Schwartz, D. L. (2016). Got game? A choice-based learning assessment of data literacy and visualization skills. *Technology, Knowledge and Learning, 21*(2), 195–210. https://doi.org/10.1007/s10758-016-9279-7

Cotič, M. (2009). Developing basic statistical literacy at the beginning of schooling. *Zunanja učna diferenciacija in čustveno-osebnostni vidik učenja,* 92.

Data to the People. (2018). Databilities: A Data Literacy Competency Framework. *Data to the People.* Retrieved from https://www.datatothepeople.org/databilities

Dierker, L., Ward, N., Alexander, J., & Donate, E. (2017). Engaging underrepresented high school students in data driven storytelling: An examination of learning experiences and outcomes for a cohort of rising seniors enrolled in the gaining early awareness and readiness for undergraduate program (GEAR UP). *Journal of Education and Training Studies, 5*(4), 54–63. https://doi.org/10.11114/jets.v5i4.2187
Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research, 133*, 285–296. https://doi.org/10.1016/j.jbusres.2021.04.070
English, L. (2013). Promoting statistical literacy through data modelling in the early school years. In E. J. Chernoff & B. S. Sriraman (Eds.), *Probabilistic thinking: Presenting plural perspectives* (pp. 441–457). Springer. https://doi.org/10.1007/978-94-007-7155-0_23
English, L. D., & Watson, J. M. (2015a). Exploring variation in measurement as a foundation for statistical thinking in the elementary school. *International Journal of STEM Education, 2*(1), 1–20. https://doi.org/10.1186/s40594-015-0016-x
English, L. D., & Watson, J. M. (2015b). Statistical literacy in the elementary school: Opportunities for problem posing. In F. Singer, N. Ellerton, & J. Cai (Eds.), *Problem posing: From research to effective practice* (pp. 241–256). Springer. https://doi.org/10.1007/978-1-4614-6258-3_11
English, L. D., & Watson, J. M. (2016). Development of probabilistic understanding in fourth grade. *Journal for Research in Mathematics Education, 47*(1), 28–62. https://doi.org/10.5951/jresematheduc.47.1.0028
English, L. D., & Watson, J. (2018). Modelling with authentic data in sixth grade. *ZDM Mathematics Education, 50*(1), 103–115. https://doi.org/10.1007/s11858-017-0896-y
Fielding-Wells, J. (2018). Dot plots and hat plots: Supporting young students emerging understandings of distribution, center and variability through modeling. *ZDM Mathematics Education, 50*(7), 1125–1138. https://doi.org/10.1007/s11858-018-0961-1
Francois, K., Monteiro, C., & Allo, P. (2020). Big-data literacy as a new vocation for statistical literacy. *Statistics Education Research Journal, 19*(1), 194–205. https://doi.org/10.52041/serj.v19i1.130
Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: A Pre-K-12 Curriculum Framework.* American Statistical Association. Retrieved from https://www.amstat.org/asa/files/pdfs/GAISE/GAISEPreK-12_Full.pdf
Friedrich, A., Schreiter, S., Lehmann, J., Mosetter, T., Malone, S., Becker, S., Kuhn, J., Brünken, R., & Vogel, M. (2021, September 3). Systematic review on the definition and fostering of statistical/data literacy in STEM school education. 10.17605/OSF.IO/DVPBN
Frischemeier, D. (2019). Primary school students' reasoning when comparing groups using modal clumps, medians, and hatplots. *Mathematics Education Research Journal, 31*(4), 485–505. https://doi.org/10.1007/s13394-019-00261-6
Frischemeier, D. (2020). Building statisticians at an early age–Statistical projects exploring meaningful data in primary school. *Statistics Education Research Journal, 19*(1), 39–56. https://doi.org/10.52041/serj.v19i1.118
Gal, I. (2002). Response: Developing statistical literacy: Towards implementing change. *International Statistical Review/revue Internationale De Statistique, 70*(1), 46–51. https://doi.org/10.2307/1403721
Gal, I. (2004). Statistical literacy: Meanings, components, responsibilities. In J. B. Garfield & D. Ben-Zvi (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 47–78). Kluwer.
Ganesan, N., & Leong, K. E. (2018). Effectiveness of Fathom on statistical reasoning among form four students. *MOJES Malaysian Online Journal of Educational Sciences, 6*(4), 12–22.
Garfield, J., & Ben-Zvi, D. (2007). How students learn statistics revisited: A current review of research on teaching and learning statistics. *International Statistical Review, 75*(3), 372–396. https://doi.org/10.1111/j.1751-5823.2007.00029.x
Gebre, E. H. (2018). Young adults' understanding and use of data: Insights for fostering secondary school students' data literacy. *Canadian Journal of Science, Mathematics and Technology Education, 18*(4), 330–341. https://doi.org/10.1007/s42330-018-0034-z
Gould, R. (2017). Data literacy is statistical literacy. *Statistics Education Research Journal, 16*(1), 22–25. https://doi.org/10.52041/serj.v16i1.209
Guler, M., Gursoy, K., & Guven, B. (2016). Critical views of 8th grade students toward statistical data in newspaper articles: Analysis in light of statistical literacy. *Cogent Education, 3*(1), 1268773. https://doi.org/10.1080/2331186X.2016.1268773
Heidrich, J., Bauer, P., & Krupka, D. (2018). *Future Skills: Ansätze zur Vermittlung von Data Literacy in der Hochschulbildung*. Hochschulforum Digitalisierung.
Höper, L., Podworny, S., Schulte, C., & Frischemeier, D. (2021). Exploration of location data: real data in the context of interaction with a cellular network. *Proceedings of the IASE 2021 Satellite Conference*. https://doi.org/10.52041/iase.nkppy
Hourigan, M., & Leavy, A. M. (2020). Interrogating a measurement conjecture to introduce the concept of statistical association in upper elementary education. *Teaching Statistics, 43*(2), 62–71. https://doi.org/10.1111/test.12249
Irish, T., Berkowitz, A., & Harris, C. (2019). Data explorations: Secondary students' knowledge, skills and attitudes toward working with data. *Eurasia Journal of Mathematics, Science and Technology Education, 15*(6), em1686. https://doi.org/10.29333/ejmste/103063
Jiang, S., & Kahn, J. (2020). Data wrangling practices and collaborative interactions with aggregated data. *International Journal of Computer-Supported Collaborative Learning, 15*(3), 257–281. https://doi.org/10.1007/s11412-020-09327-1
Kahn, J., & Jiang, S. (2020). Learning with large, complex data and visualizations: Youth data wrangling in modeling family migration. *Learning, Media and Technology, 46*(2), 128–143. https://doi.org/10.1080/17439884.2020.1826962
Kelley, T. R., & Knowles, J. G. (2016). A conceptual framework for integrated STEM education. *International Journal of STEM Education, 3*, 11. https://doi.org/10.1186/s40594-016-0046-z
Khan, H. R., Kim, J., & Chang, H. C. (2018). *Toward an understanding of data literacy.* iConference 2018 Proceedings.
Kimura, S. (1999). Toukeizyouhoukyouikuno Karikyuramuto 5-dankaino Toukeitekitankyu Purosesu [Curriculum of statistics education and five phases of statistical inquiry process] (in Japanese). In Zentouken (Ed.), *Toukeizyouhoukyouikuno Rironto Zyugyouzissenno Tenkai*. Tsukuba Syuppankai (pp. 33–46).
Kochevar, R. E., Krumhansl, R., Krumhansl, K., Peach, C. L., Bardar, E., Louie, J., Sickler, J., Mueller-Northcott, J., Busey, A., LaVita, S., & DeLisi, J. (2015). Inspiring future marine and data scientists through the lure of ocean tracks. *Marine Technology Society Journal, 49*(4), 64–75. https://doi.org/10.4031/MTSJ.49.4.4
Koparan, T., & Güven, B. (2014a). According to the M3ST model analyze of the statistical thinking levels of middle school student. *Egitim Ve Bilim, 39*(171), 37–51.
Koparan, T., & Güven, B. (2014b). The effect of project-based learning on the statistical literacy levels of student 8th grade. *European Journal of Educational Research, 3*(3), 145–157. https://doi.org/10.12973/eu-jer.3.3.145
Kultusministerkonferenz. (2004a). *Bildungsstandards im Fach Mathematik für den Primarbereich.* Beschluss vom 15.10.2004. München, Germany: Wolters-Kluwer.
Kultusministerkonferenz. (2004b). *Bildungsstandards im Fach Mathematik für den mittleren Schulabschluss.* München, Germany: Wolters Kluwer.
Kuntze, S., Lindmeier, A., & Reiss, K. (2008). "Using models and representations in statistical contexts" as a sub-competency of statistical literacy – Results from three empirical studies. *Proceedings of the 11th International Congress on Mathematical Education (ICME 11).*
Kuntze, S., Vargas, F., Martignon, L., & Engel, J. (2015). Competencies in understanding statistical information in primary and secondary school levels: An inter-cultural empirical study with German and Colombian students. *Avances De Investigación En Educación Matemática, 7*, 5–25. https://doi.org/10.35763/aiem.v1i7.103
Lee, V. R., Drake, J., Cain, R., & Thayne, J. (2021). Remembering what produced the data: Individual and social reconstruction in the context of a Quantified Self elementary data and statistics unit. *Cognition and Instruction, 39*(4), 367–408. https://doi.org/10.1080/07370008.2021.1936529
Lindmeier, A., Kuntze, S., & Reiss, K. (2007). *Representations of data and manipulations through reduction–competencies of German secondary students.* Proceedings of the IASE/ISI Satellite Conference on Statistical Education, Guimarães, Portugal.

Liu, T. C., & Lin, Y. C. (2010). The application of Simulation-Assisted Learning Statistics (SALS) for correcting misconceptions and improving understanding of correlation. *Journal of Computer Assisted Learning, 26*(2), 143–158. https://doi.org/10.1111/j.1365-2729.2009.00330.x
Liu, T. C., & Lin, Y. C. (2011). Developing two-tier diagnostic instrument for exploring students' statistical misconceptions: Take "Correlation" as the example. *Bulletin of Educational Psychology, 42*(3), 379–400. https://doi.org/10.6251/bep.20090805
Macher, D., Papousek, I., Ruggeri, K., & Paechter, M. (2015). Statistics anxiety and performance: Blessings in disguise. *Frontiers in Psychology, 6*, e1116. https://doi.org/10.3389/fpsyg.2015.01116
Makar, K., & Allmond, S. (2018). Statistical modelling and repeatable structures: Purpose, process and prediction. *ZDM, 50*(7), 1139–1150. https://doi.org/10.1007/s11858-018-0956-y
Malaspina, M., & Malaspina, U. (2020). Game invention as means to stimulate probabilistic thinking. *Statistics Education Research Journal, 19*(1), 57–72. https://doi.org/10.52041/serj.v19i1.119
Mendonca, L. D. O., & Lopes, C. E. (2011). Mathematical Modeling: A learning environment for the implementation of statistics education in high school. *Bolema-Mathematics Education Bulletin-Boletim De Educacao Matematica, 24*(40), 701–724.
Merriman, L. (2006). Using media reports to develop statistical literacy in Year 10 students. *Proceedings of the 7th International Conference on Teaching Statistics.* Auckland, New Zealand. Retrieved from https://www.stat.auckland.ac.nz/~iase/publications/17/8A3_MERR.pdf
Mills, J. D., & Holloway, C. E. (2013). The development of statistical literacy skills in the eighth grade: Exploring the TIMSS data to evaluate student achievement and teacher characteristics in the United States. *Educational Research and Evaluation, 19*(4), 323–345. https://doi.org/10.1080/13803611.2013.771110
Ministério da Educação, (2006). *Parâmetros curriculares nacionais: Matemática* (National curricular parameters: Mathematics). Brasilia, Brazil.
Ministry of Education. (2015). *The New Zealand Curriculum.* Learning Media Limited.
Mosquera, C. K., Steinmaurer, A., Eckhardt, C., & Guetl, C. (2020). Immersively learning object oriented programming concepts with sCool. In D. Economou, A. Klippel, H. Dodds, A. Pena-Rios, M. J. W. Lee, D. Beck, J. Pirker, A. Dengel, T. M. Peres, & J. Richter (Eds.), *Proceedings of 6th International Conference of the Immersive Learning Research Network, iLRN 2020* (pp. 124–131). IEEE Xplore. https://doi.org/10.23919/iLRN47897.2020.9155144
Nacarato, A. M., & Grando, R. C. (2014). The role of language in building probabilistic thinking. *Statistics Education Research Journal, 13*(2), 93–103. https://doi.org/10.52041/serj.v13i2.283
National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. NCTM.
National Research Council. (1996). *National Science Education Standards*. The National Academies Press. https://doi.org/10.17226/4962
National Research Council. (2011). *Successful K-12 STEM education: Identifying effective approaches in science, technology, engineering, and mathematics.* National Academies Press. https://doi.org/10.17226/13158
National Research Council. (2012). *A framework for K12 science education: Practices, cross cutting concepts, and core ideas*. National Academies Press. https://doi.org/10.17226/13165
New Zealand Ministry of Education. (2014). The New Zealand Curriculum Mathematics and Statistics. Retrieved from https://nzcurriculum.tki.org.nz/The-New-Zealand-Curriculum/Mathematics-and-statistics/Achievement-objectives.
Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence, 2*, 100041. https://doi.org/10.1016/j.caeai.2021.100041
NGSS Lead States. (2013). *Next generation science standards: For states, by states*. The National Academies Press.
OECD (2018). PISA for Development Mathematics Framework. In OECD (Ed.), *PISA for Development Assessment and Analytical Framework: Reading, Mathematics and Science* (pp. 49–70). OECD Publishing. https://doi.org/10.1787/9789264305274-5-en
OECD. (2021). 21st-century readers: Developing literacy skills in a digital world. *OECD Publishing*. https://doi.org/10.1787/a83d84cb-en
Oslington, G., Mulligan, J., & Van Bergen, P. (2020). Third-graders' predictive reasoning strategies. *Educational Studies in Mathematics, 104*(1), 5–24. https://doi.org/10.1007/s10649-020-09949-0
Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ, 372*, 1–9. https://doi.org/10.1136/bmj.n71
Petocz, P; Reid, A. & Gal, I. (2018). Statistics Education Research. In Ben-Zvi, D., Makar, K., & Garfeld, J. (Eds.). International handbook of research in statistics education. *Springer International Handbooks of Education.* (p. 71–100) https://doi.org/10.1007/978-3-319-66195-7_12.
Pfannkuch, M. (2005). Characterizing year 11 students' evaluation of a statistical process. *Statistics Education Research Journal, 4*, 5–26. https://doi.org/10.52041/serj.v4i2.512
Pfannkuch, M. (2018). Reimagining curriculum approaches. In Ben-Zvi, D., Makar, K., & Garfeld, J. (Eds.). International handbook of research in statistics education. *Springer International Handbooks of Education.* (p. 387–414). https://doi.org/10.1007/978-3-319-66195-7_12
Primi, C., Morsanyi, K., Donati, M. A., Galli, S., & Chiesi, F. (2017). Measuring probabilistic reasoning: The construction of a new scale applying item response theory. *Journal of Behavioral Decision Making, 30*(4), 933–950. https://doi.org/10.1002/bdm.2011
Redecker, C. & Punie, Y. (2017). *European framework for the digital competence of educators: DigCompEdu*. European Commission, Joint Research Centre. https://doi.org/10.2760/178382
Ridsdale, C., Rothwell, J., Smit, M., Ali-Hassan, H., Bliemel, M., Irvine, D., Kelley, D., Matwin, S., Wuetherick, B. (2015). *Strategies and Best Practices for Data Literacy Education: Knowledge Synthesis Report*. Dalhousie University, Canada. http://hdl.handle.net/10222/64578
Saidi, S. S., & Siew, N. M. (2019). Reliability and validity analysis of statistical reasoning test survey instrument using the Rasch measurement model. *International Electronic Journal of Mathematics Education, 14*(3), 535–546. https://doi.org/10.29333/iejme/5755
Saidi, S. S., & Siew, N. M. (2022). Assessing secondary school students' statistical reasoning, attitudes towards statistics, and statistics anxiety. *Statistics Education Research Journal, 21*(1), 1–19. https://doi.org/10.52041/serj.v21i1.67
Schield, M. (2004). Information literacy, statistical literacy and data literacy. *IASSIST Quarterly, 28*(2/3), 6–11. https://doi.org/10.29173/iq790
Schreiter, S., Friedrich, A., Fuhr, H., Malone, S., Brünken, R., Kuhn, J., & Vogel, M. (2024). Teaching for statistical and data literacy in K-12 STEM education: a systematic review on teacher variables, teacher education, and impacts on classroom practice. *ZDM, 56*, 31–45. https://doi.org/10.1007/s11858-023-01531-1
Schüller, K. & Busch, P. (2019). *Data Literacy: Ein Systematic Review zu Begriffsdefinition, Kompetenzrahmen und Testinstrumenten.* Hochschulforum Digitalisierung.
Schüller, K., Busch, P., & Hindinger, C. (2019). *Future Skills: Ein Framework für Data Literacy. Kompetenzrahmen und Forschungsbericht.* Hochschulforum für Digitalisierung.
Schüller, K., Koch, H., & Rampelt, F. (2021). *Data-Literacy-Charta. Version 1.2*. Stifterverband.
Sharma, S. V. (2005). High school students interpreting tables and graphs: Implications for research. *International Journal of Science and Mathematics Education, 4*(2), 241–268. https://doi.org/10.1007/s10763-005-9005-8
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
Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. K. Lester (Ed.), *Second handbook of research on mathemaatics teaching and learning* (pp. 957–1010). Information Age Publishing.
Sproesser, U., Kuntze, S., & Engel, J. (2014). A multilevel perspective on factors influencing students' statistical literacy. In K. Makar, B. de Sousa, & R. Gould (Eds.), *Proceedings of the Ninth International Conference on Teaching Statistics: Sustainability in Statistics Education*. International Association for Statistical Education. Retrieved from https://www.iase-web.org/icots/9/proceedings/pdfs/ICOTS9_7E2_SPROESSER.pdf

Sproesser, U., Kuntze, S., & Engel, J. (2018). Using models and representations in statistical contexts. *Journal Für Mathematik-Didaktik, 39*(2), 343–367. https://doi.org/10.1007/s13138-018-0133-4
Swan, K., & Vahey, P. (2013). Problem-based learning across the curriculum: Exploring the efficacy of a cross-curricular application of preparation for future learning. *Interdisciplinary Journal of Problem-Based Learning, 7*(1), 8. https://doi.org/10.7771/1541-5015.1307
Tedre, M., Vartiainen, H., Kahila, J., Toivonen, T., Jormanainen, I., & Valtonen, T. (2020, October). Machine Learning introduces new perspectives to data agency in K–12 computing education. *In 2020 IEEE Frontiers in Education Conference (FIE)* (pp. 1–8). IEEE. https://doi.org/10.1109/FIE44824.2020.9274138
Tivian (2021). *EFS Survey*. Retrieved from https://www.unipark.de/
Utomo, D. P. (2021). An analysis of the statistical literacy of middle school students in solving TIMSS problems. *International Journal of Education in Mathematics, Science and Technology (IJEMST)*. https://doi.org/10.46328/ijemst.1552
Utrecht University. (2021). *ASReview Lab (0.18)* [computer software]. Retrieved from https://asreview.nl
Vahey, P., Yarnall, L., Patton, C., Zalles, D., & Swan, K. (2006). *Mathematizing middle school: Results from a cross-disciplinary study of data literacy.* American Educational Research Association Annual Conference. San Francisco, USA.
Vahey, P., Rafanan, K., Patton, C., & Swan, K. (2012). A cross-disciplinary approach to teaching data literacy and proportionality. *Educational Studies in Mathematics, 81*(2), 179–205. https://doi.org/10.1007/s10649-012-9392-z
van de Schoot, R., de Bruin, J., Schram, R., Zahedi, P., de Boer, J., & Wie jdema, F., & Oberski, D. L. (2021). An open-source machine learning framework for efficient and transparent systematic reviews. *Nature Machine Intelligence, 3*(2), 125–133. https://doi.org/10.1038/s42256-020-00287-7
VERBI Software. (2021). *MAXQDA 2022* [computer software]. VERBI Software. Retrieved from https://www.maxqda.com
Vigil-Colet, A., Lorenzo-Seva, U., & Condon, L. (2008). Development and validation of the statistical anxiety scale. *Psicothema, 20*(1), 174–180.
Wallman, K. K. (1993). Enhancing statistical literacy: Enriching our society. *Journal of the American Statistical Association, 88*(421), 1–8. https://doi.org/10.2307/2290686
Watson, J. M. (1997). Assessing statistical thinking using the media. In I. Gal & J. B. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 107–122). IOS Press.
Watson, J. M. (2006). *Statistical literacy at school: Growth and goals*. Lawrence Erlbaum.
Watson, J. M., & Callingham, R. A. (2003). Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal, 2*(2), 3–46. https://doi.org/10.52041/serj.v2i2.553
Watson, J. M., Callingham, R. A., & Kelly, B. A. (2007). Students' appreciation of expectation and variation as a foundation for statistical understanding. *Mathematical Thinking and Learning, 9*(2), 83–130. https://doi.org/10.1080/10986060709336812
Watson, J. M., Chick, H., & Callingham, R. A. (2014). Average: The juxtaposition of procedure and context. *Mathematics Education Research Journal, 26*(3), 477–502. https://doi.org/10.1007/s13394-013-0113-4
Watson, J. M., & English, L. (2015). Introducing the practice of statistics: Are we environmentally friendly? *Mathematics Education Research Journal, 27*, 585–613. https://doi.org/10.1007/s13394-015-0153-z
Watson, J. M., & English, L. (2017). Reaction time in Grade 5: Data collection within the practice of statistics. *Statistics Education Research Journal, 16*(1), 262–293. https://doi.org/10.52041/serj.v16i1.231
Watson, J. M., & English, L. (2018). Eye color and the practice of statistics in Grade 6: Comparing two groups. *The Journal of Mathematical Behavior, 49*, 35–60. https://doi.org/10.1016/j.jmathb.2017.06.006
Watson, J., Fitzallen, N., English, L., & Wright, S. (2019). Introducing statistical variation in year 3 in a STEM context: Manufacturing licorice. *International Journal of Mathematical Education in Science and Technology, 51*(3), 354–387. https://doi.org/10.1080/0020739X.2018.1562117
Watson, J. M., & Kelly, B. A. (2005). Cognition and instruction: Reasoning about bias in sampling. *Mathematics Education Research Journal, 17*(1), 24–57. https://doi.org/10.1007/BF03217408
Watson, J. M., & Kelly, B. A. (2007). The development of conditional probability reasoning. *International Journal of Mathematical Education in Science and Technology, 38*(2), 213–235. https://doi.org/10.1080/00207390601002880
Watson, J. M., & Kelly, B. A. (2008). Sample, random and variation: The vocabulary of statistical literacy. *International Journal of Science and Mathematics Education, 6*(4), 741–767. https://doi.org/10.1007/s10763-007-9083-x
Watson, J. M., Kelly, B. A., & Izard, J. F. (2005). Statistical literacy over a decade. In: P. Clarkson, A. Downton, D. Gronn, M. Horne, A. McDonough, R. Pierce, & A. Roche (Eds.), *Building connections: Theory, research and practice* (Proceedings of the 28th annual conference of the Mathematics Education Research Group of Australasia, Melbourne, Vol. 2., pp. 775–782). MERGA.
Watson, J. M., Kelly, B. A., & Izard, J. (2006). A longitudinal study of student understanding of chance and data. *Mathematics Education Research Journal, 18*(2), 40–55. https://doi.org/10.1007/BF03217435
Watson, J. M., & Moritz, J. B. (2000a). Developing concepts of sampling. *Journal for Research in Mathematics Education, 31*(1), 44–70. https://doi.org/10.2307/749819
Watson, J. M., & Moritz, J. B. (2000b). Development of understanding of sampling for statistical literacy. *Journal of Mathematical Behavior, 19*(1), 109–136. https://doi.org/10.1016/S0732-3123(00)00039-0
Watson, J. M., & Moritz, J. B. (2003). The development of comprehension of chance language: Evaluation and interpretation. *School Science and Mathematics, 103*(2), 65–80. https://doi.org/10.1111/j.1949-8594.2003.tb18222.x
Wild, C., & iNZight Team. (2023). *The iNZightVIT project*. University of Auckland, inzight.nz.
Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review, 67*(3), 223–265. https://doi.org/10.1111/j.1751-5823.1999.tb00442.x
Wilkerson, M. H., Lanouette, K., & Shareff, R. L. (2021). Exploring variability during data preparation: a way to connect data, chance, and context when working with complex public datasets. *Mathematical Thinking and Learning*. https://doi.org/10.1080/10986065.2021.1922838
Wolff, A., Gooch, D., Cavero, M., Jose, J., Rashid, U., & Kortuem, G. (2017). 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
Wolff, A., Wermelinger, M., & Petre, M. (2019). Exploring design principles for data literacy activities to support children's inquiries from complex data. *International Journal of Human-Computer Studies, 129*, 41–54.
Yolcu, A. (2014). Middle school students' statistical literacy: Role of grade level and gender. *Statistics Education Research Journal, 13*(2), 118–131. https://doi.org/10.52041/serj.v13i2.285

Publisher's Note

Anja Friedrich is a postdoctoral researcher at the Department of Education at Saarland University, Campus A4 2, 66123 Saarbrücken, Germany; Her research focuses on teacher professionalization and the advancement of teachers pedagogical competencies. Furthermore, she maintains a strong focus on higher education in her research pursuits.

Saskia Schreiter is Full Professor for Mathematics and Mathematics Education at the Institute of Mathematics and Computer Science of the University of Education Schwäbisch Gmünd, Oberbettringer Str. 200, 73525 Schwäbisch Gmünd, Germany; Her research focuses on promoting data literacy and advancing teachers professional competencies. Additionally, she explores cognitive processes through eye-tracking technology and investigates the development of computational thinking in mathematics education.

Markus Vogel is Full Professor for Mathematics and Mathematics Education at the Faculty of Natural and Sociological Sciences of Heidelberg University of Education, Keplerstraße 87, 69120 Heidelberg, Germany; His research focuses on learning and teaching processes in the field of statistics and stochastics education. He also researches learning with multiple representations in multimedia learning environments in mathematics.

Sebastian Becker-Genschow is Assistant Professor for Digital Education at the Department of Didactics of Mathematics and Natural Sciences at the University of Cologne, Albertus-Magnus-Platz, 50923 Cologne, Germany; His research focuses on the empirical investigation of the effectiveness of digital support of teaching-learning processes, the analysis of learning processes by means of eye tracking and the use of AI-based technology in schools and universities.

Roland Brünken is Full Professor for Empirical Educational Research at the Department of Education at Saarland University, Campus A4 2, 66123 Saarbrücken, Germany; His research focuses on fundamental and applied exploration of cognitive processes in multimedia-based learning across various content domains, including STEM education, the development of teachers professional competencies, and traffic education.

Jochen Kuhn is LMUexcellent Full Professor of Physics Education Research and Chair of Physics Education at the Faculty of Physics of the Ludwig-Maximilians-Universität München (LMU Munich), Geschwister-Scholl-Platz 1, 80539 Munich, Germany; His research focuses on learning with multiple representations in advanced multimedia learning environments in STEM. He also uses Artificial Intelligence as well as physiological measures—such as eye tracking in learning and problem-solving.

Jessica Lehmann is a masters student in Psychology at Saarland University, Saarbrücken Campus, 66123 Saarbrücken, Germany; She holds a Bachelors degree in Psychology.

Sarah Malone is a postdoctoral researcher at the Department of Education at Saarland University, Campus A4 2, 66123 Saarbrücken, Germany; Her research focuses on learning with multiple representations and understanding, measuring, and promoting cognitive processes during technology-enhanced learning.