Literature review on the competencies of data literacy for middle-grade learners

Semi Yeom[1]

Abstract

In today's data-driven world, it is crucial for students to be data literate; able to view, understand, and reason with data in multimodal forms representing real-world phenomena. Despite its importance, data literacy is rarely integrated into K-12 curricula, and its definition remains unclear for this age group. This paper reviews existing literature to define the competencies relevant to adolescent learners and highlights those crucial for middle-grade students. A literature review of theoretical and empirical discussions on data literacy concepts, instructional practices, and assessments revealed eight key competencies. Among these, two were identified as most critical for middle-grade students: interpreting data representations and evaluating claims based on data representations. This paper aims to serve as a conceptual and practical guide to enhance data literacy in educational settings, providing a foundation for educators and researchers to collaboratively support middle-grade learners.

Keywords

Data literacy, K-12 school settings, literature review, middle-grade learners

Introduction

In today's technology-driven world, data—numerical, textual, algorithmic, and visual—shape how we perceive and interact with our surroundings. Algorithms guide online experiences, such as news articles featuring statistical insights or personalized advertisements based on browsing habits. Big data, including numbers, texts, images, times, and locations, plays a significant role in individual and organizational decision-making (Fontichiaro & Oehrli, 2016; Gould, 2017).

Data literacy involves critically analyzing, interpreting, and using data in various contexts. It encompasses understanding the social and cultural dimensions of data, which are deeply rooted in power structures and individual identities (Twidale et al., 2013). Researchers define data literacy as the ability to interact with and contextualize data, recognizing that interpretations are shaped by personal and cultural perspectives (English & Watson, 2018; Gunter, 2007; van't Hooft et al., 2012;). Students today frequently encounter multimodal representations of real-world phenomena, and data literacy may provide them with the necessary tools to engage with these forms of information thoughtfully.

With data literacy, students can study real-life problems and develop evidence-based solutions (Erwin Jr., 2015; Yates et al., 2021;). This skill enables them to recognize the factors and contexts underlying

presented data and to apply data purposefully in diverse situations. Students learn to identify messages and intentions behind data representations and evaluate their validity before drawing conclusions (Philip et al., 2016). These critical skills empower them to make informed decisions and participate in public discussions about issues affecting their lives (Gordon et al., 2016). In a world filled with streams of data, not all of which are neutral or truthful, data literacy can help students discern accurate information and evaluate sources critically (Forzani, 2018; Turton & Martin, 2020;). Beyond personal benefits, it fosters their ability to engage in civic activities and public discourse effectively (Ercegovac, 2015).

Despite its importance, data literacy remains underrepresented in K-12 education. Few instructional programs or initiatives integrate data literacy into curricula due to challenges such as limited classroom technology, insufficient teacher training, and rigid subject boundaries (Gunter, 2007; van't Hooft et al., 2012). In many schools, opportunities to explore numerical and statistical data are scarce and often confined to science or mathematics classes (Deahl, 2014). Moreover, the emphasis on meeting core requirements and preparing for high stakes testing leaves little room for incorporating data literacy practices into lesson plans (Ridsdale et al., 2015).

Another barrier to teaching data literacy is the lack of a clear, consistent definition for K-12 students. While researchers largely agree on its core elements, subtle differences in focus exist. Some emphasize posing questions about data, using tools for representation, extracting relevant information, and evaluating inferences (English & Watson, 2018; van't Hooft et al., 2012). Others define it as accessing, analyzing, generating, and assessing data-based representations and inferences (Love, 2004; Vahey et al., 2012). Organizations like the Oceans of Data Institute highlight skills such as collecting, synthesizing, and visualizing data, while others underscore ethical considerations, such as understanding privacy issues and the intent behind data collection (Gould, 2017). Critical examination of data sources, their limitations, and their reliability is another recurring theme (Lipton & Wellman, 2012; Yates et al., 2021). Collectively, these competencies enable individuals to make data-informed decisions in everyday life (Sorapure, 2019).

Although these competencies are foundational, a comprehensive framework defining data literacy and its proficiency levels remains elusive. Scholars argue that the field must clarify critical competencies to support effective instruction for K-12 students (Bowler & Shaw, 2024). The challenge is compounded by overlaps with other literacies, such as statistical or informational literacy, and the multidisciplinary perspectives brought by researchers, educators, and data professionals (Shields, 2005; Twidale et al., 2013).

Middle-grade students, in particular, would benefit from targeted data literacy education. While attention to data literacy has grown in higher education as a core learning outcome (Cubarrubia, 2019), studies focusing on middle school learners remain limited. Early exposure during this developmental stage can lay a strong foundation for advanced data literacy skills in later years. Recent efforts have emphasized the importance of introducing these concepts during secondary education (Wolff et al., 2016). However, the lack of a clear understanding of essential competencies for adolescent learners presents a significant obstacle.

Before developing robust instructional models, it is crucial to identify the specific competencies middle-grade students need to attain data literacy. Educators require a clearer understanding of these competencies to design effective curricula and integrate data literacy practices into school settings.

To address this need, this paper investigates existing literature to identify and categorize data literacy competencies relevant to adolescent learners. By synthesizing empirical and theoretical studies and cross-referencing findings with nationwide middle-grade learning standards, this study aims to clarify the essential competencies for middle-grade students and provide a foundation for data literacy education in K-12 settings. My research questions for this paper are as follows:

1. What does it mean to be data literate for adolescent learners?

2. What data literacy competencies would students in middle-grade years benefit from?

Methods

I conducted a literature review to explore theoretical and empirical discussions on the concepts, instructional practices, and assessments of data literacy. This review focuses on identifying practices relevant to data literacy for school-aged learners. I searched education, library science, anthropology, and behavioral science databases to reflect the interdisciplinary nature of the field. Using key terms aligned with my research questions, I included related terms with similar meanings. The review encompasses peer-reviewed journal articles published in English, including studies from international contexts, such as Australia (Callingham et al., 2017; English & Watson, 2018;). A flowchart detailing the search and selection process is provided in Figure 1, adapted from Surrain and Luk's (2017) procedure.

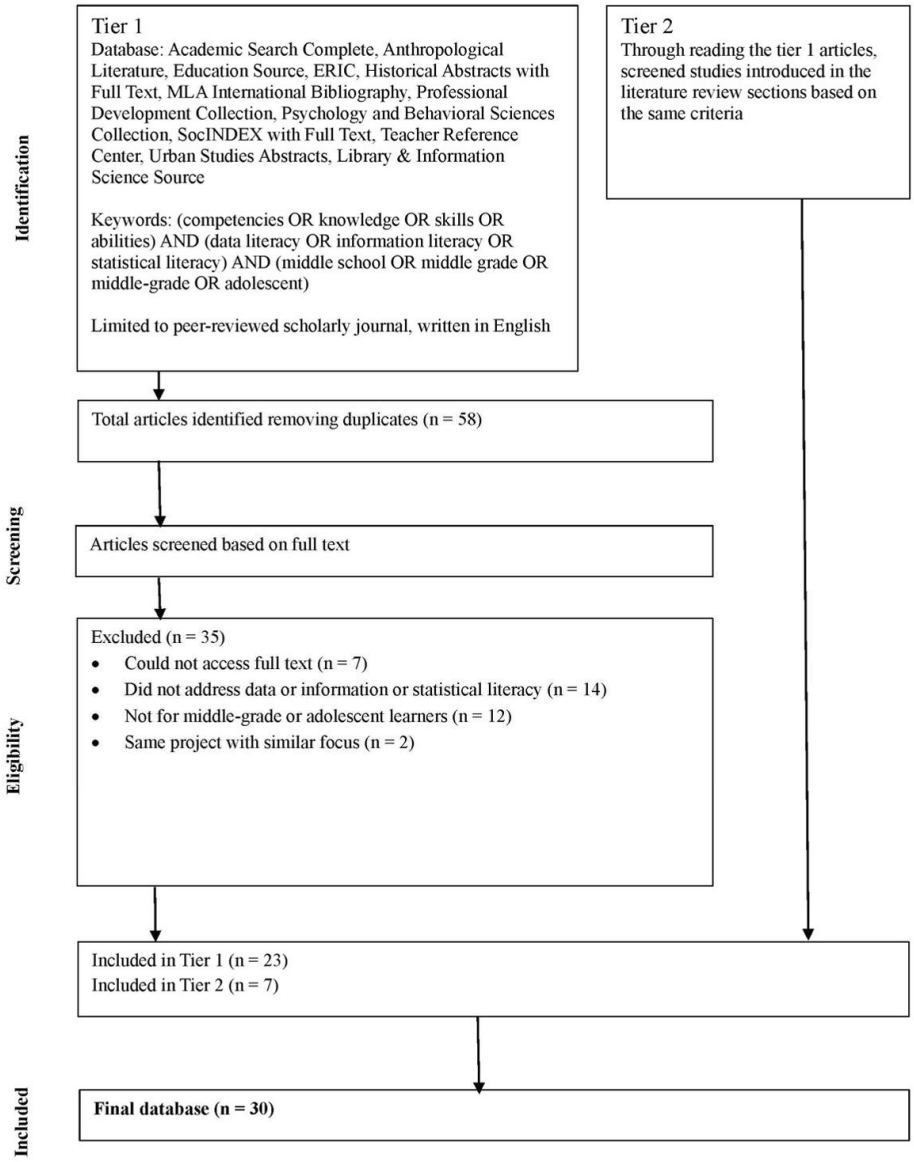
Inclusion criteria of the studies

The initial search (Tier 1) yielded 58 articles, which I screened based on inclusion criteria. I included studies addressing data literacy or its complementary fields, such as information and statistical literacy (Calzada Prado & Marzal, 2013; Shields, 2005;). I focused on articles targeting adolescent learners or K-12 students and excluded those with overlapping datasets unless they raised distinct research questions. This process narrowed the selection to 23 articles. A subsequent review of these articles' literature sections (Tier 2) identified seven additional studies meeting the same criteria.

Coding the studies

To address my first research question, I analyzed the terms and definitions of data literacy discussed in theoretical frameworks, teaching practices, and assessments. I aggregated verb-noun phrases describing knowledge, skills, and abilities (KSAs), categorized them by overarching themes, and consolidated a list of competencies. For example, discussions on learners formulating data-related questions were categorized as *Pose questions*. To guide future research, I also coded characteristics of empirical studies, such as participants' demographics, settings, and theoretical approaches.

Figure 1: Search process and selection criteria



For my second research question, I cross-referenced these competencies with middle school learning standards, including Common Core State Standards (CCSS), Next Generation Science Standards (NGSS), and Middle School Learning Standards (MSLS) for subjects like English Language Arts (ELA), Science, Social Studies, and Mathematics. Each competency was broken into specific skills and aligned with relevant subject standards. For example, the competency *Pose questions* was matched with standards on formulating research questions. I created a table linking competencies with specific standards codes and descriptions to ensure alignment and noted instances where a standard corresponded to multiple competencies.

Characterization of the studies

Table 1 summarizes the participants, contexts, theories, methods, and subjects in the 22 empirical studies analyzed. Participant numbers ranged from five to 7,000, with diverse racial, ethnic, and socioeconomic backgrounds. Most studies included both native and non-native English speakers, though some omitted demographic details. In the U.S., more studies were conducted in urban than rural settings. Studies were conducted between 2001 and 2022, with 10 based in the U.S. and 12 internationally. Settings included regular classrooms (n=11) and after-school programs (n=6).

Table 1: Analysis of characterization of the focal studies

Participants (N ranged 5 to 7,000)		*No. of studies*		*No. of studies*
	Language		*Social class*	
	English-as-L1	N=2	Middle	N=4
	English-as-Lx	N=10	Lower-middle	N=2
	Both	N=3	Lower	N=2
	Unidentified	N=7	Unidentified	N=14
	Race/Ethnicity included			
	European American	N=4		
	African American	N=4		
	Latino/a	N=4		
	Asian	N=8		
	Native American	N=1		
	Unidentified	N=7		
Contexts	*U.S.*	*N=10*	*Non-U.S.*	*N=12*
	Rural	N=1	Suburban	N=1
	Urban	N=4	Urban	N=1
	Unidentified	N=5	Unidentified	N=10
	State in the U.S.		*Name of country*	
	Mid-West	N=2	Turkey, Canada, France, Australia, Germany, Columbia, Ecuador, Indonesia, China, UK, Korea	
	North-East	N=2		
	South-West	N=2		
	West	N=3		
	Unidentified	N=1		
			In-school settings	N=11
			Regular classroom	N=6
			After-school	N=5
			Unidentified	
Major theories	Statistical literacy	N=10		
	Information literacy	N=9		
	Data literacy	N=12		

	New literacy studies/multiple literacies	N=2		
	Media literacy	N=2		
Methodology	*Qualitative*	*N=10*		
	Case study	N=3	Interview	N=1
	Testing	N=2	Participatory Action Research	N=1
	Content analysis	N=2		
			Interpretive microanalysis	N=1
	Quantitative	*N=10*	Phenomenography	N=1
	Survey	N=4		
	Testing	N=6		
	Mixed-Methods	*N=2*		
School subjects	Math	N=7	*Other subject involved*	
	Science/STEM	N=5	Digital media literacy	
	Social Studies*	N=4		
	English	N=2		
	Not specified	N=3		
	*Including history, political science			

Note. STEM: Science, Technology, Engineering, and Mathematics

The studies often referenced overlapping concepts of data, statistical, and information literacy, with some adopting broader frameworks like new literacy studies. Methodologies varied; ten studies used quantitative methods, ten studies used qualitative methods, and two studies employed mixed methods. Six quantitative and two qualitative studies used assessments as a method. Mathematics, science, and social studies were the primary subjects, although three studies did not link practices to specific disciplines.

Findings

The results of my literature review are organized according to the two research questions: 1) what are the data literacy competencies for adolescent learners; and 2) what are the benefits of developing data literacy competencies during adolescence.

What does it mean to be data literate for adolescent learners?

In response to the first research question, I synthesized KSAs that are emphasized in research on adolescents' data literacy (see Table 2) into eight competencies: *Pose questions, Access/Collect, Transform, Manage/Handle, Analyze, Interpret, Evaluate, Answer questions,* and *Present/Communicate*. Appendix A illustrates the relevant articles that mentioned each competency and corresponding learning standards.

Table 2: Data literacy by competency

Pose questions

- Investigate authentic problems
- Formulate/articulate data-based questions
- Identify the audience and context for the problem
- Define your goals and how they can be achieved with data

Access/Collect

- Know how to select research and statistical methods and tools that align to purposes
- Demonstrate an understanding of validation
- Document methods and tools
- Understand a wide variety of tools for accessing data
- Investigate the source of information
- Explore the data that are available

Transform

- Clean, transform, manipulate & synthesize data
- Understand a wide variety of tools for converting and manipulating data
- Synthesize information from multiple sources/map data across heterogeneous sources
- Understand how representations in computers can vary and why data must sometimes be altered before analysis

Manage/Handle

- Understand issues of data privacy, confidentiality, ownership, and handling process
- Understand the importance of the provenance of data and how data are stored
- Understand ethical use of personal data
- Understand ways of protecting and documenting data to maintain reproducibility
- Understand issues of data quality and credibility markers
- Assess risk and bias involved in conducting the study

Analyze

- Develop an analysis plan and conduct exploratory analyses
- Know how to use analytic software
- Understand some aspects of predictive modeling

Interpret

- Understand the data representations
- Interpret information from data
- Develop data-based inferences and explanations
- Explore patterns in data with a skeptical but open mind

- Produce explanations, comparisons and predictions based on the variability in the data
- Compare the results with other findings

Evaluate

- Understand statistical arguments
- Evaluate data-based claims, inferences and explanations
- Use appropriate data, tools, and representations to support critical thinking
- Use data as part of evidence-based analytic thinking

Answer questions

- Answer data-based questions
- Make statistically sound decisions
- Discuss and document findings
- Form a judgment and conclusions and draw arguments from data
- Discuss limitations of the study

Present/Communicate

- Create and construct basic descriptive representations and visualizations of data to answer questions about real-life processes
- Translate and present information into different data representations
- Communicate solutions and recommendations
- Understand how to do citation

Pose Questions. This competency entails identifying problems to solve based on data, with an emphasis on audience and context. It requires defining clear goals and determining how data can be used to achieve those objectives. This competency aligns with Wolff et al.'s (2016) PPDAC cycle—*Problem, Plan, Data, Analysis,* and *Conclusion*—where *posing questions* serves as a critical initial step in framing data-driven inquiries in K-12 contexts, as demonstrated in various studies (Ercegovac, 2015; van't Hooft et al., 2012;). Kim et al. (2016) also highlighted the ability to "formulate questions to search for information" (p. 445) when they asked students to self-assess their information literacy.

Access/Collect. Gathering information and data requires selecting appropriate research methods, validating sources, and documenting tools used for data access (Cunningham et al., 2018). This includes evaluating sources, understanding data relevance, and exploring available data. Ridsdale et al. (2015) highlight "collection" as a core competency for data literacy, emphasizing the ability to search for, assess, and contextualize data. Aillerie et al. (2016) explore how teenagers practice this skill on social networking sites. Programs like City Digits, where public-school students analyze state lottery data to build evidence-based arguments, demonstrate how such competencies can be applied to real-world contexts (Deahl, 2014).

Transform. This competency includes cleaning and synthesizing information from diverse sources, critical for handling raw data. It requires a deep understanding of multiple data conversion and manipulation tools, the ability to synthesize information from multiple sources, and the capability to

map data across different platforms. Additionally, one needs a grasp of varied computer representations and rationales for altering data before analysis. Data literacy projects, like those at the University of Maine, encourage students to visualize and interpret real data sets, reinforcing the importance of transformation skills (Wolff et al., 2016). Cohen et al. (2017) also emphasized that transforming data requires an understanding of technology systems, an essential part of preparing students for STEM fields.

Manage/Handle. This competency involves understanding data privacy, quality, and credibility markers. Examples of credibility markers include data relevance to users' interests and authorship. Understanding authorship requires knowledge about authorial intention, which involves selecting data or visualizations based on intended meanings, contexts, and sociocultural codes (Hullman & Diakopoulos, 2011). Understanding credibility markers, authorship, and data relevance, as discussed in the literature, equips students to handle data ethically and critically evaluate sources and interpretations (Utomo, 2021).

Analyze. This competency focuses on hands-on data analysis, creating an analysis plan, and using tools proficiently. Additionally, it includes aspects of predictive modeling, allowing students to extract meaningful insights from data. This competency is reinforced by curricula such as *Thinking with Data*, which engages students in problem-based exercises to compare learning gains in data literacy between those exposed to the curriculum and those who are not (van't Hooft et al., 2012). Zalles (2005) also highlighted students' application of analytical skills in real-world data in the EPA Phoenix performance task about air quality (Zalles, 2005).

Interpret. This competency is for understanding processed data and visualizations, exploring and extracting information from mapped data, graphs, pie charts, and emerging forms of visualizations (Gordon et al., 2016). It requires skills to extract insights from visual representations. Body of research emphasizing the necessity of interpreting variability and terminology used in data, a skill practiced in both K-12 and postsecondary contexts (Ben-Zvi & Arcavi, 2001; Mandinach & Gummer, 2012; Wolff et al., 2016; Yolcu, 2014;;). For instance, Ben-Zvi and Arcavi (2001) emphasized the importance of understanding variable values in data representations such as tables and graphs.

Evaluate. This competency encompasses assessing the trustworthiness of data claims, which is essential for informed judgments and data-driven insights (Gunter, 2007). It includes detecting loaded language or omitted facts based on the author's purpose or context. The competency allows critical investigation into arguments derived from data and identifying missing information to strengthen those arguments (Büscher, 2022). It supports critical inquiry into the evidence behind claims, connecting data to the meanings conveyed, and questioning the legitimacy of those connections (Cunnimgham et al., 2018). Seroff (2017) described exercises where students evaluate limitations in data collection methods about traffic fatalities, reinforcing the importance of questioning data validity and generalizability.

Answer Questions. This competency focuses on addressing problems or questions identified at the beginning of data-related tasks (Utomo, 2021). Drawing conclusions based on data requires, understanding the limitations of data from collection methods or unwarranted assumptions. It necessitates forming judgments and making arguments grounded in data, while also discussing limitations inherent in the study. Project-based learning further facilitates this competency by addressing authentic societal problems through the examination of data and the use of digital tools

(Erwin Jr., 2015; Wu et al., 2020). van't Hooft et al. (2012)'s Thinking with Data project assessments also include performance tasks that gauge students' ability to provide well-supported answers and suggest what additional data would be needed to answer questions.

Present/Communicate. This competency involves effectively communicating data-based understandings and findings to others (Wu et al., 2020). It includes translating information into various data representations and following proper citation practices. Creating and employing different graph types is essential for middle school students (Hunter-Thomson, 2019). For instance, City Digits and the GLOBE Integrated Investigation Assessments emphasize skills in representing data through various formats, such as graphs and maps, to communicate findings effectively (Wolff et al., 2016; Zalles, 2005).

What are the benefits of developing data literacy competencies during adolescence?

Developing data literacy competencies during adolescence, particularly *Interpret* and *Evaluate*, can help students gain essential skills for academic success, make informed decisions, and engage critically with the increasingly data-driven world. These competencies enable students to analyze, interpret, and evaluate data representations across various contexts, fostering critical thinking and problem-solving abilities. Strengthening data literacy at this stage prepares students for future educational and professional settings where data-driven reasoning is crucial.

Interpret is one of the most emphasized competencies in existing literature, with the largest number of corresponding learning standards for middle-grade students. This competency is a fundamental competency that helps students comprehend and analyze different forms of data representation, including graphs, tables, and multimedia formats. Thirteen articles, including those by Callingham et al. (2017) and Büscher (2022), identified *Interpret* as a core competency of data literacy. Yolcu (2014), in defining statistical literacy within a three-tiered framework, highlighted the ability to interpret statistical information and messages as the second tier. Projects like City Digits demonstrate the practical application of these skills by engaging students in interpreting data to analyze social issues, such as studying lottery data, thereby strengthening their interpretive abilities in meaningful contexts (Deahl, 2014).

Findings also revealed that many middle-grade learning standards across various disciplines emphasized *Interpret* as a key competency. For instance, the CCSS highlight the ability to determine central ideas, information, or conclusions in literacy for history, social studies, science, and technical texts. Similarly, NGSS underscore students' ability to interpret graphical displays of data to identify linear and nonlinear relationships. In mathematics, the MSLS focus on the ability to discuss and understand the correspondence between data sets and their graphical representations.

Evaluate is also among the competencies most frequently addressed in the studies. This competency allows students to critically analyze claims based on data and assess the validity of conclusions drawn from data-driven arguments. *Evaluate* is prominently featured in research, with 21 studies identifying it as essential for data literacy. For instance, Womack (2015) emphasizes the skills required to critically evaluate information and integrate it into students' knowledge and value systems. Chin et al. (2016) measured evaluation in their Choicelets assessment by testing students' ability to critique sources of visually presented information. Similarly, van't Hooft et al. (2012) used problem-based exercises requiring students to evaluate air quality data and make evidence-based recommendations. The importance of *Evaluate* is further underscored in today's data-driven society, where critical skills are

needed to assess the credibility of messages in media and web-based information (Bussert-Webb et al., 2017; Cuervo Sánchez et al., 2021; Leu et al., 2013).

Numerous learning standards emphasize the importance of *Evaluate* across disciplines, highlighting its multidisciplinary relevance. The NGSS stress the KSAs needed to assess the merit and validity of ideas and methods, while the CCSS call for the ability to evaluate content presented in diverse formats and media. Additionally, the CCSS underscore skills in distinguishing between facts and reasoned judgments derived from research findings, particularly in the contexts of history, social studies, and science and technical texts.

These two competencies are interrelated, with many researchers classifying them under one category. For example, *Interpret* and *Evaluate* belong to the "Conclusion" step in the PPDAC cycle defined by Wolff et al. (2016). This step encompasses interpreting data to understand patterns and evaluating the validity of explanations based on that data. Additionally, both competencies align with the "Data Evaluation" category identified by Ridsdale et al. (2015), which encompasses the KSAs needed to assess graphical representations of data. Leu et al. (2013) also emphasize that *Evaluate* should be assessed alongside a related competency, further supporting the focus on these closely linked but distinct competencies.

Discussion and Implications

This paper contributes significantly to the conceptualization of data literacy by synthesizing different arrays of definitions and competencies. Data literacy encompasses a range of KSAs, making it challenging to define and tailor for students across grade levels. To address this, I conducted a literature review of data literacy and related constructs, such as information and statistical literacy, to aggregate and categorize competencies. Among these, two key competencies—interpreting data representations and evaluating claims based on data—emerged as most emphasized in the learning standards for middle-grade students. By cross-referencing competencies with standards like CCSS and NGSS, I highlighted their relevance to core subjects such as math, science, English, and social studies.

The findings have significant implications in developing a practical framework of data literacy education applicable to middle school settings. They can guide educators in initiating data literacy instruction by providing clear target competencies aligned with learning standards. This shared understanding can help implement data literacy practices across subjects, offering insights for designing interdisciplinary curricula. Curricula focusing on the development of *Interpret* and *Evaluate* skills can be incorporated into syllabi aimed at fostering other competencies. They can also serve as a foundation for helping students develop data literacy skills throughout high school and higher education.

The study highlights the need for future research to address gaps in current data literacy education. Few studies have been conducted in rural settings or included minoritized populations, such as Native American students. Additionally, most studies employing instructional interventions lacked assessment tools to measure data literacy, limiting their ability to evaluate the impact on students' competencies. Only two studies in this review utilized mixed-method approaches, which are essential for capturing both quantitative outcomes and qualitative insights into learners' experiences. This finding underscores the need for more research focusing on underrepresented populations and incorporating validated assessments and diverse methodologies. The paper emphasizes the

importance of integrating instruction and assessment practices to enhance data literacy among diverse middle-grade students.

References

Aillerie, K., & McNicol, S. (2016). Information literacy and social networking sites: Challenges and stakes regarding teenagers' uses. *ESSACHESS–Journal for Communication Studies*, *9*(2-18), 89–100.

Ben-Zvi, D., & Arcavi, A. (2001). Junior high school students' construction of global views of data and data representations. *Educational Studies in Mathematics*, *45*, 35–65. https://doi.org/10.1023/A:1013809201228

Bowler, L., & Shaw, C. (2024). Trends in data literacy, 2018-2023: A review of the literature. *Information Research an international electronic journal*, *29*(2), 198-205. https://doi.org/10.47989/ir292822

Büscher, C. (2022). Design principles for developing statistical literacy in middle schools. *Statistics Education Research Journal*, *21*(1), 1–16. https://doi.org/10.52041/serj.v21i1.80

Bussert-Webb, K., & Henry, L. (2017). Promising digital practices for nondominant learners. *International Journal of Educational Technology, 4*(2), 43–55.

Callingham, R & Watson, J.M. (2017) The development of statistical literacy at school. *Statistics Education Research Journal, 16*(1), 181–201.

Calzada Prado, J. C., & Marzal, M. Á. (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

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

Cohen, J. D., Renken, M., & Calandra, B. (2017). Urban middle school students, twenty-first century skills, and STEM-ICT careers: Selected findings from a front-end analysis. *TechTrends*, *61*, 380-385. https://doi.org/10.1007/s11528-017-0170-8

Cubarrubia, A. P. (2019, October 18). We all need to be data people. *The Chronicle of Higher Education*, *66*(7). https://www.chronicle.com/article/we-all-need-to-be-data-people/

Cuervo Sánchez, S. L., Foronda Rojo, A., Rodriguez Martinez, A., & Medrano Samaniego, C. (2021)Media and information literacy: A measurement instrument for adolescents. *Educational Review*, *73*(4), 487-502. https://doi.org/10.1080/00131911.2019.1646708

Cunningham, V., & Williams, D. (2018). The seven voices of information literacy (IL). *Journal of information literacy*, *12*(2). http://dx.doi.org/10.11645/12.2.2332

Deahl, E. S. (2014). *Better the data you know: Developing youth data literacy in schools and informal learning environments.* [Master's Thesis, the Massachusetts Institute of Technology]. https://doi.org/10.2139/ssrn.2445621

English, L. D., & Watson, J. (2018). Modelling with authentic data in sixth grade. *ZDM Mathematics Education*, *50*(1–2), 103–115. https://doi.org/10.1007/s11858-017-0896-y

Ercegovac, Z. (2015). Data-driven society begins with data-savvy youth. *Bulletin of the Association for Information Science & Technology*, *42*(1), 42–48. https://doi.org/10.1002/bul2.2015.1720420111

Erwin, R. W., Jr. (2015). Data literacy: Real-world learning through problem-solving with data sets. *American Secondary Education*, *43*(2), 18–26. https://www.jstor.org/stable/43694208

Fontichiaro, K., & Oehrli, J. A. (2016). Why data literacy matters. *Knowledge Quest*, *44*(5), 21–27. https://eric.ed.gov/?id=EJ1099487

Forzani, E. (2018). How well can students evaluate online science information? Contributions of prior knowledge, gender, socioeconomic status, and offline reading ability. *Reading Research Quarterly*, *53*(4), 385–390. https://doi.org/10.1002/rrq.218

Gebre, E. (2022). Conceptions and perspectives of data literacy in secondary education. *British Journal of Educational Technology, 53*(5), 1080–1095. https://doi.org/10.1111/bjet.13246

Godaert, E., Aesaert, K., Voogt, J., & van Braak, J. (2022). Assessment of students' digital competences in primary school: A systematic review. *Education and Information Technologies, 27,* 9953–10011. https://doi.org/10.1007/s10639-022-11020-9

Gordon, E., Elwood, S., & Mitchell, K. (2016). Critical spatial learning: Participatory mapping, spatial histories, and youth civic engagement. *Children's Geographies*, *14*, 1–15. https://doi.org/10.1080/14733285.2015.1136736

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), http://dx.doi.org/10.1080/2331186X.2016.1268773

Gunter, G. A. (2007). Building student data literacy: An essential critical-thinking skill for the 21st Century. *MultiMedia & Internet @ Schools, 14*(3), 24–28.

Hullman, J., & Diakopoulos, N. (2011). Visualization rhetoric: Framing effects in narrative visualization. *IEEE Transactions on Visualization and Computer Graphics*, *17*(12), 2231–2240. https://doi.org/10.1109/TVCG.2011.255

Hunter-Thomson, K. (2019). Data literacy 101: How do we set up graphs in science? *Science Scope*, *42*(2), 78–82.

Kim, E. M., & Yang, S. (2016). Internet literacy and digital natives' civic engagement: Internet skill literacy or Internet information literacy? *Journal of Youth Studies*, *19*(4), 438–456. https://doi.org/10.1080/13676261.2015.1083961

Kostelnick, C. (2004). Melting-pot ideology, modernist aesthetics, and the emergence of graphical conventions: The statistical atlases of the United States, 1874–1925. In C. A. Hill & M. Helmers (Eds.), *Defining visual rhetorics* (pp. 215–242). Routledge.

Leu, D. J., Forzani, E., Burlingame, C., Kulikowich, J., Sedransk, N., Coiro, J., & Kennedy, C. (2013). The New Literacies of online research and comprehension: Assessing and preparing students for the 21st century with Common Core State Standards. In S. Neuman (Ed.), *Quality Reading Instruction in the Age of Common Core Standards* (pp. 219–236). International Reading Association.

Lipton, L., & Wellman, B. (2012). *Got data? Now what? Creating and leading cultures of inquiry.* Solution Tree Press.

Love, N. (2004). Taking data to new depths. *Journal of Staff Development*, *25*(4), 22–26. https://all4ed.org/wp-content/uploads/2013/09/lovejsdarticle1.pdf

Mandinach, E. B., & Gummer, E. S. (2012). Navigating the landscape of data literacy: It IS complex. *WestEd.* https://eric.ed.gov/?id=ED582807

Manduca, C. A., & Mogk, D. W. (2002). Using data in undergraduate science classrooms: Final report on an interdisciplinary workshop held at Carleton College. https://d32ogoqmya1dw8.cloudfront.net/files/usingdata/UsingData.pdf

Oceans of Data Institute (ODI) (2015). *Building global interest in data literacy: A dialogue*. Educational Development Center. https://oceansofdata.org/our-work/building-global-interest-data-literacy-dialogue-workshop-report

Philip, T. M., Olivares-Pasillas, M. C., & Rocha, J. (2016). Becoming racially literate about data and data-literate about race: Data visualizations in the classroom as a site of racial-ideological micro-contestations. *Cognition and Instruction*, *34*(4), 361–388. https://doi.org/10.1080/07370008.2016.1210418

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. https://dalspaceb.library.dal.ca/server/api/core/bitstreams/7c0db868-79ac-4c3c-9ad0-f2a93cf2ccc6/content

Shields, M. (2005). Information literacy, statistical literacy, data literacy. *IASSIST Quarterly, 28*(2-3), 6–11. https://doi.org/10.29173/iq790

Seroff, J. (2017). Using data in the research process. In K. Fontichiaro, J. A. Oehrli, & A. Lennex (Eds.), *Creating data literate students*. Michigan Publishing.

Sorapure, M. (2019). Text, image, data, interaction: Understanding information visualization. *Computers and Composition*, *54*, 1–16. https://doi.org/10.1016/j.compcom.2019.102519

Statewide Longitudinal Data Systems Grant Program. (2015). *SLDS data use standards: Knowledge, skills, and professional behaviors for effective data use, version 2*. U.S. Department of Education National Center for Education Statistics. https://eric.ed.gov/?id=ED595040

Surrain, S., & Luk, G. (2019). Describing bilinguals: A systematic review of labels and descriptions used in the literature between 2005–2015. *Bilingualism: Language and Cognition*, *22*(2), 401-415. https://doi.org/10.1017/S1366728917000682

Turton, W., & Martin, A. (2020, January 5). How deepfakes make disinformation more real than ever. *Bloomberg*. https://www.bloomberg.com/news/articles/2024-02-09/fighting-deepfakes-whats-being-done-biden-robocalls-to-taylor-swift-ai-images

Twidale, M. B., Blake, C., & Gant, J. (2013). Towards a data literate citizenry. In *Proceedings of iConference 2013* (pp. 247–257). https://hdl.handle.net/2142/38385

Vahey, P., Rafanan, K., Patton, C., Swan, K., Hooft, M., Kratcoski, A., & Stanford, T. (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

Vahey, P., Yarnall, L., Patton, C., Zalles, D., & Swan, K. (2006, April). Mathematizing middle school: Results from a cross-disciplinary study of data literacy. Presented at the *American Educational Research Association Annual Conference,* San Francisco, CA.

van't Hooft, M., Swan, K., Cook, D., Stanford, T., Vahey, P., Kratcoski, A., Rafanan, K., & Yarnall, L. (2012). A cross-curricular approach to the development of data literacy in the middle grades: The Thinking with Data Project. *Middle Grades Research Journal*, *7*(3), 19–33.

Viegas, F. B., & Wattenberg, M. (2006). Communication-minded visualization: A call to action. *IBM Systems Journal, 45*(4), 801–812. https://doi.org/10.1147/sj.454.0801

WestEd. (n.d.). Developing Assessments of Data Literacy. *WestEd*. https://www.wested.org/project/data-literacy-assessment-development/

Wolff, A., Gooch, D., Cavero Montaner, J. J, Rashid, U., Kortuem, G., (2016). Creating an understanding of data literacy for a data-driven society. *The Journal of Community Informatics, 12*(3), 9–26. https://doi.org/10.15353/joci.v12i3.3275

Womack, R. (2015). Data visualization and information literacy. *IASSIST Quarterly, 38*(1), 12–17. https://doi.org/10.29173/iq619

Wu, D., Yu, L., Yang, H. H., Zhu, S., & Tsai, C. C. (2020). Parents' profiles concerning ICT proficiency and their relation to adolescents' information literacy: A latent profile analysis approach. *British Journal of Educational Technology*, *51*(6), 2268-2285. https://doi.org/10.1111/bjet.12899

Yates, S., Carmi, E., Lockley, E., Wessels, B. & Pawluczuk, A. (2021). Me and My Big Data: Understanding citizens data literacies. *Nuffield Foundation*. https://openaccess.city.ac.uk/id/eprint/26756/

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

Zalles, D. (2005). Designs for assessing foundational data literacy. https://serc.carleton.edu/files/NAGTWorkshops/assess/ZallesEssay3.pdf

Zoellick, B., Schauffler, M., Flubacher, M., Weatherbee, R., & Webber, H. (2016, April). Data literacy: Assessing student understanding of variability in data. Paper presented at *the Annual Meeting of the National Association for Research in Science Teaching.* Baltimore, MD.

Endnotes

[1] Semi Yeom, University of Maryland, College Park, USA, can be reached at syeom@terpmail.umd.edu

Appendix A: Competencies of Data Literacy as a Function of Relevant Learning Standards and Literature

Competency	Relevant Standards	Specific Examples	Relevant Literature
Pose questions	• CCSS Math G6-12 • MSLS Mathematics/Social Studies/Science/ELA Standards • NGSS G6-8	• Make sense of problems and persevere in solving them. • Formulate/address questions/pose and define problems and design/execute studies. • Develop the ability to refine & refocus broad & ill-defined questions. • Use conjectures to formulate new questions & studies to answer them.	Vahey et al. (2012); van't Hooft et al. (2012); Ben-Zvi & Arcavi (2001); Erwin Jr. (2015); Chin et al. (2016); Wolff et al. (2016); Guler et al. (2016); Kim et al. (2016); Utomo (2021)
Access/collect	• CCSS Writing G6-12 • MSLS Mathematics/Social Studies/ELA • NGSS G6-8 • CCSS Math G7	• Conduct short as well as more sustained research projects based on focused questions, demonstrating understanding of the subject under investigation. • Gather relevant information from multiple print and digital sources. • Collect/gather data about a characteristic shared by two populations or different characteristics within one population. • Use random sampling to draw inferences about a population.	Erwin Jr. (2015); van't Hooft et al. (2012); Shields (2005); Deahl (2014); Calzada Prado et al. (2013); Ridsdale et al. (2015); Ercegovac (2012); Wolff et al. (2016); Aillerie et al (2016); Cuervo Sánchez et al. (2021); Cunningham et al. (2018); Kim et al. (2016); Bussert-Webb et al. (2017)
Transform	• MSLS ELA • CCSS Reading/Speaking & Listening/Writing G6-12 • CCSS Literacy in science and	• Synthesize data from a variety of sources. • Integrate content presented in diverse formats and media, including visually and quantitatively, as well as in words. • Integrate quantitative or technical information expressed in words in a text with a version of that information expressed visually.	Calzada Prado et al. (2013); Vahey et al. (2012); van't Hooft et al. (2012); English & Watson (2018); Gould (2017); Erwin Jr. (2015); Chin et al. (2016); Shields (2005); Gordon et al. (2016); Wolff et al. (2016); Cuervo

	technical texts G6-8		Sánchez et al. (2021); Cohen et al. (2017); Wu et al. (2020)
Manage/handle	• CCSS Writing G6-12	• Know how to avoid plagiarism. • Assess the credibility and accuracy of each source.	Calzada Prado et al. (2013); Gould (2017); Erwin Jr. (2015); Deahl (2014); Gordon et al. (2016); Ercegovac (2012); Utomo (2021)
Analyze	• NGSS G6-8 • CCSS Reading/Literacy G6-12 • CCSS Math G6-12 • NGSS G6-8 • CCSS Literacy in history/social studies G6-8 • MSLS Math	• Investigate chance processes and develop, use, and evaluate probability models. • Develop, use, and revise models to describe, test, and predict more abstract phenomena and design systems. • Analyze human behavior in relation to its physical and cultural environments. • Extend quantitative analysis to investigations, distinguishing between correlation and causation, and basic statistical techniques of data and error analysis. • Investigate patterns of association in bivariate data. • Identify patterns in large data sets. • Analyze the relationship between a primary and secondary source on the same topic.* • Specify relationships between variables and clarify arguments and models.* • Use observations about differences between two or more samples to make conjectures about the populations* • Develop understanding of statistical variability*	Calzada Prado et al. (2013); Gould (2017); Ben-Zvi & Arcavi (2001); English & Watson (2018); Erwin Jr. (2015); Chin et al. (2016); Shields (2005); Zalles (2005); Zoellick et al. (2016); Ridsdale et al. (2015); Guler et al. (2016); Yolcu (2014); Buscher (2022); Cuervo Sánchez et al. (2021); Bussert-Webb et al. (2017)

Interpret	• CCSS Literacy in history/social studies/science and technical texts G6-8 • MSLS Math/Social Studies/Science • CCSS Math G6-12 • NGSS G6-8 • CCSS Writing G6-12	• Determine the central ideas or information or conclusions. • Determine the meaning of keywords, symbols, domains-specific words and phrases in a specific context. • Discuss & understand the correspondence between data sets & their graphical representations. • Construct sound historical interpretations. • Compare and contrast the information gained from experiments, simulations, video, or multimedia sources with that gained from reading a text on the same topic. • Interpret graphical displays of data and/or large data sets to identify linear and nonlinear relationships. • Analyze the relationship between a primary and secondary source on the same topic.* • Use observations about differences between two or more samples to make conjectures about the populations.* • Use evidence to generate explanations.* • Use appropriate tools strategically such as diagrams, two-way tables, graphs, flowcharts and formulas to identify important quantities in a practical situation and map their relationships.*	Deahl (2014); Vahey et al. (2012); van't Hooft et al. (2012); Calzada Prado et al. (2013); Twidale et al. (2013); Zalles (2005); Ridsdale et al. (2015); Wolff et al. (2016); Guler et al. (2016); Callingham et al. (2017); Yolcu (2014); Buscher (2022); Cuervo Sánchez et al. (2021)
Evaluate	• CCSS Reading/ Speaking & listening G6-12 • NGSS G6-8 • CCSS Literacy in	• Evaluate content presented in diverse formats and media, including visually and quantitatively, as well as in words and orally. • Evaluate the merit and validity of ideas and methods.	Calzada et al. (2013); Gould (2017); English & Watson (2018); Erwin Jr. (2015); Shields (2005); Deahl (2014); Vahey et al. (2012); van't Hooft et al. (2012); Calzada Prado et al. (2013);

	history/social studies/science and technical texts G6-8 ● MSLS Social Studies/ Science	● Analyze the author's purpose in providing an explanation, describing a procedure, or discussing an experiment in a text. ● Distinguish among fact, opinion, and reasoned judgment based on research findings in a text. ● Question and identify gaps in data. ● Propose alternative explanations and critique explanations and procedures. ● Delineate the argument and specific claims in a text, including the validity of the reasoning as well as the relevance and sufficiency of the evidence.	Twidale et al. (2013); Womack (2015); Zalles (2005); Ercegovac (2012); Wolff et al. (2016); Aillerie et al (2016); Callingham et al. (2017); Yolcu (2014); Cuervo Sánchez et al. (2021); Cohen et al. (2017); Kim et al. (2016); Bussert-Webb et al. (2017)
Answer questions	● MSLS Science ● NGSS G6-8 ● CCSS writing G6-12	● Use evidence to generate explanations.* ● Use mathematical concepts to support explanations and arguments.* ● Draw evidence from literary or informational texts to support analysis, reflection, and research.* ● Construct explanations and design solutions supported by multiple sources of evidence consistent with scientific ideas, principles, and theories. ● Construct a convincing argument that supports or refutes claims for either explanations or solutions about the natural and designed world(s).	Vahey et al. (2012); van't Hooft et al. (2012); Calzada Prado et al. (2013); Gunter (2007); Twidale et al. (2013); Womack (2015); Zalles (2005); Guler et al. (2016); Callingham et al. (2017); Cohen et al. (2017); Wu et al. (2020); Utomo (2021)
Present/communicate	● MSLS Math/ELA ● CCSS Speaking & listening G6-12 ● NGSS G6-8 ● CCSS Literacy in history/social	● Prepare for and participate effectively in a range of conversations and collaborations with diverse partners, building on others' ideas and expressing their own clearly and persuasively. ● Use mathematical, computational, and/or algorithmic representations of	Vahey et al. (2012); Gould (2017); English & Watson (2018); Erwin Jr. (2015); van't Hooft et al. (2012); Chin et al. (2016); Shields (2005); Deahl (2014); Gunter (2007); Twidale et al. (2013); Womack (2015); Zalles

	studies/science and technical texts G6-10 ● CCSS Math G6-12	phenomena/provide evidence or design solutions to describe and/or support claims and/or explanations. ● Present information, findings, and supporting evidence such that listeners can follow the line of reasoning, and the organization, development, and style are appropriate to task, purpose, and audience. ● Create appropriate graphical representations of data. ● Use spoken, written, & visual language and communicate their discoveries in ways that suit their purpose & audience. ● Translate quantitative or technical information expressed in words into a visual form. ● Provide an accurate summary of the source distinct from prior knowledge or opinions. ● Cite specific textual evidence to support analysis.	(2005); Ercegovac (2012); Guler et al. (2016); Callingham et al. (2017); Cohen et al. (2017); Utomo (2021)

Note. * = Standards relevant to more than one competency. A few competencies have overlapping standards with each other because, in the standards, different competencies are clustered in one sentence.