Scientific Visualization Literacy: Assessment, Challenges, and Opportunities

Chaoli Wang[1], Patrick Phuoc Do[1], Robert S. Laramee[2], and Alark Joshi[3]

[1]Department of Computer Science and Engineering, University of Notre Dame, United States
[2]School of Computer Science, University of Nottingham, United Kingdom
[3]Department of Computer Science, University of San Francisco, United States

Abstract

Over the past decade, visualization literacy has been extensively investigated in the fields of visualization and human-computer interaction, encompassing topics such as measurement and assessment, training and learning, and design for literacy and accessibility. Existing work in this area, however, is primarily focused on information visualization (charts, graphs, and diagrams) and rarely explores scientific visualization (such as volumes and surfaces). In this position paper, we discuss the fundamental differences between information visualization literacy and scientific visualization literacy. We outline an approach to designing literacy assessment tests for scientific visualization and brief a prototype. We identify research opportunities and open challenges facing scientific visualization literacy. The goal is to examine the role of literacy in scientific visualization and to advocate for a community effort to advance scientific visualization literacy, thereby supporting an informed public discourse around science and engineering.

CCS Concepts

• ***Human-centered computing*** → ***Scientific visualization; Empirical studies in visualization;***

1. Introduction

Visualization literacy is the ability to interpret, understand, and produce visual representations of data, a crucial skill in our data-driven society for informed decision-making [BBG19, NZM*24, BFGL25]. In a recent survey on visualization literacy, Varona et al. [VBH*25] emphasized that visualization literacy must be operationalized based on specific application contexts and is organized around four key competency themes: *consumption*, *construction*, *critique*, and *connection*, as illustrated in Figure 1, Consumption involves tasks such as reading and comparing values, interpreting trends, and making decisions based on visualizations (e.g., images, videos). Construction refers to the ability to use visualization creation tools, select appropriate encodings, and apply visual hierarchy (i.e., arranging visual elements so that viewers naturally perceive and process information in a deliberate order). Critique focuses on applying critical thinking to visualizations, ranging from identifying misleading elements to assessing a visualization's neutrality or bias. Connection relates data within a visualization to broader contexts or external knowledge, including the formation of mental models or engagement in debates. We refer interested readers to additional surveys on interactive visualization literacy [FJL22], visualization literacy for standardization [BFGL25], and domain-specific visualization literacy, such as data visualization for financial literacy [DAMW25].

The four competency themes form an increasingly complex skill set. In reality, even at the beginning consumption level, many individuals face significant challenges in accurately interpreting visualizations [LKH*15, BMBH16, NZM*24]. These challenges are often categorized into three types [NZM*24]: *translation barriers* occur when users struggle to convert a data-driven question into a visual query; *encoding barriers* arise from a misunderstanding of the visual encodings within a chart; and *decoding barriers* refer to an inability to extract data values from a visualization correctly. These barriers can be conceptual (resulting from flaws in understanding) or operational (due to mistakes in interpreting values). Common problems include misreading visual channels, confusing chart labels with data values, high visual complexity, and unfamiliar visualization types [RM15, BMBH16, NZM*24].

Addressing these problems requires targeted educational interventions. Approaches include active learning, direct engagement with visual elements, and immediate feedback [PFLJ22, BFGL25]. Pedagogical strategies also focus on teaching skills like deconstruction (breaking down visualizations) and design empathy (understanding design choices) [NZM*24]. In addition, explicitly teaching how to identify deceptive visualizations [CCB*22] is a core component of visualization literacy education. There is also a great need to develop a unifying construct and protocol for visualization literacy to standardize definitions and assessment methods, while also exploring barriers in diverse contexts and populations.

DOI: 10.2312/eved.20261001

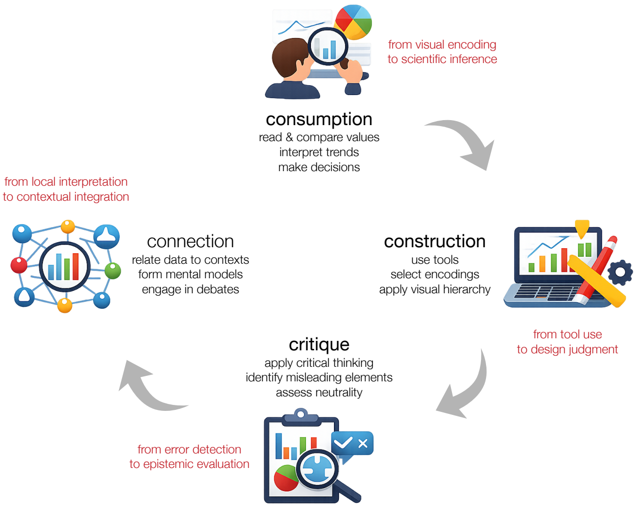

Figure 1: *The four competency themes of visualization literacy adopted from [VBH*25]. The text in red highlights opportunities and challenges of visualization literacy in the SciVis context.*

In visualization, the two primary branches are *information visualization* (InfoVis) and *scientific visualization* (SciVis). Both share areas of overlap and convergence, including hybrid data (e.g., spatial networks, geotemporal data), standard interaction techniques, increasing emphasis on explainability and accessibility, and cross-pollination of evaluation methods. Nevertheless, they are fundamentally different. InfoVis focuses on *abstract data* with *designer-assigned geometry*, supporting *sense-making for broad audiences* through charts and graphs. It is often evaluated via *user-centered studies*. In contrast, SciVis focuses on *spatially grounded data* with *inherent geometry*, supporting *understanding of physical phenomena for domain experts* through techniques such as volume rendering and isosurfaces. It is often evaluated through *scientific correctness and domain-centered validation*. For further information, refer to the work of Tory and Möller [TM04], which presents a taxonomy that clarifies the overlaps and differences between InfoVis and SciVis by categorizing visualization methods based on design models, rather than only data type.

Existing visualization literacy studies (e.g., [FIB*19, FDL20, BZP*21, FDWL22]) and assessment tests (e.g., [LKK17, GCK23, PO23, CGD*24]), however, are largely centered on InfoVis. There is a *notable lack of systematic studies on SciVis literacy, assessment, and education*. We conjecture that this is not due to neglect, but rather to a deep entanglement with domain knowledge, the difficulty of standardization and evaluation, a historical focus on algorithms over users, and limited pressure from non-expert audiences. Similar to investigating InfoVis literacy, studying SciVis literacy is essential to ensure that visualizations support accurate spatial reasoning, enable responsible decision-making, and remain interpretable as visualization techniques become increasingly complex and widespread.

We provide an example to illustrate how SciVis literacy manifests across the four competency themes [VBH*25]. Consider a mixed rendering of a climate simulation that displays 3D atmospheric temperature and wind fields over time, utilizing volume rendering and animated pathlets. At a basic level, viewers must be able to *consume* the visualization by correctly interpreting color mappings, spatial position, and motion to identify temperature gradients or dominant flow patterns. Beyond interpretation, scientists may be asked to *construct* such a visualization by selecting appropriate rendering techniques, parameter settings, and viewpoints that faithfully represent the underlying model while remaining interpretable to a target audience. As visualizations become inputs to analysis or communication, viewers must also be able to *critique* them: questioning whether the chosen encodings obscure uncertainty, exaggerate patterns, or introduce artifacts that could mislead spatial reasoning. Finally, effective use of SciVis requires the ability to *connect* what is shown to external knowledge and broader contexts, such as linking observed patterns to physical mechanisms, comparing results across models, or using the visualization to inform interdisciplinary discussion or policy decisions. This example illustrates how SciVis literacy spans all four competency themes and how increasing proficiency involves moving from basic visual decoding to integrated scientific reasoning and judgment.

This position paper advocates for a community effort for SciVis literacy research. The main contributions are as follows:

- We highlight the similarities and differences between InfoVis literacy and SciVis literacy, reasoning that SciVis literacy presents special challenges.
- We propose developing SciVis literacy assessment tests through a rigorous, multi-phase process and brief the results from a prototype implementation.
- We outline opportunities and challenges for advancing SciVis literacy across the four competency themes and provide an action plan to advance research and education in this area.

The remainder of the paper is organized as follows. Section 2 reviews related work on visualization literacy. Section 3 examines the similarities and differences between InfoVis literacy and SciVis literacy and discusses the challenges unique to SciVis literacy. Section 4 presents an assessment framework for developing and evaluating SciVis literacy. Section 5 discusses research opportunities and open challenges. Finally, Section 6 outlines an action plan to advance SciVis literacy research and education, and Section 7 concludes the paper.

2. Related Work

We review related work in visualization literacy education and assessment, as well as the study of visualization literacy in the context of foundation models. Note that prior work in these areas is squarely on InfoVis literacy, underscoring an absence in the SciVis literacy counterpart.

2.1. Visualization Literacy Education

Public proficiency in visualization literacy remains low. To address this, educational efforts on visualization literacy encompass a range of approaches. Game-based learning is a notable direction. Examples include narrative-focused role-playing games that can improve engagement and enjoyment in children without sacrificing learning [HNGC21]. Serious games, such as Iguanodon [ALE*25], enhance visualization construction literacy for adults by breaking down design choices into manageable subtasks. Interactive toolkits such as C'est la Vis [ARC*17, CRA*18] and Data is Yours [BVY*23] facilitate active, embodied learning and collaboration for younger children by allowing them to physically craft and interact with visualizations, often relying on everyday materials to make the process approachable and encourage experimentation. Additionally, design thinking methodologies, inspired by concepts like "Dear Data," [KMK19] are being integrated into higher education courses to foster divergent thinking and solution-oriented visualization creation among novices, thereby preventing common pitfalls such as tunnel vision. Despite the above advancements, challenges persist. These include overcoming user overconfidence, addressing curriculum gaps in which the critical interpretation of misleading visuals is often neglected, managing the increased time commitment required by immersive learning methods, and adapting tools to diverse audiences and learning contexts. Such challenges highlight a continuous need for more comprehensive, longitudinal studies and standardized assessment methods for visualization literacy [DC20].

2.2. Visualization Literacy Assessment

To measure visualization literacy, based on *item response theory* (IRT) [HSR91], Boy et al. [BRBF14] developed an approach [DeM10] for assessing the ability to use established visualizations like line graphs and bar charts confidently. Building on this, Lee et al. [LKK17] created the 53-item *visualization literacy assessment test* (VLAT), evaluating tasks such as retrieving values, finding extremums, and making comparisons across various chart types. Recognizing the gap in assessing critical thinking for misleading visuals, Ge et al. [GCK23] introduced *critical thinking assessment for literacy in visualizations* (CALVI), a 45-item instrument to identify and reason about erroneous visualizations. To address the length of these tests, Pandey and Ottley [PO23] developed the 12-item Mini-VLAT, a psychometrically sound short form of the VLAT with good reliability and strong correlation to the original. Similarly, Cui et al. [CGD*24] created two *computerized adaptive tests* (CATs), A-VLAT (27 items) and A-CALVI (15 items), which halve the number of questions while maintaining high test-retest reliability and validity. Further advancements include Chang et al.'s work on linking visualization literacy to visual attention [CWW*26], in which models (Lit2Sal and Sal2Lit) predict literacy levels based on distinct viewing patterns. These diverse efforts aim to provide more efficient, reliable, and nuanced measurements of visualization literacy; however, challenges such as context dependency and the need for domain-specific adaptations remain.

2.3. LLM's Visualization Literacy

Visualization literacy is increasingly used to probe the capabilities and limitations of *large language models* (LLMs) and *visual–language models* (VLMs) with respect to visual information. Early evaluations found that systems such as GPT-4 and Gemini often fall short of human performance on established benchmarks like VLAT, particularly on tasks that required careful reading of values. In many cases, the models appear to rely on general world knowledge rather than grounding their answers in the visualization itself [HSFM25]. Later work shows that these weaknesses are not immutable. Techniques such as Charts-of-Thought [DTM26] demonstrate that when models are guided to slow down and explicitly reason through extraction and verification steps, their performance can improve substantially. Even so, interpreting color encodings, making sense of dense geographic visualizations, estimating values without explicit numerical labels, and resisting misleading titles or visual framing remain challenging for current models [LMPL24, BS25, VDLE25, PO25]. More focused studies, including work based on datasets like FlowLearn [PZC*24], paint a nuanced picture in which model capabilities vary widely across task types rather than improving uniformly. Systems such as InvVis [YLLW24] point toward a different direction by attempting to recover underlying data directly from visualization images, enabling new forms of interaction and reuse. Finally, analyses of model internals using tools like AG-CAM [DC26] suggest that while VLMs can capture meaningful spatial and semantic cues, their reasoning remains fragile, often sensitive to prompt phrasing and prone to breakdown in multi-step analyses. This body of work highlights both the promise of foundation models for visualization understanding and the need for continued research toward more robust, transparent, and visually grounded reasoning.

3. Scientific Visualization Literacy

Both *InfoVis literacy* and *SciVis literacy* refer to the capacity to interpret, analyze, and critically evaluate visual representations of data to support understanding, reasoning, and decision-making, integrating *data literacy* (i.e., understanding data concepts) and *visual literacy* (i.e., understanding visual encodings and perception). The differences are that InfoVis literacy focuses on reading *abstract data representations*, while SciVis literacy focuses on reasoning *spatially grounded visual representations of physical phenomena*.

Compared to InfoVis, visualization literacy is especially challenging in SciVis due to the following reasons:

- **3D complexity and interaction needs**: SciVis processes 3D scalar, vector, or tensor fields. Unlike InfoVis, which typically represents abstract relationships in 2D visual encodings, unique 3D-intrinsic interactions in SciVis include camera navigation, arbitrary slicing, clipping, particle seeding, and region selection. Low visualization literacy can limit a user's ability to interact effectively in 3D.
- **Domain-specific knowledge assumption**: SciVis typically represents spatiotemporal physical phenomena (e.g., fluid flow, molecular structures, medical scans). Interpreting these scientific phenomena and specialized variables (e.g., pressure, vorticity) through visualization often assumes understanding the underlying science. As such, SciVis literacy typically demands stronger domain knowledge, whereas InfoVis literacy focuses more on general-purpose visual reasoning about abstract data.
- **Abstract, large, unfamiliar data**: SciVis frequently represents phenomena invisible to the human eye (e.g., electromagnetic fields, particle simulations). These data are often orders of magnitude larger than those in InfoVis. Unlike bar or line charts, which use familiar metaphors, these abstract representations can make interpretation more challenging.
- **Uncertainty and noise**: Scientific data often come with measurement errors, simulation artifacts, or incomplete sampling. Conveying and interpreting uncertainty in a 3D context is complex and requires a higher level of literacy.

- **Perceptual and cognitive load**: SciVis often involves dense data, overlapping structures, and semi-transparent layers. It also requires specialized lighting and shading techniques to perceive 3D and time-varying phenomena. This creates a higher visual and cognitive load compared to a cleaner, more abstract InfoVis design.
- **Lack of standardized literacy frameworks**: While assessment tools exist for InfoVis literacy, no equivalent standardized framework has been developed for SciVis. This makes it harder to teach, assess, and systematically improve SciVis literacy.

In summary, SciVis literacy presents additional challenges because it requires users to navigate 3D interactions, apply domain-specific knowledge, interpret abstract, large, and noisy data, and manage a higher perceptual load, all without standardized assessment tools to guide learning.

4. Development of SciVis Literacy Assessment

We propose a framework for developing a SciVis literacy assessment that draws on established principles from psychological and educational measurement [CST21] and on methodological precedents from prior visualization literacy instruments [LKK17, GCK23]. Central to this framework is a multi-stage, psychometrically informed process that connects construct definition to test blueprinting, item development, and empirical evidence gathered through pilot and large-scale tryout studies. This approach situates SciVis literacy assessment within the broader ecosystem of visualization literacy tests, which follows a process similar to previous work [LKK17, PO23, GCK23, CGD*24].

4.1. Blueprint Construction

The process begins by articulating the *construct and constraints* that define SciVis literacy, including adopting a *closed-world* assumption: each test item must be answerable using only the provided visualization and its caption. Although interpreting scientific visualizations in real practice may require domain knowledge, the proposed literacy test intentionally stays at the consumption level. Accordingly, the test assesses an audience's ability to read and interpret information directly from the visual elements (e.g., rendering, color encoding, spatial structure, legends) and textual elements (e.g., labels and captions) without relying on prior knowledge about the underlying scientific application. This design choice is particularly important because the target population includes a general public audience, for whom domain expertise cannot be assumed.

Adopting the closed-world assumption ensures that the instrument measures visualization literacy rather than disciplinary knowledge. If items required prior familiarity with specific scientific concepts, performance would confound visualization comprehension with differences in educational background. By restricting all necessary information to the visual representation and its accompanying text, the assessment isolates the ability to decode visual encodings, interpret spatial relationships, and integrate graphical and textual cues, which are core competencies of visualization literacy. This approach also reflects a common goal of SciVis communication: figures should remain interpretable to readers outside the domain when supported by clear legends and captions.

A *test blueprint* is then constructed to operationalize the construct along two complementary dimensions: representative SciVis visualization techniques and a task taxonomy that specifies the types of evidence the assessment is intended to elicit.

The technique dimension encompasses common practices in both canonical SciVis research and public-facing scientific communication. It may include the following categories:

- **Technique families** that are widely recognized as foundational to SciVis and well documented in the literature [Tel14], such as *color mapping*, *volume rendering*, *surface rendering*, *texture- or integration-based techniques*, and *mixed rendering approaches*.
- **Glyph-, plot-, and mesh-based representations**, capturing symbolic or abstract encodings frequently used alongside spatial techniques.
- **Scientific illustrations**, encompassing schematic and explanatory depictions that rely on diagrammatic conventions rather than direct data rendering [Goo05].

The task dimension is derived from an interpretation-oriented taxonomy originally developed for volume visualization analysis [LBLS15], adapted here for domain-independent assessment. It includes five high-level categories instantiated as eleven task types:

- **Search** tasks, consisting of *presence/absence* and *counting*.
- **Pattern recognition** tasks, such as identifying *trends* or *repetitions*.
- **Spatial understanding** tasks, covering *absolute*, *relative*, and *intersection-based* judgments.
- **Quantitative estimation** tasks, including *absolute estimation*, *binary relative estimation*, and *quantitative relative estimation*.
- **Shape description** tasks.

This level of specification supports systematic item authoring, balanced coverage of the construct space, and task-level diagnostic analyses during later validation stages. We note that these typical SciVis tasks *differ* from many InfoVis tasks [TM04].

4.2. Visualization Preparation

A curated set of visualizations is assembled to represent the target technique categories while avoiding over-specialization to a narrow data context. To promote generalizability, the collection is diversified across *data types*, *dimensionalities*, and *application domains*. Depending on the intended audience (e.g., college students vs. the general public), source materials may emphasize different venues, ranging from textbooks and research articles to science journalism, educational websites, and online media.

The resulting set encompasses scalar, vector, tensor, and mixed-field data, includes 2D, 3D, and time-varying cases, and spans multiple scientific disciplines to minimize reliance on domain-specific expertise. Both static images and animated sequences are incorporated: techniques that depend on temporal variation are presented as short videos, whereas others are shown as still figures. Quality control is conducted to ensure that images and videos have sufficient resolution, high-quality rendering, and clearly visible legends. For simplicity, we intentionally target *view-independent* interpretation that can be assessed from a figure and caption under our closed-world assumption. Accordingly, interaction-dependent exploration and uncertainty-specific reasoning are out of scope for this initial instrument and left for future extensions that can provide standardized interactive interfaces and uncertainty encodings.

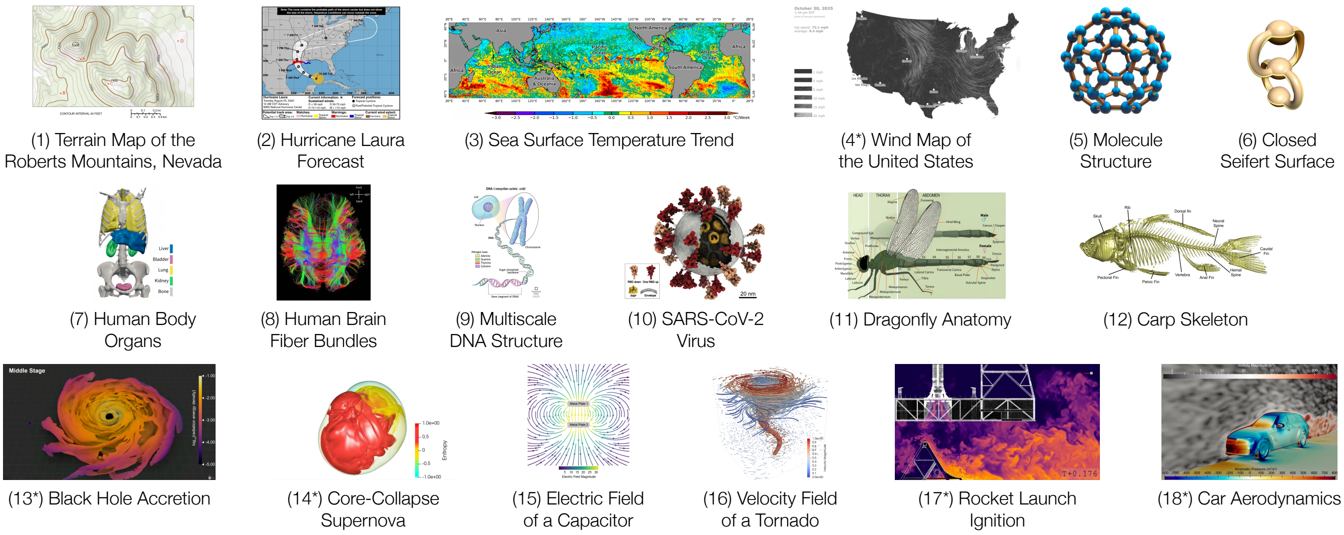

Figure 2: *The 18 visualizations selected for SVLAT. Indices marked with * indicate animations; all others are static images.*

Each visualization is accompanied by a caption that provides essential context without giving cues to the correct answer. When possible, captions are adapted from original source descriptions and rewritten for a general audience, explicitly identifying the phenomenon, visualization technique, and key encodings while avoiding technical jargon. Captions are kept concise and neutral. LLMs (e.g., ChatGPT) may assist in drafting and refinement, with all captions undergoing manual verification to ensure accuracy and clarity.

4.3. Item Generation

Item generation follows directly from the blueprint and results in selected-response items, primarily multiple-choice questions (MCQs), with some true/false (T/F) items, paired with carefully constructed distractors that reflect plausible interpretation errors. Open-ended items could also be included. However, this type of item is difficult to evaluate reliably because it requires judgmental evaluation and qualitative analysis [LKK17]. An initial item pool is created by instantiating combinations of visualization techniques and task types using the prepared visual materials. Each item is designed to target a single dominant task type, facilitating clean blueprint coverage and enabling fine-grained diagnostic analysis.

Key aspects of the item development process include:

- **Item format:** Selected-response formats support objective scoring and scalable administration. MCQs are preferred for their diagnostic richness, while T/F items are reserved for cases where MCQ options would otherwise incur excessive reading. A "Skip" option is included to reduce forced guessing.
- **Drafting and revision:** Items are iteratively drafted and refined, using LLMs to improve clarity and alignment with the construct while maintaining the closed-world constraint. Revisions aim to ensure that difficulty arises from visual interpretation rather than linguistic complexity or external knowledge.
- **Distractor design:** Distractors are item-specific, plausible, and construct-relevant, reflecting common visualization misconceptions rather than superficial wording differences.
- **Task labeling:** Multiple researchers independently assign each item to a primary task. Disagreements are resolved through discussion, with ambiguous cases reviewed by one or more SciVis experts and labeled according to the dominant intended evidence target.

Following item generation, *content validity* is evaluated through expert judgments of item essentiality. Lawshe's *content validity ratio* (CVR) is used to quantify agreement and guide item revision or removal [Law75]. This expert screening step mirrors established practice in prior visualization literacy test development efforts [LKK17, PO23, GCK23].

4.4. Pilot Study and Test Tryout

A *pilot study* is first conducted to assess feasibility and identify issues requiring early revision, including completion time, clarity of prompts and captions, technical fidelity of visual materials, and preliminary item-level performance signals.

Subsequently, a large-scale *test tryout* is administered to enable psychometric evaluation and item calibration. Participants complete a prototype containing all retained items. Depending on the target population, administration may occur in classroom settings (for student samples) or via crowdsourcing platforms such as Amazon Mechanical Turk for general adult populations. The test should be administered on a laptop or desktop computer with sufficient screen resolution; participants are not permitted to use mobile phones to complete the test. Responses are collected through a web-based system with standard data-quality controls. The resulting dataset supports subsequent item refinement, structural validation, and reliability analysis.

4.5. Item Analysis and Test Refinement

Tryout data are analyzed using IRT [DeM10]. Appropriate dichotomous models, such as the Rasch (1PL) and two-parameter logistic (2PL) models, are fitted to place items on a common latent scale. Item functioning is evaluated using diagnostics, including item characteristic curves, fit indices, and checks for local dependence. Measurement precision is assessed through item and test information functions, which guide decisions about item retention, revision, and coverage across the intended ability range [ER00, DeM10, HJ93].

4.6. Reliability Evaluation

Reliability is reported using IRT-based indices of overall and conditional precision, such as the conditional standard error of measurement derived from the test information function [ER00, DeM10]. To facilitate comparison with existing visualization literacy instruments, McDonald's ω_t [McD13] and Cronbach's α [Cro51] are reported as reliability estimates to assess internal consistency and the coherence of the item set in measuring SciVis literacy.

4.7. Prototype

Following the above guidelines, we have operationalized a *scientific visualization literacy assessment test* (SVLAT) to measure SciVis literacy across a wide range of visualization forms and interpretation tasks. SVLAT consists of 49 items based on 18 scientific visualizations and illustrations (see Figure 2) representing eight visualization techniques. The instrument was developed through a staged, psychometrically informed process. We first defined the construct and created a test blueprint, then generated candidate items and conducted an expert review using CVR with five SciVis experts. Next, we conducted a pilot study (30 participants) followed by a large-scale test tryout (485 participants) to evaluate the instrument's psychometric properties. For validation, we performed item analysis using both *classical test theory* (CTT) [AY79] and IRT to assess item performance and overall test quality. Results from the tryout sample indicate that SVLAT has high reliability (McDonald's $\omega_t = 0.82$, Cronbach's $\alpha = 0.81$). More information is provided in [DTAW26].

5. Research Opportunities and Open Challenges

In this section, we examine the opportunities and challenges associated with advancing SciVis literacy across the four competency themes [VBH*25]: consumption, construction, critique, and connection (see Figure 1). Progress in this area requires moving beyond isolated perceptual or usability studies toward integrated frameworks that capture how visualization skills develop, interact, and transfer across different contexts. A key challenge is building higher competency without requiring extensive domain expertise. A significant opportunity lies in aligning design, education, and assessment to support scalable scientific reasoning.

5.1. Consumption: From Visual Decoding to Scientific Inference

At the foundational level, consumption competency in SciVis involves accurately decoding visuals and extracting basic information. These skills align with well-studied perceptual tasks [QR22]. A key research opportunity lies in understanding how these low-level skills scale to more complex scientific representations, including multivariate and time-varying data. Perceptual limitations include occlusion, depth ambiguity, or color misinterpretation. Challenges arise when these limitations compromise interpretive reliability, particularly for novices or cross-domain audiences.

At higher competency levels, we expect consumption to extend beyond simply value-reading. It should form scientifically meaningful interpretations and support hypothesis generation. This is challenging because correct interpretations often depend on implicit domain knowledge, data provenance, or hidden model constraints. Developing literacy models that distinguish between surface-level correctness and deeper scientific understanding presents a research opportunity. Designing assessments and visual encodings that help users recognize uncertainty, artifacts, and the inherent limitations of scientific data is also crucial [BHJ*14].

5.2. Construction: From Tool Use to Design Judgment

At basic levels, construction competency in SciVis involves the ability to operate visualization tools, reproduce standard techniques, and follow established design patterns. Compared to InfoVis, there are fewer SciVis software tools, and the learning curve is steeper. Newer systems lower the technical barrier to creating visualizations [WHS*23, ATW26]. However, interpretability or scientific validity do not simply follow ease of construction. Research opportunities exist in studying how novice creators reason about default settings, parameter choices, and automated pipelines. Additionally, it is essential to determine where construction fluency may outpace interpretive understanding.

At advanced levels, construction involves making principled design decisions that align techniques with questions, data, and audience. This highlights the challenges of tacit expertise, as many effective SciVis design choices are learned informally. Opportunities include formalizing design knowledge, creating literacy-aware authoring tools, and supporting reflection on trade-offs among expressiveness, accuracy, and cognitive load.

5.3. Critique: From Error Detection to Epistemic Evaluation

At introductory levels, critique competency in SciVis literacy involves identifying obvious errors, misleading encodings, or inappropriate color scales. While these skills are increasingly emphasized in general visualization literacy [CCB*22] and in Bloom's taxonomy-based interventions [PFLJ22], SciVis presents additional challenges as technical and scientific authority is often attributed to complex graphical representations. Research opportunities lie in understanding how trust and expertise influence users' willingness to question scientific visualizations. This is especially important when recognized institutions (such as NASA, NIH, etc.) or sophisticated techniques (such as flow visualization using integral curves, topological data analysis–based feature extraction) produce them.

At higher competency levels, critique becomes epistemic [Cha23]. That is, users must evaluate whether a visualization supports valid scientific reasoning. They should question whether alternative visual representations lead to different conclusions. In addition, they should investigate whether uncertainty and assumptions are adequately communicated. Advancing literacy at this level requires addressing challenges in teaching and assessing critical reasoning about models, simulations, and derived data [MTD*10]. Opportunities include developing critique frameworks tailored to SciVis [FJDL08], designing tasks that elicit evaluative reasoning rather than factual recall, and studying how experts articulate and justify skepticism when interpreting visual evidence.

Table 1: *Research agenda mapping SciVis literacy competency themes to foundation model roles.*

Theme	Model as student	Model as tutor	Model as judge
Consumption	VLMs can be applied to interpret SciVis images and investigate potential failure modes, uncovering perceptual ambiguities and limitations in visual encoding that influence human interpretation.	Adaptive explanations, prompts, and alternative interpretations can be produced to guide users and support more thorough scientific reasoning and inference.	Visualizations can be evaluated to determine how effectively they can be interpreted, considering the specific tasks they support, the target audience, and how uncertainty is represented.
Construction	By examining human-created scientific visualizations, models can learn to recognize common design patterns and novice errors.	These tools can aid in creating visualizations, selecting appropriate parameters, and articulating design rationales, fostering a reflective understanding of the data and design choices.	Users can critique visualization designs by evaluating how well the chosen techniques align with the data's characteristics and the intended scientific objectives.
Critique	Models can simulate common misinterpretations to reveal systematic biases and potentially misleading visual conventions.	Critical reflection can be encouraged by generating counterfactual views, alternative visual encodings, or questions that highlight uncertainty.	The validity of a visualization, its potential for misrepresentation, and the associated epistemic risks can be evaluated to determine how reliably it supports scientific reasoning.
Connection	Models can learn the relationships between visual patterns and the underlying scientific phenomena across different domains.	Contextual narratives and cross-domain explanations can be provided to help users build accurate, coherent mental models.	Visualizations can be assessed to determine whether they facilitate responsible use and integration within broader scientific, policy, or public contexts.

5.4. Connection: From Local Interpretation to Contextual Integration

At early stages, connection competency in SciVis literacy involves relating graphical representations of data to the immediate context. For example, recognizing what physical phenomenon is shown or how different views correspond to one another. A key challenge is that many SciVis representations abstract away context to manage complexity. This could make it harder for inexperienced/novice users to build coherent mental models. Research opportunities include investigating how annotations [MCL*25], storytelling [MLF*12], and coordinated multiple views [Rob07] can support contextual grounding without oversimplifying the science.

At more advanced levels, connection extends to integrating visualization insights with external knowledge, prior experience, and broader scientific or societal contexts. Examples include comparing results across studies, engaging in interdisciplinary discussions, and using visualizations to inform policy or public discourse. The challenge here is that such connections often rely on both visualization literacy and domain literacy. Therefore, disentangling the two in research and assessment is difficult. Opportunities lie in studying how scientific visualizations function as boundary objects across communities. Equally important is the design of literacy interventions that foster meaningful dialogue among experts, decision-makers, and the public. For example, effective use of medical visualizations or illustrations can help explain complex concepts to patients in a clear and accessible manner [MGS*22].

5.5. Implications of Foundation Models for SciVis Literacy

Building on the above theme-by-theme discussion that focuses on human-centered competencies, the integration of LLMs and VLMs introduces a complementary, socio-technical dimension to advancing SciVis literacy. As listed in Table 1, foundation models can participate in the literacy pipeline in three distinct but interrelated roles: as *students* that expose systematic interpretive failures, as *tutors* that scaffold learning and reflection, and as *judges* that support critique and assessment.

Across the four competency themes, these roles create new opportunities to probe the limits of visual encodings. They provide adaptive and context-sensitive guidance. Additionally, they scale evaluative practices that previously relied solely on expert judgment. Meanwhile, these roles pose new challenges, including aligning model behavior with human perceptual constraints, risks of over-reliance on automated explanations or evaluations, and questions of epistemic authority when models mediate scientific interpretation. Collectively, this perspective reframes SciVis literacy as an evolving partnership between humans and AI. That is, advancing competency depends on improving visual designs and educational practices as well as integrating foundation models as diagnostic, instructional, and evaluative instruments.

5.6. Equity Considerations

SciVis literacy cannot be uniform across audiences because of their varying prior exposure to visualization conventions, interactive 3D manipulation, and spatial reasoning strategies. All of them shape how reliably people can interpret volumes, surfaces, and time-varying phenomena. Equity issues arise from perceptual and technological constraints, such as color-vision deficiency, low-contrast viewing conditions, and motion sensitivity. These constraints can systematically disadvantage less-trained viewers or those unable to disambiguate subtle color differences in color maps or to follow animation-heavy encodings, posing substantial challenges. Addressing SciVis literacy, therefore, presents opportunities not only in teaching "what the visualization shows," but more importantly, in cultivating transferable strategies for decomposing complexity. These strategies include, for example, managing occlusion and depth ambiguity, interpreting motion cues, and recognizing artifacts/uncertainty. Meanwhile, promoting design and communication practices that remain readable across diverse viewers and contexts is imperative.

6. Advancing SciVis Literacy Research and Education

Advancing SciVis literacy requires coordinated efforts across research, education, and practice. We outline several priorities that could benefit from community-wide collaboration.

- **Establish shared conceptual frameworks**: The community should work toward clearer definitions of SciVis literacy, including its core competencies, levels of expertise, and relationships to visualization use (e.g., consumption vs. production). A shared framework would help align research agendas, assessment design, and educational goals.
- **Develop open assessment instruments and repositories**: Publicly available, validated instruments, such as literacy tests, benchmark tasks, and annotated visualization datasets, would enable reproducible studies and cross-study comparisons. Maintaining open repositories of assessment materials and example visualizations would support both research and teaching.
- **Create benchmark datasets and evaluation protocols**: Standardized benchmarks for studying SciVis comprehension can help researchers systematically evaluate visualization techniques, interaction methods, and design choices. Shared evaluation protocols would improve comparability across studies and accelerate methodological progress.
- **Expand empirical studies across diverse audiences**: Our current prototype is aimed at the general public. Future research should examine a broader range of audiences, including interdisciplinary scientists and specific student populations, to better understand how SciVis literacy develops across different educational and experiential backgrounds.
- **Integrate SciVis literacy into education and training**: Visualization literacy should be incorporated into curricula in visualization, data science, and STEM disciplines. Developing shared teaching resources, such as instructional modules, example visualization case studies, and classroom activities, can help educators teach students how to interpret and critically evaluate scientific visualizations. SciVis literacy assessments, such as SVLAT, can be used in classroom settings to identify learning gaps. Instructors can focus on addressing those learning gaps in the classroom.
- **Foster interdisciplinary collaboration**: Advancing SciVis literacy will benefit from partnerships among visualization researchers, domain experts, education researchers, and cognitive scientists. Such collaborations can improve theoretical understanding, inform the design of effective assessments, and support evidence-based educational practices.

Together, these efforts can establish SciVis literacy as both a research agenda and an educational priority, strengthening the ability of scientists, students, and the public to interpret and communicate scientific information through visualization.

7. Concluding Remarks

In contrast to InfoVis, SciVis presents unique challenges due to the spatially grounded nature of data and its inherent geometric properties. Due to these reasons, users must navigate 3D spatial complexity, interpret abstract phenomena, and manage high perceptual and cognitive loads. InfoVis literacy has been extensively studied, whereas SciVis literacy remains underdeveloped. It requires urgent attention to ensure that visual representations support accurate scientific reasoning and informed public discourse around science and engineering.

In this position paper, we first propose a framework for developing SciVis literacy assessments and report preliminary results from our prototype. Previous work on InfoVis literacy in classroom settings has a long history (e.g., see Table 6 of Firat et al. [FJL22]). Future work in SciVis literacy also includes studying SciVis literacy in educational settings. We believe that the proposed literacy assessment can inform and serve as the basis for classroom tests; however, this remains a major direction for future work.

At the end, we outline a research agenda that moves beyond basic visual decoding toward integrated scientific inference and epistemic evaluation. Furthermore, we emphasize the emerging role of foundation models as students, tutors, and judges in the literacy pipeline, thereby positioning SciVis literacy as an evolving partnership between humans and AI. Ultimately, we call for a sustained community effort to advance SciVis literacy research and education. Aligning design, education, and assessment can better support both domain experts and the public in interpreting the complex visual data central to modern science.

Acknowledgements

C. Wang and P. P. Do are supported by funding from the NSF IIS-2101696, OAC-2104158, and IIS-2401144. R. S. Laramee is supported by funding from the EPSRC UKRI157, APP17227. The authors would like to thank the anonymous reviewers for their insightful comments.

Figure Credits

Figure 2 image credits: (1) Terrain Map of the Roberts Mountains, Nevada—Image adapted from USGS National Geologic Map Database. (2) Hurricane Laura Forecast—Image adapted from NOAA National Hurricane Center. (3) Sea Surface Temperature Trend—Image adapted from NOAA Coral Reef Watch. (4*) Wind Map of the United States—Animation from hint.fm. (5) Molecule Structure—Image from free3d.com. (6) Closed Seifert Surface—Image adapted from Jarke J. van Wijk (Eindhoven University of Technology, Netherlands). (7) Human Body Organs—Data from The Cancer Imaging Archive. Image created using ParaView. (8) Human Brain Fiber Bundles—Image adapted from Zeynep Saygin / MIT Koch Institute. (9) Multiscale DNA Structure—Image from Cleveland Clinic. (10) SARS-CoV-2 Virus—Image from Hangping Yao et al., Cell 183(3):730-738, 2020. (11) Dragonfly Anatomy—Image adapted from M. A. Broussard @ Wikimedia Commons. (12) Carp Skeleton—Data from Michael Scheuring (University of Erlangen, Germany). Image created using ParaView. (13*) Black Hole Accretion—Animation adapted from Yan-Fei Jiang (Flatiron Institute) and Patrick Moran (NASA Ames Research Center). (14*) Core-Collapse Supernova—Data from John M. Blondin (North Carolina State University) and Anthony Mezzacappa (Oak Ridge National Laboratory). Animation created using ParaView. (15) Electric Field of a Capacitor—Image created using Python code from scipython.com. (16) Velocity Field of a Tornado—Data produced using C code from Roger A. Crawfis (The Ohio State University). Image created using ParaView. (17*) Rocket Launch Ignition—Animation adapted from Michael Barad and Tim Sandstrom (NASA Ames Research Center). (18*) Car Aerodynamics—Animation adapted from CFDSupport @ Youtube.

References

[ALE*25] ADELBERGER P., LESOTA O., ECKELT K., SCHEDL M., STREIT M.: Iguanodon: A code-breaking game for improving visualization construction literacy. *IEEE Transactions on Visualization and*

Computer Graphics 31, 9 (2025), 5713–5725. doi:10.1109/TVCG.2024.3468948. 2

[ARC*17] ALPER B., RICHE N. H., CHEVALIER F., BOY J., SEZGIN T. M.: Visualization literacy at elementary school. In *Proceedings of ACM Conference on Human Factors in Computing Systems* (2017), pp. 5485–5497. doi:10.1145/3025453.302587. 3

[ATW26] AI K., TANG K., WANG C.: NLI4VolVis: Natural language interaction for volume visualization via multi-LLM agents and editable 3D Gaussian splatting. *IEEE Transactions on Visualization and Computer Graphics 32*, 1 (2026), 46–56. doi:10.1109/TVCG.2025.3633888. 6

[AY79] ALLEN M., YEN W. M.: *Introduction to Measurement Theory*, first ed. Wadsworth, 1979. 6

[BBG19] BÖRNER K., BUECKLE A., GINDA M.: Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments. *Proceedings of the National Academy of Sciences of the United States of America 116*, 6 (2019), 1857–1864. doi:10.1073/pnas.1807180116. 1

[BFGL25] BESCHI S., FALESSI D., GOLIA S., LOCORO A.: Characterizing data visualization literacy for standardization: A systematic literature review. *IEEE Access 13* (2025), 65704–65725. doi:10.1109/ACCESS.2025.3559298. 1

[BHJ*14] BONNEAU G.-P., HEGE H.-C., JOHNSON C. R., OLIVEIRA M. M., POTTER K., RHEINGANS P., SCHULTZ T.: Overview and state-of-the-art of uncertainty visualization. In *Scientific Visualization: Uncertainty, Multifield, Biomedical, and Scalable Visualization*, Hansen C. D., Chen M., Johnson C. R., Kaufman A. E., Hagen H., (Eds.). Springer, 2014, pp. 3–27. doi:10.1007/978-1-4471-6497-5_1. 6

[BMBH16] BÖRNER K., MALTESE A., BALLIET R. N., HEIMLICH J.: Investigating aspects of data visualization literacy using 20 information visualizations and 273 science museum visitors. *Information Visualization 15*, 3 (2016), 198–213. doi:10.1177/1473871615594652. 1

[BRBF14] BOY J., RENSINK R. A., BERTINI E., FEKETE J.-D.: A principled way of assessing visualization literacy. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (2014), 1963–1972. doi:10.1109/TVCG.2014.2346984. 3

[BS25] BENDECK A., STASKO J.: An empirical evaluation of the GPT-4 multimodal language model on visualization literacy tasks. *IEEE Transactions on Visualization and Computer Graphics 31*, 1 (2025), 1105–1115. doi:10.1109/TVCG.2024.3456155. 3

[BVY*23] BAE S. S., VANUKURU R., YANG R., GYORY P., ZHOU R., DO E. Y.-L., SZAFIR D. A.: Cultivating visualization literacy for children through curiosity and play. *IEEE Transactions on Visualization and Computer Graphics 29*, 1 (2023), 257–267. doi:10.1109/TVCG.2022.3209442. 3

[BZP*21] BISHOP F., ZAGERMANN J., PFEIL U., SANDERSON G., REITERER H., HINRICHS U.: Construct-A-Vis: Exploring the free-form visualization processes of children. *IEEE Transactions on Visualization and Computer Graphics 26*, 1 (2021), 451–460. doi:10.1109/TVCG.2019.2934804. 2

[CCB*22] CAMBA J. D., COMPANY P., BYRD V. L., SANTOS B. S., ALFORD G.: Identifying deception as a critical component of visualization literacy. *IEEE Computer Graphics and Applications 42*, 1 (2022), 116–122. doi:10.1109/MCG.2021.3132004. 1, 6

[CGD*24] CUI Y., GE L. W., DING Y., YANG F., HARRISON L., KAY M.: Adaptive assessment of visualization literacy. *IEEE Transactions on Visualization and Computer Graphics 30*, 1 (2024), 628–637. doi:10.1109/TVCG.2023.3327165. 2, 3, 4

[Cha23] CHANG H.-Y.: Scaffolding online peer critiquing to develop students' epistemic criteria of data visualization. *Computers & Education 203* (2023), 104863. doi:10.1016/J.COMPEDU.2023.104863. 6

[CRA*18] CHEVALIER F., RICHE N. H., ALPER B., PLAISANT C., BOY J., ELMQVIST N.: Observations and reflections on visualization literacy in elementary school. *IEEE Computer Graphics and Applications 38*, 3 (2018), 21–29. doi:10.1109/MCG.2018.032421650. 3

[Cro51] CRONBACH L. J.: Coefficient alpha and the internal structure of tests. *Psychometrika 16*, 3 (1951), 297–334. doi:10.1007/BF02310555. 6

[CST21] COHEN R. J., SCHNEIDER W. J., TOBIN R.: *Psychological Testing and Assessment: An Introduction to Tests and Measurement*, tenth ed. McGraw-Hill, 2021. 4

[CWW*26] CHANG M., WANG Y., WANG H. W., ZHOU Y., BULLING A., BEARFIELD C. X.: Tell me without telling me: Two-way prediction of visualization literacy and visual attention. *IEEE Transactions on Visualization and Computer Graphics 32*, 1 (2026), 1405–1415. doi:10.1109/TVCG.2025.3634815. 3

[DAMW25] DU M., AMOR R., MA K.-L., WÜNSCHE B. C.: Data visualization for improving financial literacy: A systematic review. *Visual Informatics* (2025). In Press. doi:10.1016/j.visinf.2025.100272. 1

[DC20] DONOHOE D., COSTELLO E.: Data visualisation literacy in higher education: An exploratory study of understanding of a learning dashboard tool. *International Journal of Emerging Technologies in Learning 15*, 17 (2020), 115–126. doi:10.3991/ijet.v15i17.15041. 3

[DC26] DONG L., CRISAN A.: Probing the visualization literacy of vision language models: The good, the bad, and the ugly. *IEEE Transactions on Visualization and Computer Graphics 32*, 1 (2026), 1175–1185. doi:10.1109/TVCG.2025.3634791. 3

[DeM10] DEMARS C.: *Item Response Theory: Understanding Statistics Measurement*, first ed. Oxford University Press, 2010. doi:10.1093/acprof:oso/9780195377033.001.0001. 3, 5, 6

[DTAW26] DO P. P., TANG K., AI K., WANG C.: SVLAT: Scientific visualization literacy assessment test. *arXiv preprint arXiv:2603.19000* (2026). doi:10.48550/2603.19000. 6

[DTM26] DAS A. K., TARUN M., MUELLER K.: Charts-of-Thought: Enhancing LLM visualization literacy through structured data extraction. *IEEE Transactions on Visualization and Computer Graphics 32*, 1 (2026), 1427–1437. doi:10.1109/TVCG.2025.3634813. 3

[ER00] EMBRETSON S. E., REISE S. P.: *Item Response Theory for Psychologists*, first ed. Lawrence Erlbaum Associates, 2000. doi:10.4324/9781410605269. 6

[FDL20] FIRAT E. E., DENISOVA A., LARAMEE R. S.: Treemap literacy: A classroom-based investigation. In *Proceedings of Eurographics (Education Papers)* (2020), pp. 29–38. doi:10.2312/EGED.20201032. 2

[FDWL22] FIRAT E. E., DENISOVA A., WILSON M. L., LARAMEE R. S.: P-Lite: A study of parallel coordinate plot literacy. *Visual Informatics 6*, 3 (2022). doi:10.1016/J.VISINF.2022.05.002. 2

[FIB*19] FUCHS J., ISENBERG P., BEZERIANOS A., MILLER M., KEIM D. A.: EduClust - a visualization application for teaching clustering algorithms. In *Proceedings of Eurographics (Education Papers)* (2019), pp. 9–16. doi:10.2312/EGED.20191023. 2

[FJDL08] FELIZ D. A., JACKSON C. D., DRURY F., LAIDLAW D. H.: Using visual design experts in critique-based evaluation of 2D vector visualization methods. *IEEE Transactions on Visualization and Computer Graphics 14*, 4 (2008), 877–884. doi:10.1109/TVCG.2008.29. 6

[FJL22] FIRAT E. E., JOSHI A., LARAMEE R. S.: Interactive visualization literacy: The state-of-the-art. *Information Visualization 21*, 3 (2022), 285–310. doi:10.1177/14738716221081831. 1, 8

[GCK23] GE L. W., CUI Y., KAY M.: CALVI: Critical thinking assessment for literacy in visualizations. In *Proceedings of ACM Conference on Human Factors in Computing Systems* (2023), pp. 815:1–815:18. doi:10.1145/3544548.3581406. 2, 3, 4, 5

[Goo05] GOOCH B.: Illustrative scientific visualization framework. In *Proceedings of Eurographics Workshop on Computational Aesthetics in Graphics, Visualization and Imaging* (2005), pp. 57–65. 4

[HJ93] HAMBLETON R. K., JONES R. W.: An NCME instructional module on comparison of classical test theory and item response theory and their applications to test development. *Educational Measurement: Issues and Practice 12*, 3 (1993), 38–47. doi:10.1111/j.1745-3992.1993.tb00543.x. 6

[HNGC21] HUYNH E., NYHOUT A., GANEA P., CHEVALIER F.: Designing narrative-focused role-playing games for visualization literacy in young children. *IEEE Transactions on Visualization and Computer Graphics 27*, 2 (2021), 924–934. doi:10.1109/TVCG.2020.3030464. 2

[HSFM25] HONG J., SETO C., FAN A., MACIEJEWSKI R.: 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 (2025), 7004–7018. doi:10.1109/TVCG.2025.3536358. 3

[HSR91] HAMBLETON R. K., SWAMINATHAN H., ROGERS H. J.: *Fundamentals of Item Response Theory (Measurement Methods for the Social Science)*, first ed. SAGE, 1991. 3

[KMK19] KREKHOV A., MICHALSKI M., KRÜGER J. H.: Integrating visualization literacy into computer graphics education using the example of dear data. In *Proceedings of Eurographics (Education Papers)* (2019), pp. 1–8. doi:10.2312/EGED.20191022. 3

[Law75] LAWSHE C. H.: A quantitative approach to content validity. *Personnel Psychology 28*, 4 (1975), 563–575. doi:10.1111/j.1744-6570.1975.tb01393.x. 5

[LBLS15] LAHA B., BOWMAN D. A., LAIDLAW D. H., SOCHA J. J.: A classification of user tasks in visual analysis of volume data. In *Proceedings of IEEE Scientific Visualization Conference* (2015), pp. 1–8. 4

[LKH*15] LEE S., KIM S.-H., HUNG Y.-H., LAM H., KANG Y.-A., YI J. S.: How do people make sense of unfamiliar visualizations?: A grounded model of novice's information visualization sensemaking. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (2015), 499–508. doi:10.1109/TVCG.2015.2467195. 1

[LKK17] LEE S., KIM S.-H., KWON B. C.: VLAT: Development of a visualization literacy assessment test. *IEEE Transactions on Visualization and Computer Graphics 23*, 1 (2017), 551–560. doi:10.1109/TVCG.2016.2598920. 2, 3, 4, 5

[LMPL24] LI Z., MIAO H., PASCUCCI V., LIU S.: Visualization literacy of multimodal large language models: A comparative study. *arXiv preprint arXiv:2407.10996* (2024). doi:10.48550/arXiv.2407.10996. 3

[McD13] MCDONALD R. P.: *Test Theory: A Unified Treatment*, first ed. Psychology Press, 2013. doi:10.4324/9781410601087. 6

[MCL*25] MATABOS M., COTTAIS P., LEROUX R., CENATIEMPO Y., GASNE-DESTAVILLE C., ROULLET N., SARRAZIN J., TOUROLLE J., BORREMANS C.: Deep sea spy: An online citizen science annotation platform for science and ocean literacy. *Ecological Informatics 86* (2025), 103065. doi:10.1016/j.ecoinf.2025.103065. 7

[MGS*22] MEUSCHKE M., GARRISON L. A., SMIT N. N., BACH B., MITTENENTZWEI S., WEISS V., BRUCKNER S., LAWONN K., PREIM B.: Narrative medical visualization to communicate disease data. *Computers & Graphics 107* (2022), 144–157. doi:10.1016/j.cag.2022.07.017. 7

[MLF*12] MA K.-L., LIAO I., FRAZIER J., HAUSER H., KOSTIS H.-N.: Scientific storytelling using visualization. *IEEE Computer Graphics and Applications 32*, 1 (2012), 12–19. doi:10.1109/MCG.2012.24. 7

[MTD*10] MEYER J., THOMAS J. J., DIEHL S., FISHER B. D., KEIM D. A.: From visualization to visually enabled reasoning. In *Scientific Visualization: Advanced Concepts*, Hagen H., (Ed.). Springer, 2010, pp. 227–245. doi:10.4230/DFU.SCIVIZ.2010.227. 6

[NZM*24] NOBRE C., ZHU K., MÖRTH E., PFISTER H., BEYER J.: Reading between the pixels: Investigating the barriers to visualization literacy. In *Proceedings of ACM Conference on Human Factors in Computing Systems* (2024), pp. 197:1–197:17. doi:10.1145/3613904.3642760. 1

[PFLJ22] PENG I., FIRAT E. E., LARAMEE R. S., JOSHI A.: Evaluating Bloom's taxonomy-based learning modules for parallel coordinates literacy. In *Proceedings of Eurographics (Education Papers)* (2022), pp. 21–29. doi:10.2312/eged.20221042. 1, 6

[PO23] PANDEY S., OTTLEY A.: Mini-VLAT: A short and effective measure of visualization literacy. *Computer Graphics Forum 42*, 3 (2023), 1–11. doi:10.1111/cgf.14809. 2, 3, 4, 5

[PO25] PANDEY S., OTTLEY A.: Benchmarking visual language models on standardized visualization literacy tests. *Computer Graphics Forum 44*, 3 (2025), e70137. doi:10.1111/cgf.70137. 3

[PZC*24] PAN H., ZHANG Q., CARAGEA C., DRAGUT E. C., LATECKI L. J.: FlowLearn: Evaluating large vision-language models on flowchart understanding. In *Proceedings of European Conference on Artificial Intelligence* (2024), pp. 73–80. doi:10.3233/FAIA240473. 3

[QR22] QUADRI G. J., ROSEN P.: A survey of perception-based visualization studies by task. *IEEE Transactions on Visualization and Computer Graphics 28*, 12 (2022), 5026–5048. doi:10.1109/TVCG.2021.3098240. 6

[RM15] RUCHIKACHORN P., MUELLER K.: Learning visualizations by analogy: Promoting visual literacy through visualization morphing. *IEEE Transactions on Visualization and Computer Graphics 21*, 9 (2015), 1028–1044. doi:10.1109/TVCG.2015.2413786. 1

[Rob07] ROBERTS J. C.: State of the art: Coordinated & multiple views in exploratory visualization. In *Proceedings of International Conference on Coordinated and Multiple Views in Exploratory Visualization* (2007), pp. 61–71. doi:10.1109/CMV.2007.20. 7

[Tel14] TELEA A. C.: *Data Visualization: Principles and Practice*, second ed. CRC Press, 2014. doi:10.1201/b17217. 4

[TM04] TORY M., MÖLLER T.: Rethinking visualization: A high-level taxonomy. In *Proceedings of IEEE Symposium on Information Visualization* (2004), pp. 151–158. doi:10.1109/INFVIS.2004.59. 2, 4

[VBH*25] VARONA M., BONILLA K., HEDAYATI M., JOSHI A., KAY M., HARRISON L., NOBRE C.: The state of the art in visualization literacy. *arXiv preprint arXiv:2509.01018* (2025). doi:10.48550/arXiv.2509.01018. 1, 2, 6

[VDLE25] VALENTIM M., DHANOA V., LEÓN G. M., ELMQVIST N.: The plot thickens: Quantitative part-by-part exploration of MLLM visualization literacy. *arXiv preprint arXiv:2504.02217* (2025). doi:10.48550/arXiv.2504.02217. 3

[WHS*23] WANG Y., HOU Z., SHEN L., WU T., WANG J., HUANG H., ZHANG H., ZHANG D.: Towards natural language-based visualization authoring. *IEEE Transactions on Visualization and Computer Graphics 29*, 1 (2023), 1222–1232. doi:10.1109/TVCG.2022.3209357. 6

[YLLW24] YE H., LI C., LI Y., WANG C.: InvVis: Large-scale data embedding for invertible visualization. *IEEE Transactions on Visualization and Computer Graphics 30*, 1 (2024), 1139–1149. doi:10.1109/TVCG.2023.3326597. 3