CROSS-CURRICULAR APPROACHES TO PROMOTING STATISTICAL AND DATA LITERACY IN STEM SCHOOL EDUCATION: A SYSTEMATIC REVIEW

Saskia Schreiter[1], Markus Vogel[1], Anja Friedrich[2], Sarah Malone[2], Roland Brünken[2], Sebastian Becker-Genschow[3], Jessica Lehmann[2], Tobias Mosetter[1], Stephanie Haaß[1], and Jochen Kuhn[4]
[1]Heidelberg University of Education, Germany
[2]Saarland University, Saarbrücken, Germany
[3]University of Cologne, Germany
[4]Ludwig Maximilian University, Munich, Germany
schreiter@ph-heidelberg.de

Statistical and data literacy are counted among the key competencies of the 21st century and are often set out as an expected outcome of school education. Accordingly, curricular adaptions of science, technology, engineering, and mathematics (STEM) education have been made in many countries with the aim of developing competencies that promote competent handling of data, statistical information, and their forms of representation. In this paper, we report first results of an ongoing systematic review on the definition and fostering of statistical and data literacy in STEM school education. This paper focuses explicitly on cross-curricular approaches to promoting statistical and data literacy in STEM school education, which have been described and empirically investigated in international research.

INTRODUCTION

Competent use of data, statistics and their forms of representation is of great importance in our data-driven society and is often set out as an expected outcome of school education (Chalkiadaki, 2018). The relevance of such skills, which are overarchingly categorized as statistical literacy and data literacy (e.g., Gal, 2002; Gould, 2017) appears to be increasingly important in broader public perception and research (OECD, 2021). Technological progress makes it possible to collect, store, and analyze more and more data, and important decisions in society, business, and science are increasingly made based on data (Schüller & Busch, 2019). Not least the current global pandemic situation has shown the relevance of citizens being able to critically engage with interpretations and representations of data in the media.

The increasing importance of statistics and data in daily life has led to adaptions of school curricula in STEM domains in many countries, with the aim of developing competencies that promote competent use of data, statistics, and their forms of representation. Existing literature reviews identified promising methods for promoting statistical and data literacy in higher education (e.g., Schüller & Busch, 2019). What has not been explicitly addressed so far is the question of what role school-based STEM education plays in building important foundational skills for statistical and data literacy.

The main objectives of the ongoing systematic review described in this paper are to compile, systematize, and interpret relevant findings from international research on the role of STEM subjects in fostering statistical and data literacy. The question addressed within the framework of this paper is: "What cross-curricular approaches to promote statistical/data literacy in STEM school education are empirically investigated?". In this contribution, we refer to "cross-curricular" when more than one STEM school subject is addressed.

STATISTICAL AND DATA LITERACY IN STEM SCHOOL EDUCATION

Many statistics educators and scholars have tried to define and describe statistical literacy (e.g., Gal, 2002) and data literacy (e.g., Schüller & Busch, 2019). According to Gal (2002), statistical literacy is the ability to interpret, critically evaluate, and communicate statistical information. Gould (2017) found that knowledge about data collection, privacy, and the creation of data representations is also gaining relevance. Thus, additional skills were included in his definition, and statistical literacy was expanded to data literacy, which was understood as an overarching concept. However, it should be noted that there is no consensus on the definitions of these terms, and often, the terms are defined for adults as data recipients (e.g., Gal, 2002).

The analysis of various competency frameworks and educational standards (e.g., National Council of Teachers of Mathematics, 2000) shows that some of the above listed competency areas related to statistical and data literacy are already listed in STEM education curricula (e.g., measuring, collecting data, experimenting, presenting data, interpreting data representations, evaluating data,

In S. A. Peters, L. Zapata-Cardona, F. Bonafini, & A. Fan (Eds.), *Bridging the Gap: Empowering & Educating Today's Learners in Statistics. Proceedings of the 11th International Conference on Teaching Statistics (ICOTS11 2022), Rosario, Argentina.* International Association for Statistical Education. iase-web.org

concept of measurement uncertainty). According to the educational plans, such skills should be developed in mathematics and science classes from primary school age and in all STEM subjects in secondary schools. What is missing so far, however, is an overview of cross-curricular approaches that have been empirically studied for their effectiveness in promoting statistical and data literacy in STEM education at the primary and secondary levels. The ongoing systematic review presented in this paper contributes to fill this research gap.

METHOD

Search Strategy

The systematic review presented in this paper is planned, conducted, and reported according to PRISMA 2020 guidelines (Page et al., 2021). It is preregistered at the open-source platform OSF (Friedrich et al., 2021, September 3). A broad systematic search of three literature databases (PsychInfo, ERIC, Web of Science) was conducted in September 2021 that resulted in 18,243 records (see Figure 1). The search string included the terms ("data," "quantitative," "probabil*," "data oriented," "data based," "statistic*), supplemented with keywords ("literacy," "literate," "reason*," "competen*," "think," "understand," "comprehend*," "argu*," "mastery") in every possible combination and order, as well as the terms ("data and chance," "uncertainty and data") in every possible order.

Furthermore, a systematic search of 16 relevant conferences in psychological and educational contexts was conducted that resulted in 51 records. Please note that the dataset is not yet complete, because a for- and backwards search for further possibly relevant articles based on the included articles is still pending.

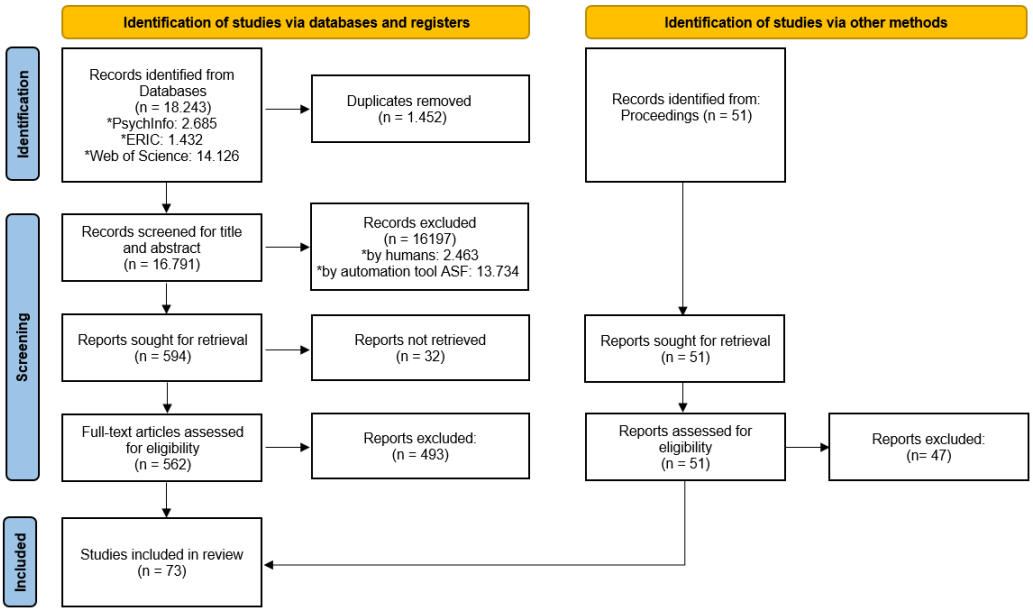

Figure 1. PRISMA flowchart

Study Selection, Eligibility Criteria, and Data Extraction

After duplicates were removed, two independent reviewers performed a title–abstract screening against eligibility criteria using *ASReview* (van Haastrecht et al., 2021), an active learning software based on machine learning. Discrepancies were resolved by a third researcher. All potentially relevant full-text articles were retrieved and independently screened against eligibility criteria by two researchers. All included studies met the following eligibility criteria:

- At least an abstract in the English language
- Original research
- Investigation scope relates to school education
- Investigation scope relates to STEM domains (physics, biology, chemistry, mathematics, (computer) science, technology, general studies)
- Peer reviewed

In each eligible full-text publication, findings relevant to the research questions were coded, extracted, and synthesized using qualitative content analysis. For the research question presented in this paper, study details in terms of information about general study characteristics, sample size, addressed STEM domains, school career (primary/secondary), as well as detailed information on the investigated approach to promote statistical/data literacy were extracted.

RESULTS

Of all studies included in the overall systematic review, eight were identified that empirically examined cross-curricular approaches to promoting statistical/data literacy in STEM education (see Table 1). Several of these studies used interdisciplinary learning environments addressing authentic problems and including real-world datasets. This approach has been shown to be effective in developing students' knowledge of statistical concepts (Andre et al., 2020) and their ability to ask and answer data-based questions (Swan et al., 2013; Vahey et al., 2012). Furthermore, increases in students' confidence in data analysis and statistical skills (such as data management and data analysis; Dierker et al., 2017) as well as positive effects on students' interest in data (Kochevar et al., 2015) have been found. Although most studies used large existing data sets, one study examined the effectiveness of an approach in which students collected their own data (Wolff et al., 2021). The results showed that experience with data collection promotes a critical approach to data among students.

A common feature of most of the approaches listed in Table 1 is the integration of digital media. The usage of specific software (e.g., *Gapminder* in Andre et al., 2020; Jiang et al., 2020) and novice-friendly web-based interfaces (Kochevar et al., 2015; Mosquera et al., 2020; Wolff et al., 2019) has been demonstrated to be effective in helping students deal with large and complex real-world data sets.

Table 1. Summary of included studies on cross-curricular approaches

Study	Domains	School career	Sample size	Approach to promote statistical/data literacy
Andre et al. (2020)	Interdisciplinary STEM	Secondary (L)	45	Real-world problem-based learning
Dierker et al. (2017)	Mathematics, Computer Science	Secondary (U)	33	Real-world project-based learning
Jiang et al. (2020)	Interdisciplinary STEM	Secondary (L, U)	17	Collaborative interactions; personal contexts
Kochevar et al. (2015)	Mathematics, Science	Secondary (U)	195	Real-world project-based learning
Mosquera et al. (2020)	Interdisciplinary STEM	Secondary (L)	39	Game-based learning
Swan et al. (2013)	Mathematics, Science	Secondary (L)	576	Real-world problem-based learning
Vahey et al. (2012)	Mathematics, Science	Secondary (L)	606	Real-world problem-based learning
Wolff et al. (2019)	Interdisciplinary STEM	Primary, Secondary (L)	67	Collecting own data

Note: The School career is divided into Primary (grade level 1–4), Lower Secondary (L) (grade level 5–8) and Upper Secondary (U) Education (grade level 9–13).

CONCLUSION

In this paper we reviewed findings from international research on cross-curricular approaches to promoting statistical/data literacy in STEM education. Using authentic problems and incorporating real-world data has been shown in several studies to be an effective cross-curricular approach to increase students' knowledge, interest, and motivation in statistics and data literacy. The use of specific software and novice-friendly web-based interfaces has proven helpful for students to deal with large complex data sets. In the next step, the review of international research on the promotion of statistical and data literacy in STEM school education will also be extended to subject-specific approaches.

REFERENCES

(*Included in systematic review)

*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)*. ISI/IASE. http://iase-web.org/documents/papers/rt2020/IASE2020%20Roundtable%2012_ANDRE.pdf?1610923749

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://www.e-iji.net/dosyalar/iji_2018_3_1.pdf

*Dierker, L., Ward, N., Alexander, J., & Donate, E. (2017). Engaging underrepresented high school students in data driven storytelling. *Journal of Education and Training Studies, 5*(4), 54–63. https://doi.org/10.11114/jets.v5i4.2187

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. https://doi.org/10.17605/OSF.IO/DVPBN

Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, *70*(1), 1–25. https://doi.org/10.1111/j.1751-5823.2002.tb00336.x

Gould, R. (2017). Data literacy is statistical literacy. *Statistics Education Research Journal*, *16*(1), 22–25. https://doi.org/10.52041/serj.v16i1.209

*Jiang, S., & Kahn, J. (2020). Data wrangling practices and collaborative interactions with aggregated data. *International Journal of Computer-Supported Collaborative Learning, 15*, 257–281. https://doi.org/10.1007/s11412-020-09327-1

*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

*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. Peña-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)*. ILRN. https://doi.org/10.23919/iLRN47897.2020.9155144

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*.

Organization for Economic Co-operation and Development. (2021). *21st-century readers: Developing literacy skills in a digital world*. OECD Publishing. https://doi.org/10.1787/a83d84cb-en.

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., May-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* 2021, *372*(71). https://doi.org/10.1136/bmj.n71

Schüller, K. & Busch, P. (2019). *Data literacy: Ein systematic review.* Hochschulforum Digitalisierung.

*Swan, K., Vahey, P., van 't Hooft, M., Kratcoski, A., Rafanan, K., Stanford, T., Yarnall, L., & Cook, D. (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), 91–110. https://doi.org/10.7771/1541-5015.1307

*Vahey, P., Rafanan, K., Patton, C., Swan, K., van 't Hooft, M., Kratcoski, A., & Stanford, T. (2012). A cross-disciplinary approach to teaching data literacy and proportionality. *Educational Studies in Mathematics, 81*, 179–205. https://www.jstor.org/stable/23254237

van Haastrecht, M., Sarhan, I., Yigit Ozkan, B., Brinkhuis, M., & Spruit, M. (2021). SYMBALS: A systematic review methodology blending active learning and snowballing. *Frontiers in research Metrics and Analytics*, *6*, Article 685591. https://doi.org/10.3389/frma.2021.685591

*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. https://doi.org/10.1016/j.ijhcs.2019.03.006