Frontiers in Educational Practice and Research
Vol. 2 (2026), No. 3, pp. 154-159
DOI: https://doi.org/10.55578/fepr.2607.011

PERSPECTIVE

Beyond Pattern Recognition: Rethinking Graph Comprehension Through Data Provenance

Christiane Lieb [1,2], Wolfgang Mueller [2], Carola Pickhardt [3]*

[1] *Institute for Sustainable Teaching and Learning, Albstadt-Sigmaringen University, Sigmaringen, Germany*

[2] *Department of Media Education, Weingarten University of Education, Weingarten, Germany*

[3] *Department of Life Sciences, Albstadt-Sigmaringen University, Sigmaringen, Germany*

*Corresponding author. Email: lieb@hs-albsig.de

Received: May 30 2026, Revised: June 17 2026, Accepted: June 22 2026, Published: 3 July 2026

Keywords: Visual Literacy, Data Literacy, Data Provenance, Graph Comprehension, Artificial Intelligence

1. INTRODUCTION: SEEING PATTERNS IS NOT ENOUGH

The ability to interpret and critically reflect on data-based visualizations is regarded as a key competence in data literacy. As AI-generated visualizations equire critical evaluation, visual literacy is becoming increasingly important. Studies indicate that learners often struggle to interpret visualized data and critcally reflect on their origin and informational value [1-6]. Visualizations are frequently perceived as objective representations of reality.

However, prevailing understandings of visual literacy often focus on the ability to identify patterns, describe trends, and draw conclusions. This understanding seems increasingly inadequate in AI-mediated data environments. Little attention has been paid to how data are produced and how this influences visualization interpretation. As a result, a fundamental prerequisite of critical graph comprehension often remains implicit. This prerequisite is the ability to reflect on the origin, creation, and processing of the underlying data.

The central argument of this paper is that, although pattern recognition is a necessary condition for critical graph comprehension, it is not a sufficient one. Against this background, we argue that graph comprehension should not be reduced to the identification of patterns. Instead, it should incorporate reflection on data provenance as a core dimension of visual literacy. Building on this perspective, the paper discusses the theoretical and educational implications for the conceptualization of visual literacy in data-driven and AI-mediated information environments.

2. WHAT MAKES GRAPH COMPREHENSION? EXISTING PERSPECTIVES AND LIMITATIONS

Understanding data visualizations is not a trivial process but relies on complex cognitive mechanisms. Foundational studies on graph comprehension describe the interpretation of graphs as an interplay between perceptual and conceptual processes [7,8]. The former involves the perception of visual features such as trends, patterns, or relationships. The latter encompass the integration of meaning into existing knowledge structures and therefore extend beyond mere visual perception.

Many approaches regard pattern recognition as fundamental to structuring and interpreting visual information. Yet pattern recognition alone does not ensure critical understanding of visual representations.

https://doi.org/10.55578/fepr.2607.011

Furthermore, existing research emphasizes the importance of contextual knowledge for successful interpretation [4,7]. Misinterpretations often result from missing information about the underlying data rather than deficiencies in visual perception. This highlights a dimension that remains largely implicit in existing models: data provenance. Recent work on provenance visualization similarly emphasizes that understanding visualizations requires consideration of the people, processes, and transformations underlying the displayed data [9].

More recent approaches extend the understanding of visual literacy by incorporating additional cognitive processes [1]. These include an understanding of how data are collected, processed, and represented. They also include the ability to identify potential biases and inconsistencies in visualizations. Consequently, visual literacy is understood as a multidimensional construct encompassing analytical, interpretative, and reflective aspects, including reflection on data provenance. Recent systematic reviews likewise conclude that existing conceptualizations comprise multiple, partly overlapping competence dimensions and highlight the need for a more coherent and standardized conceptualization of visual literacy [10].

Taken together, existing research is critical to understanding the role of pattern recognition and cognitive processes in the interpretation of visualized data. At the same time, a research gap remains. While pattern recognition has frequently been the focus of scholarly attention, the role of data provenance in the critical interpretation of visualized data has received comparatively little investigation.

3. A CONCEPTUAL SHIFT: DATA PROVENANCE AS A MISSING DIMENSION OF VISUAL LITERACY

As outlined in Section 2, current understandings of visual literacy frequently emphasize the ability to identify patterns, interpret graphical representations, and derive relevant information from visualized data. Existing models of graph comprehension primarily focus on identifying and interpreting visible structures in visualizations [7-9].

In AI-mediated data environments, systems automatically generate visualizations, transform datasets, and create plausible synthetic data. This can lead recipients to recognize patterns without critically on data origin, quality, or trustworthiness of the underlying data. Correct interpretation therefore does not necessarily imply a critical understanding.

This development raises a fundamental question: What does it mean to understand a graph? We argue that graph comprehension should go beyond identifying patterns and interpreting visual structures to include reflection on how the underlying data are generated, processed, and transformed. Visual literacy should thus incorporate reflection on data provenance as a central competence dimension.

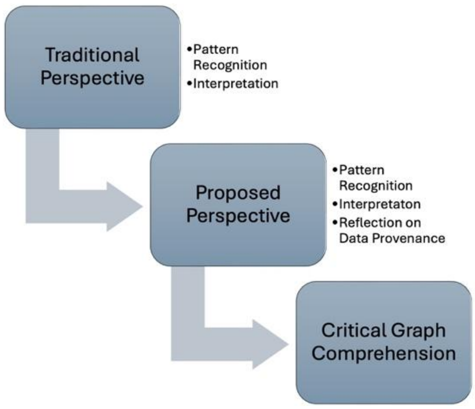

Figure 1. Conceptual extension of graph comprehension through reflection on data provenance

Data provenance refers to the origin, context, and transformation processes of data. In the context of visualized data, data provenance includes questions concerning data collection, data selection, aggregation, and processing. It also involves potential uncertainties, biases, and the use of AI-based procedures. These questions go beyond the visible representation of the data. They address the epistemic conditions that shape the creation of data visualizations.

Reflection on data provenance involves awareness of data origin, understanding data collection and processing procedures, evaluating data quality and biases, and considering AI-mediated influences on data generation and visualization. These facets, informed by ongoing work on the operationalization of data provenance reflection, may manifest in activities such as evaluating information, communicating findings, deriving recommendations, or making informed decisions based on visualized data. While not constituting a validated competence model, they may provide a preliminary analytical framework for future theoretical and empirical work.

Table 1. Preliminary conceptualization of reflection on data provenance

Facet of Data Provenance Reflection	Guiding Question Example	Example Reflective Action
Data origin	Where do the data come from?	Evaluate the credibility of the data source
Data collection	How were the data collected?	Reflect on data collection procedures
Data processing and transformation	How were the data processed, aggregated, or transformed?	Assess the effects of aggregation or preprocessing
Data quality and bias	Which uncertainties, omissions, or biases may affect the data?	Critically evaluate limitations of the dataset
AI mediation	Which AI-supported processes influenced the data or its viusalization?	Reflect on AI-generated modifications or representations

Against this background, we propose a conceptual extension of visual literacy. This extension explicitly integrates reflection on data provenance as a core component of graph comprehension. Understanding visualized data means more than recognizing what a graph shows. It also involves critically questioning how and why it was created. Visual literacy is thus conceptualized as a metacognitive competence that combines interpretation, contextualization, epistemic reflection, and reflection on data provenance.

The proposed perspective does not suggest replacing related constructs like statistical literacy, contextual knowledge, or critical thinking on data provenance. Rather, we argue that it represents a distinct dimension of visual literacy. Statistical literacy centers on interpreting quantitative data characteristics, contextual knowledge aids in understanding the visualization's subject matter, and critical thinking encompasses broader information analyses. Reflection on data provenance focuses specifically on the origin, generation, processing, and transformation of data. It addresses a dimension of graph comprehension not fully captured by existing constructs and particularly relevant in AI-mediated information environments.

Data provenance should not be regarded as a peripheral aspect of visual literacy, but as a necessary component of critical graph comprehension in data-driven and AI-mediated information environments.

4. DISCUSSION

The proposed perspective shifts attention from the interpretation of visible structures toward reflection on the conditions under which visualized information is produced and interpreted. Visual literacy encompasses not only the ability to extract information from graphical representations. It also includes the ability to critically reflect on the conditions under which such information is produced and to evaluate its significance within a given context.

Data generation and transformation processes often remain invisible to recipients, causing visualizations to appear as direct representation of reality.

The increasing use of artificial intelligence further intensifies this challenge by reducing the transparency of data generation and processing procedures. The challenge lies not only in correctly interpreting visual information. It also requires critical reflection on the context in which the information was generated. Recent research highlights the growing role of AI in interpreting and explaining visual information [11]. At the same time, studies evaluating the visualization literacy in large language models demonstrate the increasing importance of critical graph comprehension in AI-mediated environments

[12]. Together, these developments underscore the need to understand how visualized data are produced, transformed, and communicated.

From this perspective, greater emphasis on reflective competencies within visual literacy appears necessary. Graph comprehension requires consideration not only of visible patterns but also of the conditions under which they emerged. Critical engagement with data sources, data generation processes, and potential biases thus becomes an integral component of understanding visualized information. Consistent with this view, recent visualization research argues that provenance information should itself become visible within visual analysis, enabling users to better understand how visual representations are generated and transformed [13].

These developments also have implications for education. Recent work on visual literacy in the context of generative AI emphasizes combining the interpretation of visual representations with critical reflection on their production processes. In particular engaging students in the creation and analysis of AI-generated visual artefacts can foster reflection on how visual content is generated and shaped [14]. This supports the argument that visual literacy should address not only what visualizations show but also how they are produced.

Overall, this perspective suggests that visual literacy should be understood as a broader capability for critically engaging with visualized knowledge. In data-driven and AI-mediated information environments, reflection on data provenance may constitute a key prerequisite for understanding visualized information. It may also be essential for evaluating such information in an informed and critical manner.

5. IMPLICATIONS FOR RESEARCH AND EDUCATION

From the perspective proposed in this paper, implications can be derived both for future research and for the design of teaching and learning environments. Educational interventions should support both pattern identification and interpretation in light of the underlying data.

This includes, in particular, critical engagement with data collection processes, potential biases, and the conditions under which data are generated. Learners should not only identify patterns but also evaluate their significance in light of the underlying data. This requires reflection on the completeness of the data and the assumptions underlying their processing and representation.

For higher education, this implies the need to view visualizations not merely as the end products of analytical processes. It also requires making the generation and transformation processes of the underlying data more visible. This perspective is consistent with recent work highlighting that visualization education should foster competencies extending beyond the interpretation of graphical representations to broader forms of reasoning about data and visualizations [15]. Teaching and learning settings can support these goals by encouraging comparisons of data sources, identification of biases, and discussion of alternative interpretations.

The critical evaluation of data provenance becomes particularly important in the context of automated data analyses and AI-supported visualization systems. Consequently, the promotion of visual literacy should not focus solely on the interpretation of visible information. It should also enable learners to critically question the conditions under which visualized data are produced and to assess their trustworthiness in a reflective manner. Against this background, reflection on data provenance may be considered a key prerequisite for critical graph comprehension in the future.

6. OUTLOOK

The perspective presented in this paper is based on a methodological investigation of the relationship between pattern recognition and the understanding of data provenance. The empirical findings of this investigation will be reported in a subsequent publication. At the same time, it opens up further avenues for research regarding the theoretical positioning, empirical assessment, and educational promotion of reflection on data provenance as a potential dimension of visual literacy.

Future research should clarify whether reflection on data provenance constitutes an independent dimension of visual literacy, how it relates to broader concepts such as data literacy, and how it can be distinguished from existing facets of graph comprehension.

From a methodological perspective, further research is required concerning the development and validation of appropriate measurement instruments. Recent work highlights ongoing challenges in establishing valid assessment measures and defining the competence dimensions underlying visualization literacy [16]. Future research should therefore investigate how reflection on data provenance can be assessed empirically, whether robust competence models can be derived from such assessments, and which indicators are suitable for capturing this competence validly and reliably.

Given the prevalence of data-driven and AI-supported information systems, future research should examine the role of data provenance reflection in the competent use of visualized data. Further research is needed to investigate how learners and other user groups take into account the origin and conditions of data generation within automated processes of data analysis and visualization. Future studies should also examine which educational approaches are most effective in fostering this reflective capability.

The perspective presented here is intended as a contribution to the further development of existing understandings of visual literacy. It is also intended as a starting point for research that defines graph comprehension not solely in terms of the interpretation of visible patterns but places greater emphasis on the critical examination of the origin and generation of visualized data.

Acknowledgements and Use of AI: All written works within this article are the original works of the authors. While preparing this manuscript, the author employed AI-assisted tools. The first author originally wrote the manuscript in German and then translated it into English using ChatGPT-5.5. The output generated was carefully reviewed and revised by the authors, who assumes full responsibility for the accuracy and integrity of the final content presented in this publication.

Funding Statement: The authors declare that no financial support was received for the research and/or publication of this article.

Author Contrubution: The authors contributed to the research and writing of this article and have read/agreed to the published version of the manuscript.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data available on request from the corresponding author.

Conflict of Interest Statement: The authors declare no conflict of interest.

REFERENCES

[1] Börner, K., Bueckle, A., & Ginda, M. (2019). Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments. Proceedings of the National Academy of Sciences, 116(6), 1857-1864. http://dx.doi.org/10.1073/pnas.1807180116

[2] Driessen, J. E. P., Vos, D. A. C., Smeets, I. & Albers, C. J. (2022) Misleading graphs in context: Less misleading than expected. PLoS ONE 17(6): e0265823. https://doi.org/10.1371/journal.pone.0265823

[3] Lo, L. Y. H., Gupta, A., Shigyo, K., Wu, A., Bertini, E. & Qu, H. (2022). Misinformed by visualization: What do we learn from misinformative visualizations? Computer Graphics Forum 41(3), 515-525. https://doi.org/10.48550/arXiv.2204.09548

[4] Friel, S. N., Curcio, F. R. & Bright, G. W. (2001). Making Sense of Graphs: Critical Factors Influencing Comprehension and Instructional Implications. Journal for Research in Mathematics Education 32(2), 124-158. https://doi.org/10.2307/749671

[5] Firat, E. E., Denisova, A. & Laramee, R. S. (2020). Treemap Literacy: A Classroom-Based Investigation. In M. Romero & B. Sousa Santos (Hrsg.), Eurographics 2020 – Education Papers, 29-38. https://doi.org/10.2312/eged.20201032

[6] Lieb, C. & Pickhardt, C. (2025). A didactic intervention to strengthen critical thinking in the interpretation of visualized data. Front. Educ. 10:1680396. https://doi.org/10.3389/feduc.2025.1680396
[7] Shah, P., & Hoeffner, J. (2002). Review of graph comprehension research: Implications for instruction. Educational Psychology Review, 14(1), 47–69. https://doi.org/10.1023/A:1013180410169
[8] Carpenter, P. A., & Shah, P. (1998.). A Model of the Perceptual and Conceptual Processes in Graph Comprehension. Journal of Experimental Psychology: Applied, 4(2), 75-100. https://doi.org/10.1037/1076-898x.4.2.75
[9] Vancisin, T., Clarke, L. & Hinrichs, M. O. (2023). Provenance visualization: Tracing people, processes, and practices through a data-dricen approach to provenance. Digital Scholarship in the Humanities, 38(3), 1-18. https://doi.org/10.1093/llc/fqad020
[10] Beschi, S., Falessi, D., Golia, S. and Locoro, A. (2025). Characterizing Data Visualization Literacy for Standardization: A Systematic Literature Review, IEEE Access, 13, 65704-65725. https://doi.org/10.1109/ACCESS.2025.3559298
[11] Li, Z., Miao, H., Pascucci, V. and Liu, S. (2024). Visualization Literacy of Multimodal Large Language Models: A Comparative Study. arXiv preprint arXiv:2407.10996. https://doi.org/10.48550/arXiv.2407.10996
[12] Hong, J., Seto, C., Fan, A. & Maciejewski, R. (2025). Do LLMs Have Visualization Literacy? An Evaluation on Modified Visualizations to Test Generalization in Data Interpretation. IEEE Transactions on Visualization and Computer Graphics, 31(10), 7004-7018. https://doi.org/10.1109/TVCG.2025.3536358
[13] Narechania, A., Guo, S., Koh, E., Endert, A. & Hoffswell, J. (2025). Utilizing Provenance as an Attribute for Visual Data Analysis: A Design Probe With ProvenanceLens. IEEE Transactions on Visualization and Computer Graphics, 31, 8452-8465. https://doi.org/10.1109/TVCG.2025.3571708
[14] Källström, K. (2025). GIFT-AI: Damn! Jpg-visual literacy through image-making with generative AI. Front. Educ. 10:1621207. https://doi.org/10.3389/feduc.2025.1621207
[15] Hedayati, M. & Kay, M. (2024). What University Students Learn In Visualization Classes. IEEE Transactions on Visualization and Computer Graphics, 31(1), 1072-1082. https://doi.org/10.1109/TVCG.2024.3456291
[16] Brockbank, E., Verma, A., Lloyd, H., Huey, H., Padilla, L. & Fan, J. E. (2025). Evaluating convergence between two data visualization literacy assessments. Cognitive Research: Principles and Implications, 10(15). https://doi.org/10.1186/s41235-025-00622-9