Scroll or Stroll: Evaluating Scrollytelling and Free Exploration for Data Stories

SEJAL AGARWAL and SAMIHA ANUVA, University of Waterloo, Canada

Interactive data stories increasingly rely on visual narratives to communicate complex information to broad audiences. Yet, little empirical evidence exists comparing how presentation paradigms affect comprehension, efficiency, and user experience. This study presents a within-subjects experiment comparing two dominant approaches to data storytelling: scrollytelling, a guided narrative that synchronizes text and visualization through scrolling, and free exploration, which allows readers to navigate charts independently. Using two matched Pew Research Center articles implemented in both paradigms, eight participants completed comprehension, usability, and workload assessments. Results revealed no significant difference in comprehension accuracy, but free exploration enabled faster information retrieval, while scrollytelling was perceived as more coherent, aesthetic, and engaging. Preferences split evenly, reflecting distinct reader goals such as learning and context-building versus quick lookup and control. We discuss design implications for selecting or combining paradigms based on user intent and offer guidance for balancing narrative coherence. We found that scrollytelling interfaces are better for learning information whereas free exploration is better for fast retrieval of information.

Additional Key Words and Phrases: Scrollytelling, Data storytelling, Visualization, User studies, Interaction techniques, Human-computer interaction

1 Introduction

Data-driven storytelling has become a central technique for communicating complex information to broad public audiences. As datasets grow in scale and sophistication, organizations increasingly rely on visual narratives to contextualize trends, convey uncertainty, and articulate evidence-based arguments. Newsrooms, policy institutes, and scientific communicators now routinely pair explanatory text with interactive graphics to help readers make sense of unfamiliar or multifaceted topics. Yet despite the widespread adoption of interactive data stories, there remains limited empirical understanding of how different mode of presentation shape comprehension, engagement, and the overall user experience.

Two paradigms dominate contemporary narrative visualization: scrollytelling and free exploration. **Scrollytelling**, popularized in journal articles like *The New York Times' Snow Fall* [4], presents a linear, guided narrative in which visualizations are pinned or animated in synchrony with progressive blocks of text. Scrolling acts as the primary interaction event, advancing the reader through predetermined steps that highlight specific insights, annotations, or transitions. Prior work characterizes this form as highly author-driven, emphasizing narrative coherence and minimizing decision-making burden placed on the reader [3, 25]. These works suggest that constraining the path and pacing of the story through scrollytelling can reduce cognitive load, making it easier for readers to follow the intended message.

In contrast, **free exploration** embodies a more reader-driven or exploratory philosophy. Rather than forcing the user through a fixed sequence, exploratory interfaces pan, zoom, and hover to allow readers to navigate the visual information space according to their own goals and curiosity [32]. Systems such as Voyager [32] and interactive dashboards in tools like Tableau [30] illustrate this paradigm, providing flexible interaction but offering few guarantees that users will encounter the narrative structure intended by the author. As Lee et al. (2015) [16] argues, such experiences can support agency and discovery but may risk fragmented understanding if readers overlook key elements of the story.

Authors' Contact Information: Sejal Agarwal, s97agarw@uwaterloo.ca; Samiha Anuva, ssanuva@uwaterloo.ca, University of Waterloo, Waterloo, Ontario, Canada.

Although both paradigms are well theorized and widely deployed, there is very little empirical evidence comparing them. Existing literature typically focuses on characterizing narrative structures, interaction patterns, or design taxonomies [13, 17, 25] rather than measuring how presentation modes influence user outcomes. Scrollytelling research has examined engagement and attention [3], while exploratory visualization studies emphasize usability, flexibility, and analytical workflows [33]. However no controlled study has directly compared scrollytelling and free exploration, leaving open questions about their trade-offs in comprehension, accuracy, efficiency, workload, usability, and preference. Understanding these trade-offs is increasingly important as content creators must decide whether to favour authorial guidance, reader autonomy, or some hybrid of the two.

This paper addresses this gap through a controlled within-subjects experiment comparing scrollytelling and free exploration for interactive data stories. We designed two interfaces that present matched visual narratives derived from two Pew Research Center short-form data journalism articles [6]. Each article was implemented in both paradigms: a scrollytelling interface in which readers progress through explanatory text and synchronized visualizations using scroll-only interaction, and a free exploration interface in which readers can explore charts through panning, zooming, and hover-triggered annotations. Participants encountered both paradigms and both stories, with interface-story pairings counterbalanced across four possible sequences to control for order effects. After interacting with each story, participants completed comprehension questions and post-conditions usability and workload questionnaires.

Across findings, we did not find statistically significant differences in comprehension between the two interfaces; however, time to answer was shorter with the free exploration interface. Within our participant pool, preferences for the two interfaces were split evenly. Users who preferred scrollytelling said they liked it because it provided context and was less intimidating whereas users who preferred free exploration found it easier and faster to lookup information.

In summary, this work contributes:

- **The first controlled empirical comparison of scrollytelling and free exploration for data stories**, using matched narratives from real-world journalism.
- **Quantitative and qualitative evidence characterizing the strengths and limitations of each paradigm**, particularly in comprehension, usability, workload, and preference.
- **Design guidance for creating storytelling interfaces that balance narrative coherence with user agency**, informing practitioners who must choose between or integrate these paradigms in future visualization systems.

By grounding these findings in real journalistic content and by directly contrasting two dominant design philosophies, our study advances understanding of how interactive narrative structures shape reader experience in modern data storytelling.

2 Related Work

Our literature review mainly covers the following three areas of research in data storytelling - linking storytelling with data visualization, exploring different strategies for data storytelling, and lastly, going through the comparative study of different data visualization techniques.

2.1 Data Visualization and Data Storytelling

Data storytelling has emerged as a crucial technique for communicating facts, ideas, and insights through visualization. Segel et al (2025) [25] introduced the concept of narrative visualization and defined data storytelling as a set of related visual stories that include a narrative. They categorized storytelling approaches into author-driven, reader-driven, and hybrid forms, and analyzed early examples in journalism, infographics, and interactive visualization. Building on this foundation, Lee et al. (2015) [16] distinguished data stories from general visualizations and highlighted how

shared stories can incorporate interactivity. Knaflic (2015) [20] further advanced the practice of data storytelling by offering concrete strategies for designing clear, effective visual narratives aimed at business and analytical contexts. Complementing this, Ryan (2016) [24] argued for cultivating a broader visual culture of data discovery within organizations, positioning storytelling as a key mechanism for making insights accessible and actionable. Together, these perspectives highlight storytelling not merely as a stylistic choice, but as an essential component of modern data visualization practice.

2.2 Different Data Storytelling Strategies

Early works on data storytelling methods explored rhetorical framing to demonstrate how design choices steer audience understanding [12], and analyzed how different sequencing strategies shape comprehension (2013) [14]. Boy et al. (2016) [3] conducted one of the first empirical studies of author-driven scrollytelling, showing that step-by-step scrolling techniques can increase user engagement. High-profile projects such as The New York Times' Snow Fall (2012) [4] and frameworks such as Scrollama.js [9] exemplify how scrollytelling guides audiences through a linear narrative enriched with animations and annotations. While its structured approach ensures coherence, it often limits the user's agency over the data. In contrast, exploratory interfaces, such as Voyager (Wongsuphasawat et al., 2016) [32], and storytelling features in Tableau [30] prioritize user agency by allowing users to interact directly with charts. Although these interfaces provide flexibility, they risk sacrificing narrative coherence and may lead users to overlook key insights.

Subsequent work has expanded storytelling toward more flexible or interactive forms. Dykes (2020) [7] emphasized the importance of narrative framing, while the study conducted by Milesi et al. [19] showed that storytelling benefits diminish as task complexity increases. Shi et al. (2023) [26] introduced interaction-driven storytelling that invites users to affect the narrative directly, and Srinivasan et al. (2019) [27] demonstrated how embedding data facts can support understanding without enforcing a fixed reading order.

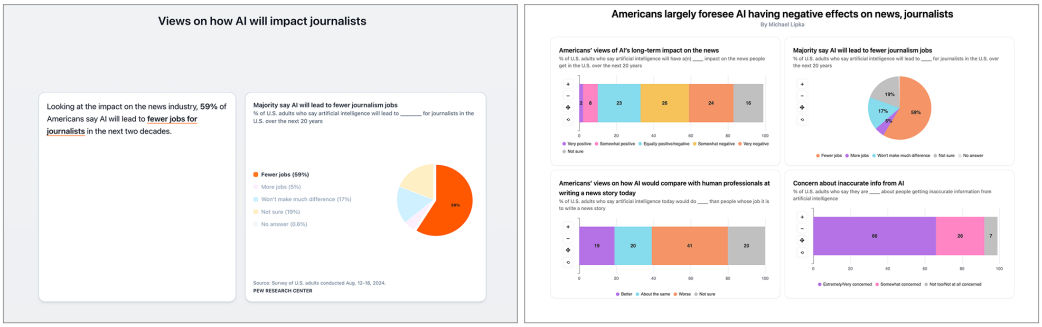
A parallel line of research examines how text and visuals should be integrated within stories. Stokes and Hearst (2022) [28] found that readers often prefer more textual explanation alongside charts, and Stokes et al. (2023) [29] showed that balancing text and visual density influences reader satisfaction. Together, these studies map a progression from highly structured narratives to increasingly interactive and text-rich designs, highlighting the core tradeoff between coherence and agency that motivates our comparative study.

2.3 Comparative study of Visualization Techniques

Prior research has examined how different visualization systems and interfaces influence interpretation, learning, and user outcomes. Echeverria et al. (2018) [8] compared exploratory and explanatory visual learning analytics tools, showing that narrative scaffolding can redirect users' attention and improve sensemaking. More recent work also explores interactive support within analytic systems; for example, Jin et al. (2025) [15] demonstrated how generative-AI based scaffolding can shape users' interactions with dashboards and influence their interpretation processes.

In the context of storytelling systems, Heravi (2025) [11] analyzed 118 narratives to characterize exploratory and explanatory structures, but focused on structural properties rather than assessing reader performance or comprehension.

Despite these contributions, no empirical study has directly compared scrollytelling and free-exploration interfaces with respect to comprehension, efficiency, workload, and usability. Our work addresses this gap through a controlled comparative evaluation of the two storytelling approaches.



(A) (B)

Fig. 1. The two storytelling interfaces used in the study. Figure (A) shows an example frame from the scrollytelling interface with DS1 where text corresponding to a graph is shown and the relevant pie slice is highlighted, and figure (B) shows Free exploration interface with DS1 where the user can pan, zoom or hover over charts to view information.

3 Research Questions and Hypotheses

Our study examines how two common data storytelling paradigms, scrollytelling and free exploration, affect users' comprehension, task performance, and subjective experience when interacting with matched data stories. While prior work has investigated different presentation methods for data storytelling, there remains limited empirical evidence comparing these paradigms directly. Building on the distinctions outlined in the literature, we focused on how presentation mode shapes measurable user outcomes.

We investigated four research questions (RQs):

- **RQ1 (Comprehension):** *How does interface paradigm influence users' comprehension accuracy?*
- **RQ2 (Efficiency):** *How do scrollytelling and free exploration differ in the response time required to answer post-story questions?*
- **RQ3 (Usability and Workload):** *How do participants evaluate the usability and perceived workload of each interface?*
- **RQ4 (Preferences):** *Which paradigm do participants prefer and why?*

Grounded in narrative-visualization theory and prior work on exploratory interaction, we formulated the following hypotheses:

- **H1 (Comprehension):** Scrollytelling will support higher comprehension accuracy due to its guided structure.
- **H2 (Efficiency):** Free exploration will produce faster response times than scrollytelling.
- **H3 (Usability and Workload):** Participants will report higher usability and lower workload with scrollytelling rather than free exploration.
- **H4 (Preferences):** Participants will express an overall preference for scrollytelling for structured data-story consumption.

4 Technical Approach

This project involved building two interactive web interfaces, one scrollytelling-based and one free exploration-based, to present matched data stories using different interaction paradigms. Although the two interfaces differed in how users navigated the content, they shared identical datasets, chart

components, and visual encodings to ensure that differences in user outcomes could be attributed to the presentation paradigm alone.

4.1 Interface Implementation

Both interfaces were implemented in React [18] and used Vite [34] as the build tool. The two systems diverged in their visualization libraries and interaction frameworks. The free exploration interfaces were built using Echarts [1], while the scrolltelling interfaces used D3.js [2] for custom SVG rendering combined with react-scrollama [9] to coordinate scroll-driven text and visualization updates. Styling across both systems was handled through Tailwind CSS [31], and all data was loaded from static JSON files to maintain controlled performance across conditions.

4.2 Data Stories

We selected two short-form articles from the Pew Research Center. As data story 1 (DS1) we had the article *Americans largely foresee AI having negative effects on news, journalists* [21] and as our second data story (DS2) we had *Relatively few Americans are getting news from AI chatbots like ChatGPT* [22]. These stories were chosen for three reasons.

Comparable article type and length. Both articles are concise data journalism pieces published by the same organization and written in a similar narrative style. Their comparable length and structure help standardize reading time and cognitive demands across conditions, making them appropriate for a controlled experimental setup.

Similarity in visualization formats. Each article contains multiple stacked bar charts and a pie chart, with the only difference being that DS1 includes one additional stacked bar chart. The consistency in chart types and visual style helps equalize content complexity while still providing some variation across stories.

Topical coherence. Both stories address public attitudes toward AI in news production and consumption. This shared topical focus minimizes content-driven confounds and ensures that observed differences in user outcomes can be attributed to interface design rather than subject matter.

Together, these characteristics ensured that the two stories provided comparable bases for evaluating how interaction paradigms influence comprehension and user experience.

4.3 Design Rationale

To isolate interaction paradigm effects, we ensured that each interface conveyed the same set of factual statements but presented them differently.

In the scrollytelling interface, all sentences from the original article were included. We only excluded sentences that contained numerical results but did not have a graph associated with it, because we would not be able to represent that data in the free exploration interface for comparable analysis.

In the free-exploration interface, we reproduced or adapted only the sentences that referred to individual, non-aggregated data values. When an article sentence aggregated across multiple subgroups or chart we replaced it with a pair of sentences that described the underlying individual data points. As illustrated in Figure 2, the original article sometimes described aggregated quantities in text. In this case, the sentence “Just 10% say it will have a very (2%) or somewhat (8%) positive effect” combines two subgroups into a single 10% summary. In the free-exploration interface, we did not reproduce such aggregate statements. Instead, the interface only exposed the underlying individual data values. When users hovered over the relevant regions of the chart, they saw separate

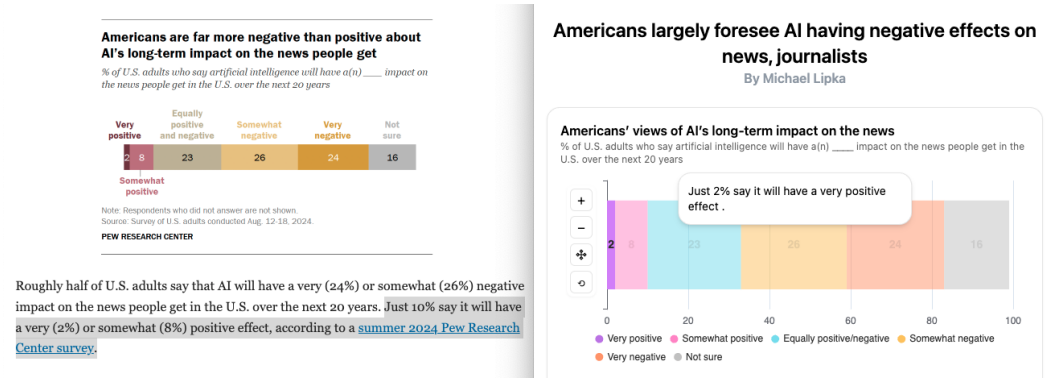


Fig. 2. Example of how aggregate statements were handled. In the original article text (left), the author writes that "Just 10% say it will have a very (2%) or somewhat (8%) positive effect," explicitly aggregating the two positive categories. In the free-exploration interface (right), this aggregate 10% statement is omitted; users instead see only the individual values (e.g., "2% say it will have a very positive effect" and "8% say it will have a somewhat positive effect") when hovering over the corresponding chart segments.

descriptions such as "Just 2% say AI will have a very positive effect" with no textual mention of the combined 10%.

4.4 Scrollytelling Interface Implementation

The scrollytelling interface was developed with React [18], D3.js [2], and react-scrollama [9]. Each chart was rendered as an SVG, and each fact-bearing sentence was represented as a scrollytelling "step", as shown in Figure 1.

Scroll-driven sequencing. As the user scrolled, the visualization remained pinned on the right side of the screen while the sentence on the left updated to reflect the next step. The previous sentence faded out before the new one appeared, creating a smooth narrative progression. When transitioning to a different chart, the user scrolled to a new panel where the next visualization replaced the previous one, and the stepwise sequence continued.

Highlighting and focus. To emphasize the portion of the chart referenced by the current sentence, we used an opacity-based highlighting technique. Relevant bars or slices were shown at full opacity, while non-relevant elements were dimmed. This approach draws attention without adding distracting motion or animation.

Layout. For steps containing only text, content was centered on the page. When paired with a chart, text was placed on the left and visualization on the right, following conventions in scrollytelling design. This implementation reflects the canonical guided-narrative paradigm as the author determines the sequence, pacing, and focus of the reader's experience.

4.5 Free Exploration Interface Implementation

The free-exploration interface was built using React [18] and ECharts [1], which provides built-in support for interactive panning, zooming, and responsive tooltips. Figure 1 illustrates how the free exploration interface looks.

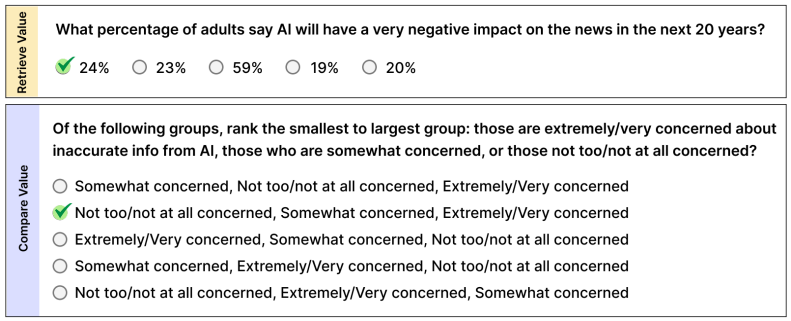


Fig. 3. Two types of multiple choice questions were used in the study - (1) Retrieve value, (2) Compare value or higher order thinking.

All charts visible at once. Unlike the scrollytelling version, all charts for a given story were displayed simultaneously on a single page. Users could move freely between charts without being constrained to a predefined reading order.

Interaction controls. Each chart included a control bar allowing users to:

- Zoom in or out
- Pan the chart by dragging
- Reset the view to the default framing.

These interactions foregrounded user agency and mirrored common exploratory visualization workflows.

Hover-based annotations. Hovering over a chart element revealed a tooltip containing the associated sentence or value. The hovered area was also visually emphasized using opacity adjustments similar to those used in scrollytelling. These annotations were derived from either verbatim article sentences when they references a single data value, or researcher-written equivalents, when the article aggregated multiple values.

Layout. Two charts appeared in each row with consistent titles, axes, and legends. No explanatory text was displayed unless triggered by hover, ensuring that insight discovery emerged from user-driven exploration rather than a predetermined narrative.

4.6 Ensuring Parity Across Conditions

Although the two interfaces differed substantially in structure and interaction model, we controlled for several factors to ensure that participants received equivalent informational content across conditions. Specifically, we standardized the underlying datasets, the chart encodings (e.g., axes, categories, value mappings), labeling conventions, the set of factual statements conveyed in each story, and the highlighting logic used to draw attention to specific chart elements.

By holding these elements constant, differences in comprehension, workload, and perceived usability can be attributed to the interaction paradigm itself rather than inconsistencies in visualization content or information availability.

5 Methods

Here we describe the methods used to conduct our study.

5.1 Participants

Eight participants were recruited from the University of Waterloo. All of them were graduate students aged between 19 to 35. Participation was voluntary and individuals received no compensation.

5.2 Apparatus

The study was conducted using web-based interfaces that ran on the researcher's Macbook laptops to ensure that all participants were viewing the interfaces on the same hardware. Participants answered all of the questions in Qualtrics [23] using their personal laptop.

5.3 Study Design

We used a 2 × 2 within-subjects design crossing interaction paradigm (scrollytelling vs. free exploration) with data story (DS1 vs. DS2). Each participant completed both interaction paradigms and both stories, but each story was encountered exactly once and each paradigm exactly once. Participants were randomly assigned to one of four counterbalanced presentation orders:

(1) DS1–Scrollytelling → DS2–Exploration
(2) DS2–Scrollytelling → DS1–Exploration
(3) DS1–Exploration → DS2–Scrollytelling
(4) DS2–Exploration → DS1–Scrollytelling

This design controlled for potential order effects and allowed comparisons between paradigms independent of story content.

5.4 Tasks and Procedure

After providing informed consent and completing a brief demographics questionnaire, participants were informed that they would read two data stories presented through two different interaction paradigms. Each participant completed two study blocks, one for each story and interface pairing, following the counterbalanced order described above.

At the start of each block, participants were given brief instructions explaining the available interactions for the assigned interface (scrolling for scrollytelling; panning, zooming, and hovering for exploration). They were then given up to three minutes to freely familiarize themselves with the interface and read through the assigned data story.

Immediately afterward, they completed an untimed free-response summary describing the story and articulating the main takeaway. They then answered a set of comprehension questions consisting of two retrieval questions and one comparison question per graph (figure 3). The order of these questions was randomized, and response times were recorded for each item.

Following the comprehension task, participants completed the NASA-TLX [10] workload assessment and the System Usability Scale (SUS) [5] to evaluate perceived workload and usability for the interface just used.

After completing both study blocks, participants selected the interface they preferred overall and provided a written explanation for their choice. The full session lasted approximately 40 minutes.

5.5 Measures

We collected both objective and subjective measures to evaluate participants' performance and experience with each interaction paradigm.

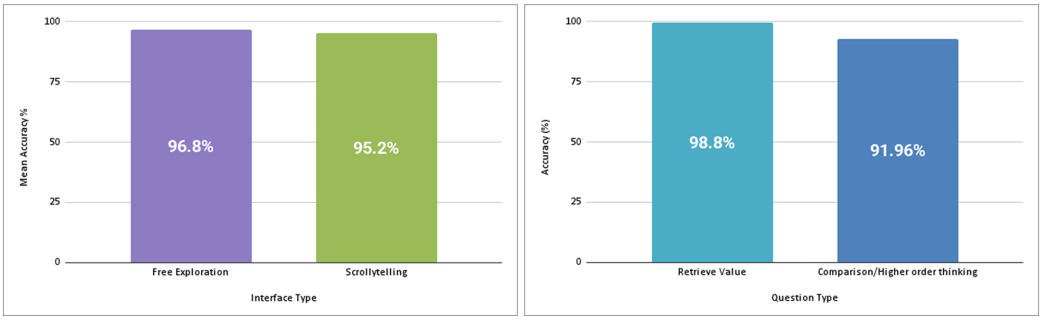


Fig. 4. Figure (A) shows the mean task performance accuracy per participant across the two interfaces. Figure (B) shows the average accuracy in task performance per question (task) type.

Comprehension Accuracy. Accuracy was calculated as the proportion of correctly answered multiple-choice items. Each graph contributed three questions—two retrieval questions targeting specific values and one comparison question requiring interpretation across categories.

Response Time. For each comprehension item, we recorded the time elapsed from question onset to the participant's final submission. These values were analyzed to assess efficiency across interaction paradigms.

Workload (NASA–TLX). Participants completed the NASA-TLX after each story–interface block. We collected ratings for mental demand, effort, and frustration to capture perceived workload.

Usability (SUS). Usability was assessed using the System Usability Scale (SUS), which measures participants' perceived ease of use, clarity, and overall usability of the interface.

Preference Ratings. At the conclusion of the study, participants indicated which interface they preferred overall and provided a brief justification for their choice.

Qualitative Feedback. Participants' open-ended responses for story summaries, interface impressions, and preference explanations were recorded to provide additional context for interpreting quantitative results.

6 Result Analysis

In this section, we analyze participants' performance, efficiency, usability perceptions, and workload ratings across the two interface conditions using a combination of quantitative and qualitative methods. Task performance was evaluated through accuracy of multiple choice question answers for retrieval and comparison questions, followed by a 2 × 2 within-subjects repeated-measures ANOVA and Wilcoxon signed-rank tests to examine the effects of interface and question type. Efficiency was assessed using task completion time, with normalization applied to account for differences in data story length, and analyzed using paired Wilcoxon tests. Usability perceptions were assessed using the System Usability Scale (SUS), while cognitive workload was evaluated using NASA-TLX ratings. Finally, participants' subjective comments and interface preferences were examined to contextualize the quantitative findings. Together, these analyses provide a comprehensive view of how each interface supported comprehension, usability, and workload.

Table 1. Repeated-measures ANOVA results for accuracy. Here, Interface = interface type; Question = question type.

Measure	Effect	F(df)	p
Accuracy	Interface	F(1,7) = 0.5	.501
	Question	F(1,7) = 7.20	.031
	Interface X Question	F(1,7) = 0.03	.878

6.1 Task Performance (Comprehension Accuracy)

Participants completed both retrieval (RV) and comparison (CV) questions in each interface (see figure 3). Comprehension accuracy (figure 4) was calculated as the proportion of correct responses out of the total number of questions per participant.

A 2 × 2 repeated-measures ANOVA was conducted with interface (scrollytelling, free exploration) and question type (RV, CV) as within-subject factors (see table 1).

Effect of Interface

There was no significant effect of interface on accuracy ($F(1, 7) = 0.50$, $p = .501$), indicating that participants performed similarly in scrollytelling ($M = .952$) and free exploration ($M = .968$). Thus, neither interface conferred a measurable performance advantage in terms of overall correctness (Figure 4).

Effect of Question Type

There was a significant effect of question type ($F(1, 7) = 7.20$, $p = .031$), demonstrating that RV questions were answered significantly more accurately than CV questions, regardless of interface.

Interaction Effect

The interface × question type interaction was not significant ($F(1, 7) = 0.03$, $p = .878$). This indicates that the difference between RV and CV questions was consistent across interfaces; that is, RV questions were easier in both interfaces, and the interface did not modulate that difficulty gap.

Interface-Level Accuracy

To complement the ANOVA, a Wilcoxon signed-rank test on overall accuracy across the two interfaces also indicated no significant difference ($W = 8.0$, $p = .590$). This reinforces the conclusion that the interface did not affect participants' performance across different question types. Thus, the task performance results do not support our hypothesis **(H1)**.

Qualitative Analysis of Summarization Answers

We conducted a qualitative examination of participants' data story summarization responses to assess the nature and variability of the produced summaries. Overall, the summaries were comparable in both length and content, with little variation across participants. Most of the participants directly rephrased section titles of the data stories in the summary or discussed the largest percentage slice of the stacked bar graphs and pie charts in their responses.

6.2 Efficiency

Task efficiency was evaluated using task completion time for each interface. Mean completion times were 225.27 seconds (SD = 107.33) for Scrollytelling and 203.61 seconds (SD = 100.75) for free

exploration. Because the two data stories differed in length, with DS1 requiring substantially more time to read than DS2, raw completion times were not appropriate for direct comparison across participants. To control for this confound, completion times were normalized within each story by dividing each participant's time by the median time for the corresponding data story.

A Wilcoxon signed-rank test on these normalized times revealed no significant difference in efficiency between the two interfaces ($W = 12.0$, $p = .461$). This indicates that participants completed the tasks at comparable speeds in scrollytelling and free exploration after adjusting for differences in story length.

For completeness, we also did a Wilcoxon test on the unnormalized times. However, as expected, this comparison was biased by story length differences and likewise did not yield a significant effect ($W = 14.0$, $p = .641$).

Overall, the results indicate that neither interface provided a clear efficiency advantage, and participants completed the tasks at comparable speeds in both conditions once differences in story length were controlled. This outcome contrasts with our initial expectation that participants would require less time in the Free Exploration interface. Thus, the efficiency findings do not support our original hypothesis **(H2)**.

6.3 Usability and Workload

To evaluate participants' subjective experience with the two interfaces, we analyzed both perceived usability and cognitive workload (see figure 5). Usability was assessed using the SUS and cognitive workload was measured using the NASA-TLX instrument. Together, these measures provide complementary insight into how intuitive, demanding, and comfortable each interface was to use from the participants' perspective.

System Usability (SUS Score)

Participants completed the SUS after interacting with each interface. The SUS consists of 10 items, each rated on a 5-point Likert scale, which are then converted into a 0–100 usability score following the standard scoring procedure: odd-numbered items contribute positively, even-numbered items are reverse-scored, and the summed value is multiplied by 2.5 to yield the final score. Higher SUS values indicate greater perceived usability.

Scrollytelling received a mean SUS score of 85.0 (SD = 9.82), while Free Exploration received 78.1 (SD = 13.72). Although scrollytelling scored slightly higher, both interfaces fall within the "Good–Excellent" usability range, indicating strong overall user acceptance (figure 5).

A Wilcoxon signed-rank test revealed no significant difference in perceived usability between the two interfaces ($W = 15.0$, $p = .742$).

Since the p-value is well above 0.05, there is no significant difference between the usability of the two interfaces. This finding does not align with our initial hypothesis that participants would report higher usability for the scrollytelling interface.

Perceived Workload (NASA-TLX)

NASA-TLX responses were analyzed on a 1–7 scale. Scrollytelling resulted in numerically lower overall workload (M = 2.79, SD = 0.72) compared to free exploration (M = 3.38, SD = 1.02).

A Wilcoxon signed-rank test confirmed that this difference was not statistically significant ($W = 10.5$,$p = .383$).

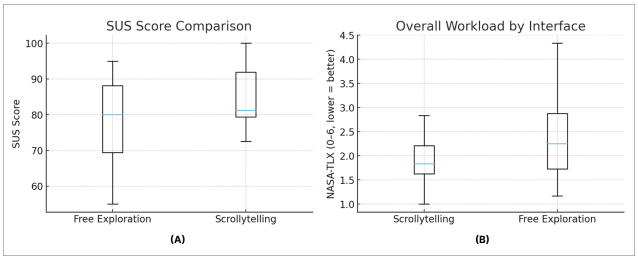


Fig. 5. Comparison of both interfaces in terms of usability and workload.

Across subscales, scrollytelling showed consistently lower (better) scores for mental, physical, temporal, effort, and frustration, but none reached significance. Notably, physical demand approached significance ($p = .084$), suggesting that the free exploration may put more workload on users. Thus, the findings does not completely align with our hypothesis **(H3)** that the scrollytelling interface will have lower worload.

6.4 Thematic Analysis of User Preferences

After completing the tasks in both interfaces, participants were asked to indicate which interface they preferred and to explain the reason for their choice. Preferences were evenly split, with half favouring the scrollytelling interface and half favouring the free exploration interface. This finding contradicts with our hypothesis **(H4)**. We conducted a thematic analysis of their responses and came across several recurring themes that characterized their rationale.

Information Retrieval. Participants reported that the free exploration interface helped them to retrieve answers from the story faster compared to the scrollytelling interface. Since they could view the whole story on the screen together and just hover to see detailed information if required, it was easier to scan for particular data points without needing to navigate step-by-step through the narrative. Although the scrollytelling interface imposed lower cognitive load, the need to scroll back and forth to revisit earlier sections was described as cumbersome and time-consuming.

> "...The free exploration interface presented a lot of information at once and so was more overwhelming initially, but once I was familiar with what information was there, I found it much easier to quickly skim and find answers to the questions..." (P7)

Background Context. Participants reported that the explanatory characteristic of the scrollytelling interface helped them to understand the context of the data story. The gradual reveal of information was less intimidating, especially at the beginning of the task, compared to the overwhelming amount of information they were exposed to in the free exploration interface at once.

> "...I can engage more quickly with storytelling way as it has provided more interactions and more background context. I think with free exploration I can easily lose focus..." (P8)

Comprehension and Recall. Several participants indicated that the scrollytelling interface improved their comprehension of the data story and supported better recall of key ideas. The guided progression and segmented presentation were helpful for memory and overall understanding.

> "...I prefered the scrollytelling interface because it gave me a better comprehension of the data and allowed me to memorize/ recall information faster..." (P5)

Learning experience. The participants thought scrollytelling provided more enjoyable and structured learning experience. Some viewed the interface as visually engaging and conducive to deeper learning, but acknowledged a trade-off between understanding and efficiency.

> "...It's so much more fun! I feel like I get to learn while scrolling (way better than reels) and the way the information is visually presented is more appealing to me..." (P3)

Relationship Analysis. Participants who favoured the free exploration interface reported that seeing all charts simultaneously made it easier to compare visualizations, verify assumptions, and identify relationships across multiple data visualizations. In general, the user driven free exploratory interface made it possible to discover new correlations and relationships between the data, which was a bit difficult in the scrollytelling interface.

> "Getting to see all the charts at once makes it easier to make sense of the overall material. For examples, 42% people saying they are not sure whether AI news is true can be concerning, but the chart besides says 75% never gets news from AI, which put things into perspective..." (P4)

Perceived Control. Finally, participants valued the sense of control offered by the free exploration interface. They appreciated being able to navigate the content at their own pace and in their own order, whereas the scrollytelling interface was perceived as more restrictive due to its author-driven flow.

> "...For free Exploration, I have more control and it would cost less to get the data information that I want..." (P6)

7 Discussion

Our within-subjects comparison of scrollytelling and free exploration suggests that the two paradigms support *different reader goals* rather than producing strong, uniform advantages on accuracy or perceived experience. With eight participants, we found no statistically significant differences in overall comprehension accuracy between interfaces, whether examined via a participant-level nonparametric test or a two-way repeated-measures ANOVA. Instead, question type (retrieval vs. comparison) showed the clearest performance effect, independent of interface. These patterns indicate that the format of the information need can dominate interface choice for accuracy at small scales.

7.1 Two paradigms, two goals

Despite comparable accuracy, participants moved noticeably faster through comprehension items in the free-exploration condition, aligning with our a priori expectation that direct control, immediate tooltips, and simultaneous access to all charts would streamline information lookup. This speed advantage is consistent with free exploration's affordances for quick navigation and localized value retrieval. In contrast, scrollytelling appeared to better support reflective reading and meaning-making: half of participants preferred it for providing context and learning, describing it as less intimidating and more coherent as a narrative experience. Taken together, these results point to a pragmatic division of labour: free exploration is better for time-constrained lookup or side-by-side comparison tasks, and scrollytelling is better for tasks that aim to build understanding, context, and confidence.

7.2 Experience quality converges, but "feel" differs

Subjective measures reinforce this complementarity. SUS scores for both interfaces fell in the "Good–Excellent" range with no significant difference, and overall NASA–TLX workload also did

not differ significantly. Participants therefore judged both paradigms as usable and not overly taxing, even as their qualitative comments framed scrollytelling as more "guided," "aesthetic," and "fun," while free exploration felt "faster," "in control," and "better for finding things quickly." In other words, the *feel* of interaction diverged while overall usability converged. Designers can leverage this since when both paradigms are comparably usable, aligning the interface with the user's *intent* (learn vs. retrieve) becomes the key lever.

7.3 Why preferences split

Preference split evenly, and the reasons map cleanly onto the paradigms' affordances. Participants favoring scrollytelling cited background context, comprehension, and a better learning experience; those favouring free exploration emphasized finding information, user control, speed, and relationship analysis. These justifications echo design theory that positions author-driven narrative as reducing decision burden while reader-driven exploration maximizes agency. Our empirical observations suggest that, for short journalistic stories with stacked-bar and pie charts, neither approach *by itself* guarantees superior comprehension. Instead, each interface channels attention differently: scrollytelling curates a sequence and focal highlights; free exploration broadens the visible information horizon and reduces navigation friction.

7.4 Design Implications

Our findings highlight that scrollytelling and free exploration serve distinct reader intentions, and that designers should choose or blend these paradigms carefully depending on the primary goal of the experience.

7.4.1 Using scrollytelling. Scrollytelling is best suited for contexts where the goal is to teach, explain, or build understanding. Its guided progression helps reduce cognitive load by controlling pacing, focus, and the order in which key insights appear. This makes it ideal for narrative journalism, educational explainers, or policy communication settings where authors want to ensure that readers grasp the intended storyline. The trade-off is reduced flexibility and slower information retrieval once readers know what they are looking for.

7.4.2 Using free exploration. Free exploration, in contrast, fits situations where the goal is to retrieve, compare, or investigate. Having all charts visible at once and offering direct manipulation allows readers to access information quickly and according to their own curiosity. It suits dashboards, reference materials, and analysis tools where efficiency and control matter more than narrative cohesion. However, the cognitive burden is higher since readers must determine where to look and how to interpret relationships themselves.

Balancing both paradigms. While combining the two approaches can be powerful, it is also a delicate balance. Adding exploratory controls to a narrative interface can easily increase complexity and distract from the story. A more effective approach is sequential where we use scrollytelling to guide newcomers through the key message, then offer a lightweight "explore more" mode for readers who want to dive deeper. This layering preserves clarity while supporting both comprehension and agency without overwhelming users.
In summary, scrollytelling supports learning and coherence, free exploration supports control and speed, and blending them requires careful restraint to avoid cognitive overload.

8 Limitations

This study provides preliminary insights into how scrollytelling and free exploration influence comprehension and user experience, but several limitations constrain the strength and generalizability

of our findings. Our participant pool was small (N=8) and relatively homogeneous, consisting mostly of university students. As a result, statistical power was limited, and our sample may not reflect the diversity of real-world readers who vary in age, technical background, and familiarity with interactive data stories. A larger and more demographically diverse group would help determine whether the observed patterns hold across different audience types and expertise levels.

The scope of our materials and tasks also narrows interpretation. Both data stories were short-form news pieces from the Pew Research Center, featuring simple bar and pie charts with straightforward factual statements. These choices allowed for experimental control but limit the applicability of our results to more complex or explanatory storytelling contexts that use richer datasets, dynamic filtering, or more interpretive narratives. Moreover, while our free-exploration interface included panning, zooming, and hover interactions, participants rarely used these features because all charts were visible at once. This suggests that our implementation may not have fully captured the benefits or cognitive demands of exploration in denser visual environments. Finally, our comprehension questions focused primarily on factual recall and basic comparisons which limited our understanding of comprehension when viewing data stories under these paradigms. Together, these limitations underscore the need for larger samples, more varied story types, and broader task designs to more fully characterize the trade-offs between guided and exploratory data storytelling.

9 Future Work

Future research should build on these findings by involving larger and more diverse participant samples and by testing a broader range of story types, visualization formats, and interaction complexities. Expanding beyond short-form news pieces to include explanatory, analytical, or persuasive narratives would help reveal how story intent shapes the effectiveness of each paradigm. Future studies could also examine hybrid designs that blend scrollytelling's narrative guidance with selective exploration opportunities, identifying thresholds where added interactivity enhances rather than overwhelms readers. Finally, incorporating higher-level reasoning tasks and longitudinal measures of engagement or recall would deepen understanding of how narrative structure and interaction design jointly influence learning, efficiency, and user satisfaction in data storytelling.

10 Conclusion

This study provides one of the first direct comparisons between scrollytelling and free exploration for interactive data stories. While our small-scale experiment did not reveal strong performance differences in comprehension or usability, participants' qualitative feedback illuminated meaningful distinctions in intent and experience. Scrollytelling better supported understanding, focus, and narrative engagement, whereas free exploration enabled faster access, greater control, and more flexible navigation. Together, these findings suggest that the effectiveness of a storytelling paradigm depends less on the interface itself than on the reader's goal, whether to learn or to retrieve. For designers and communicators, this demonstrates the importance of aligning interaction design with user intent rather than seeking a single "best" format. In practice, hybrid designs that guide first-time readers through a structured narrative and then open space for exploration may best balance coherence and autonomy. As interactive data storytelling continues to evolve, understanding how structure and freedom interact remains key to crafting engaging, informative visual narratives.

Author Contributions

Both authors made equal contributions to this project.

Sejal contributed to the study design, developed the two scrollytelling interfaces, and created the question set for DS1. She also designed the survey flow, conducted qualitative analyses, recruited and conducted user studies with 4 participants, and authored sections of the final report, including

the Abstract, Introduction, Research Questions and Hypotheses, Technical Approach, Methods, Discussion, Limitations, and Conclusions.

Samiha contributed to the study design, developed the two free exploration interfaces, and created the question set for DS2. She also found the data stories, conducted the literature review, added DS2, NASA-TLX and SUS questions to the survey, recruited and conducted user studies with 4 participants, performed result analysis (quantitative and qualitative), and wrote Related Works and Results sections, as well as adding figures.

References

[1] Apache Software Foundation. ECharts: An open-source interactive charting library. https://echarts.apache.org/, 2025. Accessed 2025.

[2] Bostock, M., Ogievetsky, V., and Heer, J. D3: Data-driven documents. In *Proceedings of the IEEE Symposium on Information Visualization* (2011), pp. 1–8.

[3] Boy, J., Detienne, F., and Fekete, J.-D. Storytelling in information visualizations: Does it engage users to explore data? CHI '15, Association for Computing Machinery.

[4] Branch, J. Snow fall: The avalanche at tunnel creek - multimedia feature, Dec 2012.

[5] Brooke, J. Sus: A quick and dirty usability scale. Tech. Rep. TR.246, Digital Equipment Corporation, Reading, UK, 1996. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, I. L. McClelland (Eds.), *Usability Evaluation in Industry*, Taylor & Francis, London, 1996.

[6] Center, P. R. Numbers, facts and trends shaping your world. Website, 2025. Accessed: 2025-11-18.

[7] Dykes, B. *Effective data storytelling: how to drive change with data, narrative and visuals.* John Wiley & Sons, 2019.

[8] Echeverria, V., Martinez-Maldonado, R., Shum, S. B., Chiluiza, K., Granda, R., and Conati, C. Exploratory versus explanatory visual learning analytics: Driving teachers' attention through educational data storytelling. *Journal of Learning Analytics 5*, 3 (2018), 73–97.

[9] Goldenberg, R. react-scrollama: React wrapper for scrollama. https://github.com/russellgoldenberg/react-scrollama, 2025. Accessed 2025.

[10] Hart, S. G., and Staveland, L. E. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in Psychology* (1988), vol. 52, North-Holland, pp. 139–183.

[11] Heravi, B. Exploratory and explanationory features in data storytelling: untangling the interplay and associations with linearity, interactivity, and structure. *Journal of Computational Social Science 8*, 4 (2025), 94.

[12] Hullman, J., and Diakopoulos, N. Visualization rhetoric: Framing effects in narrative visualization. *IEEE transactions on visualization and computer graphics 17*, 12 (2011), 2231–2240.

[13] Hullman, J., Drucker, S., Riche, N. H., Lee, B., Fisher, D., and Adar, E. A deeper understanding of sequence in narrative visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2013), ACM, pp. 91–100.

[14] Hullman, J., Drucker, S., Riche, N. H., Lee, B., Fisher, D., and Adar, E. A deeper understanding of sequence in narrative visualization. *IEEE Transactions on visualization and computer graphics 19*, 12 (2013), 2406–2415.

[15] Jin, Y., Yang, K., Yan, L., Echeverria, V., Zhao, L., Alfredo, R., Milesi, M., Fan, J. X., Li, X., Gasevic, D., et al. Chatting with a learning analytics dashboard: The role of generative ai literacy on learner interaction with conventional and scaffolding chatbots. In *Proceedings of the 15th International Learning Analytics and Knowledge Conference* (2025), pp. 579–590.

[16] Lee, B., Riche, N. H., Isenberg, P., and Carpendale, S. More than telling a story: Transforming data into visually shared stories. *IEEE Computer Graphics and Applications 35*, 5 (2015), 84–90.

[17] McKenna, S., Staheli, K., Munzner, T., and Correll, M. A design activity framework for narrative visualization. *IEEE Transactions on Visualization and Computer Graphics 23*, 1 (2017), 365–375.

[18] Meta Platforms, Inc. React: A javascript library for building user interfaces. https://react.dev/, 2025. Accessed 2025.

[19] Milesi, M. E., Mejia-Domenzain, P., Brandl, L., Echeverria, V., Jin, Y., Gasevic, D., Tsai, Y.-S., Käser, T., and Martinez-Maldonado, R. " piecing data connections together like a puzzle": Effects of increasing task complexity on the effectiveness of data storytelling enhanced visualisations. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (2025), pp. 1–18.

[20] Milesi, M. E., Mejia-Domenzain, P., Brandl, L., Echeverria, V., Jin, Y., Gasevic, D., Tsai, Y.-S., Käser, T., and Martinez-Maldonado, R. "piecing data connections together like a puzzle": Effects of increasing task complexity on the effectiveness of data storytelling enhanced visualisations. CHI '25, Association for Computing Machinery.

[21] Pew Research Center. Americans largely foresee AI having negative effects on news, journalists, 2025. Retrieved

from https://www.pewresearch.org/short-reads/2025/04/28/americans-largely-foresee-ai-having-negative-effects-on-news-journalists/.

[22] Pew Research Center. Relatively few americans are getting news from AI chatbots like ChatGPT, 2025. Retrieved from https://www.pewresearch.org/short-reads/2025/10/01/relatively-few-americans-are-getting-news-from-ai-chatbots-like-chatgpt/.

[23] Qualtrics, LLC. Qualtrics xm platform [computer software]. https://www.qualtrics.com/, 2025. Accessed 2025.

[24] Ryan, L. *The visual imperative: Creating a visual culture of data discovery.* Morgan kaufmann, 2016.

[25] Segel, E., and Heer, J. Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics 16*, 6 (2010), 1139–1148.

[26] Shi, Y., Gao, T., Jiao, X., and Cao, N. Breaking the fourth wall of data stories through interaction. *IEEE Transactions on Visualization and Computer Graphics 29*, 1 (2022), 972–982.

[27] Srinivasan, A., Drucker, S. M., Endert, A., and Stasko, J. Augmenting visualizations with interactive data facts to facilitate interpretation and communication. *IEEE transactions on visualization and computer graphics 25*, 1 (2018), 672–681.

[28] Stokes, C., and Hearst, M. Why more text is (often) better: Themes from reader preferences for integration of charts and text. *arXiv preprint arXiv:2209.10789* (2022).

[29] Stokes, C., Setlur, V., Cogley, B., Satyanarayan, A., and Hearst, M. A. Striking a balance: Reader takeaways and preferences when integrating text and charts. *IEEE Transactions on Visualization and Computer Graphics 29*, 1 (2022), 1233–1243.

[30] Tableau Software. Tableau. https://www.tableau.com/, 2025. Accessed: 2025-10-02.

[31] Tailwind Labs. Tailwind CSS: Utility-first css framework. https://tailwindcss.com/, 2025. Accessed 2025.

[32] Wongsuphasawat, K., Moritz, D., Anand, A., Mackinlay, J., Howe, B., and Heer, J. Voyager: Exploratory analysis via faceted browsing of visualization recommendations. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (2016), 649–658.

[33] Yi, J. S., Kang, Y. a., Stasko, J., and Jacko, J. Toward a deeper understanding of the role of interaction in information visualization. *IEEE transactions on visualization and computer graphics 13*, 6 (2007), 1224–1231.

[34] You, E. Vite: Next generation frontend tooling. https://vitejs.dev/, 2025. Accessed 2025.