Position paper: A case to study the relationship between data visualization readability and visualization literacy

ANNE-FLORE CABOUAT, TINGYING HE, FLORENT CABRIC, TOBIAS ISENBERG, and PETRA ISENBERG, Université Paris-Saclay, CNRS, Inria, LISN, France

In this position paper we argue that improving visualization literacy evaluation tools is important for defining and understanding the concept of *readability* in data visualizations. Only with reliable and relevant measures can we assess how a potential factor affects a reader's performance; accordingly, only with appropriate measuring instruments can we start to investigate the tight web of interactions between individual characteristics, features of the visual design, and reading tasks requirements. As we slowly progress in our understanding of how people process information from data visualization, and based on these improved tools and other developments, we can further develop theoretical foundations in data visualization.

CCS Concepts: • **Human-centered computing** → **Visualization**; **Empirical studies in visualization**; **Visualization design and evaluation methods**; **Visualization theory, concepts and paradigms**.

Additional Key Words and Phrases: Readability, Visualization literacy, Assessment tests

ACM Reference Format:

Anne-Flore Cabouat, Tingying He, Florent Cabric, Tobias Isenberg, and Petra Isenberg. 2024. Position paper: A case to study the relationship between data visualization readability and visualization literacy. In *Proceedings of the Workshop: Toward a More Comprehensive Understanding of Visualization Literacy (ACM CHI'24).* 10 pages. https://hal.science/hal-04523790

1 THE READING CONUNDRUM

We—as researchers—sometimes use terms that, while they are seemingly intuitive in everyday language, lack formal definitions within our field of expertise and thus make it difficult to clearly express findings or arguments. *Read* and *readability* are examples of such words in the context of data visualization [10]. Although the notion of being able to *read* is pervasive to the literature in graph comprehension (e. g., [23, 34]) and visualization usability studies (e. g., [28, 45, 51]), there is no formal definition of what constitutes *reading a data visualization.*

In this paper, we assume that the act of *reading* covers the range of cognitive processes that allow readers to retrieve meaningful pieces of information from a visual representation of data. *Readability*, then, describes how easy and effective this information retrieval is for a person. From models of graph comprehension [20, 22, 30, 44], we derive that readability in data visualization is influenced by three interacting factors: the display, the reading task, and the individual characteristics of the reader, as we show in Fig. 1. These interactions make readability a difficult construct to observe and quantify in data visualization—which echoes well-known issues in measuring text readability [5, 8].

As *readability* includes personal factors, it is also closely connected to the concept of visualization literacy. Visualization literacy, as an individual factor, encompasses the reader's ability to decode a visual representation of data; it is thus likely to have a significant impact on how readable individuals find a particular visualization for a given task.

[1]In Fig. 1, we mainly refer to low-level tasks as they were defined in Amar *et al.*'s taxonomy [3].

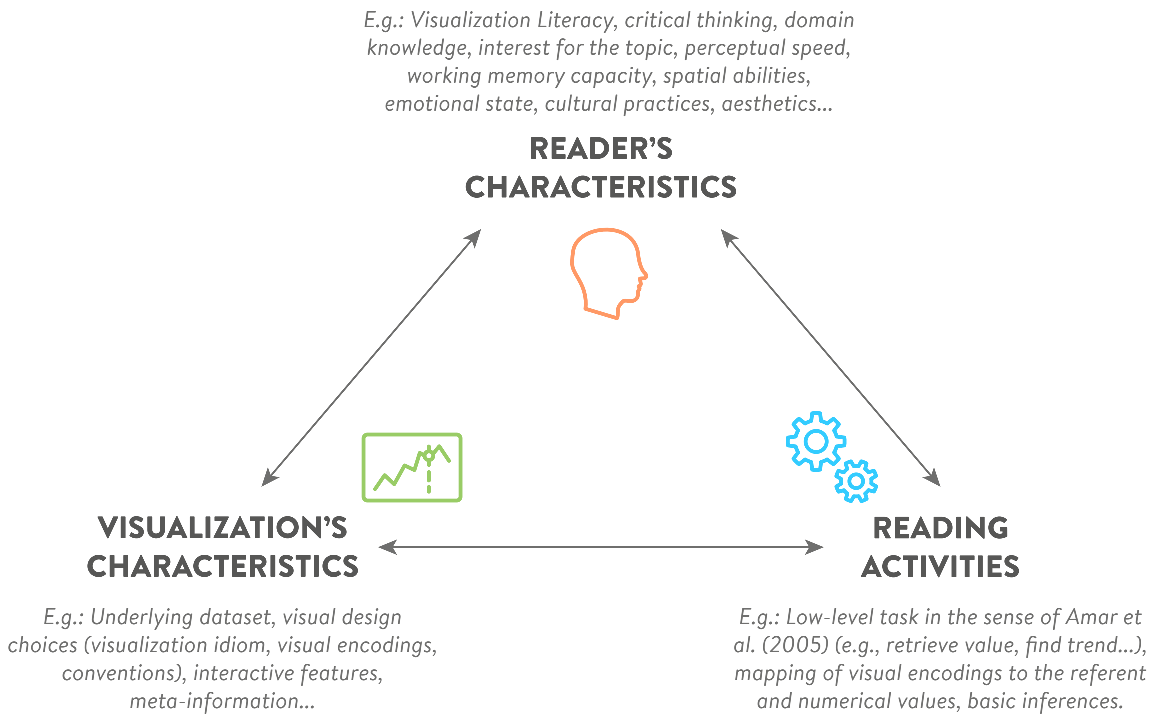

Fig. 1. A triadic reciprocal representation of components influencing reading task[1] performance in data visualization.

We are currently working on a larger research project to understand and formalize the concept of *readability*. Here, we posit that any attempt to assess readability should integrate measures of the readers' visualization literacy—that is, at least until we can derive from such measures a cognitive model of reading. Such a model should be comprehensive enough to roughly predict readers' performance, while accounting for the role of visualization literacy in reading activities. In the following sections, we first lay out how we think readability and visualization literacy are conceptually intertwined, in spite of our current lack of empirical proof to support such claims. We then focus on two possible pitfalls that hinder the utility of current visualization literacy test scores when studying readability. Finally, we make the case that bridging—not merging!—visualization literacy and visualization readability research would benefit both domains.

Note on the scope of this position paper: We acknowledge the broadly accepted definition of visualization literacy as encompassing the ability to read, interpret, design and create data visualizations [4, 7]. While we see no reason why our call to implement visualization literacy measures in empirical studies should not also apply to creation-oriented work such as elicitation studies, here we mainly focus on the "read and interpret" components. This orientation aligns with our interest in studying data visualization readability, and most usability studies relevant to our topic are conducted on already developed visualization systems.

2 VISUALIZATION LITERACY AND INDIVIDUAL DIFFERENCES

There is increasing evidence that individual characteristics might have been overlooked in seminal data visualization perception studies. For example, Davis et al. [17] recently showed how limited previous assumptions were on canonical rankings of the effectiveness of visualization *idioms*—i. e., types of representations [38]. Indeed, only 22% of the 118 participants in their study performed according to Cleveland and McGill's established ranking of accuracy [14]: BAR > PIE > STACKED BARS > BUBBLE. Their results reinforce the growing agreement that there is no "one-size-fits-all" set of rules for data visualization [40, 49]. In fact, the scope of individual factors that affect effectiveness of designs for a given task appears extremely broad [36]: visualization literacy [7, 9, 35, 48] and meta-cognitive skills such as critical thinking [27], of course; but also domain knowledge [47], prior beliefs [32, 54], manifold cognitive abilities (e. g., spatial abilities [15, 41], working memory capacity [52], or verbal abilities [56]), personal semantic associations [1], personality traits [2], aesthetic preferences [11, 29], and cultural practices [50] are all examples of good candidates. The complex interplay of these factors, and their effects on perceptual, reading, and reasoning processes, remain largely unexplored. As a result, we do not know, on a theoretical level, how visualization literacy affects readability. Reciprocally, we do not know either how visualization readability might affect respondents' answers in visualization literacy assessment tests.

3 VISUALIZATION LITERACY IN THE VISUALIZATION INFORMATION PROCESSING FRAMEWORK

Visualization researchers have undertaken efforts to theorize how people process information from data visualization. There are calls to develop cognitive models [33, 42] that explain how people perceive, read, and interpret visual representations of data, and how readers can leverage information retrieved from data visualizations for higher-level tasks such as decision making, learning, or data analysis. To inform such models, we need to understand how factors of the visual display, the individual reader, and the reading task, interact when people engage with a visualization [21].

Early theoretical work in visualization and education research domains pointed out components that we now relate to visualization literacy: we can, for example, think of Pinker's *visualization schema* [44] and Freedman and Shah's *graphical knowledge* [22]. Researchers outside of the data visualization and education communities also produced empirical work that relates to effects of skills akin to visualization literacy, such as those found in the context of health risk communication [24–26] and cognitive psychology [13]. In both domains, researchers refer to the ability to read visual representations of data as *graphicacy*. While researchers in the health communication domain focused on evaluating the effects of graphicacy on health risk information understanding and recall, cognitive psychologists focused on formally establishing trend judgment as a *perceptual building block* of the ability to read scatterplots [13].

Therefore, we have solid reasons to expect an effect of visualization literacy on reading task performance in visualization. Yet, we could not find supporting empirical evidence of such a relationship in the visualization literature.

4 VISUALIZATION LITERACY TEST SCORES AND READING PERFORMANCE: THE ELUSIVE LINK

Despite efforts of the research community to develop visualization literacy assessment tests [9, 27, 35], we found only one study [37] that investigated the relationship between visualization literacy test scores and reading performance beyond the testing context. This study by Mansoor et al. [37] partially replicated the comparison reading experiment from Cleveland and McGill [14] in a crowd-sourced online study with 29 participants. Their goal was to assess how participants' scores on a Visualization Literacy Assessment Test (VLAT [35]) related to their performance in a value comparison task on similar visualizations. Participants were given an abbreviated version of the VLAT (comprising the BAR CHART, BUBBLE CHART, PIE CHART, LINE CHART, and TREEMAP VLAT visualizations) and 15 comparison task trials for each of 3 visualization types (BAR CHARTS, PIE CHARTS, and TREEMAPS).

The authors found no relation between the participants' average reading task accuracy and their score on the abbreviated VLAT. Instead, the results showed a correlation between the variance in participant's errors and the abbreviated VLAT score. These findings suggest that people with higher visualization literacy might be more *consistent* in their reading of a data visualization—but not more *accurate*. Interestingly, the authors also indicate that they did not find correlations between accuracy in reading performance or abbreviated VLAT score and self-reported visualization experience, statistical experience or other demographics such as age and educational attainment.

For our purpose of studying the effect of visualization literacy on readability, we see at least two potential pitfalls in current visualization literacy measurement tools that require attention: the aggregate nature of the scores they yield and the lack of readability variety in the visualization situations they propose. We discuss these matters below.

4.1 The problem of visualization literacy aggregate scores

Table 1. Examples of existing general-purpose Visualization Literacy assessment tools yielding an aggregate ability score

Tool	Types of visualizations	Tasks	Number of items[2] in a test
VLAT [35]	**12 idioms:** line chart, bar chart, stacked bar chart, 100% stacked bar chart, pie chart, histogram, scatterplot, bubble chart, area chart, stacked area chart, choropleth map, and treemap	**8 tasks:** retrieve value, find extremum, determine range, characterize distribution, find anomalies, find clusters, find correlations/trends, and make comparisons	55 (up to 7 per visualization)
Mini-VLAT [43]	**12 idioms** (all idioms from VLAT, with slightly different designs)	**8 tasks** (all tasks from VLAT)	12 (1 per visualization)
CALVI [27]	**9 idioms:** line chart, bar chart, stacked bar chart, 100% stacked bar chart, pie chart, scatterplot, area chart, stacked area chart, and choropleth map **11 misleaders:** cherry picking, concealed uncertainty, inappropriate aggregation, inappropriate scale range, arbitrary use of non-linear scales, unconventioanl scale directions, misleading annotations, mising data, missin normalization, and overplotting	**6 tasks:** retrieve value, find extremum, find correlations/trends, make comparisons, make predictions, and aggregate values	30 (1 per visualization) (15 *normal* items and 15 trick items to pick from of a 45-items bank)
A-VLAT and **A-CALVI** [16]	**A-VLAT:** 12 original visualizations **A-CALVI:** 11 original misleaders	**A-VLAT:** 8 original tasks **A-CALVI:** no specification for tasks	**A-VLAT:** 27 **A-CALVI:** 15 (11 *trick* items and 4 *normal* items)

Literacy considered broadly is a complex construct, and literacy assessments are usually multi-dimensional [31, 46]. For instance, the 2022 OECD PISA's mathematical literacy assessment [39] scores differently on four categories of mathematical skills, namely: "change and relationships", "space and shape", "quantity", "uncertainty", and "data." Similarly, as shown in Table 1, visualization literacy test items[2] encompass a variety of different tasks and visualization *idioms*—i. e., types of visual representations [38]. Consequently, it is likely that visualization literacy is a multi-dimensional construct [35]. To our knowledge, unfortunately, most of existing general-purpose visualization literacy tests result in one aggregate score: a single value output which encapsulates and hides levels in possible sub-skills and knowledge components of visualization literacy.

Aggregate scores from current visualization literacy assessment tools make it difficult to explore how visualization literacy relates to data visualization readability, because they hide the presumable diversity of sub-skills that form visualization literacy. Some sub-components of visualization literacy might be task-dependent: as an example, Firat *et al.* [19] identified 7 categories of possible difficulties for readers in a literacy test that focused on Parallel Coordinate Plots (PCPs, an example of which is shown in Fig. 2). Out of the 7 groups of potential difficulties, 2 are directly named after tasks: identifying *correlations*, and *path tracing*—a low-level task in graphs which consists of visually following a line. The authors then characterize their PCP literacy test's items according to these categories. They observe that any item generally belongs to 2–3 of these categories, which they identify as *cognitive processes*. We find that their groups incorporate visualization literacy components together with (unrelated) readability components:

[2] In this work, we refer to a test *item* as a combination of one question and one specific visualization.

- We can consider elements such as "understanding of multivariate data attributes" or "statistical terminology (i. e., correlation)" as required *knowledge* for the individual, and thus part of visualization literacy.
- Similarly, we can recognize elements describing abilities relevant to reading a plot and that a person can learn as *skills* belonging to visualization literacy. The "use of parallel axes" or the "path tracing" tasks are examples.
- However, elements like "overplotting" or "placement of axes" are characteristics of the display, and thus not directly part of visualization literacy. There is no evidence that a cognitive skill exists to read overplotted data points or lines—it is more likely a problem to solve through visualization system design, e. g., *via* interactive features.

We opened this section by emphasizing how detailed scores for visualization literacy components would participate in explaining visualization readability; in the same way, a better understanding of visualization readability is needed to support researchers in categorizing visualization literacy test items.

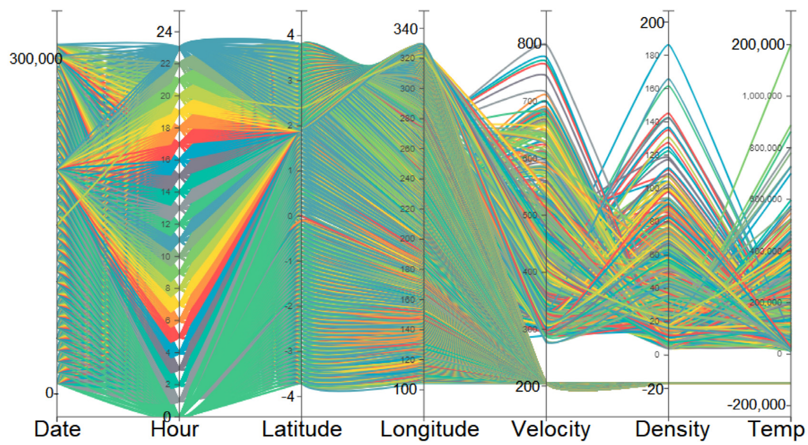

Fig. 2. An example of a Parallel Coordinate Plot image used in P-Lite, as shown by Firat *et al.* [19].
Image used under the CC BY 4.0 license.

The multi-dimensional nature of visualization literacy might also relate to the type of visualization. For instance, being able to infer the original value of a mark on a logarithmic scale can be a part of visualization literacy for some idioms, such as a DOT PLOT or a SCATTER PLOT that may sometimes use logarithm scales. Conversely, the inability to deduce a value from a position on a logarithm scale would not impede a reader in other visualization contexts (e. g., being able to estimate the value of a population encoded as circle areas in a BUBBLE CHART).

If a test only yields an aggregate visualization literacy score for respondents, this score does not inform us about which visualization literacy sub-skills or knowledge components are involved for specific reading tasks and visualization types. Being able to observe correlation patterns between measured sub-skill levels in readers could inform a framework of visualization information processing. Moreover, if we combine such a correlation analysis with data from eye-tracking devices we can start to shed light on how sub-skills may rely on common perceptual or visual processes.

Identifying the multiple dimensions and underlying skills and knowledge components of visualization literacy is crucial to better understand the cognitive processes at stake in *reading* data visualizations. Therefore, we renew Lee et al.'s [35] call for multi-dimensional factor analysis of visualization literacy test results, and encourage researchers to pursue this direction. We add that frameworks providing typologies of visualization literacy skills and knowledge such as Börner *et al.*'s [6] can provide a starting point to conduct Confirmatory Factor Analysis on test results, or to specify prior distributions in Bayesian models in both test development and test results analysis.

4.2 The lack of readability variance in visualization literacy assessment tests

Another barrier to the use of existing visualization literacy tests in empirical studies on visualization readability is the relatively low variance of visualization readability itself in the tests. This poses a problem as visualization literacy scores measured from questions on simple datasets with very clear designs do not tell us how "resilient" a person might be to low-readability situations.

Expanding from the pioneering work of Boy *et al.* [9], Lee *et al.* [35] aimed to develop "a comprehensive visualization literacy assessment test containing various data visualizations, covering inclusive data visualization tasks, and following the whole procedure of test development." To that end, they used a set of 12 different idioms (as we show in Fig. 3). If we examine the visualizations through the lens of readability factors, however, we notice that these represent relatively simple datasets and use clean visual encodings, both characteristics which maximize the display's clarity and therefore increase the ease with which readers can extract information from the visualizations (= our current definition of *readability*). This observation may shed some light on the lack of results that Mansoor *et al.* reported [37] when they attempted to link partial VLAT scores to performance in a comparison task beyond the test's context (as we describe with more details in the beginning of Section 4).

The high visual clarity of the VLAT visualizations is certainly in line with the authors' goal to provide an assessment test for novice visualization readers. Meanwhile, as Wang *et al.* noted in their work on the comprehension of scatterplots [53], VLAT items do not fully reflect the reality of what makes a data visualization easy or difficult to read. A similar observation led Ge *et al.* [27] to develop the CALVI test, which assesses respondents' ability to apply critical reasoning while reading visualizations. The authors created a pool of realistic and diverse test items. They applied 11 *misleaders* (e. g.,

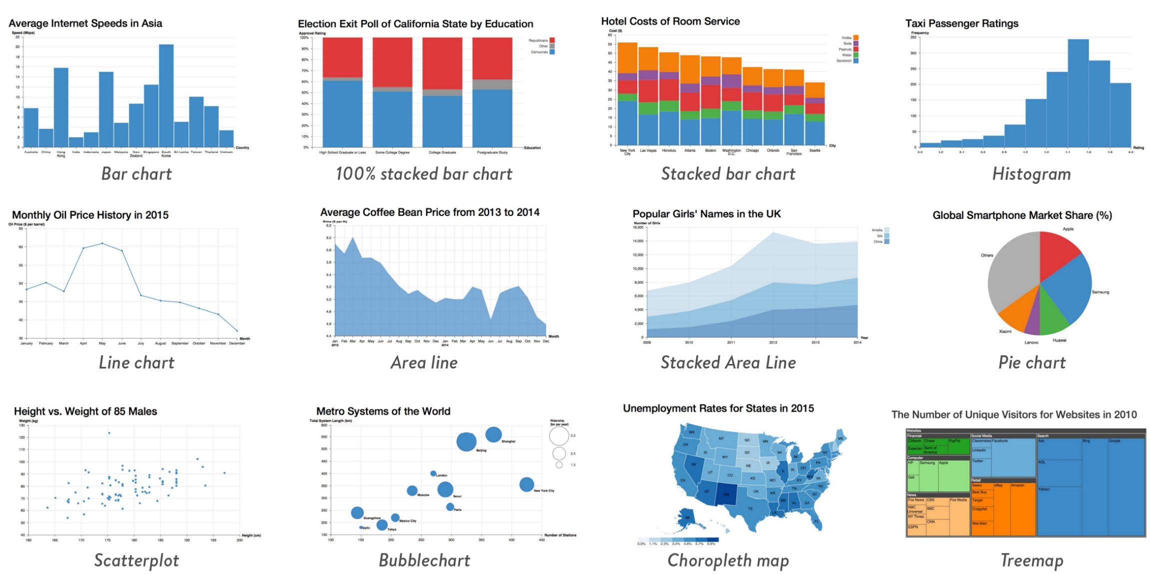

Fig. 3. The 12 data visualizations that compose the VLAT [35], as found on the ReVISit [18] project repository.
The ReVISit software is distributed under the BSD 3-Clause License.

inappropriate aggregation, *missing data*, different *manipulations of scales*, or *overplotting*) to create *trick* items—questions that require the reader to think critically about the visualization in order to provide a correct answer.

As we are still progressing in our definition of readability, it is unclear for now which of CALVI's misleaders directly relate to the ease of reading, and which depend solely on the lack of something we might want to call "faithful interpretability" of designs. For example, let us consider the practice of *concealing uncertainty*, which consists of plotting uncertain data without visually encoding a measure of uncertainty. This is a case of design-induced misguidance for which the outcome relates less to the ease of reading and more to the falseness of a reader's sense of confidence in their interpretation of the visualization. A truncated axis, in contrast, is a form of *scale manipulation* that directly affects the ease with which a reader can perform a bar comparison task (in the sense of Cleveland and McGill [14]; Fig. 4).

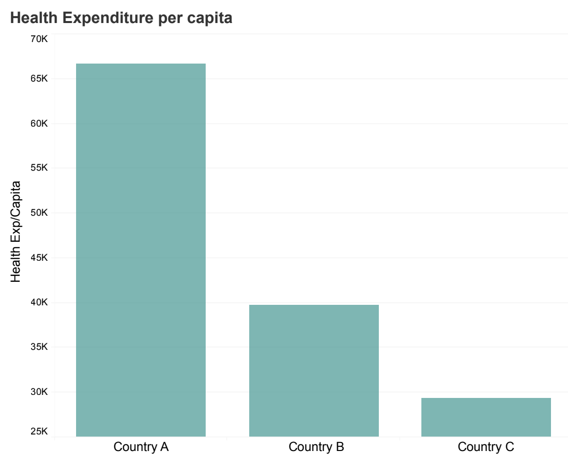

Fig. 4. A bar chart with a truncated y-axis makes it difficult to rely solely on *reading* for answering the question: "Is it true that the average health expenditure per capita in country A is more than twice the amount in country B?"

In fact, to obtain a correct answer, such manipulation of the y-axis might force the reader to perform two visual *readings* of values, and then to perform a *calculation* of the values' proportional relationship—not a *visualization reading* task anymore. The visualization would technically not be *readable* for this task; but readers with high enough visualization literacy could leverage their knowledge to reason that direct reading of the answer is impossible, and adapt their strategy accordingly for answering the question. In many ways, the CALVI items provide us with a precious resource in studying and reflecting on readability.

Other factors of readability, such as variations in data complexity or readers' domain knowledge, remain unexplored in existing general-purpose visualization literacy assessment tests. An example of an open question is: "Can a reader's visualization literacy affect the amount and quality of insights they can derive from a visualization in the absence of domain knowledge?" In such a situation, what is the place of readability?

Meanwhile, we still lack standardized tools to evaluate readability in visualizations. Researchers aiming to develop readability evaluation tools should strive to apply the principled approach demonstrated in the visualization literacy test development research, including its most recent advances such as the use of Bayesian models [17, 27].

5 CONCLUSION

In this paper, we make the case that visualization literacy and visualization readability research have overlapping topics of interest but different focus. Visualization literacy research aims to investigate the knowledge and skills that people require in order to effectively understand, use and create data visualizations. Meanwhile, readability research is targeted at explaining how different characteristics of readers, tasks and data display contribute in making a specific visualization easy or hard to read. We do not argue that one stream of research should be merged into the other, but rather that the development or refinement of evaluation tools in one domain can support the community's efforts to develop and improve evaluation tools in the other. Explorations of the dimensions and sub-skills of visualization literacy will help define and explain readability in data visualization. Reciprocally, a deeper understanding of readability would help expand the scope of situations in which visualization literacy is assessed, and support researchers in their exploration of sub-components of visualization literacy. Therefore, communication and iterations between our respective fields should prove fruitful in the future. Finally, we emphasize that empirical findings from both streams of research are key to inform a more comprehensive framework of how people process information in data visualization. Such a framework should then be taught as part of visualization design curricula, and be included in models for recommender [55] and linter [12] systems.

REFERENCES

[1] Jarryullah Ahmad, Elaine Huynh, and Fanny Chevalier. 2021. When red means good, bad, or Canada: Exploring people's reasoning for choosing color palettes. In *Proc. VIS*. IEEE CS, Los Alamitos, 56–60. doi: 10/gtjwqr

[2] Tomás Alves, Bárbara Ramalho, Daniel Gonçalves, Sandra Gama, and Joana Henriques-Calado. 2020. Exploring how personality models information visualization preferences. In *Proc. VIS*. IEEE CS, Los Alamitos, 201–205. doi: 10/gtjwqq

[3] R. Amar, J. Eagan, and J. Stasko. 2005. Low-level components of analytic activity in information visualization. In *Proc. InfoVis*. IEEE CS, Los Alamitos, 111–117. doi: 10/bwrm27

[4] Benjamin Bach, Mandy Keck, Fateme Rajabiyazdi, Tatiana Losev, Isabel Meirelles, Jason Dykes, Robert S. Laramee, Mashael AlKadi, Christina Stoiber, Samuel Huron, Charles Perin, Luiz Morais, Wolfgang Aigner, Doris Kosminsky, Magdalena Boucher, Søren Knudsen, Areti Manataki, Jan Aerts, Uta Hinrichs, Jonathan C. Roberts, and Sheelagh Carpendale. 2024. Challenges and opportunities in data visualization education: A call to action. *IEEE Trans Vis Comput Graph* 30, 1 (2024), 649–660. doi: 10/gtjwqj

[5] Alan Bailin and Ann Grafstein. 2001. The linguistic assumptions underlying readability formulae: a critique. *Lang Commun* 21, 3 (2001), 285–301. doi: 10/d6jngm

[6] Katy Börner, Andreas Bueckle, and Michael Ginda. 2019. Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments. *PNAS* 116, 6 (2019), 1857–1864. doi: 10/gfw9xb

[7] Katy Börner, Adam Maltese, Russell Nelson Balliet, and Joe Heimlich. 2016. Investigating aspects of data visualization literacy using 20 information visualizations and 273 science museum visitors. *Inf Vis* 15, 3 (2016), 198–213. doi: 10/f8sgws

[8] Rebekah George Benjamin. 2012. Reconstructing readability: Recent developments and recommendations in the analysis of text difficulty. *Educ Psychol Rev* 24, 1 (2012), 63–88. doi: 10/bdjfkd

[9] Jeremy Boy, Ronald A. Rensink, Enrico Bertini, and Jean-Daniel Fekete. 2014. A principled way of assessing visualization literacy. *IEEE Trans Vis Comput Graph* 20, 12 (2014), 1963–1972. doi: 10/f6qjv6

[10] Anne-Flore Cabouat, Tingying He, Petra Isenberg, and Tobias Isenberg. 2023. Pondering the reading of visual representations. https://inria.hal.science/hal-04240900

[11] Nick Cawthon and Andrew Vande Moere. 2007. The effect of aesthetic on the usability of data visualization. In *Proc. IV*. IEEE CS, Los Alamitos, 637–648. doi: 10/c8v6s8

[12] Qing Chen, Fuling Sun, Xinyue Xu, Zui Chen, Jiazhe Wang, and Nan Cao. 2022. VizLinter: A linter and fixer framework for data visualization. *IEEE Trans Vis Comput Graph* 28, 1 (2022), 206–216. doi: 10/grj2ft

[13] Lorenzo Ciccione, Mathias Sablé-Meyer, Esther Boissin, Mathilde Josserand, Cassandra Potier-Watkins, Serge Caparos, and Stanislas Dehaene. 2023. Trend judgment as a perceptual building block of graphicacy and mathematics, across age, education, and culture. *Sci Rep* 13, 1 (2023), 10266. doi: 10/gtkckx

[14] William S. Cleveland and Robert McGill. 1984. Graphical perception: Theory, experimentation, and application to the development of graphical methods. *J Am Stat Assoc* 79, 387 (1984), 531–554. doi: 10/cbtq7v

[15] Cheryl A. Cohen and Mary Hegarty. 2007. Individual differences in use of external visualisations to perform an internal visualisation task. *Appl Cognit Psychol* 21, 6 (2007), 701–711. doi: 10/bsjr9z

[16] Yuan Cui, Lily W. Ge, Yiren Ding, Fumeng Yang, Lane Harrison, and Matthew Kay. 2024. Adaptive assessment of visualization literacy. *IEEE Trans Vis Comput Graph* 30, 1 (2024), 628–637. doi: 10/gtjwqh

[17] Russell Davis, Xiaoying Pu, Yiren Ding, Brian D. Hall, Karen Bonilla, Mi Feng, Matthew Kay, and Lane Harrison. 2024. The risks of ranking: Revisiting graphical perception to model individual differences in visualization performance. *IEEE Trans Vis Comput Graph* 30, 3 (2024), 1756–1771. doi: 10/gtjwqp

[18] Yiren Ding, Jack Wilburn, Hilson Shrestha, Akim Ndlovu, Kiran Gadhave, Carolina Nobre, Alexander Lex, and Lane Harrison. 2023. reVISit: Supporting Scalable Evaluation of Interactive Visualizations. In *2023 IEEE Visualization and Visual Analytics (VIS)*. 31–35. doi: 10.1109/VIS54172.2023.00015

[19] Elif E. Firat, Alena Denisova, Max L. Wilson, and Robert S. Laramee. 2022. P-Lite: A study of parallel coordinate plot literacy. *Visual Inf* 6, 3 (2022), 81–99. doi: 10/gs28q5

[20] Amy Rae Fox. 2023. Theories and models in graph comprehension. In *Visualization Psychology*, Danielle Albers Szafir, Rita Borgo, Min Chen, Darren J. Edwards, Brian Fisher, and Lace Padilla (Eds.). Springer, Cham, 39–64. doi: 10/gtgz78

[21] Amy Rae Fox and James D. Hollan. 2023. Visualization Psychology: Foundations for an interdisciplinary research program. In *Visualization Psychology*, Danielle Albers Szafir, Rita Borgo, Min Chen, Darren J. Edwards, Brian Fisher, and Lace Padilla (Eds.). Springer, Cham, 217–242. doi: 10/gtgz76

[22] Eric G. Freedman and Priti Shah. 2002. Toward a model of knowledge-based graph comprehension. In *Proc. Diagrams*. Springer, Berlin, 18–30. doi: 10/fr56t9

[23] Susan N. Friel, Frances R. Curcio, and George W. Bright. 2001. Making sense of graphs: Critical factors influencing comprehension and instructional implications. *J Res Math Educ* 32, 2 (2001), 124–158. doi: 10/bmpjc8

[24] Wolfgang Gaissmaier, Odette Wegwarth, David Skopec, Ann-Sophie Müller, Sebastian Broschinski, and Mary C. Politi. 2012. Numbers can be worth a thousand pictures: Individual differences in understanding graphical and numerical representations of health-related information. *Health Psychol* 31, 3 (2012), 286–296. doi: 10/dmfpv5

[25] Mirta Galesic and Rocio Garcia-Retamero. 2011. Graph literacy: A cross-cultural comparison. *Med Decis Making* 31, 3 (2011), 444–457. doi: 10/fqxhjn

[26] Rocio Garcia-Retamero and Edward T. Cokely. 2013. Communicating health risks with visual aids. *Curr Dir Psychol Sci* 22, 5 (2013), 392–399. doi: 10/b2pk

[27] Lily W. Ge, Yuan Cui, and Matthew Kay. 2023. CALVI: Critical thinking assessment for literacy in visualizations. In *Proc. CHI*. ACM, New York, 1–18. doi: 10/gtjwqc

[28] Mohammad Ghoniem, Jean-Daniel Fekete, and Philippe Castagliola. 2005. On the Readability of Graphs Using Node-Link and Matrix-Based Representations: A Controlled Experiment and Statistical Analysis. *Inf Vis* 4, 2 (2005), 114–135. doi: 10/bncqv6

[29] Christopher G. Healey and James T. Enns. 2002. Perception and painting: A search for effective, engaging visualizations. *IEEE Comput Graph Appl* 22, 2 (2002), 10–15. doi: 10/cqbq7k

[30] Mary Hegarty. 2011. The cognitive science of visual-spatial displays: Implications for design. *Top Cognit Sci* 3, 3 (2011), 446–474. doi: 10/c3274w

[31] Peter Johnston and Paula Costello. 2005. Principles for literacy assessment. *Reading Res Q* 40, 2 (2005), 256–267. doi: 10/cg8hjq

[32] Yea-Seul Kim, Logan A. Walls, Peter Krafft, and Jessica Hullman. 2019. A Bayesian cognition approach to improve data visualization. In *Proc. CHI*. ACM, New York, 682:1–682:14. doi: 10/ghpvgm

[33] Robert Kosara. 2016. An empire built on sand: Reexamining what we think we know about visualization. In *Proc. BELIV*. ACM, New York, 162–168. doi: 10/gfz5kr

[34] Stephen M. Kosslyn. 1989. Understanding charts and graphs. *Appl Cognit Psychol* 3, 3 (1989), 185–225. doi: 10/cvdzr5

[35] Sukwon Lee, Sung-Hee Kim, and Bum Chul Kwon. 2017. VLAT: Development of a visualization literacy assessment test. *IEEE Trans Vis Comput Graph* 23, 1 (2017), 551–560. doi: 10/f92d38

[36] Zhengliang Liu, R. Jordan Crouser, and Alvitta Ottley. 2020. Survey on individual differences in visualization. *Comput Graph Forum* 39, 3 (2020), 693–712. doi: 10/gg6cr5

[37] Hamid Mansoor, Kartik Vasu, and Lane Harrison. 2017. Linking performance on graphical perception tasks to visualization literacy. In *IEEE VAST Posters*.

[38] Tamara Munzner. 2014. *Visualization Analysis and Design*. CRC Press, New York. doi: 10/gd3xgq

[39] OECD. 2023. *PISA 2022 Assessment and Analytical Framework*. OECD Publishing, Paris. doi: 10/gtkcpp

[40] Nels Oscar, Shannon Mejía, Ronald Metoyer, and Karen Hooker. 2017. Towards personalized visualization: Information granularity, situation, and personality. In *Proc. DIS*. ACM, New York, 811–819. doi: 10/ggqq3b

[41] Alvitta Ottley, Evan M. Peck, Lane T. Harrison, Daniel Afergan, Caroline Ziemkiewicz, Holly A. Taylor, Paul K. J. Han, and Remco Chang. 2016. Improving Bayesian reasoning: The effects of phrasing, visualization, and spatial ability. *IEEE Trans Vis Comput Graph* 22, 1 (2016), 529–538. doi: 10/gf65g9

[42] Lace M. Padilla. 2018. A case for cognitive models in visualization research: Position paper. In *Proc. BELIV*. IEEE CS, Los Alamitos, 69–77. doi: 10/gf7cng

[43] Saugat Pandey and Alvitta Ottley. 2023. Mini-VLAT: A short and effective measure of visualization literacy. *Comput Graph Forum* 42, 3 (2023), 1–11. doi: 10/gtjwqb

[44] Steven Pinker. 1990. A theory of graph comprehension. In *Artificial Intelligence and the Future of Testing*, Roy Freedle (Ed.). Lawrence Erlbaum Associates, Inc., Hillsdale, Chapter 4, 73–126.

[45] Puripant Ruchikachorn and Klaus Mueller. 2015. Learning visualizations by analogy: Promoting visual literacy through visualization morphing. *IEEE Trans Vis Comput Graph* 21, 9 (2015), 1028–1044. doi: 10/ggjbtd

[46] Kristine Sørensen, Stephan Van den Broucke, James Fullam, Gerardine Doyle, Jürgen Pelikan, Zofia Slonska, Helmut Brand, and (HLS-EU) Consortium Health Literacy Project European. 2012. Health literacy and public health: A systematic review and integration of definitions and models. *BMC Public Health* 12, 1 (2012), 80. doi: 10/fzhzds

[47] Priti Shah and Eric G. Freedman. 2011. Bar and line graph comprehension: An interaction of top-down and bottom-up processes. *Top Cogn Sci* 3, 3 (2011), 560–578. doi: 10/frczb8

[48] Priti Shah, Eric G. Freedman, and Ioanna Vekiri. 2005. The comprehension of quantitative information in graphical displays. In *The Cambridge Handbook of Visuospatial Thinking* (1 ed.), Priti Shah and Akira Miyake (Eds.). Cambridge University Press, 426–476. doi: 10/crwhhf

[49] Ben Steichen, Giuseppe Carenini, and Cristina Conati. 2013. User-adaptive information visualization: Using eye gaze data to infer visualization tasks and user cognitive abilities. In *Proc. IUI*. ACM, New York, 317–328. doi: 10/gtgz8c

[50] Sharifa Sultana, Syed Ishtiaque Ahmed, and Jeffrey M. Rzeszotarski. 2021. Seeing in context: Traditional visual communication practices in rural Bangladesh. *Proc ACM Hum-Comput Interact* 4, CSCW3 (2021), 214:1–214:31. doi: 10/gtjwqs

[51] Alice Thudt, Jagoda Walny, Charles Perin, Fateme Rajabiyazdi, Lindsay MacDonald, Riane Vardeleon, Saul Greenberg, and Sheelagh Carpendale. 2016. Assessing the readability of stacked graphs. In *Proc. GI*. CHCCS, Waterloo, 167–174. doi: 10/gtgz8d

[52] Dereck Toker, Cristina Conati, Giuseppe Carenini, and Mona Haraty. 2012. Towards adaptive information visualization: On the influence of user characteristics. In *User Modeling, Adaptation, and Personalization*, Judith Masthoff, Bamshad Mobasher, Michel C. Desmarais, and Roger Nkambou (Eds.). Springer, Berlin, 274–285. doi: 10/gtjwqt

[53] Jiachen Wang, Xiwen Cai, Jiajie Su, Yu Liao, and Yingcai Wu. 2022. What makes a scatterplot hard to comprehend: data size and pattern salience matter. *J Vis* 25, 1 (2022), 59–75. doi: 10/gtjwph

[54] Cindy Xiong, Chase Stokes, Yea-Seul Kim, and Steven Franconeri. 2023. Seeing what you believe or believing what you see? Belief biases correlation estimation. *IEEE Trans Vis Comput Graph* 29, 1 (2023), 493–503. doi: 10/gtg7mh

[55] Sujia Zhu, Guodao Sun, Qi Jiang, Meng Zha, and Ronghua Liang. 2020. A survey on automatic infographics and visualization recommendations. *Visual Inf* 4, 3 (2020), 24–40. doi: 10/ghhxst

[56] Caroline Ziemkiewicz and Robert Kosara. 2009. Preconceptions and individual differences in understanding visual metaphors. *Comput Graph Forum* 28, 3 (2009), 911–918. doi: 10/bdghfc

Received 01 March 2024; accepted 25 March 2024