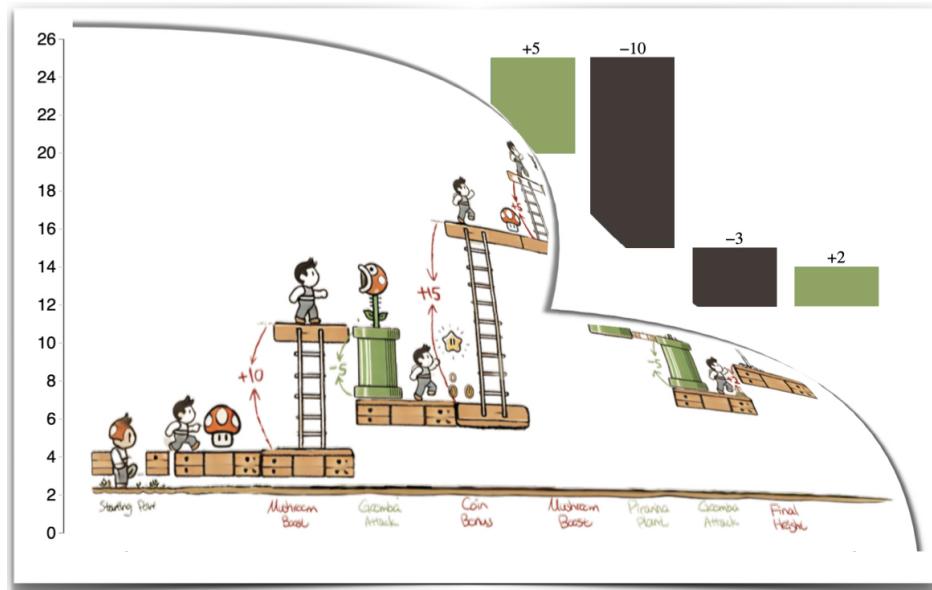


# From Reality to Recognition: Evaluating Visualization Analogies for Novice Chart Comprehension

O. Huang<sup>ID</sup>, P.Y.K. Lee<sup>ID</sup>, and C. Nobre<sup>ID</sup>

Department of Computer Science, University of Toronto, Canada



**Figure 1:** Example use of a visualization analogy (top) providing an accessible entry point to understanding the underlying data visualization, a waterfall chart (bottom). The Mario analogy preserves the main features of a waterfall chart (common baseline in adjacent bars) to teach novice viewers how to read the chart.

## Abstract

Novice learners often have difficulty learning new visualization types because they tend to interpret novel visualizations through the mental models of simpler charts they have previously encountered. Traditional visualization teaching methods, which usually rely on directly translating conceptual aspects of data into concrete data visualizations, often fail to attend to the needs of novice learners navigating this tension. To address this, we conducted an empirical exploration of how analogies can be used to help novices with chart comprehension. We introduced visualization analogies: visualizations that map data structures to real-world contexts to facilitate an intuitive understanding of novel chart types. We evaluated this pedagogical technique using a within-subject study ( $N=128$ ) where we taught 8 chart types using visualization analogies. Our findings show that visualization analogies improve visual analysis skills and help learners transfer their understanding to actual charts. They effectively introduce visual embellishments, cater to diverse learning preferences, and are preferred by novice learners over traditional chart visualizations. This study offers empirical insights and open-source tools to advance visualization education through analogical reasoning.

## CCS Concepts

- Human-centered computing → Visualization design and evaluation methods; Empirical studies in visualization;

## 1. Introduction

As visualizations increasingly mediate our understanding of everything from scientific discoveries to daily news, the ability to interpret them effectively has become a fundamental skill [NZM\*24]. However, while most people are familiar with basic charts like bar and line graphs, learning to work with more advanced visualizations remains challenging [BKR\*23]. A key barrier is that new visualization types often appear arbitrary and disconnected from learners' existing knowledge [RM15].

One promising but understudied approach to bridging this gap is the use of analogies - mapping new concepts to familiar ones to facilitate learning. While analogies have proven effective in STEM education [HT06] and computer science teaching [HGA\*16], their application to visualization education remains mostly unexplored.

In this work, we investigate the effectiveness of visualization analogies: visualizations that map data structures (e.g., histogram bins) to familiar real-world objects (e.g., stacks of exam papers). These analogies aim to make abstract visualization concepts more concrete by grounding them in learners' existing experiences. The aim of our work is to provide an empirical account of whether visualization analogies can be used as an intuitive pedagogical tool in visualization education. Specifically, we conducted a survey with 128 visualization novices [BLC\*23] on their performance, preference, and learning of one of eight new chart types ranging from the simple (e.g., bar charts) to the complex (e.g., sunburst charts).

We focused on **5 key dimensions** to assess the role of visualization analogies in learning: the **effectiveness** of visualization analogies on visual analysis tasks; the ability of visualization analogies to support **learning transfer** to baseline chart visualizations; the influence of **visual embellishments** in visualization analogies on user attention and potential misinterpretation of charts [BMG\*10, AHL\*21]; the **generalizability** of visualization analogies to different learning styles (as captured through participant VARK preferences [ACCL24]); and the self-reported **preferences** of participants towards visualization analogies. Our results show that visualization analogies are effective in helping visualization novices in identifying key data encodings while maintaining clarity and interpretability.

Overall, the contributions of our work are two-fold:

**1. Empirical Evaluation of Visualization Analogies.** We conducted an empirical study with 128 participants to investigate how mapping chart structures to familiar, real-world objects affects novices' understanding of data. Our findings show that visualization analogies significantly improve comprehension compared to standard chart representations.

**2. Open Sourcing of Visualization Analogies** We develop a set of visualization analogies, refining them through feedback from visualization experts. The final set of visualization analogies are made publicly available through an open-source repository ([visanalogy.github.io](https://github.com/visanalogy/visanalogy.github.io)), enabling educators and researchers to adapt and extend these techniques.

## 2. Related Work

In this section, we survey two key research areas which inform our approach: efforts to make visualization education more engaging and accessible, and the use of analogies to aid data comprehension.

### 2.1. Visualization Education for Novices

Devising pedagogical strategies to help novice learners of visualization techniques adapt to unfamiliar visualization types can be challenging. Chang et al. [CCT24] observed that students often chose familiar visualizations for decision-making tasks rather than those best suited to the task. This aligns with Grammel et al.'s findings that novices rely on familiar visualization types when addressing new data visualization needs [GTS10]. At the educator level, this means that visualizations that seem intuitive to work with can be highly dependent on the learner's prior experiences. For example, Chevalier et al. [CRA\*18] identified a discrepancy among elementary school educators, who see visualizations as intuitive learning tools, yet feel their students are not skilled enough to interpret and create them.

Researchers have explored various approaches to overcome this barrier, including: Interactive tools and games (e.g., Roboviz [ALR23], Diagram Safari [GWL\*19]), hands-on activities with data physicalization [Per21], and narrative-based approaches [HNGC21]. These techniques emphasize models of learning that are exploratory, experiential, and curiosity-driven to heighten engagement over a traditional didactic learning model.

While these approaches have shown promise in engaging learners, they often require significant development effort and time investment from learners. Additionally, most focus on a limited range of visualization types, making them difficult to scale across the broad spectrum of charts that novices need to learn.

These limitations in existing approaches highlight the need for a teaching method that is both scalable and intuitive. Analogies—a well-established technique in other educational domains—offer a promising direction. By mapping visualization structures to familiar real-world concepts, we can potentially help novices grasp new chart types without the substantial onboarding overhead of games or physical tools. This potential synergy between analogical thinking and visualization education remains largely unexplored, despite its successful application in other learning contexts.

### 2.2. Analogies in Data Comprehension

Analogies have long been a popular pedagogical technique in other disciplines [HT06], but their use in visualization education has been limited. Studies have demonstrated how abstract numerical data can be made more accessible by re-expressing it through familiar objects and scenarios [CVG, HKN\*18]. While visual analogies [Spe10] have proven particularly effective for novices by enabling intuitive interpretation of complex data values, these approaches typically focus on single data points rather than complete datasets or visualization structures.

More promising developments have emerged in spatial contexts, where researchers have successfully used extended analogies to represent multiple data values. Kim et al. [KHA16] showed how analogies about area and distance enhance understanding of spatial data. Similarly, Riederer et al. [RHG18] demonstrated that comparing unfamiliar statistics to familiar object measurements—such as relating country-level statistics to city areas—significantly improved comprehension. Their empirical results confirmed that such perspective-based analogies make complex numerical information more relatable and meaningful. However, scaling these techniques to heterogeneous datasets while maintaining their intuitive nature remains challenging [HKN\*18].

Visual analogies [LSL96] have further advanced analogical learning by mapping abstract data onto physical forms of everyday objects, making data relationships more apparent [MPP01]. One study found that presenting health data as visual analogies led to better patient understanding compared to conventional line graphs [TGM\*20]. Another showed that bar chart analogies improved comprehension over traditional representations [Mor17]. While these studies reveal the potential of visual analogies for clarifying complex relationships, they remain limited to simple chart types and lack comprehensive empirical validation for building broader visualization literacy.

Recent advances in this space include *AnalogyMate* [CSZ\*24], which systematically links numerical values to familiar objects through a structured design space. Their framework organizes analogy creation strategies into categories like comparison, proportion, unitization, and accumulation—demonstrated through examples such as comparing bar heights to familiar objects. This work extends analogical representations beyond individual measurements, suggesting the potential for whole-dataset comprehension.

Our research builds upon these foundations while addressing a specific challenge in visualization education. Rather than focusing on individual numerical measurements, we develop visualization analogies that encompass entire datasets. By mapping visualization structures to real-world objects, we leverage people's existing knowledge to help them understand novel visualizations directly, instead of requiring comprehension through simpler chart types. Our work provides empirical evidence for the effectiveness of this approach in developing visual analysis skills.

### 3. Visualization Analogies Design

#### 3.1. Design Criteria

To ensure meaningful evaluations, we design visualization analogies that effectively support novice learners in understanding complex concepts. Our design objective is to guide learners to envision concepts in practical situations while preserving the analytical strength of the original charts. Below, we outline the design requirements for the analogies, emphasizing how they contribute to creating intuitive and accessible visualizations.

**Representation of Data Attributes:** Analogies should maintain the data attributes of the original dataset while presenting them in a more accessible and familiar real-world context [CSZ\*24]. This helps learners connect abstract chart representations to everyday scenarios, clearly communicating key relationships, trends, and patterns in the data.

**Visual Reconstruction:** Adapting from the concept of re-expression [CVG], the visual structure of an analogy should closely resemble the original chart. This requires careful attention to layout, shape, and data point representation. The primary aim is to maintain a strong visual correspondence with the original chart.

**Analysis Potential:** An effective analogy does more than map data into the shape of physical entities. It should maintain meaningful contexts that relate closely to learners' real-world experiences. The chosen scenarios should align with the data's characteristics and maintain conceptual relevance. This connection helps learners link the analogy context to the original data.

**Scalability:** Lastly, visualization analogies must support the same data complexity as the original chart [RPA\*24]. Achieving

this requires selecting meaningful and scalable objects as contexts. These choices should naturally accommodate more data points without overwhelming the learner. We chose physical objects that support proportional mapping to achieve this goal. This ensures that larger datasets can be represented intuitively.

#### 3.2. Chart Selection

We selected a diverse set of representative charts from the [Visual Vocabulary](#) to evaluate the effectiveness of analogies (Figure 2). Each chart type represents a unique data relationship, ensuring our analogies can scale across different datasets and analytical needs.

	Visualization Data Characteristics						Analogy Design Space															
	Data Relationships			Dimension	Data Scale		Analogy Strategies		Data-Binding													
	Deviation	Correlation	Ranking	Distribution	Temporal	Part-to-Whole	Magnitude	Flow	Univariate	Bivariate	Multivariate	Nominal	Ordinal	Interval	Ratio	Comparison	Unitization	Accumulation	Proportion	Length	Area	Time
Butterfly Chart	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Calendar Heatmap	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Bubble Chart	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Histogram	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Stacked Area Chart	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Sunburst Plot	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Bar Chart	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Waterfall Plot	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■

Figure 2: Chart Selection Framework for Analogy Design

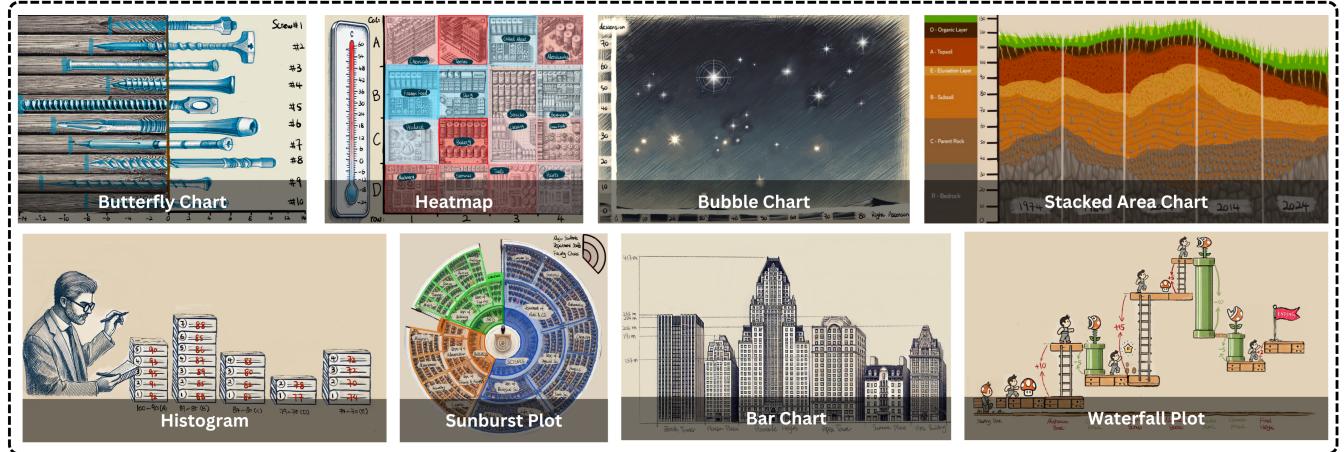
Dimensionality affects the complexity of data visualization. To expose learners to various visual complexities, we included charts for univariate, bivariate, and multivariate datasets [Den18], making our analogy framework adaptable to diverse data structures.

Data scales also shape data visualization and comprehension [BBG19]. We assessed visualization analogies across scales such as categorical distinctions, ranks, quantitative differences, and proportional relationships. Our chart selection aligns with the established design space for data analogies [CSZ\*24], focusing on comparisons, accumulations, and proportions. We increased chart adaptability using data-binding elements like length or area for quantitative data. Spatial data relationships (e.g. choropleth maps) were excluded as they naturally align with real-world scenarios [BVB\*13] and could dilute our unique analogical design approach, despite their commercial popularity [HHS20].

#### 3.3. Visualization Analogy Creation Process

**Challenges in analogy creation:** The creation of our visualization analogies (Figure 3) followed the design criteria outlined in Section 3.1, with the overarching goal of ensuring clarity, relevance, and scalability across various datasets. Choosing a suitable analogy for each chart type proved challenging. Each proposed analogy had to faithfully capture the chart's core analytical structure while remaining both conceptually engaging and capable of accommodating larger or more complex datasets. For example, an analogy might excel at explaining quantitative differences but struggle to depict proportional relationships. We adopted an iterative design process to address these trade-offs by systematically evaluating candidate analogies for mapping fidelity (i.e., preserving data attributes), contextual relevance (for novice learners), and scalability (for complex data). This approach ensured that our final designs upheld the design criteria outlined in Section 3.1, while still offering an intuitive learning experience.

**Iterative design and expert review:** Our iterative process involved initial brainstorming sessions followed by multiple refine-



**Figure 3:** Overview of the 8 selected visualization analogies for the study, chosen based on their alignment with chart characteristics, contextual relevance, and scalability across diverse datasets.

ment cycles to ensure each visualization analogy met our design requirements. The analogies were then presented to five established visualization researchers, all tenure-track faculty, for formal expert review. These reviewers used a rubric derived from the design criteria 3.1 to evaluate each analogy. Their feedback surfaced issues about the need to preserve proportional relationships, layout clarity, and potential visual clutter when scaled to larger datasets. Insights from the expert reviews were integrated into subsequent design iterations, leading to the final analogies in Figure 3.

We have open-sourced the final visualization analogies as a novel teaching technique for educational use and future research. They can be found at [visanalogy.github.io](https://visanalogy.github.io)

#### 4. Study

Building upon the selected visualization analogies, we conducted a multi-dimensional within-subjects study with 128 visualization novices to evaluate the effectiveness of analogies compared to unmodified charts. The study investigates their potential as an effective teaching technique for understanding and analyzing novel visualizations. In this context, we use the term **baseline** charts to refer to conventional visualizations without any visual embellishments, serving as a standard for comparison. Our goal was to evaluate the effectiveness of visualization analogies in the following dimensions:

- **RQ1 [Performance & Load]:** How do visualization analogies impact understanding, performance, and cognitive load compared to baseline charts?
- **RQ2 [Learning Transfer]:** Can visualization analogies aid in transferring knowledge on how to interpret similar baseline charts?
- **RQ3 [Visual Embellishments]:** How do visual embellishments in analogies affect users' attention to data encodings?
- **RQ4 [Generalizability]:** Are visualization analogies effective across different learning preferences?
- **RQ5 [Preferences]:** How do learner preferences affect the effectiveness of visualization analogies in educational contexts?

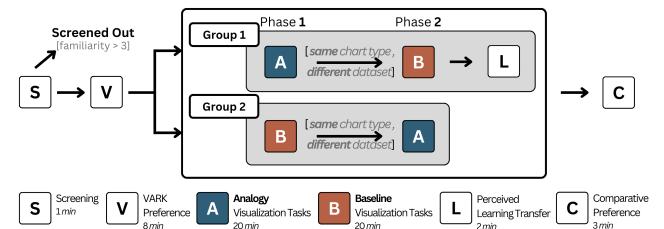
#### 4.1. Study Deployment

The study was conducted on Qualtrics. We implemented a think-aloud protocol to gather insights into participants' sentiments, cognitive processes, and decision-making strategies. This allowed participants to record up to two minutes of verbal input for each question. Voice input was captured using the Phonic AI API and integrated into the survey platform.

Additionally, participants completed reproduction tasks by sketching each chart's key characteristics to demonstrate understanding using a custom drawing interface that offered adjustable pen width and color selection. All study materials —including the Qualtrics questionnaires, sketching platform code, visualization tasks, and evaluation rubric— are publicly available on GitHub, which also contains our expert reviews, deployed visualization analogies, unprocessed think-aloud data, and analysis code for reproducibility.

#### 4.2. Procedure

An overview of the study procedure is displayed in Figure 4.



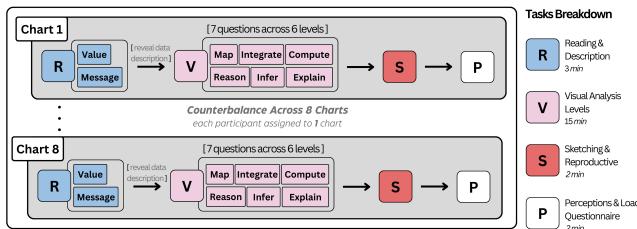
**Figure 4:** Study procedure overview: participant progression through screening, VARK questionnaire, division into 2 balanced groups, and finally preference questionnaire.

The study employed a within-subjects design, with participants evenly distributed into two groups to counterbalance the order of visualization exposure. Group 1 viewed visualization analogies first, followed by baseline charts, while Group 2 experienced the reverse order. The datasets used in each phase were designed to share iden-

tical contexts but not the same data, reducing the potential for learning effects from repeated exposure to identical datasets.

Initially, each participant was assigned to a specific chart type. They then self-rated their familiarity with the assigned chart on a scale from 1 to 5. To ensure a novice-level comprehension of the given data visualization, only the participants who rated their familiarity at 3 or below were allowed to proceed. To ensure balanced representation and reduce biases associated with specific chart types, the study distributed the selected participants evenly across the 8 chart types.

Figure 5 illustrates the procedure for each phase. Each phase began with participants viewing a visualization without any accompanying title or description. They were asked to interpret the chart's purpose and identify key data points based solely on its visual representation (**Reading & Description**). These questions were adapted from the framework established by Bateman et al. [BMG\*10] to determine whether visual embellishments might lead to misinterpretation. After that, the chart title and dataset description were provided, and participants proceeded to perform a series of **Visual Analysis** tasks. These tasks were organized in ascending analytic levels [LFM21], progressing through stages of Mapping, Integrating, Reasoning, Inferring, and Explaining.



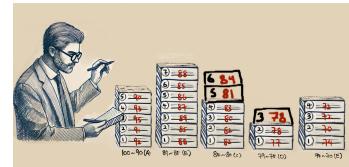
**Figure 5:** Breakdown of visualization tasks: reading & description, visual analysis, sketching, and post-task ratings.

Participants were then prompted with a **Sketching** task (Figure 6) where we aim to evaluate their comprehension and retention of the presented data. Participants ended the survey with a post-task questionnaire to evaluate their perceived learning and confidence levels. Additionally, the NASA-TLX was employed in the Perceptions & Load section to assess the cognitive load - explicitly focusing on mental demand, effort, and frustration encountered while interpreting each visualization.

For participants in Group 1 (i.e., for those who saw the analogy first and then the baseline), we included additional questions to assess their self-rated learning transfer to the baseline chart across interpretation, context, and data relationships. Finally, all participants compared the two visualization representations, evaluating them in terms of intuitiveness, engagement, and clarity.

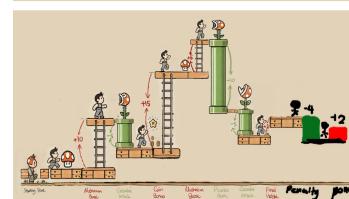
### 4.3. Participants

After screening for minimal familiarity with the tested chart types, we recruited 128 participants through **Prolific** (79 males and 49 females), aged 19–68 ( $M = 31.4$ ). The final sample had an average familiarity rating of 2.09 out of 5, reflecting a novice-level understanding of data visualization. They completed the VARK questionnaire to determine their learning preferences, focusing on their dominant modality [PM14]: Visual ( $n = 30$ ), Auditory ( $n = 28$ ),



### Histogram

**Q:** Sketch how the chart would change if **three** additional scores of **81, 84, and 78** were added.



### Waterfall Plot

**Q:** Sketch how the chart would look if the **next phases** include a **-4 penalty event** and **+2 bonus event**.

**Figure 6:** Example sketch questions showcasing visual recreations of a histogram chart (top) and a waterfall plot (bottom)

Reading/Writing ( $n = 31$ ), and Kinesthetic ( $n = 39$ ). The study received approval from our institute's ethics review board, and informed consent was obtained from all participants before the study. We paid participants \$15 to participate in the study. On average, each participant took 1 hour to complete the study.

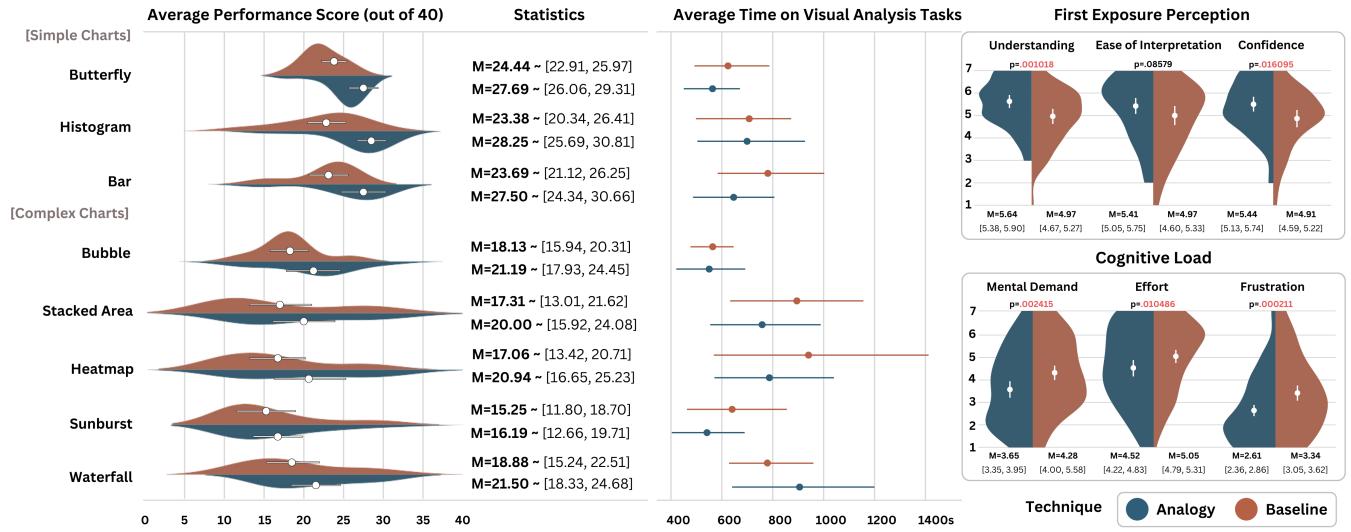
### 4.4. Data Analysis

Each visualization task was given equal weight. We assessed participants' responses using a detailed rubric focused on the completeness and justification of their reasoning. We conducted an intra-rater reliability test, where the first author re-evaluated a randomly selected subset of responses after a 3-week interval. To confirm the consistency of the rubric, we measured the agreement between the initial and re-evaluated scores using Cohen's Kappa. Each question was scored out of 5, resulting in a maximum possible score of 40 points for the visualization task, and 10 points for description tasks. We also recorded the time spent on each question to analyze cognitive engagement and compare the efficiency of visualization analogies to baseline charts.

For performance scores, we employed a mixed-effects model to control for fixed effects, such as visualization technique, order of exposure, chart difficulty, and random effects associated with individual participant differences. For statistical analysis, we employed a one-way ANOVA for normally distributed data with equal variances and applied the Kruskal-Wallis H test to non-normal data. We conducted pairwise comparisons using Welch's t-tests for normally distributed data and Mann-Whitney U tests for non-normal data.

## 5. Results

In this section we report on the effectiveness of visualization analogies across the 5 measured dimensions. First, we evaluate how these analogies impact performance in visual analysis tasks compared to baselines [RQ1]. Subsequently, we investigate their role as a teaching tool by exploring the learning transfer effects from analogy charts to baselines [RQ2]. We also examine how visual embellishments in analogies influence users' attention to data encodings [RQ3]. Furthermore, we assess the generalizability of visualization analogies across VARK learning preferences [RQ4]. Finally, we capture learners' perceptions and analyze their engagement and comprehension preferences regarding these visualization methods [RQ5].



**Figure 7:** An overview of visual analysis tasks results (from left to right): the average performance score for each technique; the average time spent on each technique; self-reported evaluations on perceptions of understanding, ease of interpretation, and confidence from initial exposure; and cognitive load evaluations of mental demand, effort, and frustration based on the NASA-TLX assessment.

### 5.1. RQ1 [Performance & Load]: Effectiveness of Visualization Analogies

We investigated the average performance score across all visualization tasks (Figure 7). The result indicates **visualization analogies significantly enhance users' performance in visual analysis tasks compared to baselines**. Participants who used visualization analogies achieved higher performance scores ( $M=22.91$ ,  $SD=7.30$ ) than those who used baselines ( $M=19.77$ ,  $SD=6.65$ ). A paired t-test revealed a statistically significant difference ( $p=.0004$ ).

The selection of charts used in this study include charts across a spectrum of complexities. In order to investigate the effect of analogies as a function of complexity, we performed a cluster analysis on the performance scores to categorize the charts based on difficulty levels. The optimal clustering resulted in two groups, allowing us to separate the charts into **simple** and **complex** categories.

Building on these categorizations, we found that **the effectiveness of analogies can vary based on chart complexity**. For simple charts, participants using analogy charts had a superior mean performance score of  $27.81$  ( $SD=4.66$ ), compared to  $23.83$  ( $SD=4.54$ ) for baselines. This difference was highly significant ( $p<.0001$ ). For complex charts, the analogy condition yielded a mean score of  $19.96$  ( $SD=7.03$ ), while the baseline condition had a mean of  $17.33$  ( $SD=6.54$ ). While the difference in performance between the two conditions was not as high for complex charts, it was still a significant difference ( $p = .015$ ).

Our mixed-effects model further confirmed these findings. The fixed factors included visualization technique, chart difficulty, average time spent, and order of exposure. The model showed that analogy charts led to an average of 11% higher performance scores compared to baselines. Simpler charts outperformed complex ones by 19%. Notably, the average time spent on tasks was not a signif-

icant factor, indicating that participants could understand analogy charts quickly and still demonstrate a high level of comprehension.

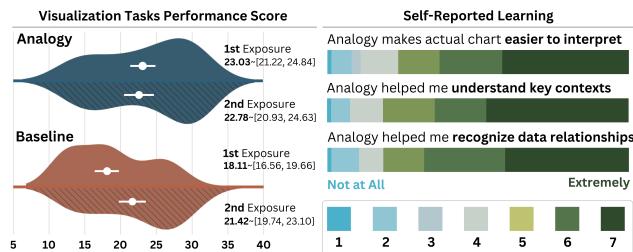
In regards to the self-reported dimensions, **visualization analogies led to significantly higher comprehension and confidence compared to baselines**. Participants felt they understood the charts better when first exposed to analogy charts. Comprehension ratings were significantly higher ( $p=0.001$ ). Confidence ratings were also significantly higher ( $p=0.016$ ) for analogy charts ( $M=5.41$ ) than baselines ( $M=4.97$ ). Ratings for ease of interpretation did not reach statistical significance ( $p=0.086$ ), suggesting that both methods maintain a certain level of cognitive friction necessary for extracting insights.

To further explore the overall load introduced by analogies, we examined cognitive load across three dimensions: mental demand, effort, and frustration. The result indicates **visualization analogies reduce cognitive load across all three measured dimensions**. Specifically, participants experienced lower mental demand when using analogy charts ( $p=0.002$ ). They reported exerting less effort ( $p=0.010$ ) and feeling less frustration ( $p<0.001$ ) compared to baseline charts. These measures suggest that visualization analogies improve performance and make analysis tasks less cognitively demanding.

### 5.2. RQ2 [Learning Transfer]: Enhancing Comprehension from Analogy to Baseline

In order to establish the ability of analogies to promote learning transfer, we first examined the performance scores for analogy charts across both groups with varying exposure orders, finding that **analogies are effective independently, regardless of prior exposure to baselines**. Participants performed similarly on analogy charts, regardless of whether or not they had previously seen the baseline charts. As shown in Figure 8, participants who were first exposed to analogy charts had a mean performance score of

23.03 ( $SD=7.22$ ), while those who viewed analogy charts after the baselines had a mean score of 22.78 ( $SD=7.44$ ). This difference was insignificant ( $p=0.929$ ).



**Figure 8:** Overview of learning transfer result. (left) performance score by first and second exposure to each technique; (right) self-rated perceived learning transfer on a scale of 1 to 7

Given that analogy charts are not affected by potential learning effects from the baseline charts, we examined how prior exposure to analogies influenced performance to understand the impact on learning transfer. Our result indicates **visualization analogies promote learning transfer to baseline charts**. Participants who saw analogy charts before baselines performed significantly better on the baselines. The mean performance score for baselines after analogy exposure was 21.42 ( $SD=6.73$ ), compared to 18.11 ( $SD=6.19$ ) for those who saw baselines first. This improvement was statistically significant ( $p=0.007$ ). This indicates that analogy charts enhance understanding and aid in interpreting baseline charts.

Aligning with the performance score, **participants reported high levels of perceived learning transfer from analogy to baselines**. A majority of participants felt that analogy charts helped them interpret baselines more efficiently. Specifically, 76.56% reported that analogies made it easier to interpret the baseline. Additionally, 81.25% said analogies helped them understand key contexts of the baseline, and 81.25% felt analogies helped them recognize data relationships more clearly.

### 5.3. RQ3 [User Attention]: Understanding the Effects of Visual Embellishment

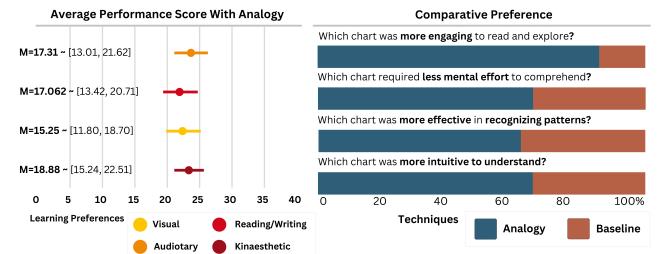
In order to investigate whether visual embellishments (an intrinsic part of analogies) detracted from viewer attention and comprehension of the charts, we investigated users' ability to interpret the assigned chart before being given any information on the underlying dataset. We call this metric 'interpretation accuracy'. Our findings indicate that **visualization analogies do not compromise users' attention to data encoding**. Participants' descriptions of the analogy charts were as accurate as those of the baseline charts. Even without the chart title or dataset context, they could identify the core idea using both visualization techniques. The mean description scores for the analogy ( $M=9.01$ ) and baseline ( $M=8.36$ ) were similar, and a pairwise t-test showed no significant difference between the techniques ( $p=0.872$ ).

Further analysis examined the impact of chart complexity on interpretation accuracy. **For complex charts, visualization analogies significantly enhanced interpretation accuracy compared to baselines**. Participants scored higher using visualization analogies ( $M=8.51$ ,  $SD=3.05$ ) than with baselines ( $M=7.16$ ,  $SD=2.95$ ) when describing charts with more challenging interpretations. This

statistically significant difference ( $p=0.047$ ) suggests that the additional visual elements provided by the analogies facilitated better interpretation of complex charts. In contrast, for simpler charts, the average scores were comparable between analogy ( $M=9.56$ ,  $SD=1.79$ ) and baseline ( $M=9.56$ ,  $SD=1.79$ ). The pairwise t-test showed no significant difference ( $p=0.087$ ), indicating that both techniques were equally effective in conveying key information in simpler visualizations.

### 5.4. RQ4 [Generalizability]: Effectiveness Across Diverse Learning Preferences

We conducted an ANOVA F-test to compare the performance scores grouped by different learning preferences. We found that **visualization analogies are generalizable teaching techniques, effective across various learning preferences**. There was no significant difference in performance scores among the VARK learning preference categories when using analogy charts ( $p=0.826$ ). Mean scores were similar (Figure 9): Auditory (22.63,  $SD=7.27$ ), Kinesthetic (21.87,  $SD=7.08$ ), Reading/Writing (20.16,  $SD=7.06$ ), and Visual (20.72,  $SD=7.12$ ). These results indicate that visualization analogies are equally effective regardless of individual learning preferences.



**Figure 9:** (left) Performance score with analogies across all learning preferences. (right) Comparative preference results: each bar displays the percentage of participants favoring analogies versus baseline charts for each of the displayed questions.

### 5.5. RQ5 [Preference]: Engagement and Comprehension Preferences for Visualization Analogies

We examined participants' perceptions of visualization analogies in terms of engagement and found that **visualization analogies are perceived as more engaging**. In terms of engagement, 83.59% of participants found visualization analogies more engaging than standard charts, with one noting that "it feels like playing a game". This playful quality appears to resonate across diverse learning preferences. In regards to comprehension, 70.31% reported that analogies required less mental effort to interpret, citing clearer visuals that streamline data processing and reduce cognitive load. Additionally, 63.28% of participants also highlighted that analogies aided in pattern recognition, further indicating their potential to clarify complex data through intuitive design elements.

## 6. Discussion

Our study demonstrates that visualization analogies can effectively enhance novice performance in visual analysis tasks while facilitating learning transfer to baseline charts. We highlight four main points for discussion.

## Generalizability Across Learning Styles

Visualization analogies maintained their effectiveness across different VARK learning preferences, as evidenced by consistent performance scores regardless of self-reported learning styles. This suggests visualization analogies can have broad appeal in classroom environments, supporting diverse learners without privileging any particular learning approach. This universality is especially valuable in visualization education, where students often possess vastly different prior experiences and levels of familiarity with visualizations. Our results indicate that visualization analogies can leverage what appears to be the inherently intuitive nature of visualizations as teaching tools [CRA\*18] while actually translating this into developed visual literacy skills across the learning spectrum. We also speculate that analogies can be applied differently depending on learning style to target a learner's particular preferences and bolster the efficacy of visualization analogies. For example, kinesthetic learners may prefer visualization analogies that explicitly incorporate an analogy to real-world movement (e.g., Mario analogy for a waterfall chart), so they can more comfortably map from their preferred learning style to an abstract concept.

## Strengths and Limitations in Skill Development

Visualization analogies proved most effective when used for descriptive or diagnostic reasoning tasks, making them particularly valuable for developing fundamental chart reading and interpretation skills. For instance, participants excelled at chart reading tasks (e.g., *"How many storage sections have a temperature above 30°C?"* in the heatmap) but struggled with questions requiring extrapolation beyond the visible analogy (e.g., *"If aisle C's temperature increased to match aisle A, how might this change the overall temperature distribution in the storage area and what items may no longer be suitable in aisle C?"*). This suggests that while analogies excel at building foundational understanding, additional development may be needed to support more complex analytical skills. One promising direction is the potential use of animated visualization analogies to help learners develop extrapolation abilities with datasets.

## The Double-Edged Nature of Visual Embellishments

While our visualization analogies generally avoided introducing unnecessary chart junk [BMG\*10], we found that the effectiveness of visual embellishments varies significantly with learners' prior experiences. Depending on a learner's familiarity with the analogy context, there is a risk that visual elements intended to support the analogy might be misinterpreted as meaningful data channels. For example, in our Mario Game visualization analogy of the waterfall plot, some participants interpreted the chart as being about *"how to get a better score in Mario,"* confusing the contextual elements with the visualization's actual purpose.

Even subtle design choices can generate confusion. In our histogram analogy, the decision to stack exam booklets in descending order—despite histograms having no inherent ordering requirement—led some participants to incorrectly conclude that the visualization represented *"the sorted ranking of marks for students [...] organized from the students who scored the highest marks to the lowest marks."* These observations highlight how visual embellishments can significantly impact interpretation based on viewers' prior experiences. Success requires carefully balancing the elements needed to create meaningful analogical connections while avoiding superfluous details that might confuse novice learners.

## Challenges with Complex Visualizations

We observed particular challenges when applying analogies to multivariate charts like sunburst diagrams. While our goal was to help novices understand novel visualizations without referencing simpler chart types, participants often reverted to interpreting complex visualizations in terms of more familiar ones when faced with completely unfamiliar structures. For instance, some participants described the sunburst chart as a *"pie chart [that] has colours to better represent the sections visually"* likely due to visual similarities with familiar pie charts. Similarly, participants often described histograms in terms of bar charts, noting that the histogram *"is a bar graph with no labeled axis, but from left to right, the bars are starting from down changing to upwards."*

Despite these challenges, visualization analogies demonstrated clear benefits, particularly when chart titles and labels were omitted. Participants showed better immediate comprehension with analogies compared to baseline charts. For example, one participant could describe how a stacked area chart analogy *"probably [showed] how much of this material increase or decrease over the years and how the ground is later shaped"* while being unable to interpret the baseline version (*"I can't tell"*). This suggests that even when observers are completely unfamiliar with a visualization type, analogies can provide sufficient contextual information to extract meaningful information without external assistance.

These findings indicate that while visualization analogies offer significant promise for visualization education, their successful implementation requires careful attention to: (1) balancing analogical elements with data clarity; (2) supporting the progression from basic comprehension to complex analysis; (3) selecting appropriate and accessible real-world contexts; (4) managing visual complexity based on learner expertise; and (5) providing additional support for complex multivariate visualizations.

Our work provides empirical evidence for the effectiveness of visualization analogies while highlighting important considerations for their practical implementation in visualization education.

## 7. Conclusion

This paper introduced visualization analogies as a novel teaching technique that maps datasets to real-life objects, sharing visual and conceptual characteristics with the data. We present an initial exploration of their effectiveness as tools for enhancing comprehension in visual analysis tasks. The goal of making abstract information more intuitive and relatable is to improve user understanding when learning new chart types.

Our empirical investigation revealed that visualization analogies significantly improve user performance, comprehension, and confidence while being more engaging and less cognitively demanding than traditional charts. The visual embellishments in the analogies were particularly beneficial for navigating complex charts, did not mislead users, and were generalizable across diverse learning styles. However, learners still require support when using visualization analogies for complex multivariate charts, such as sunburst charts. Future work should explore their long-term impact on retention and explore self-supervision and friction-reduction methods to enhance understanding of complex analogies. We hope this work will inspire further research into the development and application of visualization analogies, enhancing their role as standalone and complementary tools in educational contexts.

## References

- [ACCL24] AUNHABUNDIT N., CHENG CHIEH LU A.: Visualizing success: Amplifying learning outcomes across vark learning preferences through visual notes. *Science, Technology, and Social Sciences Procedia* 2024, 6 (Jul. 2024), ICCM03. 2
- [AHL\*21] ANDRY T., HURTER C., LAMBOTTE F., FASTREZ P., TELEA A.: Interpreting the effect of embellishment on chart visualizations. In *Proceedings of CHI Conference on Human Factors in Computing Systems* (2021), ACM. [doi:10.1145/3411764.3445739](https://doi.org/10.1145/3411764.3445739). 2
- [ALAR23] ADAR E., LEE-ROBBINS E.: Roboviz: A Game-Centered Project for Information Visualization Education . *IEEE Transactions on Visualization & Computer Graphics* (Jan. 2023). [doi:10.1109/TVCG.2022.3209402](https://doi.org/10.1109/TVCG.2022.3209402). 2
- [BBG19] BORNER K., BUECKLE A., GINDA M.: Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments. *Proceedings of the National Academy of Sciences* 116 (2019). [doi:10.1073/pnas.1807180116](https://doi.org/10.1073/pnas.1807180116). 3
- [BKR\*23] BACH B., KECK M., RAJABIYAZDI F., LOSEV T., MEIRELLES I., DYKES J., LARAMEE R. S., ALKADI M., STOIBER C., HURON S.: Challenges and opportunities in data visualization education: A call to action. *IEEE Transactions on Visualization & Computer Graphics* (2023). [doi:10.1109/TVCG.2023.3327378](https://doi.org/10.1109/TVCG.2023.3327378). 2
- [BLC\*23] BURNS A., LEE C., CHAWLA R., PECK E., MAHYAR N.: Who do we mean when we talk about visualization novices? In *Proceedings of CHI Conference on Human Factors in Computing Systems* (2023), Association for Computing Machinery, pp. 1–16. [doi:10.1145/3544548.3581524](https://doi.org/10.1145/3544548.3581524). 2
- [BMG\*10] BATEMAN S., MANDRYK R. L., GUTWIN C., GENEST A., McDINE D., BROOKS C.: Useful junk?: the effects of visual embellishment on comprehension and memorability of charts. In *Proceedings of CHI Conference on Human Factors in Computing Systems* (2010), ACM. [doi:10.1145/1753326.1753716](https://doi.org/10.1145/1753326.1753716). 2, 5, 8
- [BVB\*13] BORKIN M. A., VO A. A., BYLINSKII Z., ISOLA P., SUNKAVALLI S., OLIVA A., PFISTER H.: What makes a visualization memorable? *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (2013). [doi:10.1109/TVCG.2013.234](https://doi.org/10.1109/TVCG.2013.234). 3
- [CCT24] CHANG H.-Y., CHANG Y.-J., TSAI M.-J.: Strategies and difficulties during students' construction of data visualizations. *International Journal of STEM Education* (02 2024). [doi:10.1186/s40594-024-00463-w](https://doi.org/10.1186/s40594-024-00463-w). 2
- [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 & Applications* (2018). [doi:10.1109/MCG.2018.032421650](https://doi.org/10.1109/MCG.2018.032421650). 2, 8
- [CSZ\*24] CHEN Q., SHUAI W., ZHANG J., SUN Z., CAO N.: Beyond numbers: Creating analogies to enhance data comprehension and communication with generative AI. In *Proceedings of CHI Conference on Human Factors in Computing Systems* (5 2024). [doi:10.1145/3613904.3642480](https://doi.org/10.1145/3613904.3642480). 3
- [CVG] CHEVALIER F., VUILLEMOT R., GALI G.: Using concrete scales: A practical framework for effective visual depiction of complex measures. 2426–2435. [doi:10.1109/TVCG.2013.210](https://doi.org/10.1109/TVCG.2013.210). 2, 3
- [Den18] DENIS D. J.: *SPSS Data Analysis for Univariate, Bivariate, and Multivariate Statistics*. John Wiley & Sons, 2018. [doi:DOI:10.1002/9781119465775](https://doi.org/10.1002/9781119465775). 3
- [GTM10] GRAMMEL L., TORY M., STOREY M.-A.: How information visualization novices construct visualizations. *IEEE transactions on Visualization & Computer Graphics* 16, 6 (2010). 2
- [GWL\*19] GÄBLER J., WINKLER C., LENGYEL N., AIGNER W., STOIBER C., WALLNER G., KRIGLSTEIN S.: Diagram safari: A visualization literacy game for young children. In *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion* (2019). [doi:10.1145/3341215.3356283](https://doi.org/10.1145/3341215.3356283). 2
- [HGA\*16] HARSLEY R., GREEN N., ALIZADEH M., ACHARYA S., FOSSATI D., DI EUGENIO B., ALZOUBI O.: Incorporating analogies and worked out examples as pedagogical strategies in a computer science tutoring system. In *Proceedings of the 47th acm technical symposium on computing science education* (2016), pp. 675–680. 2
- [HHS20] HOGRÄFER M., HEITZLER M., SCHULZ H.-J.: The state of the art in map-like visualization. *Computer Graphics Forum* 39, 3 (2020), 647–674. [doi:https://doi.org/10.1111/cgf.14031](https://doi.org/10.1111/cgf.14031). 3
- [HKN\*18] HULLMAN J., KIM Y.-S., NGUYEN F., SPEERS L., AGRAWALA M.: Improving comprehension of measurements using concrete re-expression strategies. In *Proceedings of CHI Conference on Human Factors in Computing Systems* (2018), pp. 1–12. [doi:10.1145/3173574.3173608](https://doi.org/10.1145/3173574.3173608). 2
- [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 & Computer Graphics* 27, 02 (Feb. 2021). [doi:10.1109/TVCG.2020.3030464](https://doi.org/10.1109/TVCG.2020.3030464). 2
- [HT06] HARRISON A. G., TREAGUST D. F.: Teaching and learning with analogies: Friend or foe? *Metaphor and analogy in science education* (2006), 11–24. [doi:10.1007/1-4020-3830-5\\_2](https://doi.org/10.1007/1-4020-3830-5_2)
- [KHA16] KIM Y.-S., HULLMAN J., AGRAWALA M.: Generating personalized spatial analogies for distances and areas. In *Proceedings of CHI Conference on Human Factors in Computing Systems* (2016), Association for Computing Machinery, pp. 38–48. [doi:10.1145/2858036.2858440](https://doi.org/10.1145/2858036.2858440). 2
- [LFM21] LOCORO A., FISHER JR W., MARI L.: Visual information literacy: Definition, construct modeling and assessment. *IEEE Access* (05 2021). [doi:10.1109/ACCESS.2021.3078429](https://doi.org/10.1109/ACCESS.2021.3078429). 5
- [LSL96] LIN H.-s., SHIAU B.-r., LAWRENZ F.: The effectiveness of teaching science with pictorial analogies. *Research in Science Education* 26, 4 (1996), 495–511. [doi:10.1007/BF02357457](https://doi.org/10.1007/BF02357457). 3
- [Mor17] MORRILL R.: Breaking the bar chart: why chart types are holding us back and how metaphors can help. *Northeastern University* (2017). [doi:10.17760/D20257369.3](https://doi.org/10.17760/D20257369.3)
- [MPP01] MARZANO R. J., PICKERING D. J., POLLOCK J. E.: *Classroom Instruction That Works: Research-Based Strategies for Increasing Student Achievement*. Association for Supervision and Curriculum Development, 2001. 3
- [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 the CHI Conference on Human Factors in Computing Systems* (2024), pp. 1–17. [doi:10.1145/3613904.3642760](https://doi.org/10.1145/3613904.3642760). 2
- [Per21] PERIN C.: What students learn with personal data physicalization. *IEEE Computer Graphics and Applications* 41, 6 (2021). 2
- [PM14] PRITHISHKUMAR I. J., MICHAEL S. A.: Understanding your student: using the vark model. *Journal of postgraduate medicine* (2014). [doi:10.4103/0022-3859.132337](https://doi.org/10.4103/0022-3859.132337). 5
- [RHG18] RIEDERER C., HOFMAN J. M., GOLDSTEIN D. G.: To put that in perspective: Generating analogies that make numbers easier to understand. In *Proceedings of CHI Conference on Human Factors in Computing Systems* (2018), pp. 1–10. [doi:10.1145/3173574.3174122](https://doi.org/10.1145/3173574.3174122). 2
- [RM15] RUCHIKACHORN P., MUELLER K.: Learning visualizations by analogy: Promoting visual literacy through visualization morphing. *IEEE Transactions on Visualization and Computer Graphics* 21 (2015). [doi:10.1109/TVCG.2015.2413786](https://doi.org/10.1109/TVCG.2015.2413786). 2
- [RPA\*24] RICHER G., PISTER A., ABDELAAL M., FEKETE J.-D., SEDLMAIR M., WEISKOPF D.: Scalability in visualization. *IEEE Transactions on Visualization and Computer Graphics* 30, 7 (July 2024). [doi:10.1109/TVCG.2022.3231230](https://doi.org/10.1109/TVCG.2022.3231230). 3
- [Spe10] SPEZZINI S.: Effects of visual analogies on learner outcomes: Bridging from the known to the unknown. *International Journal for the Scholarship of Teaching and Learning* 4 (07 2010). [doi:10.20429/ijstl.2010.040211](https://doi.org/10.20429/ijstl.2010.040211). 2
- [TGM\*20] TURCHIOE M., GROSSMAN L., MYERS A., BAIK D., GOYAL P., MASTERTSON CREBER R.: Visual analogies, not graphs, increase patients' comprehension of changes in their health status. *Journal of the American Medical Informatics Association* 27 (01 2020). [doi:10.1093/jamia/ocz217](https://doi.org/10.1093/jamia/ocz217). 3