Multidisciplinary Reviews

BIBLIOMETRIC RESEARCH
Published Online: September 9, 2025
https://doi.org/10.31893/multirev.2026171

Statistical literacy: A hybrid systematic literature review and bibliometric analysis

Iesyah Rodliyah[ab] ✉ | **I Ketut Budayasa**[a] | **Siti Khabibah**[a]

[a]Mathematics Education, Universitas Negeri Surabaya, Surabaya, Indonesia.
[b]Mathematics Education, Universitas Hasyim Asy'ari, Jombang, Indonesia.

Abstract This study presents a hybrid systematic literature review (SLR) and bibliometric analysis aimed at examining the development, trends, and theoretical implications of statistical literacy research. Using the PRISMA protocol, 89 relevant articles published between 2002 and 2025 were analysed to address three research questions concerning the significance of statistical literacy as a future research domain, the distribution of scholarly investigations, and their theoretical and practical implications. The findings show a consistent increase in research interest over the past decade, especially after 2017, with dominant themes including statistical literacy, statistics education, and data literacy. The analysis also reveals an imbalance in research allocation, with higher education receiving disproportionate attention compared to other educational levels. Moreover, the results underscore the integrative role of statistical thinking and reasoning as essential components of statistical literacy, with growing emphasis on digital learning, real-world data contexts, and interdisciplinary approaches. This review highlights both the progress and the existing gaps in the field, providing a roadmap for future research that prioritizes inclusive, context-sensitive, and pedagogically sound models for enhancing statistical literacy in a data-driven society. In addition to thematic and theoretical insights, this study also explores the most influential journals, authors, and citation networks, providing a clearer picture of the academic landscape surrounding statistical literacy. The bibliometric data help reveal publication patterns, collaboration trends, and emerging research clusters. These findings are valuable for identifying research gaps, informing curriculum development, and fostering cross-institutional and interdisciplinary partnerships. By highlighting these dimensions, the review contributes to a deeper understanding of how the field has evolved and where it is heading.

Keywords: statistics education, research trends, data literacy, PRISMA, future research

1. Introduction

In today's data-driven society, statistical literacy has become an essential competency for individuals to interpret, evaluate, and make informed decisions on the basis of statistical information (Hariyanti et al, 2025). This is especially significant in educational contexts, where there is a growing expectation for students to interact with data in purposeful and insightful ways (Sharma, 2017; Shobikhah et al., 2025). Despite its growing importance, there remains significant variation in how statistical literacy is defined, taught, and assessed across different educational levels and contexts. This inconsistency highlights the need for a more unified understanding of the concept and its application in learning environments (Aziz & Rosli, 2021; Sharma, 2017)

A systematic literature review (SLR) is a valuable methodological approach for synthesizing existing research, mapping conceptual frameworks, and identifying gaps in the current literature on statistical literacy. Through an SLR, it is possible to evaluate the effectiveness of various instructional strategies—ranging from traditional teacher-centered approaches to more innovative, student-centered pedagogies—in enhancing students' statistical understanding (Aziz & Rosli, 2021; Schreiter et al., 2024). Moreover, the findings from an SLR can inform the development of curriculum and educational policies that better integrate statistical literacy in formal education, especially in the context of STEM (Pascual et al., 2025; Schreiter et al., 2024).

In addition to improving instructional design, an SLR can also uncover common challenges and misconceptions in the teaching and application of statistical concepts, such as the misuse of p values, which remain prevalent among both students and professionals (Sharma, 2017). Furthermore, by identifying dominant research trends and synthesizing insights across disciplines, an SLR provides a comprehensive evidence base to support effective teaching and learning practices (Ghodoosi et al., 2023; Gallou-Guyot et al., 2024). Therefore, conducting a systematic literature review on statistical literacy is not only timely but also necessary to advance educational research, inform practice, and promote a statistically literate society.

Recent studies have shown that students' statistical literacy skills remain generally low, especially in the area of drawing conclusions (Shobikhah et al., 2025; Sharma, 2017). Among the components of statistical literacy, data production has been identified as the most developed topic, although it still falls within a moderate proficiency level (Shobikhah et al., 2025; Sharma, 2017). Interestingly, no significant differences in statistical literacy were observed at the gender or class level (Shobikhah et al.,

 Received: June 19, 2025 | Accepted: August 12, 2025

2025; Sharma, 2017). However, the field of education has emerged as a differentiating factor, with significant disparities in statistical literacy levels across disciplines (Shobikhah et al., 2025; Sharma, 2017). In response to these findings, researchers emphasize the need for targeted training and instructional models to effectively enhance students' statistical literacy (Aziz & Rosli, 2021; Sharma, 2017).

A study by Lestari et al. (2024) found that the integration of e-campus platforms during the COVID-19 pandemic significantly improved statistical literacy among female students, thereby helping to reduce gender disparities in this area. Additionally, research indicates that female students often outperform male students in statistical reasoning tasks, challenging prevailing stereotypes about gender and quantitative reasoning (Risqi & Ekawati, 2020). While male students with low statistical literacy generally manage basic statistical calculations, they tend to face difficulties with more complex inferential tasks (Apino et al., 2024; Retnawati et al., 2024). These differences suggest that gender influences how students approach and engage with statistical inference, highlighting the importance of differentiated teaching strategies to support learners more effectively. This aligns with the findings of Hafiyusholeh et al. (2018), who reported significant differences in statistical literacy abilities between male and female students.

Statistical literacy is a critical skill in the 21st century and is influenced by a combination of educational, personal, and socioeconomic factors. Educationally, the teaching approach plays a central role; research shows that materials-based teaching is more effective in developing students' statistical literacy than teacher-centered or student-centered methods are (Aziz & Rosli, 2021). Moreover, the use of digital platforms tailored to younger generations can enhance engagement, although their effectiveness tends to decline with increased statistical complexity (Zavarrone, 2017). Curricular frameworks such as the GAISE report promote multivariate thinking and confounding concepts, which are essential for developing deeper statistical understanding (Schield, 2017). Institutional collaboration among schools, universities, and national statistical offices, along with fostering recognition and motivation, also supports the advancement of statistical literacy (Macfeely et al., 2017). On a personal level, student attitudes and self-efficacy are strong predictors; higher self-efficacy can mediate the relationship between past academic achievement and current interest, although this does not always translate to sustained engagement (Carmichael et al., 2010). Practical experience, motivation, and learning attitudes further reinforce students' ability to internalize statistical concepts (Lateh, 2024). While gender and grade level have no significant influence, a student's field of study has been identified as a key differentiator of statistical literacy levels (Riwayani et al., 2024). Access to resources, such as laptops, also correlates positively with statistical literacy (Retnawati et al., 2024).

From a socioeconomic standpoint, family income, parental education, and occupation indirectly impact statistical literacy by shaping executive functions and home learning environments (Buckingham et al., 2013; Harwell et al., 2017; Khan et al., 2024; Jasinska et al., 2022). The cultural context also plays a role; incorporating culturally relevant content, such as hip hopping, into instruction can align learning with students' social identities and improve statistical cognition (Ortiz et al., 2018). Furthermore, cultural capital—especially parents' educational background—has been shown to mediate students' academic achievement in literacy-related domains (Myrberg & Rosen, 2008). In addition, the cognitive styles of field dependence and field independence play crucial roles in statistical literacy. Field-independent students excel in solving statistical problems because of their analytical approach and consistent use of mathematical language (Afifah, et al., 2019). Field-dependent students may benefit from instructional strategies that help them recognize and adjust for their cognitive style's influence on comprehension (Hite, 2004). Understanding these multifaceted influences allows educators and policymakers to design more targeted, equitable interventions to strengthen statistical literacy across diverse student populations.

The identification of factors that influence statistical literacy is an important foundation in the formulation of a relevant pedagogical approach. However, to deepen the analysis and provide a strong theoretical context, a clear conceptual understanding of what is meant by statistical literacy is needed. Therefore, before further discussion related to statistical literacy, the following section outlines the definition of statistical literacy on the basis of various scientific perspectives. Explanations related to the concept of statistical literacy can be found in Table 1.

2. Materials and Methods

A systematic literature review employing bibliometric analysis quantitatively evaluates literature to identify trends, patterns, and key research entities within a field. The use of the PRISMA protocol ensures methodological rigor and transparency in the selection and evaluation of sources (Page et al., 2021; Szomszor et al., 2021). In this study, the inclusion criteria were as follows: (1) articles published up to May 31, 2025; (2) articles written in English; and (3) a focus on statistical literacy. Bibliometric analysis—conducted via VOSviewer—enabled the visualization of citation networks, author collaborations, and keyword cooccurrences, providing insight into the intellectual structure and dynamics of the research domain (Chen & Song, 2019). Combining bibliometrics with a systematic review allows researchers to synthesize empirical findings while mapping the landscape of scholarly activity and identifying key contributors and emerging themes (Szomszor et al., 2021; Chen & Song, 2019). Moreover, the integration of both methods offers a comprehensive view of the development, historical trajectory, and future directions of the research field, which is particularly beneficial in interdisciplinary studies aiming for deeper insights (Chen & Song, 2019; Page et al., 2021). Finally, bibliometric analysis has been used since at least the 1970s

for strategic evaluation of scholarly publications, including assessing journal influence through citation and economic metrics (Szomszor et al., 2021).

Table 1 Defining the elements of statistical literacy.

No.	Definition of Statistical Literacy	Reference
1.	Statistical literacy is defined as the ability of individuals equipped with statistical knowledge to manage data effectively and make informed decisions about their future. It is emphasized as a crucial skill that enables individuals to evaluate data-driven results and make sound judgments regarding their validity	Batur, A., Baki, A., 2022
2.	Statistical literacy is defined as the ability to model natural and social phenomena into statistical models, allowing individuals to determine population characteristics based on sample data analysis. It encompasses interpreting and criticizing statistical information produced by others and using statistical tools effectively	Lukman, Wahyudin, Suryadi, D., Dasari, D., Prabawanto, S. , 2022
3.	Statistical literacy is defined as the ability to read and interpret statistical information, enabling individuals to assess the validity of data they encounter in daily life, such as in news articles and advertisements. It encompasses basic statistical skills and critical thinking, allowing for informed decision-making based on statistical reports	Lilly, K., Conway, B.M., 2025
4.	Statistical literacy involves the ability to understand and engage with data management processes, which is critical for developing grounded, critical thinking regarding public data practices	Towse, J., Davies, R., Ball, E., Ivory, M., Gooding, B., 2022
5.	Statistical literacy is defined as the ability to interpret data correctly and make informed decisions based on statistical information, which is increasingly important in a data-driven world	Sole, M.A, 2025
6.	Statistical literacy is defined as the ability to read and interpret statistics used in everyday life, particularly by K-12 students, undergraduates, and the general public, as explored in statistical education literature. It emphasizes understanding statistical concepts and applying them to real-world scenarios, especially in media reports and websites	Brearley, A.M., Rott, K.W., Le, L.J., 2023
7.	Statistical literacy is defined as the ability to understand and critically evaluate data and statistical arguments, which is essential for informed citizenship in a democratic society. It involves not only interpreting data but also engaging with it to make sense of arguments presented in media and society	Weiland, T., & Sundrani, A., 2022
8.	Statistical literacy is defined as the ability to read, interpret, and evaluate information using statistical language, which includes understanding symbols, vocabulary, and statistical concepts.	Rodríguez-Alveal, F., et al 2024; Pascual, R.F., et al., 2025
9.	Statistical literacy is defined as the critical thinking needed for individuals to make judgments when presented with results and claims from investigations, emphasizing the ability to apply statistical tools and contextual knowledge in everyday situations	Watson, J., & Smith, C., 2022
10.	Statistical literacy (SL) is defined as the ability to read, understand, and interpret statistical information, which includes data interpretation, presentation, and critical thinking skills. It involves evaluating the relevance, accuracy, and reliability of statistics and effectively communicating findings to others. SL is essential for making informed decisions based on statistical evidence in various contexts, including higher education and the workplace	Lestari, K.E., et al, 2024
11.	Statistical literacy is defined as the capability to engage with statistical and mathematical information in media, which includes understanding and critically analysing statistical data, evaluating the credibility of news reports, and making informed judgments based on evidence. This concept encompasses various knowledge bases, such as literacy skills, mathematical skills, statistical skills, contextual knowledge, and critical skills	Gal, I., & Geige, V., 2022
12.	Statistical Literacy is an ability that has several main components, namely, 1) Understanding Basic Statistical Concepts: Understanding basic concepts such as the mean, median, mode, variance, and distribution, 2) Ability to Interpret Data: Being able to read and understand data presented in various formats such as graphs, tables, and diagrams, 3) Statistical Reasoning: Using statistical reasoning to create	Elfitra,Siregar, T.M., 2020; Ziegler, L.,Garfield, J., 2018; Schield, 2011; Shobikha, et al, 2025; Wahab & Mahmud, 2018; Riwayani, et al, 2024; Hahs-Vaughn, D.L, et al, 2017

	hypotheses, test hypotheses, and make decisions based on data, 4) Statistical Communication: Being able to communicate statistical findings effectively to different audiences.

The preliminary phase in this scholarly examination involves the selection of keywords, which can be accomplished through a macro methodology (top-down), progressing from expansive search trajectories to more narrowly defined studies and topics. Consequently, after the limitations inherent in prior research and the scarcity of studies addressing statistical literacy are evaluated, the "statistical literacy" is incorporated as a focal point within the title, abstract, and keyword sections of the article. Furthermore, the Scopus database is employed by researchers for a myriad of investigational purposes, including the execution of literature reviews, identifying subject-matter experts, and monitoring research trends.

According to the search outcomes retrieved on May 31, 2025, from the Scopus database utilizing the article title, abstract, and keywords "Statistical AND Literacy" across diverse academic disciplines, spanning from the earliest publication in 2002 to the most recent in 2025, the total number of articles about statistical literacy is 6223 documents (refer to Figure 1). Following these findings, screening process filters are documented according to their classification. Articles are eliminated on the basis of the document type: Book chapter (15), Review (8), Editorial (4), Letter (1), Erratum (1), non-English (28), non-Open Access (137), non-Social Sciences and Mathematics (24). The screening results, categorized by document type, yielded 89 articles. This document is then further analysed in this study to answer RQ1: Is the exploration of statistical literacy a subject that continues to hold significance for future scholarly inquiry? RQ2: What is the allocation of research investigations related to statistical literacy? RQ3: What are the potential contributions to theory and practice that could inform future research?

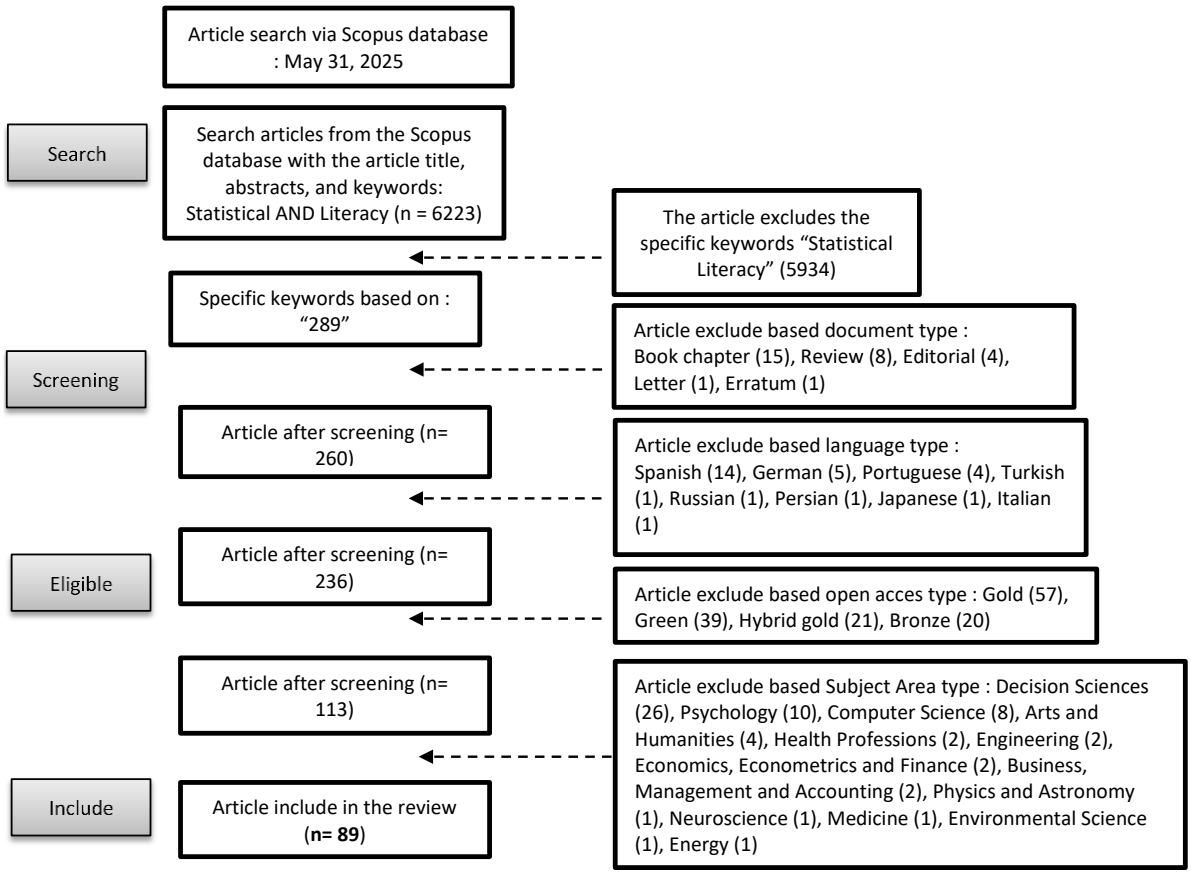

Figure 1 Systematic literature review information flow using PRISMA.

3. Results

The results of this study focus on findings from 89 articles in the Scopus database on statistical literacy. These data are sourced from identifying the number of articles published, publications throughout the years, and journal sources. This study will also highlight the most influential elements in statistical literacy, including the authors, affiliations, and countries involved.

RQ1: Is the exploration of statistical literacy a subject that continues to hold significance for future scholarly inquiry?

According to the data retrieved from the Scopus database, over four decades, scholarly work on statistical literacy comprises 89 articles; this suggests that investigations into statistical literacy remain comparatively rare, as illustrated in Figure 1. The exploration of statistical literacy commenced with its progressive development during the last decade, specifically since

2010. The inaugural study in 2002 was executed by Rumsey (2002) and was entitled "Statistical literacy as a goal for introductory statistics courses," which signified the advent of the term now recognized as Islamic leadership. At present, the evolution of research on statistical literacy has begun to capture a substantial number of scholars, concentrating on descriptions of statistical literacy, attitudes toward statistics, and statistical self-efficacy among applied linguistics research students (Zhang & Han, 2024); investigations of factors and conditions that influence students' statistical literacy (Retnawati et al., 2024); descriptions of students' statistical literacy by gender, class, and field of education (Maryati et al., 2024; Riwayani, et al, 2024); and descriptions of statistical literacy skills through the moodle learning management system (Maryati et al., 2024). Furthermore, statistical literacy contributes significantly to essential skills for the general public in today's big data era for future educators, especially prospective mathematics teachers (Hariyanti et al., 2025). Proficiency in statistical literacy is essential for these teachers because it improves students' statistical literacy in schools. Insufficient statistical literacy among prospective mathematics teachers may negatively impact their students' abilities, potentially resulting in a future society less equipped to interpret and use statistical data (see Figure 2).

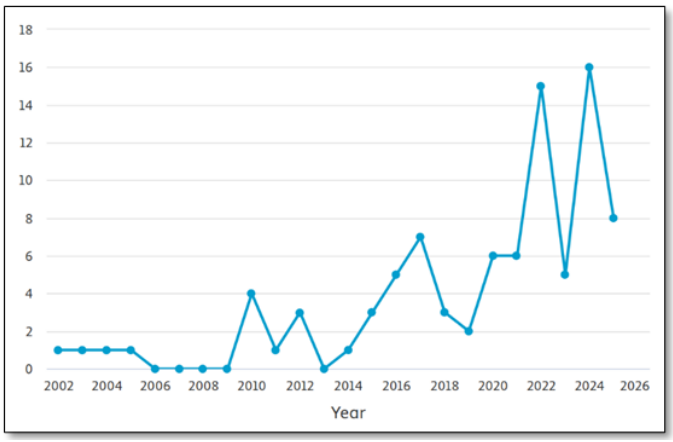

Figure 2 Number of statistical literacy publications. *Source:* Scopuse Database.

Since 2002, there has been limited literature on statistical literacy due to the lack of research published in reputable journals, creating an opportunity for future researchers to fill this gap. This research is significant for advancing insights into statistical literacy, which affects both personal conduct and the evolution of the statistical literacy framework. This may facilitate a better understanding of the practical and suitable application of statistical literacy in diverse sectors.

RQ2: What is the allocation of research investigations related to statistical literacy?

The analysis of the distribution of statistical literacy research in the 89 articles was executed by categorizing the articles according to classifications such as nation, region, affiliation, source, and author, with a constraint of solely the top 10 articles in each classification. Acumen regarding the allocation of scholars pertinent to statistical literacy will be advantageous for scholars and practitioners in elucidating the forthcoming research agenda, particularly in the sustainable advancement of the statistical literacy paradigm.

First, the allocation of scholarly inquiry pertinent to statistical literacy categorized by nation or geographical area is dominated by the United States with 28 articles, Indonesia with 10 articles, Australia with 8 articles, Germany with 7 articles, the United Kingdom with 7 articles, Spain with 5 articles, Greece with 4 articles, Turkey with 4 articles, Israel with 3 articles, and South Africa with 3 articles (see Figure 3).

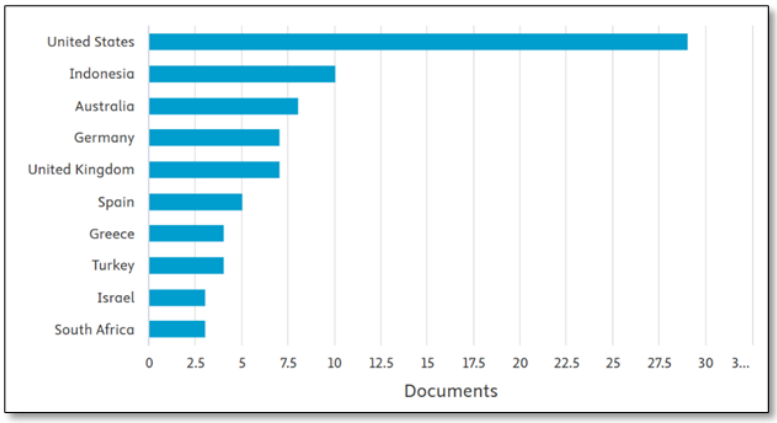

Figure 3 Number of articles by country or territory (top 10 countries). *Source:* Scopuse Database.

The distribution of scientific research focusing on statistical literacy shows that the United States occupies the leading position in academic contributions in this field, with a total of 28 published scientific papers. The second position is occupied by Indonesia, with 10 papers, indicating an increase in the attention and involvement of researchers from Southeast Asia in the issue of statistical literacy. Significant contributions also came from several other countries, such as Australia, which contributed 8 articles, and Germany and the United Kingdom, each with 7 articles. Spain also actively participated through 5 publications, followed by Greece and Turkey, which each contributed 4 articles. Moreover, Israel and South Africa were each recorded as contributing 3 scientific papers. In general, these data indicate that statistical literacy has become a global concern with dominance from developed countries but is also starting to gain a place in the research agenda of developing countries such as Indonesia. This indicates a growing recognition of the value of statistical literacy in enabling informed, data-driven decisions across different sectors. These findings indicate that the issue of statistical literacy has gained attention not only in countries with Muslim-majority populations but also in Western countries and other countries, reflecting the global relevance of the topic. Researchers will also analyse the relationships between the countries involved in statistical literacy research via VOSviewer software. This phase is imperative in formulating a systematic prospective research agenda. The examination's VOSviewer findings demonstrate the interrelations among nations in investigating the subject of statistical literacy (see Figure 4).

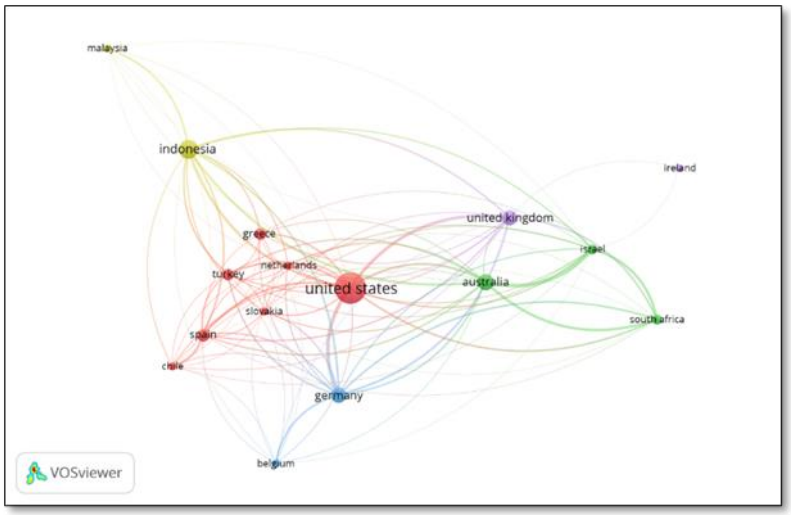

Figure 4 Network country visualization. *Source:* Output VOSviewer Software.

Mapping the allocation of scientific research that examines statistical literacy shows a strong concentration in certain countries, indicating that academic centers are active in this field. The United States stands out as the main contributor with 28 scientific papers, demonstrating a strong research tradition and institutional support for the development of statistical literacy. Interestingly, Indonesia is in second place with 10 papers, an achievement that shows the dynamics and growth of interest in mathematics education research, especially those focused on statistical literacy. Other countries that also show commitment to the development of statistical literacy include Australia (8 papers), Germany and the United Kingdom (7 papers each), and Spain (5 papers). This geographical diversity indicates that statistical literacy has become a global issue that crosses regional boundaries, although the number of contributions still shows disparities between countries. The presence of countries such as Greece, Turkey, Israel, and South Africa, each contributing 3–4 papers, also shows that attention to statistical literacy is not only centered in English-speaking countries but also has spread to southern Europe, the Middle East, and Africa.

Second, the allocation of scholarships related to statistical literacy predicated on institutional affiliations is predominantly characterized by Universitas Negeri Yogyakarta (Indonesia), with 4 articles; the University of Haifa (Israel), with 3 articles; the University of Houston (United Kingdom), with 2 articles; Universitas de Girona (Spain), with 2 articles; Australian Catholic University (Australia), with 2 articles; the University of Tasmania (Australia), with 2 articles; the University of Birmingham (United Kingdom), with 2 articles; Durham University (United Kingdom), with 2 articles; Karadeniz Technical University (Turkey), with 2 articles; and Lancaster University (United Kingdom), with 2 articles (See Figure 5).

Mapping institutional affiliations in statistical literacy research shows that contributions come not only from institutions in developed countries but also from developing countries such as Indonesia. Yogyakarta State University appears most prominently with 4 articles, showing strong institutional capacity in this field and indicating the active role of this university in developing statistical literacy studies at the national and international levels. Furthermore, the University of Haifa (Israel) follows with 3 articles, which shows the consistency of this institution in supporting statistical literacy-based research. Several other universities that provide equal contributions, each with 2 articles, include University of Houston (United States), Universitat de Girona (Spain), Australian Catholic University and University of Tasmania (Australia), University of Birmingham, Durham University, Lancaster University (United Kingdom), and Karadeniz Technical University (Turkey). The presence of

institutions from various regions of Asia, Europe, Australia, and America shows that statistical literacy is a cross-country and cross-cultural research area. The dominance of certain institutions, such as Yogyakarta State University, also reflects the existence of strong local initiatives in mainstreaming statistical literacy as part of mathematics education.

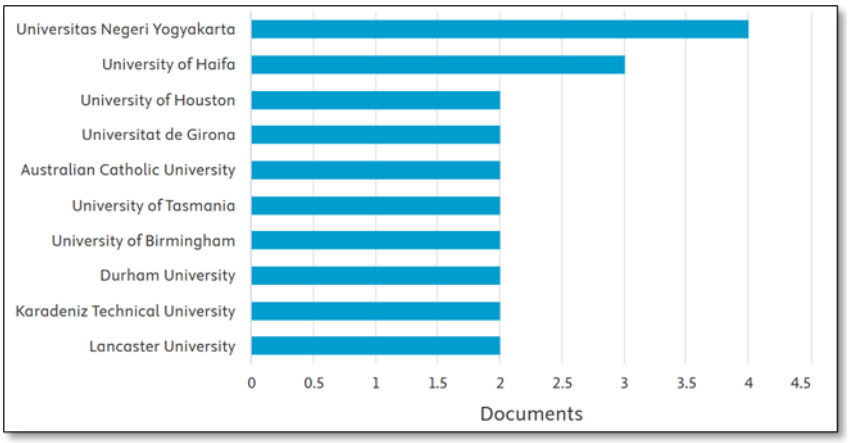

Figure 5 Number of articles by Affiliation (top 10 Affiliations). *Source:* Scopuse Database.

Third, the allocation of inquiries about statistical literacy on the basis of the source is characterized by Journal of Statistics Education with 9 articles, Journal of Statistics and Data Science Education with 7 articles, Statistics Education Research Journal with 5 articles, Numeracy with 4 articles, Educational Studies in Mathematics with 3 articles, mathematics with 3 articles, Zdm Mathematics Education with 3 articles, American Statistician with 2 articles, Educaton Sciences with 2 articles, and Egitim Ve Billim with 2 articles (see Figure 6).

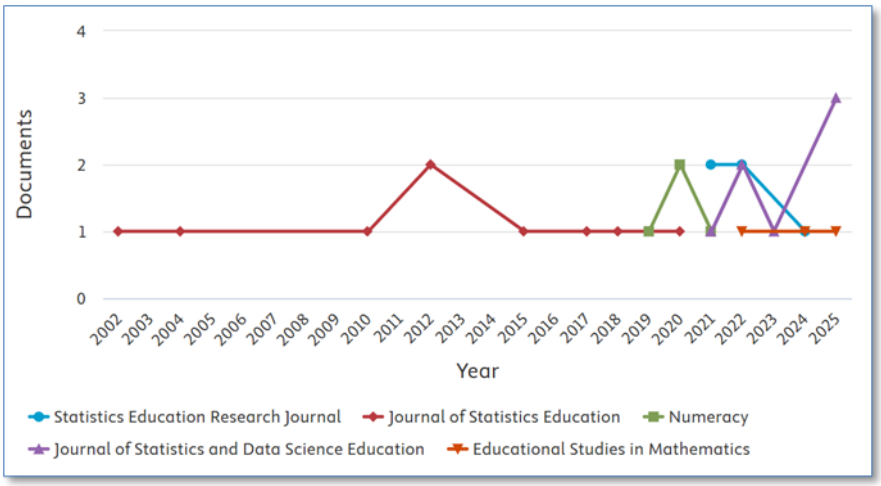

Figure 6 Number of articles by source (top 10 sources). *Source:* Scopuse Database.

The allocation of articles containing questions or studies on statistical literacy on the basis of their publication sources shows a fairly dominant pattern in specific journals in the fields of statistics and education. The Journal of Statistics Education and the Journal of Statistics and Data Science Education are the two main sources, with 9 and 7 articles, respectively, which strengthens the role of these journals as academic information centers on the topic of statistical literacy. The dominance of these journals also shows that statistical literacy is starting to be positioned as a stand-alone topic in education, no longer just a subtopic of general mathematics education. However, journals such as Numeracy, Zdm Mathematics Education, and Educational Studies in Mathematics also show that pedagogical approaches to statistics remain an important part of this research landscape. In addition, the presence of international journals such as American Statistician and Egitim ve Bilim indicates that statistical literacy studies also have a place in the global scientific community, not just local or national.

Fourth, the distribution of research related to statistical literacy based on the authors reveals no clear dominance. Among the top 10 authors, 9 (Abt, M; Alsina, A; Apino, E; Baki, A; Buscher, C; Geiger, V; Hidayati, K; Leuders, T; Loibl, K) have each written 2 articles, whereas only 1 author (Gal, I) has written 3 articles (see Figure 7).

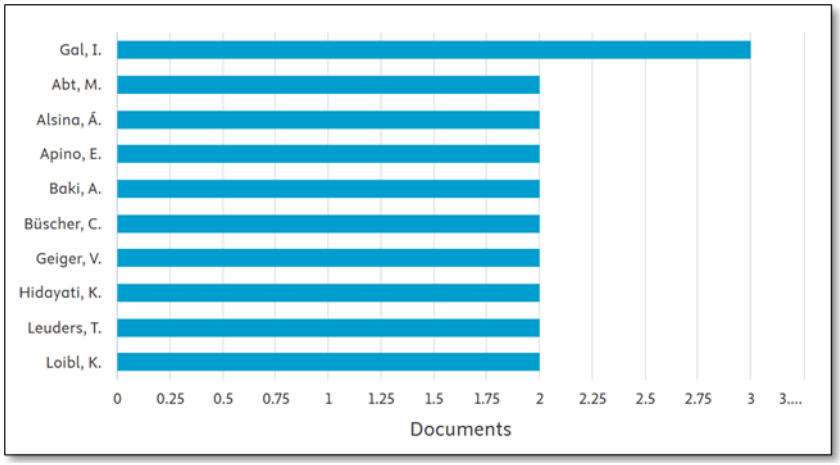

Figure 7 Count of publications by author (top 10 authors). *Source:* Scopuse Database.

The distribution of authors in research related to statistical literacy shows a relatively even pattern of contribution without any significant dominance from a particular individual. From the data analysed, there are ten top authors who are the most productive in this field. Nine of them each contributed two publications. Moreover, only one author has a greater contribution, with three publications. This finding indicates that the topic of statistical literacy has received attention from various researchers without any concentration of publications on one or two dominant figures. This can be interpreted as a sign that this field is still developing collectively and collaboratively and is open to contributions from various academic backgrounds.

RQ 3: What are the potential contributions to theory and practice that could inform future research?

The examination was undertaken on 89 manuscripts amassed from the Scopus repository. VOSviewer was employed to illustrate that the results may possess theoretical and pragmatic ramifications for forthcoming inquiries into statistical literacy. The results of the metadata analysis via VOSviewer can reveal which variables have been extensively researched by previous researchers and which variables have not been explored much, serving as a foundation for future studies. From a practitioner's perspective, the literature analysis results using VOSviewer will assist practitioners in implementing statistical literacy sustainably in the future and promoting the statistical literacy style for organizations worldwide.

Figure 8 shows the occurrence of statistical literacy (89), statistics education (9), statistics (6), statistics education research (6), education (5), data literacy (4), higher education (3), data visualization (3), mathematics education (3), gender (3), big data (2), primary education (2), critical thinking (2), cognition (1), and statistical illiteracy (1).

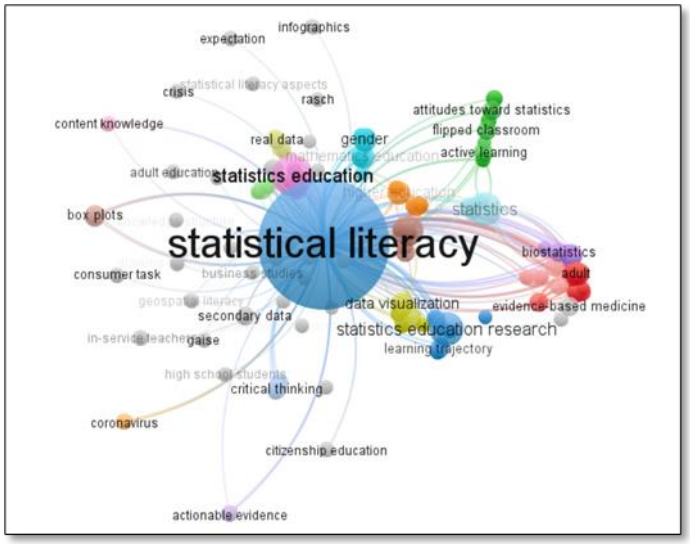

Figure 8 Co-occurrence framework and representation of key terms. *Source:* Output VOSviewer Software.

On the basis of the analysis of keywords in research related to statistical literacy, the frequency of the appearance of various main terms reflects the focus of the study in this field. The keyword "statistical literacy" is predominant, with 89 occurrences, indicating that this topic is the main focus of attention. Other terms that also frequently appear are "statistical

education" 9 times and "statistics" and "statistical education research" 6 times each, indicating an interest in the educational and research aspects of statistics.

In addition, keywords related to the educational context, such as "education" (5), "higher education" (3), and "mathematics education" (3), also appear quite significantly, revealing the relationship between statistical literacy and the realm of formal education. Keywords such as "data literacy" (4) and "data visualization" (3) indicate attention to the ability to interpret and present data effectively.

Additional topics such as “gender” (3), “big data” (2), “primary education” (2), “critical thinking” (2), “cognition” (1), and “statistical illiteracy” (1) also emerged, indicating a more specific and diverse focus in statistical literacy research, ranging from demographic aspects and teaching challenges to cognitive aspects and gaps in statistical understanding. Overall, this distribution of keywords illustrates the diversity of themes in statistical literacy research, with a primary emphasis on statistical literacy development and education at various levels of education and attention to relevant contextual and methodological issues (see Table 2).

Table 2 Keywords by authors.

Rank	Keyword	Total link strength
1	Statistical Literacy	404
2	Statistics	50
3	Statistics Education	35
4	Education	27
5	Statistical illiteracy	26
6	Cognition	25
7	Statistics Education Research	24
8	Data Literacy	22
9	Higher Education	19
10	Mathematics Education	14

Source: Output VoSviewer software.

This table presents the top ten keywords most frequently appearing in the literature on statistical literacy, along with their total link strength values, which indicate the intensity of connection between these keywords within the research network. The keyword "statistical literacy" dominates the field, with a total link strength of 404, underscoring its central role as the main subject of study and its extensive interrelation with other concepts. Following this, "Statistics" and "Statistics Education" hold significant positions with values of 50 and 35, respectively, highlighting the importance of both statistical content and its pedagogical aspects in the discourse.

The inclusion of "education" (27) emphasizes the broader educational context in which statistical literacy is situated, whereas "statistical illiteracy" (26) draws attention to the challenges and gaps that the field aims to address. The presence of "cognition" (25) reflects the consideration of mental processes involved in understanding and applying statistical concepts. Furthermore, keywords such as "Statistics Education Research" (24) and "Data Literacy" (22) reveal a growing interest in empirical studies on teaching methods and the complementary skill of data literacy, which is increasingly relevant in a data-driven society.

The terms "Higher Education" (19) and "Mathematics Education" (14) suggest that research extends into specific educational levels and disciplinary intersections, indicating a multidisciplinary approach to enhancing statistical literacy. Collectively, this keyword network not only maps the thematic landscape of statistical literacy research but also highlights critical areas such as education quality, cognitive engagement, and the evolving demands for data competence in various learning environments.

4. Discussion

The findings of this hybrid systematic literature review and bibliometric analysis confirm that statistical literacy remains a highly relevant and actively evolving area of scholarly investigation (RQ1). The consistent upwards trend in publication frequency, particularly after 2017, indicates a growing awareness of the importance of statistical literacy in education, data-informed decision-making, and critical citizenship (Sharma, 2017; Schield, 2017). Bibliometric mapping also highlights dominant keywords such as statistical literacy, statistics education, and data literacy, emphasizing the multidimensional nature of the field and its interconnection with broader educational and cognitive domains (Aziz & Rosli, 2021; Friedrich et al., 2024).

With respect to the allocation of research (RQ2), the analysis reveals a strong focus on higher education settings, with few studies targeting secondary or informal education sectors. This concentration suggests a potential research gap and aligns with previous reviews that call for more inclusive approaches across educational levels (Schreiter et al., 2024; Sharma, 2017). Moreover, regional disparities in publication outputs suggest an imbalance in the global research distribution, pointing to the need for more cross-cultural and context-sensitive studies.

From a theoretical and practical perspective (RQ3), this review underscores the necessity of integrating statistical thinking and reasoning within the broader construct of statistical literacy (Wild & Pfannkuch, 1999; Garfield & Ben-Zvi, 2007). Many studies still treat these constructs in isolation, whereas the literature increasingly supports their integration into cohesive educational models. Additionally, the review highlights the role of digital platforms and e-learning in improving statistical literacy, especially during the pandemic, although the long-term efficacy of such interventions requires further empirical validation (Zavarrone, 2017; Pinto et al., 2023).

The bibliometric analysis also revealed the prominence of keywords such as statistics education and data literacy, indicating a trend toward embedding statistical literacy within broader educational and digital literacies. This echoes findings from recent studies, which emphasize the importance of real-world data contexts and statistical modelling in fostering critical and reflective statistical thinking (Hasim et al., 2024; Wild & Pfannkuch, 1999). Consequently, incorporating contextualized data analysis and model-based reasoning into instruction is key to nurturing students' statistical competence and critical thinking.

Additionally, several studies support the growing significance of online learning environments in developing statistical literacy, particularly during the COVID-19 pandemic (Sefriani et al., 2024; Hussain et al., 2024). Platforms such as Edmodo and LMS have shown promise in facilitating access to statistics instruction, although challenges remain, including students' statistical anxiety and disparities in digital access (Blackburn, 2015; Flores et al., 2023). Hence, future teaching strategies must account not only for instructional content and delivery but also for emotional and technological equity to promote inclusive and sustainable statistical literacy.

Finally, this review uncovers underexplored areas within statistical literacy research, including secondary education, informal learning settings, and interdisciplinary approaches that connect statistics with STEM fields, cognitive development, and data ethics. These gaps suggest opportunities for future research to integrate active learning pedagogies, sociocultural responsiveness, and cross-curricular collaboration. This recommendation is supported by findings from large-scale STEM meta-analyses, which show that active learning significantly improves students' conceptual understanding and lowers failure rates (Hake, 1998). A hybrid, interdisciplinary model of statistical literacy may thus offer greater relevance and adaptability in increasingly data-driven and diverse educational contexts.

5. Final Considerations

This hybrid systematic literature review and bibliometric analysis demonstrates that statistical literacy continues to be a crucial and evolving area of scholarly focus. The increasing volume of publications and diversity of research topics underscore its relevance in addressing contemporary educational and societal challenges. Statistical literacy not only encompasses the ability to interpret and use statistical data but also involves the integration of statistical thinking and reasoning, forming a comprehensive framework for informed decision-making.

The study reveals significant gaps in research distribution, with an overrepresentation of studies in higher education and underrepresentation in secondary and informal education contexts. Furthermore, the integration of digital learning tools, sociocultural factors, and real-world data applications are emerging themes that offer promising directions for future inquiry. Theoretical developments also point toward the need for holistic and interdisciplinary approaches that bridge cognitive, pedagogical, and technological dimensions.

Overall, this review advocates for a renewed commitment to statistical literacy as a key component of education in the 21st century. Future research should aim to close identified gaps, ensure inclusivity, and develop effective instructional models that address the varying needs of learners in dynamic, data-intensive contexts.

Ethical Considerations

Not applicable.

Conflict of Interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

References

Afifah, D. S. N., & Nafi'an, M. I. (2019). An onto-semiotic approach: Analyzing of field-independent and field-dependent students' understanding in solving statistical problems. *Journal of Physics: Conference Series, 1175*, 012148. https://doi.org/10.1088/1742-6596/1175/1/012148

Apino, E., Retnawati, H., Purbani, W., & Hidayati, K. (2024). The statistical literacy of mathematics education students: An investigation on understanding the margin of error. *TEM Journal, 13*(1), 293–302. https://doi.org/10.18421/TEM131-31

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

Batur, A., & Baki, A. (2022). Lise öğrencilerinin istatistik okuryazarlık düzeyleri ile istatistik okuryazarlık öz yeterlik algıları arasındaki ilişkinin incelenmesi. *Eğitim ve Bilim, 47*(209), 171–205. https://doi.org/10.15390/EB.2022.9970

Blackburn, H. (2015). The status of women in STEM in higher education: A review of the literature 2004–2014. *Science & Technology Libraries, 34*(3), 235–273. https://doi.org/10.1080/0194262X.2015.1114816

Brearley, A. M., Rott, K. W., & Le, L. J. (2023). A biostatistical literacy course: Teaching medical and public health professionals to read and interpret statistics in the published literature. *Journal of Statistics and Data Science Education, 31*(3), 286–294. https://doi.org/10.1080/26939169.2023.2165987

Buckingham, J., Wheldall, K., & Beaman-Wheldall, R. (2013). Why poor children are more likely to become poor readers: The school years. *Australian Journal of Education, 57*(3), 190–213. https://doi.org/10.1080/00131911.2013.795129

Camilo, C., & Vaz Garrido, M. (2019). Systematic review in psychology: Challenges and guidelines. *Analise Psicologica, 37*(3), 391–403. https://doi.org/10.14417/ap.1546

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

Chen, C., & Song, M. (2019). Visualizing a field of research: A methodology of systematic scientometric reviews. *PLoS ONE, 14*(10), e0223994. https://doi.org/10.1371/journal.pone.0223994

Elfitra, & Siregar, T. M. (2020). Statistical literacy analysis of mathematics education students through KKNI assignments. *Journal of Physics: Conference Series, 1462*, 012028. https://doi.org/10.1088/1742-6596/1462/1/012028

Flores, J. R., Cueva, C. E., Rodríguez, S., & Guerra, J. (2023). The use of LMS platforms in the development of statistical skills: A systematic review. *Education and Information Technologies, 28*(1), 103–123. https://doi.org/10.1007/s10639-022-11182-3

Friedrich, A., Schreiter, S., Vogel, M., & Malone, S. (2024). What shapes statistical and data literacy research in K-12 STEM education? A systematic review of metrics and instructional strategies. *International Journal of STEM Education, 11*(1), 1–24. https://doi.org/10.1186/s40594-024-00517-z

Gal, I., & Geige, V. (2022). Welcome to the era of vague news: A study of the demands of statistical and mathematical products in the COVID-19 pandemic media. *Educational Studies in Mathematics, 111*(1), 5–28. https://doi.org/10.1007/s10649-022-10151-7

Gallou-Guyot, M., Rousseau, C., & Perrochon, A. (2024). The limits of systematic literature reviews – when too much information becomes deleterious. *Kinesitherapie, 24*(267), 60–65. https://doi.org/10.1016/j.kine.2023.11.004

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

Ghodoosi, B., West, T., Li, Q., & Dey, S. (2023). A systematic literature review of data literacy education. *Journal of Business and Finance Librarianship, 28*(2), 87–104. https://doi.org/10.1080/08963568.2023.2171552

Hafiyusholeh, M., Budayasa, K., & Siswono, T. Y. E. (2018). Statistical literacy: High school students in reading, interpreting, and presenting data. *Journal of Physics: Conference Series, 947*, 012036. https://doi.org/10.1088/1742-6596/947/1/012036

Hahs-Vaughn, D. L., Acquaye, H., Griffith, M. D., Jo, H., Matthews, K., & Acharya, P. (2017). Statistical literacy as a function of online versus hybrid course delivery format for an introductory graduate statistics course. *Journal of Statistics Education, 25*(3), 112–121. https://doi.org/10.1080/10691898.2017.1370363

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics, 66*(1), 64–74. https://doi.org/10.1119/1.18809

Hariyanti, F., Budayasa, I. K., & Setianingsih, R. (2025). A portrait of prospective mathematics teachers' readiness in statistical literacy of school students. *Perspektivy Nauki i Obrazovania, 73*(1), 190–201. https://doi.org/10.32744/pse.2025.1.12

Hariyanti, F., Budayasa, I. K., & Setianingsih, R. (2025). The role of AI in enhancing statistical literacy: A systematic review in education. *Multidisciplinary Reviews, 8*(12), 2025376. https://doi.org/10.31893/multirev.2025376

Harwell, M., Maeda, Y., Bishop, K., & Xie, A. (2017). The surprisingly modest relationship between SES and educational achievement. *Journal of Experimental Education, 85*(2), 197–214. https://doi.org/10.1080/00220973.2015.1123668

Hasim, M. S., Daud, M. F., & Sulaiman, T. (2024). Contextual teaching and learning: Enhancing students' statistical reasoning through real-world data. *Malaysian Journal of Learning and Instruction, 21*(1), 135–152. https://doi.org/10.32890/mjli2024.21.1.6

Hite, C. E. (2004). Expository content area texts, cognitive style and gender: Their effects on reading comprehension. *Reading Research and Instruction, 43*(4), 41–74. https://doi.org/10.1080/19388070409558416

Holmberg, C. (2024). Toward a better understanding of statistical significance and p-values in nursing. *Nursing Forum, 59*(1), 40–48. https://doi.org/10.1155/2024/7263781

Hussain, S., Khan, A., & Ramzan, M. (2024). The impact of online learning platforms on students' achievement in statistics education during COVID-19. *Education and Information Technologies, 29*(1), 456–470. https://doi.org/10.1007/s10639-023-11871-2

Jasinska, K., Zinszer, B., Xu, Z., & Akpé, H. (2022). Home learning environment and physical development impact children's executive function development and literacy in rural Côte d'Ivoire. *Cognitive Development, 62*, 101149. https://doi.org/10.1016/j.cogdev.2022.101265

Khan, F., Wortsman, B., Whitehead, H. L., & Jasińska, K. K. (2024). Modeling the associations between socioeconomic risk factors, executive function components, and reading among children in rural Côte d'Ivoire. *Cognitive Development, 67*, 101210. https://doi.org/10.1016/j.cogdev.2024.101436

Lateh, A. (2024). Exploring statistical literacy deficiencies and causal associations among master and doctoral graduates in Thailand. *Pakistan Journal of Life and Social Sciences, 22*(1), 10–18. https://doi.org/10.57239/PJLSS-2024-22.2.00322

Lestari, K. E., Risnawita, Yudhanegara, M. R., Nugraha, E. S., & Sylviani, S. (2024). Correspondence analysis on statistical literacy and gender: Embedding e-campus platform with random assignment of matched subject in explanatory analysis. *BAREKENG: Jurnal Matematika dan Aplikasi, 18*(3), 1975–1988. https://doi.org/10.30598/barekengvol18iss3pp1975-1988

Lilly, K., & Conway, B. M. (2025). A case study of an evaluation of pen-and-paper homework and project-based learning of statistical literacy in an introductory statistics course. *Journal of Statistics and Data Science Education, 00*(0), 1–7. https://doi.org/10.1080/26939169.2025.2462604

Lukman, L., Wahyudin, W., Suryadi, D., Dasari, D., & Prabawanto, S. (2022). Studying student statistical literacy in statistics lectures on higher education using grounded theory approach. *Infinity, 11*(1), 163–176. https://doi.org/10.22460/infinity.v11i1.p163-176

Macfeely, S., Campos, P., & Helenius, R. (2017). Key success factors for statistical literacy poster competitions. *Statistics Education Research Journal, 16*(1), 68–75. https://doi.org/10.52041/serj.v16i1.224

Maryati, I., Fisher, D., Yatim, S. A. M., & Mauladaniyati, R. (2024). Statistical literacy ability of students through virtual learning environment based on Moodle-Learning Management System. *International Journal of Information and Education Technology, 14*(1), 99–106. https://doi.org/10.18178/ijiet.2024.14.1.2029

Myrberg, E., & Rosén, M. (2008). A path model with mediating factors of parents' education on students' reading achievement in seven countries. *Educational Research and Evaluation, 14*(6), 507–520. https://doi.org/10.1080/13803610802576742

Ortiz, N. A., Capraro, M. M., & Capraro, R. M. (2018). Does it really matter? Exploring cultural relevance within a majority white classroom. *Journal of Negro Education, 87*(1), 66–78. https://doi.org/10.7709/jnegroeducation.87.4.0404

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *PLoS Medicine, 18*(3), e1003583. https://doi.org/10.1136/bmj.n71

Pascual, R. F., Caballero Mariscal, D., Pinto, M., & Marín-Jiménez, A. E. (2025). Attitudes of university students toward statistics as a pathway to data literacy: A meta-analysis. *Journal of Statistics and Data Science Education, 00*(0), 1–18. https://doi.org/10.1080/26939169.2025.2475765

Pinto, M., Caballero-Mariscal, D., García, F.-J., & Gómez-Camarero, C. (2023). A strategic approach to information literacy: Data literacy. A systematic review. *Profesional de la Información, 32*(6). https://doi.org/10.3145/epi.2023.nov.09

Retnawati, H., Hidayati, K., Apino, E., Rafi, I., & Rosyada, M. N. (2024). Exploring influential factors and conditions shaping statistical literacy among undergraduate students in mathematics education. *International Journal of Cognitive Research in Science, Engineering and Education, 12*(1), 1–17. https://doi.org/10.23947/2334-8496-2024-12-1-1-17

Risqi, E. N., & Ekawati, R. (2020). How is the statistical literacy of upper secondary students based on gender differences? *Jurnal Riset Pendidikan dan Inovasi Pembelajaran Matematika, 4*(1), 53–67. https://doi.org/10.26740/jrpipm.v4n1.p53-67

Riwayani, R., Istiyono, E., Supahar, S., & Soeharto, S. (2024). Analyzing students' statistical literacy skills based on gender, grade, and educational field. *International Journal of Evaluation and Research in Education, 13*(1), 1–9. http://doi.org/10.11591/ijere.v13i2.26299

Rodríguez-Alveal, F., Maldonado-Fuentes, A. C., & Díaz-Levicoy, D. (2024). Lexical ambiguities in statistics declared by in-training and in-service teachers. *EURASIA Journal of Mathematics, Science and Technology Education, 20*(4), em2422. https://doi.org/10.29333/ejmste/14359

Rumsey, D. J. (2002). Statistical literacy as a goal for introductory statistics courses. *Journal of Statistics Education, 10*(3), 1–12. https://doi.org/10.1080/10691898.2002.11910678

Schield, M. (2011). Statistical literacy: A new mission for data producers. *Statistical Journal of the IAOS, 27*(3–4), 173–183. https://doi.org/10.3233/SJI-2011-0732

Schield, M. (2017). GAISE 2016 promotes statistical literacy. *Statistics Education Research Journal, 16*(1), 31–37. https://doi.org/10.52041/serj.v16i1.214

Schreiter, S., Friedrich, A., Fuhr, H., & 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 – Mathematics Education, 56*(1), 105–122. https://doi.org/10.1007/s11858-023-01531-1

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 – Mathematics Education, 56*(1), 31–45. https://doi.org/10.1007/s11858-023-01531-1

Sefriani, R., Abidin, N. Z., & Irawan, Y. (2024). Effectiveness of Edmodo in improving students' statistical literacy: A quasi-experimental study. *Journal of Educational Technology & Online Learning, 5*(1), 45–59. https://doi.org/10.1080/jetol.2024.11857

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

Shobikhah, A., Sukestiyarno, Y. L., Agoestanto, A., & Cahyono, A. N. (2025). Bibliometrics on the development of students' statistical literacy: A scoping review of research between the years 2000–2024. *TEM Journal, 14*(1), 871–886. https://doi.org/10.18421/TEM141-77

Sole, M. A. (2025). An investigation designed to teach statistical thinking in the midst of the COVID-19 pandemic: Are teens living like vampires? *Journal of Statistics and Data Science Education, 00*(0), 1–11. https://doi.org/10.1080/26939169.2025.2455197

Szomszor, M., Adams, J., Fry, R., Gebert, C., & Pendlebury, D. A. (2021). Interpreting bibliometric data. *Frontiers in Research Metrics and Analytics, 5*, 628703. https://doi.org/10.3389/frma.2020.628703

Towse, J., Davies, R., Ball, E., James, R., Gooding, B., & Ivory, M. (2022). Lustre: An online data management and student project resource. *Journal of Statistics and Data Science Education, 30*(3), 266–273. https://doi.org/10.1080/26939169.2022.2118645

Wahab, A., & Mahmud, A. (2018). The effectiveness of a learning module for statistical literacy. *New Educational Review, 53*(3), 187–200. https://doi.org/10.15804/tner.2018.53.3.16

Watson, J., & Smith, C. (2022). Statistics education at a time of global disruption and crises: A growing challenge for the curriculum, classroom and beyond. *Curriculum Perspectives, 42*, 171–179. https://doi.org/10.1007/s41297-022-00167-7

Weiland, T. (2017). Problematizing statistical literacy: An intersection of critical and statistical literacies. *Educational Studies in Mathematics, 96*, 33–47. https://doi.org/10.1007/s10649-017-9764-5

Weiland, T., & Sundrani, A. (2022). Opportunities for K–8 students to learn statistics created by states' standards in the United States. *Journal of Statistics and Data Science Education, 30*(2), 165–178. https://doi.org/10.1080/26939169.2022.2075814

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review, 67*(3), 223–248. https://doi.org/10.2307/1403699

Zavarrone, E. (2017). Latent growth and statistical literacy. In *Studies in Classification, Data Analysis, and Knowledge Organization* (pp. 389–396). Springer.

Zhang, P., & Han, C. (2024). Examining statistical literacy, attitudes toward statistics, and statistics self-efficacy among applied linguistics research students in China. *International Journal of Applied Linguistics, 34*(2), 433–449. https://doi.org/10.1111/ijal.12500

Ziegler, L., & Garfield, J. (2018). Developing a statistical literacy assessment for the modern introductory statistics course. *Statistics Education Research Journal, 17*(2), 161–178. https://doi.org/10.52041/serj.v17i2.164