

Ego vs. Exo and Active vs. Passive: Investigating the Effects of Viewpoint and Navigation on Spatial Immersion and Understanding in Immersive Storytelling

Tao Lu
Georgia Institute of Technology
Atlanta, GA, USA
luttul@gatech.edu

Wong Kam-Kwai
The Hong Kong University of Science
and Technology
Hong Kong, China
kkwongar@connect.ust.hk

Qian Zhu
Renmin University of China
Beijing, China
qzhual@connect.ust.hk

Anlan Xie
Georgia Institute of Technology
Atlanta, GA, USA
anlan.xie@gmail.com

Tiffany Ma
Georgia Institute of Technology
Atlanta, GA, USA
tiffany.ma@duke.edu

Alex Endert
Georgia Institute of Technology
Atlanta, GA, USA
endert@gatech.edu

Yalong Yang
Georgia Institute of Technology
Atlanta, GA, USA
yalong.yang@gatech.edu

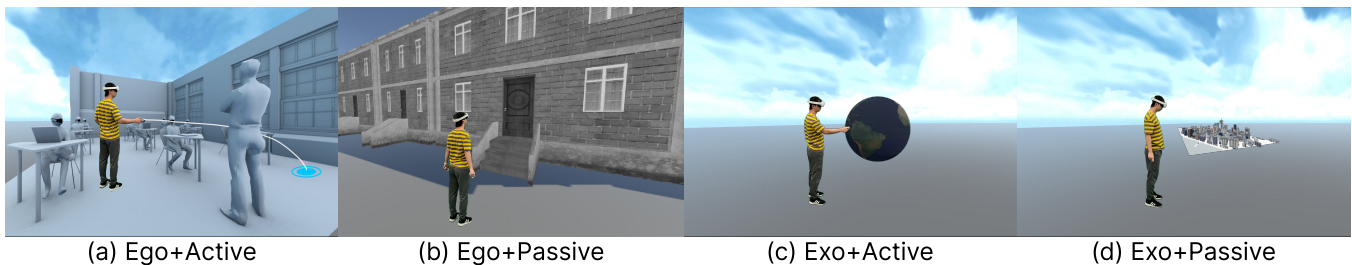


Figure 1: We investigated the effects of different combinations of viewpoints and navigation in immersive storytelling, as demonstrated in four implemented scenarios: (a), (b), (c), and (d). In (a) and (b), the audience is fully immersed in the story scene from an egocentric perspective. Conversely, in scenarios (c) and (d), the audience maintains an exocentric perspective, remaining independent of the story. In (a) and (c), the audience navigates the story manually and actively, following specific instructions. In contrast, (b) and (d) allow audience clicking a controller button, which triggers the story and allows navigation to proceed automatically.

Abstract

Visual storytelling combines visuals and narratives to communicate important insights. While web-based visual storytelling is well-established, leveraging the next generation of digital technologies for visual storytelling, specifically immersive technologies, remains underexplored. We investigated the impact of the story viewpoint (from the audience’s perspective) and navigation (when progressing through the story) on spatial immersion and understanding. First, we collected web-based 3D stories and elicited design considerations from three VR developers. We then adapted four selected

web-based stories to an immersive format. Finally, we conducted a user study (N=24) to examine egocentric and exocentric viewpoints, active and passive navigation, and the combinations they form. Our results indicated significantly higher preferences for egocentric+active (higher agency and engagement) and exocentric+passive (higher focus on content). We also found a marginal significance of viewpoints on story understanding and a strong significance of navigation on spatial immersion.

CCS Concepts

• Human-centered computing → Empirical studies in interaction design; Empirical studies in HCI; Virtual reality.

Keywords

Immersive storytelling, Story navigation, Story viewpoint in immersive environments



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

CHI '25, Yokohama, Japan

© 2025 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-1394-1/25/04

<https://doi.org/10.1145/3706598.3713849>

ACM Reference Format:

Tao Lu, Qian Zhu, Tiffany Ma, Wong Kam-Kwai, Anlan Xie, Alex Endert, and Yalong Yang. 2025. Ego vs. Exo and Active vs. Passive: Investigating the Effects of Viewpoint and Navigation on Spatial Immersion and Understanding in Immersive Storytelling. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26–May 1, 2025, Yokohama, Japan*. ACM, New York, NY, USA, 19 pages. <https://doi.org/10.1145/3706598.3713849>

1 Introduction

Visual storytelling uses the combination of visual elements and narratives to communicate complex ideas (e.g., data, facts, and opinions) in a way that is more engaging and descriptive than traditional text-only stories [37]. This approach has been widely used in various fields, including journalism [22, 60, 77], education [6, 81], commercial activities [53, 94], and scientific communication [46]. It plays a crucial role in enhancing the readers' understanding of content, as well as increasing their emotional engagement and connection to the story.

The evolution of digital technologies has greatly expanded the possibilities of visual storytelling, offering new genres and more diverse ways to deliver content [65]. The introduction of digital displays brought dynamic and interactive visual content, changing how readers interact and engage with stories [21, 91]. The rise of mobile devices further transformed access and interaction modalities, enabling readers to experience stories anywhere with new input modalities such as touch-based interactions [43, 67]. As digital technologies continue to evolve, the pressing question now is: "What is next for visual storytelling, and how can it offer novel experiences?"

As immersive technologies rapidly develop, they are becoming increasingly accessible and affordable, positioning them as promising candidates for the next generation of major public digital platforms. Immersive devices (e.g., Virtual Reality (VR) headsets) can render 2D and 3D information in the surrounding space or even replace the users' entire visual field, offering unprecedented opportunities for visual storytelling [31, 34]. A prominent use of immersive technology for storytelling is 360-degree videos [20, 28]. However, these videos often limit the readers' navigation, providing little control over story progression or detail exploration due to their fully author-driven nature. Beyond static videos, major news outlets like The New York Times and The Washington Post have embraced interactive formats in their web-based articles to present public events (e.g., COVID-19 and elections) more dynamically. They also leverage immersive technologies for more interactive experiences, such as mobile AR stories created by The New York Times [76]. However, these narratives tend to rely on basic touchscreen inputs and offer limited immersion. They are often confined to a single scene without transitions between multiple narrative elements. In this work, we focus on improving immersive storytelling specifically for public journalism. We seek to extend the design of these journalistic stories into immersive environments, where interactivity and complete storylines can foster deeper engagement and understanding.

Our goal is to leverage the unique display and interaction affordances of immersive technologies in storytelling. A major advantage of these technologies is their ability to render 3D spatial information at its original scale [31]. Building on this characteristic,

we investigated how people visually perceive 3D spatial representations (i.e., ego- vs. exocentric viewpoints) and how they can interactively navigate in the 3D space (i.e., active vs. passive navigation) as an initial exploration. Specifically, readers can experience an immersive story from an **exocentric** (exo) viewpoint to gain an overview of a 3D scene or from an **egocentric** (ego) viewpoint to fully immerse themselves with life-sized artifacts. Additionally, readers can either **actively** move in space and interact with digital artifacts with greater embodiment and agency or **passively** follow predefined transitions. We anticipate that these two factors (i.e., viewpoint and navigation) will have a significant effect on user-perceived spatial immersion and understanding in immersive storytelling. **Our research goal is to understand the advantages and limitations of these design choices in spatial immersion and understanding in immersive VR stories.**

To study the effects of viewpoints and navigation in immersive storytelling, we needed to develop immersive stories that were controlled for these factors. While there are many web-based stories for public journalism, immersive versions are still rare. Since we aimed to focus on viewpoint and navigation rather than the story content itself, we leveraged existing web-based stories and created versions in immersive environments to cover the design variations we needed for the study. We collected web-based stories from major news outlets, particularly those with rich 3D spatial content, such as a New York Times story about how COVID-19 spreads in a classroom [8]. These visual stories are in single scene and follow a linear structure. We consider this atomicity as a starting point for studying our intended effects. Meanwhile, this basic structure serves as the foundational unit for more complex stories. We then conducted a formative study with three VR experts to elicit potential designs of 117 selected stories in an immersive format, with an intended focus on viewpoint and navigation. Based on the design considerations from this formative study, we adapted four representative stories into immersive formats. Finally, we conducted a user study with 24 participants to compare the 2×2 combinations (ego vs. exo viewpoint, and active vs. passive navigation). Our paper structure follows the steps we took to prepare and conduct the study, as illustrated in Figure 2.

We found that viewpoints and navigations showed more significant impact on the story understanding and spatial immersion, respectively. Exo demonstrated benefits on story content comprehension and ego performed better on the spatial information memorability. Active navigation created a higher level of presence with a cost of more perceived workload, but it is acceptable in a ego viewpoint. Overall, participants predominantly preferred Ego+Active or Exo+Passive, depending on their priority on contents or experiences.

In summary, our primary contributions are:

- Four design considerations elicited for adapting web-based visual stories to an immersive format.
- A user study empirically investigates the effects of viewpoint and navigation on perceived spatial immersion and understanding across four adapted immersive story instances.

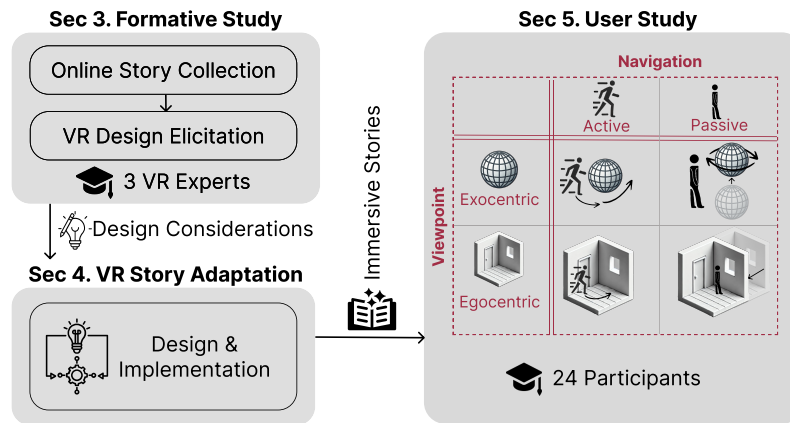


Figure 2: Overall Procedure of This Work. We started by a formative study analyzing the web-based stories in terms of their viewpoint and navigation designs. It elicited some design considerations, which we used to design and implemented four story cases. We finally investigated the effects of viewpoints and navigations on 24 participants.

2 Related Works

Visual Storytelling. Visual storytelling is widely used in journalism to help audiences understand data-backed facts through static or dynamic visual elements, such as images, videos, and animations. Segal and Heer [65] were among the first to summarize emerging narrative visualization techniques in web-based storytelling, discussing design variations across different genres. Hullman et al. [27] studied how information prioritization affects user interpretation. Kosara and Mackinlay [33] further highlighted the potential of storytelling in visualizations, emphasizing its promising future. Subsequent research explored design considerations in visual storytelling, proposing guidelines to improve content representation and interaction. For instance, Morais et al. [54] proposed a design framework for anthropographic data stories, while Dasu et al. [14] examined the use of characters in data stories to enhance enjoyment and persuasion. Yang et al. [87] also explored the application of narrative structures in data stories. These studies connect design practices with human cognition theories, offering valuable insights for visual storytelling design.

Beyond narrative structure, audience engagement and interaction are critical to the story’s impact, persuasion, and memorability. With advancements in technology, visual storytelling has expanded to include videos, games, and VR/AR platforms [51, 92, 95], offering audiences more interactive experiences than static texts and images alone. VR/AR, in particular, immerses users in virtual spaces, enhancing spatial coherence and unlocking new design possibilities. This paper focuses on two key elements in VR/AR and visual storytelling: viewpoints and navigation, which shape user perspective and agency, directly influencing immersion, comprehension, and emotional engagement. We aim to identify effective design practices for integrating these aspects into immersive storytelling.

Immersive Experiences. The goals of visual storytelling are multifaceted [2]. For authors, common objectives include (a) communicating facts and insights, and (b) persuading audiences of the opinions conveyed through the story. For audiences, the primary

goal is to gain new knowledge, such as news, insights, and perspectives, while entertainment also plays a significant role, as stories may be consumed for enjoyment.

While 2D screens are effective for visual storytelling, immersive technologies like VR/AR extend stories into 3D, closing the gap between audiences and narratives by placing them in the same space. Immersive experiences offer several advantages. First, they enhance presence and immersion [66]. VR/AR can simulate real-world scenarios [93] with real-time rendering and embodied interactions [26, 30, 96], or even go beyond reality (e.g., ground-level scaling [1]). Second, they improve understanding and retention [55, 90]. VR/AR is frequently used in scientific visualizations [49], helping users extract and communicate insights more effectively. Third, they enhance emotional communication [32], narrowing the emotional distance between authors and audiences and fostering resonance and reflection on the story’s themes [7].

These benefits align with the core goals of visual storytelling. Therefore, this paper investigates the impact of immersive environments on understanding and spatial immersion.

Immersive Storytelling. Immersive storytelling, a broader topic of integrating immersive experiences to different storytellings, has been studied in various other contexts. In AR, Shin et al. [69] studied the two factors: room size and furniture density on a AR crime-solving game, and they suggested more context-awareness such as using virtual objects as substitutes in large and empty rooms. A later survey [68] furthered this idea, proposing to balance virtual and real experiences through situated user interactions. In VR, Lee et al. leveraged VR to materialize abstract measure and units so that audiences could connect with their natural experiences [40]. Hall et al. focused on synchronous immersive data presentation and communication [23]. VR storytelling is particularly relevant in cinematic. Collen et al. investigated the cinematic techniques used in immersive narrative visualizations of 3D spatial information [12]. Practically, Wu and Karwas [83] developed a technique that allows free viewpoint change and with time-based interactivity. Other immersive storytelling works focusing on specific applications, such as situated awareness of social problems [4, 97] or emotional enjoyment of visualizations [63]. These works contribute to the design

of immersive storytelling by highlighting the benefits of immersive technologies we should leverage in specific topics or contexts, but none explored the representation and interaction design on visual journalism (e.g. scrollytelling) in immersive environment.

There are some studies focusing on how traditional storytellings on 2D screens can be transported to VR/AR, which are primarily based on the video format [56, 86, 98]. The adaptation of interactive visual stories, a popular form of visual storytelling, are still underexplored.

Viewpoint and Navigation in Immersive Environments. In traditional digital media (e.g., computer monitors or mobile devices), content is typically viewed from an exocentric (third-person) viewpoint, offering a comprehensive overview of spatial relationships and layout. Immersive technology, while capable of maintaining an exocentric view, uniquely allows for a 360-degree stereoscopic experience from an egocentric (first-person) viewpoint. Despite this capability, the egocentric view is not widely adopted due to mixed findings on its effectiveness. Yang et al. [89] found that exocentric views significantly outperformed egocentric views for geographic analysis tasks. Similarly, Kraus et al. [35] and Yang et al. [88] noted limitations in egocentric 3D scatterplot visualizations, such as lack of overview and targets being out of view. However, the egocentric view has shown promise in multi-view management within immersive environments [15, 29, 45, 64], and studies have demonstrated improved memory and knowledge retention in egocentric settings [36, 85]. Notably, Hoppe et al. [25] argued for a perspective continuum in immersive environments, finding no difference in users' perceived workload, sense of presence, and engagement between egocentric and exocentric viewpoints in a VR combat game they developed. However, it remains questionable whether this perspective continuum is applicable in immersive storytelling, as storytelling involves different perception and cognitive processes compared to games. Given these mixed results, our research explores the advantages and drawbacks of both exocentric and egocentric viewpoints in the context of visual storytelling.

Navigation in immersive environments refers to the process of moving to the target location within 3D virtual environments. It is a crucial interaction for exploring 3D spaces in these settings. In the context of immersive storytelling, navigation plays a vital role in allowing users to visually explore different parts of the story and advance the narrative (e.g., a story progresses to the next stage once the viewer reaches a specific location). Significant efforts have been devoted to developing various navigation techniques, including gesture design [78], utilizing different input modalities [72], leveraging perception and space manipulation [17], and experimenting with different scales [1, 52]. Luca et al. [16] conducted an extensive survey of these techniques and compiled a Locomotion Vault. However, some of these innovative techniques require additional hardware, while others demand a steep learning curve. As a result, the mainstream and default navigation techniques adopted by major platforms remain natural walking and teleportation. Considering the accessibility aspect of storytelling, we decided to focus on using mainstream navigation techniques for a broader user group. In many immersive applications, users need to actively navigate the 3D space to complete tasks. However, in storytelling, we have the option to let users passively follow a predefined path without explicit navigation actions (such as walking or teleportation). While

active navigation provides more interactivity, it carries the risk of disorienting the user. In contrast, passive navigation maintains narrative control but may reduce the sense of agency and induce motion sickness. Deciding which method to use is a unique and important consideration for immersive storytelling. In relation to this topic, Lages and Bowman [38] compared physical walking (where the user moves themselves) with grab-and-move techniques (where the user moves the view) in VR, demonstrating that performance may depend on the user's spatial abilities. Although both methods they studied were active, we are particularly interested in the performance of passive navigation when combined with a predefined storyline.

3 Eliciting Design Considerations for Immersive Story Adaptations

This project investigates the effects of viewpoint (ego vs. exo) and navigation (active vs. passive) in immersive storytelling through a user study. Preparing suitable study materials was the crucial first step.

Visual storytelling can be characterized by four dimensions: narratives, transitions, structure, and the balance between explorability and explainability [71, 73]. With various options in each category, visual storytelling can become quite complex. Since immersive storytelling is an emerging field with limited publicly available resources and few established design guidelines, we decided to develop our immersive stories. Creating high-quality stories is a complex and challenging task. Therefore, instead of creating stories from scratch, we chose to adapt web-based stories from major news outlets. This approach allows us to ensure story quality while reducing potential confounding factors. Considering the aforementioned four dimensions, we found that most stories from major news outlets follow a single-scene linear narrative structure with simple and continuous transitions and predominantly feature author-driven explainability. This simplicity is ideal for our initial exploration into immersive storytelling. We further focused on stories with significant 3D spatial content due to their greater synergy with immersive environments. This approach enabled us to concentrate on the study itself, maintaining story quality while avoiding potential design biases.

To guide our adaptations, we collected existing 3D spatial stories from major news outlets and sought input from three VR experts. We compiled their insights into four generic design considerations, which were used for the adaptations discussed in section 4. An overview of the formative study process is shown in Figure 3.

In summary, the formative study and adaptation process, described in section 4, are essential steps for conducting the user study detailed in section 5. Figure 2 illustrates the relationship between these steps.

The formative study involved three steps: 1) initial story collection, 2) filtering stories based on spatial information, and 3) eliciting design considerations for VR adaptation. Details of our collected stories are provided in the supplementary material.

Initial story collection. Our first step is to collect abundant web-based stories. To ensure the quality of our visual story dataset,

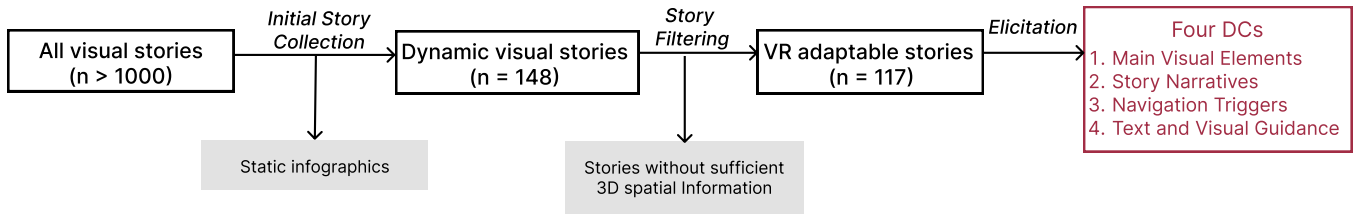


Figure 3: Formative Study Flow Diagram [57]. We first collected dynamic visual stories from major news outlets. We then filtered out those without significant 3D spatial informations. With the remaining VR adaptable stories, we elicited four design considerations (DCs) for our later adaptations.

we collected stories from major news outlets and popular online media platforms, including The New York Times (NYT) [77], The Washington Post [60], Bloomberg [11], The Guardian [22], Reuters [62], National Geographic [19], Financial Times [74], Los Angeles Times [75] and The Pudding [61]. Some platforms feature dedicated visual story sections, which we prioritized, such as the Graphics section in NYT [77] and the Visual Stories section in The Washington Post [60].

We focused on stories published between 2020 and 2024, as recent content tends to incorporate more interactive visuals. We excluded static stories without interactivity or visual transitions, as these did not align with the goals of our user study. From the sources above, we browsed 1017 visual stories in total during the search process. Ultimately, our initial corpus comprised 148 visual stories that are either interactive or feature visual transitions. Most of them from The New York Times (102), followed by Reuters (16), The Washington Post (13), The Pudding (7), Financial Times (4), Los Angeles Times (2) and The Guardian (1).

Story filtering based on spatial information. Not all visual stories were suitable for our study. For example, animated bar charts may not offer a significantly different experience in an immersive environment compared to a desktop.

Since our goal was to investigate how the unique 3D capabilities of immersive environments impact the storytelling experience, we applied further filtering based on the presence of 3D spatial information. Stories were evaluated using two criteria: 1) sufficient 3D spatial content and 2) coherent spatial transitions within a unified environment. The second requirement limited our scope to single-scene stories.

For instance, a NYT story explaining COVID-19 transmission in a classroom was deemed suitable, as it contains rich 3D spatial information and smoothly transitions the viewpoint across different parts of the environment. In contrast, the NYT story “Ukraine’s Race to Hold the Line” uses 2D visuals without enough 3D spatial content. Similarly, the story “Seeing Earth from Outer Space” includes 3D elements but presents multiple disjointed spaces, making each new scene appear completely separate from the previous one, complicating our intended study of immersive effects.

By applying these two criteria, we excluded 31 additional stories, resulting in a final set of 117 stories for our VR adaptation study.

Elicitation of design considerations for VR adaptation. One of the key goals of our formative study was to derive design considerations that could inform the VR adaptation process. To accomplish this, we conducted an elicitation study with three experienced VR developers whose backgrounds were medical VR, computer

graphics in VR, and immersive animation design, respectively. Each developer was assigned a roughly equal number of stories from our collection. They were asked to review their assigned stories individually and describe their intended adaptation in response to three specific questions related to our study factors, along with an open-ended question to explore additional design ideas. Developers were encouraged to provide detailed feedback.

- Q1: What is the best way to view this story in an immersive environment? And why?
- Q2: Do you prefer to actively move in the 3D space or stand still and let the content move for you? And why?
- Q3: What other design considerations do you consider important?

We analyzed their responses using affinity mapping to identify common themes and patterns in their feedback. Regarding the viewpoint (Q1), *information density* emerged as a key factor in choosing between egocentric and exocentric viewpoints. Developers generally favored an exocentric viewpoint for stories with dense spatial information—where a large amount of detail is concentrated within a limited area. For example, a story about the tall buildings in Manhattan, where the buildings are closely packed in a small space, would benefit from an exocentric perspective. Conversely, an egocentric viewpoint was preferred for stories featuring more dispersed spatial information, where the relative distances between objects are much greater than their individual sizes—such as a story about players on a football field. The developers largely agreed on the criteria for selecting viewpoints.

In contrast, there was no clear consensus on navigation design (Q2). For the same story, developers expressed differing preferences for active versus passive navigation. Some noted that passive navigation, which resembles the “scroll-to-progress” interaction commonly used on webpages, offers familiarity and ease of use. Others argued that active navigation in VR feels more intuitive and immersive, enhancing user engagement. These differing opinions were spread across various stories, without any strong patterns emerging.

In summary, while the developers reached a broad consensus on viewpoint selection, the decision between active and passive navigation remained inconclusive in our formative study.

Overall, we identified in the affinity diagram that experts frequently mentioned four major story components in their responses:

- DC1: Main Visual Element. The main visual element serves as the centerpiece of the story, often represented by the primary 3D model or scene.

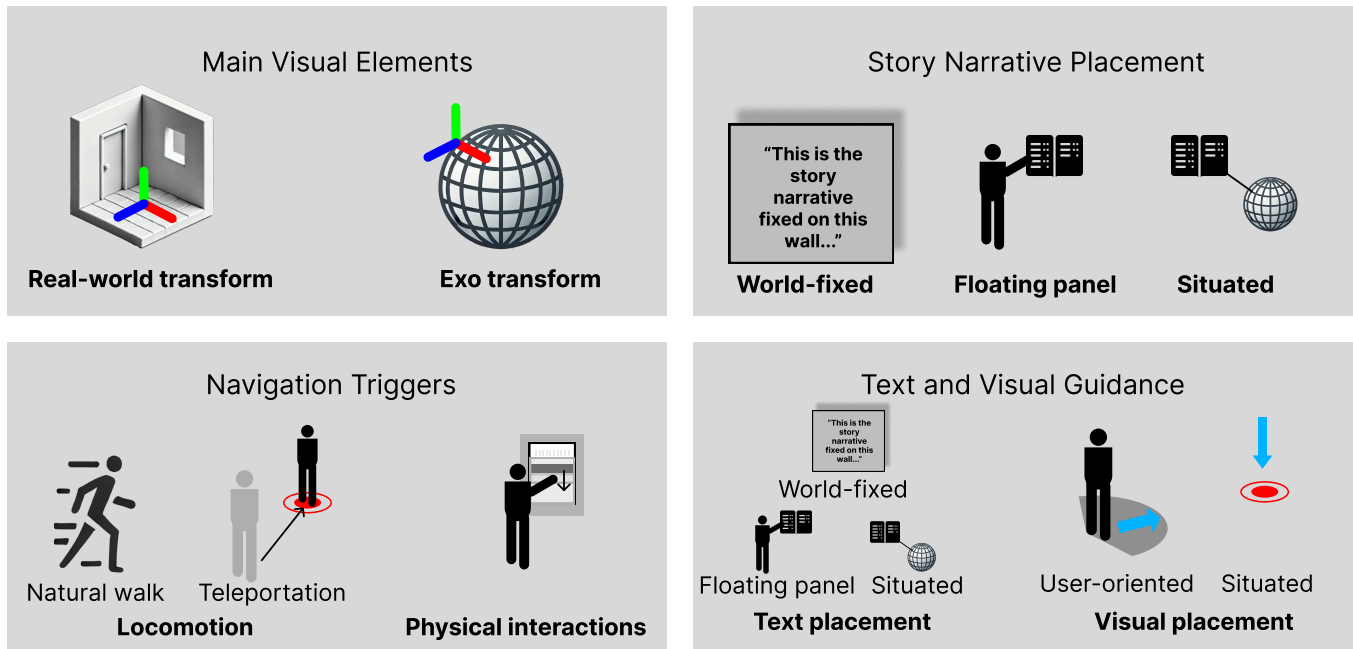


Figure 4: Four Major Story Components and Corresponding Available Options.

- DC2: Story Narratives. Story narratives refer to the structured sequence of textual content that guides the user through the story.
- DC3: Navigation Trigger. To trigger the navigation, the user needs to either move to a specific location or perform certain interactions.
- DC4: Text and Visual Guidance. Instructions regarding location and interaction information are necessary to guide the user in progressing through the story.

Among these four design considerations, DC3 and DC4 hold unique importance for immersive stories. While web-based stories typically use straightforward navigation triggers, such as scrolling and clicking that require little to no guidance, immersive stories demand more explicit navigation and guidance cues due to the opened interactions audiences can perform. Therefore, it is essential to emphasize specific navigation triggers and provide corresponding guidance when adapting stories for immersive environments. These findings of our formative study provide structured guidelines for adapting web-based stories to immersive environments.

4 VR Story Adaptation for User Study

This section outlines the process of adapting well-established web-based stories into immersive environments for use as user study materials. We planned a within-subjects user study to minimize the influence of individual differences on the results. Given the four testing conditions in our study (as shown in Figure 2), we required at least four immersive stories to mitigate the learning effect.

From the formative study results, we determined that viewpoint (ego vs. exo) was story-dependent, so we selected two stories for each viewpoint. Navigation (active vs. passive), however, was adaptable to any story, so we created both active and passive versions

for each. A Latin square design was used to balance the presentation order of stories and navigation methods for participants. Four web-based stories were ultimately adapted into immersive formats for the study.

4.1 Design Consideration Exploration

Our goal was to explore design options for studying the effects of viewpoints and navigation. We examined the four design considerations identified in the formative study and selected the most appropriate options for adaptation.

For the DC1 (*main visual content*), we reused high-quality 3D models from the original stories. In egocentric stories, real-world scale was used for full immersion, while exocentric stories featured scaled-down models to fit within the user’s field of view.

For the DC2 (*story narrative*), we reused the original text, considering several placement options: 1) Fixed placement: text appears in a consistent location (e.g., on a wall), though this may require frequent context-switching. 2) Floating panels: text is placed on movable panels, ensuring easy access with minimal effort. 3) Situated text: embedding text near the relevant objects to enhance semantic understanding.

DC3 (*navigation trigger*) in passive designs was triggered by simple actions like button presses. In active designs, it involved either locomotions [16] (e.g., walking or teleporting) or physical interactions (e.g., opening a window in VR).

For DC4 (*text and visual guidance*), concise textual instructions were provided, and visual indicators (e.g., arrows or hotspots) were used to guide navigation and interactions.

The design space of immersive stories is expansive, and building a comprehensive one is not the focus of this work and is beyond its scope. Our exploration of design options is not an exhaustive list of all possible alternatives but serves as structured guidelines

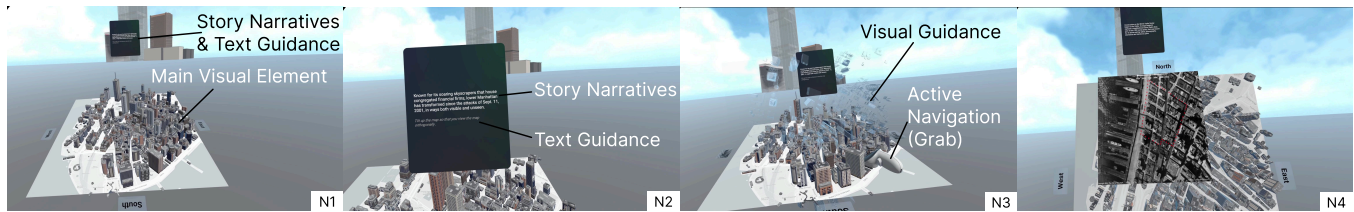


Figure 5: Illustration of Case 1. The main visual element is a Manhattan landscape with 3D buildings (N1). The story narratives and the text guidance for navigation are in a text panel above the main visual element (N2). A transparent copy of the main visual element periodically animates to a designated perspective, visualizing the text guidance (N3). Audiences follow the guidance and the story proceeds to the next paragraphs (N4)

for the necessary and critical adaptations. In the remainder of this section, we detail how we adapted the four selected stories using these design options.

4.2 Case 1: How 911 changed Financial District in Manhattan

This story covers the societal changes in Manhattan’s Financial District post-9/11 [9]. The original story featured a 3D model of Manhattan’s densely packed buildings, making it ideal for an exocentric VR adaptation. The story has 15 steps and 611 words in total. A step includes a paragraph of story narratives with corresponding visual elements. Audiences proceed to next step by following a specific navigation trigger.

Main Visual Element. We built a 3D landscape of Manhattan in VR (Figure 5-N1) through Mapbox [48]. The landscape is 2m-by-2m square base map by default. The current building information came from mapbox.3d-buildings tileset. When the story narratives mentioned a particular group of buildings (e.g. WTC buildings), they are highlighted in blue.

Story Narratives Placement. Story narratives are displayed on an interactive panel in VR (Figure 5-N2), which supports repositioning, resizing and scrolling. We utilized a floating panel design because the narratives are often associated with the whole main visual element. It is not situated to a specific parts.

Text and Visual Guidance. Text guidance appeared in smaller gray italics to minimize distraction. Visual guidance, in the form of a transparent 3D model, periodically animated to show the next perspective audience should have to look at the landscape (Figure 5-N3).

Navigation Triggers. Active navigation involved rotating and zooming in/out of the landscape to trigger changes in the scene. In passive navigation, the landscape automatically adjusted as participants progressed.

4.3 Case 2: In the Atlantic Ocean, Subtle Shifts Hint at Dramatic Dangers

This story explores the impact of climate change on the Gulf Stream [80]. We only selected the first half of the story, which has the most concentrated 3D spatial information. The main visualization of the original story was a globe with animated current circulations in Atlantic Ocean. Thus, we implemented a corresponding exocentric VR story. We summarized the text-only portion (484 words) into 2

paragraphs (100 words) to keep the length consistency. The story has 15 steps and 503 words in total.

Main Visual Element. We adapted the original animation of ocean currents onto a 3D globe in VR. The globe’s size is adjustable to suit user preferences.

Story Narratives Placement. The text portion of the story is in the same interactable panel as Case 1. However, in this case the panel is fixed around the globe surface and centered between audiences’ view and the globe centroid (Figure 6-A1). This is because the main visual element (i.e. the globe) occupies a much larger space than a normal human body and it would easily interfere with the panel if they were placed separately. To reduce the occlusion caused by the current placement, we make the panel background translucent ($\alpha = 0.1$, almost transparent) so that audiences can see underlying visualizations and keep track of the panel simultaneously.

Text and Visual Guidance. The visual guidance of this case consists of 2 animated arrows (Figure 6-A3). A bended arrow is placed in front of the audience’s view on the surface and pointed to the next designated area mentioned in narratives. Another straight arrow is at that area and pointed to the globe centroid, indicating the direction audiences should look at the globe. It shows the next step 10 seconds after entering the current step. The text narrative is at the bottom of the panel and follows the same style as Case 1.

Navigation Triggers. The active navigation of this case is primarily grabbing and rotating, as the original story rotated the globe along the Gulf Stream. For the 2 summarized text-only paragraphs, we adopt a step-away interaction: after audiences step away from the globe, the text panel leaves the surface and attaches to the audiences’ view (0.5 meters from the eyes). Audiences can press a controller button to read the next paragraph and step closer to the globe after finishing 2 paragraphs. In the passive navigation, audiences only need to press a controller button and the globe would rotate and change the visualizations automatically.

4.4 Case 3: What the Tulsa Race Massacre Destroyed

The story of this case illustrates the business achievement of black Americans in Tulsa, Oklahoma in the early 1900s and how the town was destroyed by racist mobs [58]. We only selected the first half of the story, which has the most concentrated 3D information. We summarized the second half of the story into a paragraph (130 words) and put it at the end so that story narratives are logically complete while keeping similar length with previous cases. The

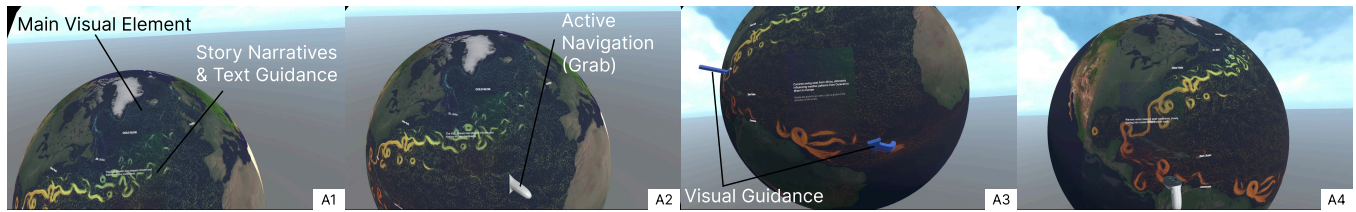


Figure 6: Illustration of Case 2. The main visual element of the story is a globe with animated current circulation in Atlantic Ocean (A1). The story panel with translucent background is fixed around the globe surface and centered between the audiences' view and the globe centroid (A1). Audiences actively navigate the story by grabbing the globe and rotating to the designated area mentioned in narratives (A2, A4), following the text and visual guidance (A1, A3).

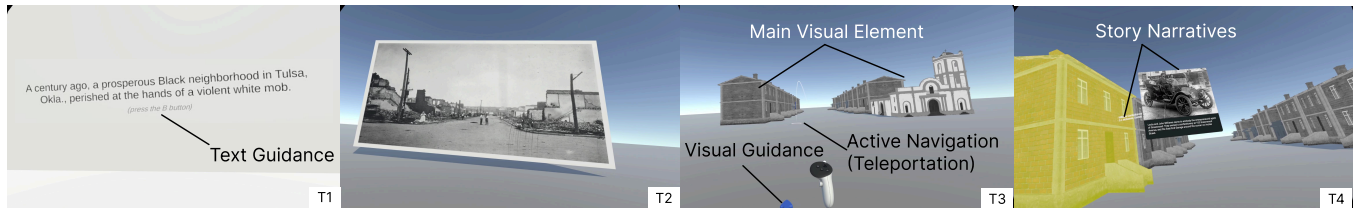


Figure 7: Illustration of Case 3. T1 and T2 replicated the signposting sentence and image at the beginning of the story. Audiences then actively navigate the story on a main avenue of Tulsa before the massacre, following the visual guidance of arrows and hotspots (T3). After they arrive at a building, the corresponding story narratives would appear in a panel and a label nearby (T4).

main visualization of the original story comprises a 3D reconstruction of Tulsa before the massacre, which leads to our design of a VR story with an egocentric viewpoint in a main avenue (Greenwood Ave) of Tulsa (Figure 7-T3). The story has 18 steps and 618 words in total.

Main Visual Elements. Greenwood Ave of Tulsa included a number of black businesses alongside the road. Audiences follow a predefined route, in which one building would be highlighted at each step. At the beginning of the story, we keep the signposting sentence and image in the original story (Figure 7-T1, T2) as they concisely describe the main event of the story.

Story Narratives Placement. The narratives of this case consist of two parts: business labels and introductory paragraphs. In the original story, they were presented in two passes of the avenue. In our VR story, we combine these two types of narratives into one pass because it is more natural to follow a linear order, which ends the story when audiences reached the end of the avenue. Each building has a label listing the businesses inside, which is shown close to the building. The introductions are not on every building as not every business was introduced in the original story. They would be shown in front of the audiences' views.

Text and Visual Guidance. The text guidance of this case only exists at the beginning (Figure 7-T1), which asks audiences press a controller button to proceed. After audiences enter the Greenwood Ave, only the visual guidance indicates the building audiences should go. It shows the next step 10 seconds after entering the current step. It includes an arrow fixed at the bottom of the field of view (FOV), a hotspot highlight of the destination and another arrow above the hotspot pointing downwards. We fix an arrow in the FOV because it ensures audiences to find the hotspot even if it is behind them.

Navigation Triggers. The active navigation of this case is predominantly teleportation. Audiences follow the visual guidance and teleport to the front of each building and gradually to the end of the avenue. In the passive navigation, audiences only need to press a controller button and they are translated to the destination automatically.

4.5 Case 4: Why Opening Windows Is a Key to Reopening Schools

The story of this case illustrated the importance of ventilation in classroom during COVID pandemic by comparing the concentrations of contamination in different ventilation conditions. The original story visualized the concentrations in both airflow and heatmap. We only selected the airflow visualization as it contains more 3D spatial information than the heatmap. Although the original story adopted an exocentric viewpoint, we decided to implement an egocentric VR story because it would be more immersive to be in a life-sized room rather than viewing a smaller room model. The story has 13 steps and 513 words in total.

Main Visual Element. The classroom includes some models of students and a teacher. An air circulation animation is shown to depict the inhale and exhale flow in the classroom. A student is then highlighted as infected (in cyan in Figure 8-C3), and the airflow becomes yellow and red to indicate the concentration differences. The interactable objects (e.g. the cyan window in Figure 8-C4) are also highlighted when interacting with them to proceed through the story.

Story Narratives Placement. The story narratives are mainly in the interactive panel. In this case, the panel is attached to the left controller, as a large portion of the narratives are about the airflow, which occupy the whole classroom and do not have a specific area to attach to.

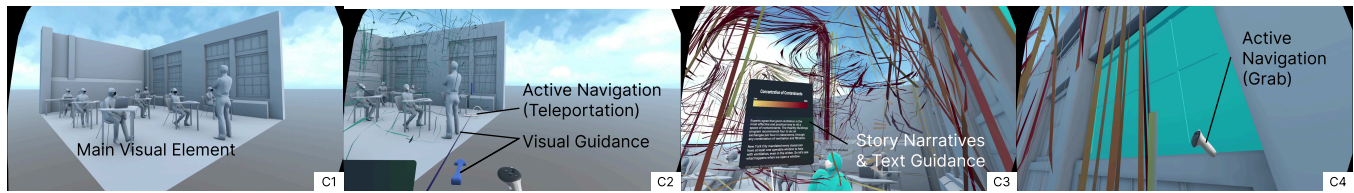


Figure 8: Illustration of Case 4. The main visual element was a 3D classroom model (C1). Audiences actively navigated to the next designated area mentioned in narratives by teleportation (C2). After they arrived, they saw a visualization of airflow with corresponding narratives on the panel (C3). They were asked to drag down the window to see a visualization of a different airflow (C4).

Text and Visual Guidance. Following Case 1 and 2, the text guidance is also at the bottom of the panel with the same font. The visual element follows the same design as Case 3, and we add an extra line connecting the hotspot and the bottom of the panel because the model of students and teacher might occlude the hotspot on the ground. Similarly, it shows the next step 10 seconds after entering the current step.

Navigation Triggers. The active navigation of this case involves teleportation and grabbing. Audiences are asked to teleport to different locations of the classroom to view the airflows. They can also drag down the window and put on the fan to change the ventilation conditions. In the passive navigation, audiences only need to press a controller button and their viewpoint, ventilation condition and corresponding visualizations would change automatically.

4.6 Story Calibrations

The original four stories have different story length, visual complexity and spatial information, which is hard to compare directly. We calibrated four cases based on the first three design considerations (DC1-3 in section 3): main visual element, story narratives, and navigation triggers. For text and visual guidance (DC4), a “one-size-fits-all” design could negatively impact some story instances. Therefore, we opted to use designs optimized for each individual story instance.

DC1: *Main Visual Element(s).* We calibrated the size of main visual element between two exo cases (Case 1 and 2) and two ego cases (Case 3 and 4), respectively. For exo cases, the main visual element was bounded by a cube with 2m edges. For egocentric cases, the main visual elements spread in a 20m-by-20m square arena, and the size of each elements was adjusted accordingly. Due to the different stories of each case, we were not able to have similar number of visual elements between exo and ego cases. But we guaranteed that the number of visual elements audiences needed to pay attention to at each step was no larger than 5.

DC2: *Story Narratives.* For each case, we divided the story into about 15 steps. The steps were created based on the original story narrative setups, which consisted of several short paragraphs in a scrollytelling format. We combined neighboring short paragraphs and summarized long paragraphs to make each step about 30 to 50 words long. Thus, four cases had the similar lengths of narratives (15 ± 3 steps and 550 ± 50 words).

DC3: *Navigation Triggers.* The navigation triggers in two exo cases were looking at the main visual element from a specific angle. In two ego cases, they were teleporting to a specific location. Due

to the different scene and visual element settings, we cannot use the same navigation triggers between exo and ego cases.

DC4: *Text and Visual Guidance.* The effectiveness of text and visual guidance was closely related to the other 3 DCs, particularly the main visual elements. Thus, we tried to optimize the guidance by designing the guidance separately based on visual element setup of each story case. For example, to indicate the manipulations audience needed to do (primarily rotation), Case 1 used the animations of a translucent “ghost” map and Case 2 used two arrows. Animation provided the clearest instructions of the next step, but it was not applicable to Case 2 whose main visual element was a globe. Similarly, both Case 3 and 4 used arrows and an on-ground hotspot to indicate the position audiences needed to teleport to. However, the hotspot could be easily occluded by student and desk models in Case 4. So we used an extra line connecting the hotspot to the left controller so that audiences would easily find the next position to teleport. We acknowledged that such inconsistency might cause confounding factors, but no single text and visual guidance design could serve as a clear and effective road sign in all four cases.

5 User Study

With the four adapted VR stories, we conducted a user study to evaluate the performance of spatial immersion and understanding among different settings of viewpoints and navigations. Our study was approved by our institution’s IRB. It implements a within-subject comparison where each participant experiences active and passive navigation methods in both egocentric and exocentric viewpoints. Each session lasted about 2 hours and participants were compensated with \$20 for their time.

5.1 Study Factors and Conditions

We studied the Ego and Exo viewpoints as well as active and passive navigations, including their individual and combined effects. Figure 9 shows the characteristics of the four conditions.

Ego vs. Exo Viewpoint. The differences between Ego and Exo lies primarily in the visual representations. For example, Ego stories have surrounding visual element around audiences, while visual elements in Exo stories are primarily in front of the participants. There are also different scales of story scene, as Ego stories tend to have a much larger range compared to the space a normal person can occupy. Additionally, the information density in audiences’ view is low, but there are out-of-view objects that they may leave out. Exo stories, in contrast, have everything in audiences’ view, which increases the information density.

Characteristics	<i>Ego</i>	<i>Exo</i>	Characteristics	<i>Active</i>	<i>Passive</i>
Surrounding Visual Elements	✓	✗	Required Embodied Interactions	✓	✗
Scale of Story Scenes	●	●	Precise Interaction	✓	✗
Spatial Information Density	●	●	Level of Control	●	●
Out-Of-View Target	✓	✗	Required Visual Search	●	●

(a) (b)

Figure 9: Characteristics of Four Conditions. A tick mark indicates that the condition has corresponding characteristics, and the cross mark indicates the absence. If both conditions have the characteristics, they are labeled with circles with size indicating the extent (more or less).

Active vs. Passive Navigation. The differences between active and passive are more distinct. Active stories require embodied interactions to push forward the story, and oftentimes the interactions requires precision (e.g. going to a specific location or looking from a specific angle). In passive stories, only a minimum effort (i.e. clicking a button on the controller) is required. Moreover, active stories give audiences a higher level of control by asking them to proceed with story manually. This also increases the amount of visual search necessary for identifying the next step. In passive stories, visual elements transition to their next predefined state automatically, and audiences have less control and visual search in this process.

To systematically test the main factors (i.e., viewpoint and navigation), we include four conditions in our study: Ego+Active, Ego+Passive, Exo+Active, and Exo+Passive (illustrated in Figure 1).

5.2 Study Setup

Our user study was conducted in our lab space at the university. Participants experienced our stories using a Meta Quest 3 virtual reality headset, which has a pixel resolution of 2064 x 2208 per eye and 90Hz refresh rate. The headset was connected to PC via Air Link, allowing it to wirelessly leverage PC computing power. Participants were allowed to physically move and interact with stories in a 3x3m² area. During the study, we set the controllers as the only input devices for stability. We mapped the grip press to the grab interaction and thumbstick forward push to teleportation for both controllers. This study setup was introduced at the very beginning of the study session for each participant.

5.3 Participants

We recruited participants from our university by sending out recruitment emails to mailing lists and Slack channels and selected 24 participants (11 females, 13 males, aged 21 to 31) from all responses. 11 participants were experienced in XR, and 5 were experienced in visualization/storytelling. 3 participants had both expertise. Participants came from various backgrounds, including computer science, UX design, aerospace engineering, and Human-Computer Interaction.

As subsection 4.5 mentioned, all four cases were calibrated in story length and overall design languages. Each participant experienced the four cases, with the order counterbalanced using a Latin

square matrix (ego vs. exo)×(active vs. passive). Since each story has both an active and a passive version, there were a total of eight possible sequences. Further details can be found in our supplementary materials. We acknowledged that this within-subjects setting could cause possible learning effects or fatigue, but we believe the difference in the nature of each cases minimizes these confounding factors.

5.4 Study Procedure

Our study procedure includes four major steps:

1. Introduction: We first showed participants the consent form for the study. After they read and signed the form, we introduced the concept of immersive storytelling, ego/exocentric viewpoints, and active/passive navigation methods.

2. Training: We built a training scene that helps participants learn necessary interactions (e.g., grab, scale, teleportation) and story setup (e.g., narrative text panels, visual guidance design) in our VR stories. The training scene is a mockup of four real stories that use the same design language.

3. Experiencing the story: After the training session, participants were to experience the four VR stories. None of them had read the original story before. After participants finished each story, we asked them to complete 3 tasks: (a) completing a survey of presence and perceived workload of experiencing the VR story; (b) describing their understanding of the story’s main idea and important details; and (c) drawing the story scene. The data we want to collect through these tasks will be described in Sec 5.5.

4. Final interview: Finally, after participants experienced all the stories, they were asked to discuss the differences among the four stories in terms of their engagement, help of understanding and other user experiences. The purpose of this interview was to collect subjective feedback on each story and find out their common considerations.

5.5 Task Data Collection

Through the study, we aimed to investigate the impact of ego/exocentric viewpoints and active/passive navigation on spatial immersion and understanding of VR stories. For spatial immersion, we collected responses to the Presence Questionnaire [10] and

NASA-TLX [24] to measure spatial presence and perceived workload, respectively.

To evaluate the understanding, we focused on the story content comprehension and spatial information depiction. The total and categorical points of these two tasks for each story was the same since the their story lengths are similar. For each story, we designed a set of comprehension rubrics that incorporated three sets: (a) *main idea memorization* (b) *detail memorization* (e.g. numbers, names, quotes or definitions, etc.); and (c) *random inspirations* (e.g. personal opinions, analyses or conclusions that are not presented directly in the story content). The rubrics is attached in the supplementary materials.

We asked participants to describe their understanding of the main idea and important details, which was then compared to our rubrics. We counted the number of facts in the rubrics that were mentioned by participants as their scores. We also recorded the facts that do not appear in the rubrics. If a fact was frequently mentioned (e.g., more than half of the participants mentioned it), we reconsidered the importance of this fact and added it to the rubrics if the majority of the study team thought it was important.

As for the spatial information depiction task, we asked participants to draw a picture of the story scene to measure their memorability of spatial details. We specifically mentioned 2 requirements: (a) *object existence*: participants should capture as many spatial elements in the story as possible, including the objects, annotations, and the viewer path (the change of attention); and (b) *object transform*: for each spatial element, provide its position, rotation/orientation and scale/size as accurate as possible. The rubrics is attached in the supplementary materials.

We did not expect participants to draw the exact shape of elements. They were encouraged to use basic geometries (e.g. lines and rectangles) and text descriptions to depict the elements.

Finally, we logged the story session duration. This was to understand if the four combinations of conditions would cause any differences in the time audiences stayed in the immersive stories.

5.6 Hypotheses

Based on the characteristics in Figure 9, we proposed the following hypotheses.

Story Session Duration. We hypothesize that active navigation will result in longer story session durations compared to passive navigation. The interactive elements in active navigation require users to engage physically with the environment, adding time to the experience. Prior studies have shown that increased interaction in VR generally extends session time due to higher engagement levels and the need for precise actions in immersive environments [38, 66]. This additional engagement likely leads participants to spend more time exploring the environment and interacting with the content, especially in active conditions.

Additionally, Ego viewpoints are likely to further increase session duration, as spatial navigation in egocentric environments tends to be more complex and challenging, requiring users to explore the scene more actively, which can lead to extended engagement [79].

Story Understanding. We hypothesize that active navigation will outperform passive navigation on both content comprehension and spatial information depiction, as embodied interaction has been shown to improve cognitive engagement and memory retention [47, 82]. Active navigation encourages physical movement, which enhances attention to both content and spatial information, leading to better comprehension and recall [66].

We also expect that exocentric viewpoints will result in better content comprehension because they provide an overview of the entire scene, facilitating the integration of key story elements and improving factual recall [89]. Egocentric viewpoints will lead to better spatial understanding, as being immersed in the scene helps participants build stronger spatial awareness and memory, which is supported by the "memory palace" technique [36, 41].

Spatial Immersion. We expect that Ego viewpoints combined with active navigation will result in the highest levels of spatial immersion. Active navigation increases the sense of presence and control, while Ego viewpoints create a fully immersive experience by surrounding participants with 3D elements [70]. Prior studies suggest that immersion is enhanced when users feel directly involved in the environment, with Ego viewpoints offering a more embodied experience [13].

While active navigation may increase perceived workload due to the effort required to interact with the environment, we hypothesize that the enhanced sense of presence and engagement outweighs the additional cognitive demand, especially when paired with the Ego viewpoint [24]. This combination offers an optimal balance between immersion and interactivity, despite a potential trade-off with workload [66].

6 Results

In this section, we reported the study results. We performed the statistical analysis to test significances on the comprehension and rating data collected from three after-session tasks and subjective comments from final interviews. Significance values were reported for $p < 0.1$ (.) (dashed line in figures), $p < 0.05$ (*), $p < 0.01$ (**), and $p < 0.001$ (***) .

6.1 Story Session Durations

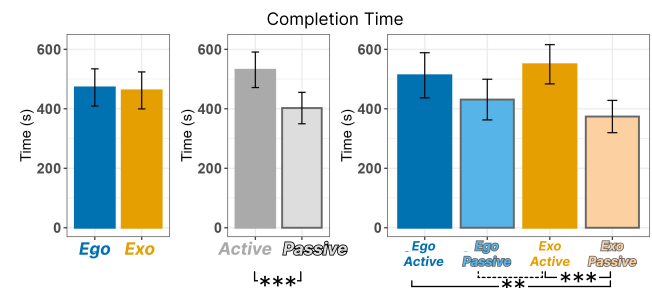


Figure 10: Average Story Session Durations. The dashed line indicates the marginal significance ($p < 0.1$) and the solid line indicates statistical significance ($p < 0.05$).

We collected the duration of each story session, defined as the total time the headset was worn, excluding any pauses. It is shown

in Figure 10. **Participants spent significantly more time on active navigation cases than passive ones, while the duration between ego and exo viewpoints was not salient.** A Shapiro-Wilk normality test confirmed the data followed a normal distribution ($p < 0.05$). We ran a General Linear Mixed Model (GLMM) with a Gaussian link function, using *Viewpoint* (ego/exo), *Navigation* (active/passive), their interaction (*Viewpoint* \times *Navigation*), and participant *Group* as fixed effects, with participant ID as a random effect. Results showed significant effects only for *Viewpoint* ($p < 0.001$), and a pairwise post-hoc Tukey's Honest Significant Difference (HSD) test confirmed a significant difference in duration between active and passive conditions in the exo viewpoint.

6.2 Story Understandings

We assessed story understanding through content comprehension and spatial information depiction tasks. We found that **ego viewpoint showed significantly better performance on spatial information depiction.**

Content Comprehension. The content comprehension task consists of three sets of questions: main idea memorization, detail memorization, and random inspirations. They have a total score of 3, 5 and 1, respectively. Each score corresponds to an idea or fact in the first 2 sets. For the random inspiration, participants would get the score if they mentioned anything derived from the story (e.g., linking to their own experiences or adding their own opinions). The only case in which they would lose this point is when they simply restated the story content (e.g. "I learned [idea/fact of the story] is important").

As Figure 11 shows, the exo and passive conditions had a higher memorization performance, while the active condition achieved better random inspiration. In terms of the four combinations, exo+passive was the best on two memorizations, whereas ego+active had the highest score in random inspirations. We then tested the normality of the data using a Shapiro-Wilk normality test, which showed that the data followed a normal distribution ($p < 0.05$). Thus, we ran a GLMM with a Gaussian link function. The results showed only a marginal significance of *Viewpoint* on detail memorization ($p = 0.09$).

Spatial Information Depiction. We evaluated participants' drawings for object existence and object transform. Figure 12 showed ego outperformed exo, especially in the ego+active condition. We also observed that participants had the worst memorability of spatial information in exo+active. We then tested the normality of the data using Shapiro-Wilk normality test, which showed that the data followed a normal distribution ($p < 0.05$). Thus, we ran a GLMM with a Gaussian link function. Results showed that only *Viewpoint* has marginal significance on object existence ($p = 0.09$) and significance on object transform ($p < 0.05$). It also showed the significance of *Viewpoint* \times *Navigation* on object transform ($p < 0.05$). A post-hoc Tukey's HSD test of *Viewpoint* \times *Navigation* on object transform showed that ego+active has a significantly higher score than exo+active.

Result Summary. Our hypothesis was partially supported by the results. We observed significantly better spatial information depiction in cases using the ego viewpoint. However, the results did not support the hypothesis that the exo viewpoint leads to better

content comprehension, and navigation did not play a significant role in story understanding.

6.3 Spatial Immersion

We collected the subjective ratings on spatial immersion for each case using a Presence Questionnaire (PQ) and NASA-TLX. We found that **active navigation enhances the sense of presence compared to passive navigation. It does not greatly increase workload when combined with ego viewpoint.**

Sense of Presence The 19 PQ questions can be classified into 5 sub-categories: Realism (question 3, 4, 5, 6, 7, 10, 13), Possibility to Act (question 1, 2, 8, 9; Act for short), Quality of Interface (question 14, 17, 18, with scores reversed; Quality for short), Possibility to Examine (question 11, 12, 19; Examine for short) and Self-Evaluation of Performance (question 15, 16; Evaluation for short).

We ran a Shapiro-Wilk normality test on each sub-category, which showed that the PQ responses did not follow a normal distribution on any sub-category. A follow-up Kolmogorov-Smirnov test showed that the data followed a Gamma distribution on both viewpoint and navigation factors. Therefore, we used a GLMM model with a Log-Gamma link function. Results showed the significance of *Navigation* and *Viewpoint* \times *Navigation* on 4 sub-categories (Realism, Act, Quality and Examine), and no significance of *Viewpoint* and *Group*. A pairwise post-hoc Tukey's HSD test on *Viewpoint* \times *Navigation* further showed that active has significantly higher realism, possibility to act and examine scores than passive, particularly in ego viewpoint. Specifically, ego+active received the highest average score on four sub-categories (Realism, Act, Examine and Evaluation). On Act and Evaluation, the post-hoc Tukey test showed that ego+active is significantly better than two other combinations.

Perceived Workload We summarized the NASA-TLX responses in Figure 14. Expectedly, active had a higher overall workload than passive. We did the same step to test the data distribution and found that it also followed a Gamma distribution. A GLMM with Log-Gamma link function confirmed our observation with the significance on *Navigation* ($p < 0.001$) and *Viewpoint* \times *Navigation* ($p < 0.05$). It did not show any significance of other effects. A follow-up post-hoc Tukey's HSD test on *Viewpoint* \times *Navigation* further showed the significantly larger workload of exo+active than others, especially in physical demand.

Result Summary. Our hypothesis was supported by the results. The Ego+Active condition produced the strongest sense of presence without resulting in the highest overall workload. However, we also observed that the heavy workload associated with active navigation might not be justified. In the Exo+Active condition, participants felt the least confident about their performance during story sessions. This suggests that the combined effects of viewpoint and navigation are more influential on story immersion than each factor individually.

6.4 Interview Analysis

Finally, we collected and analyzed the final interviews about users' subjective experiences of reading stories in VR. Overall, We found a trend that **most participants preferred either ego+active or exo+passive stories and reported less favorable opinions**

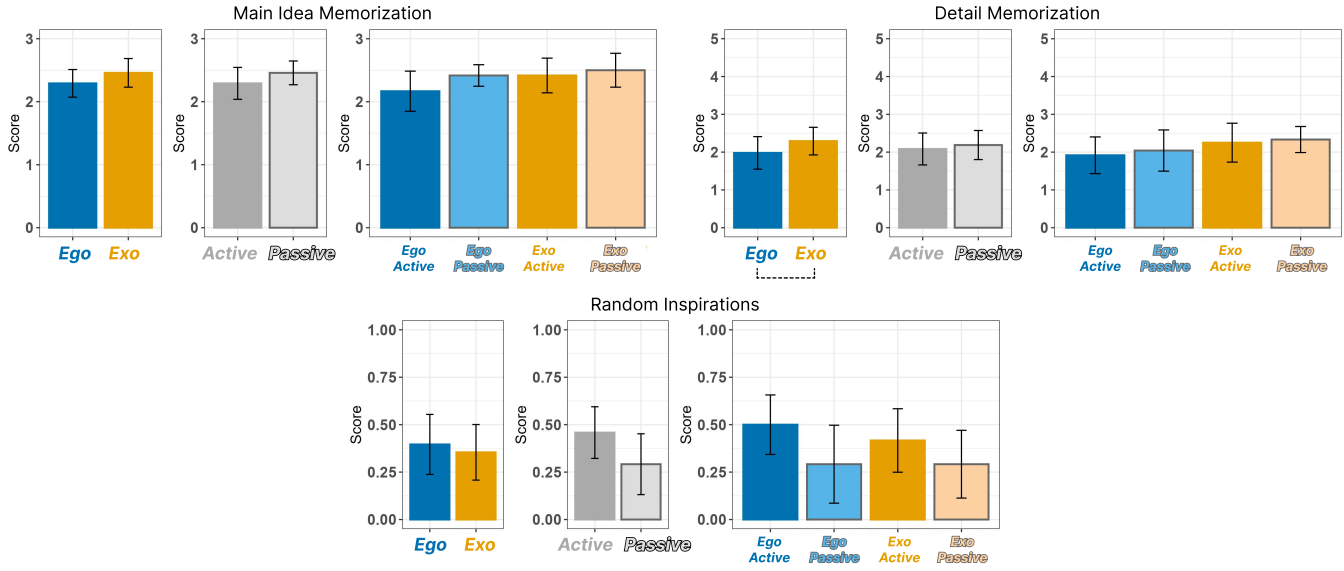


Figure 11: Average Sub-Scores of Content Comprehensions. The dashed line indicates the marginal significance ($p < 0.1$).

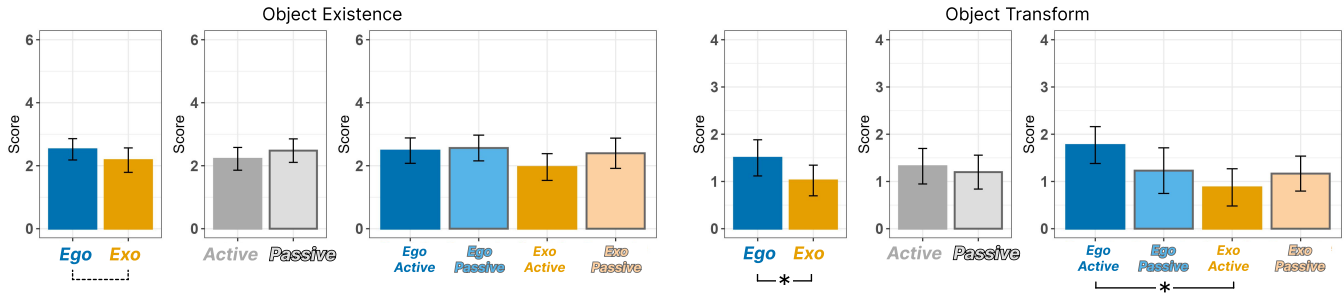


Figure 12: Average Sub-Scores of Spatial Information Depiction. The dashed line indicates the marginal significance ($p < 0.1$) and the solid line indicates statistical significance ($p < 0.05$).

towards the other two combinations. We summarized the following reasons based on their comments.

Active navigation provides freedom for explorations in the ego viewpoint. Many participants favored active navigation primarily because of the ability to explore and play with the story elements. This is particularly true in two ego stories. For example, P4, P7, P21 and P23 all mentioned that the COVID story (Case 4) allowed them to freely move in the scene so that they could view the visualizations of airflows from different perspectives and understand the spread route of contaminants in the air. Also, P5, P6, P11 and P21 strongly endorsed the window open and fan installation operations as it gave them a sense of control over the story. Similar comments appeared in the Tulsa story (Case 3) as well. Participants like the active navigation as “it is like an immersive exhibition in the museum” (P4, P12, P20). This matches the highest PQ score of ego+active, as it greatly enhanced the diversity of interaction and range of movement in the story scene.

Passive navigation is focused and efficient when reading stories in exo viewpoint. While a considerable number of participants paid attention to user experiences, some prioritized the understanding and memorability of story content. Thus, they preferred the exo+passive condition because “information was concentrated in exo (viewpoint), and passive (navigation) had less workload than active (navigation)” (P1, P9, P13, P14, P19). We observed that this group of participants did not have too much VR expertise, which increased their difficulty of active navigations and caused too much workload that impeded their focus (P19). With passive navigation, participants could quickly move to the next narratives without being trapped by VR interactions.

Passive navigation resulted in overall negative user experiences in ego viewpoint. The ego+passive condition received the most amount of critiques. Moving passively in an egocentric viewpoint (e.g. in a room) constrained the exploration participants could have. More seriously, it caused strong motionsickness for some participants (P1, P7). P7 even requested to pause the session and take off the VR headset for some rest. This indicates that our

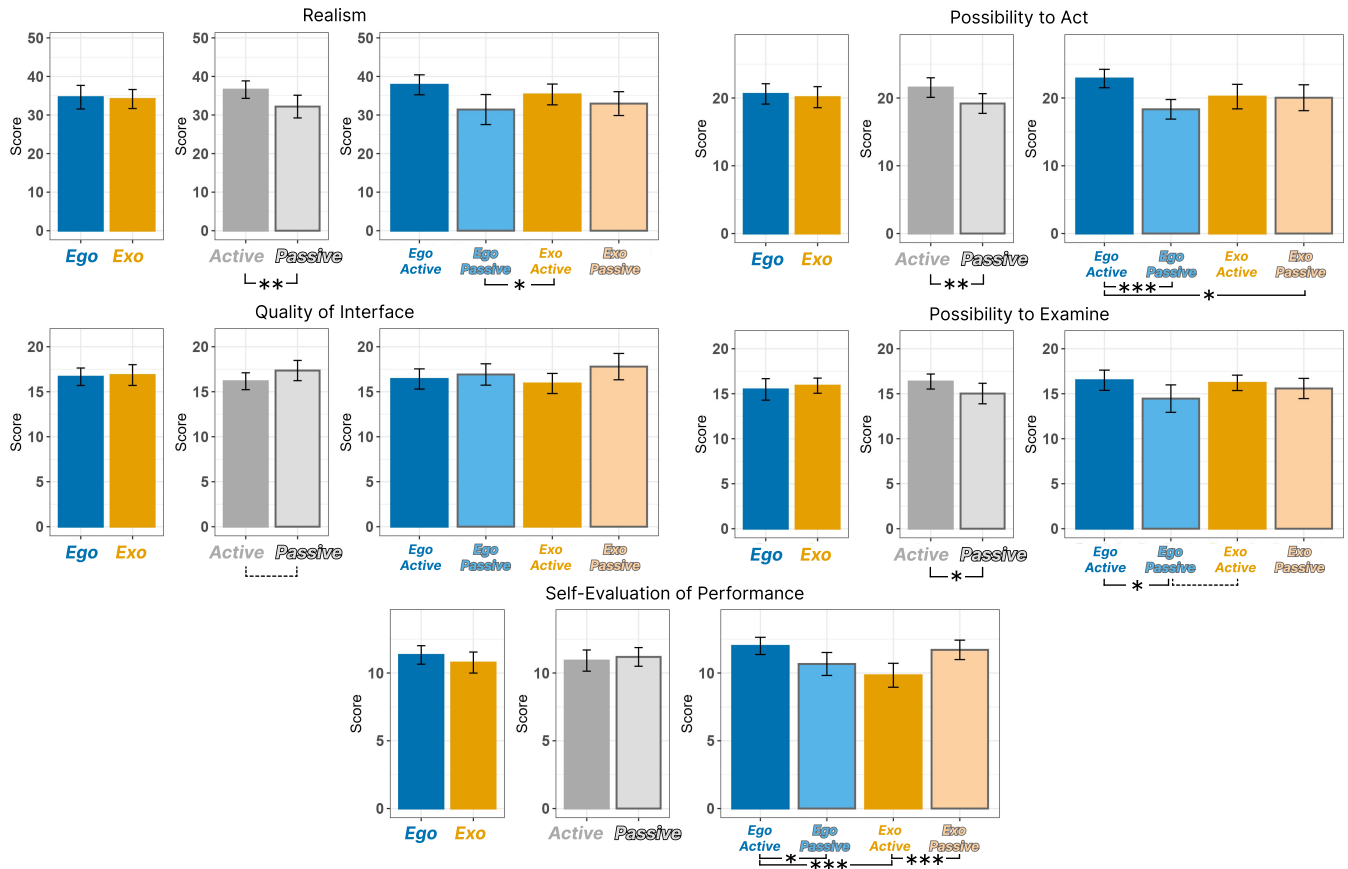


Figure 13: Average Sub-Scores of Presence Questionnaire. The dashed line indicates the marginal significance ($p < 0.1$) and the solid line indicates statistical significance ($p < 0.05$).

implementation of FOV reduction is not enough to alleviate motionsickness. Other techniques are necessary to be incorporated. Still, we received some positive feedback from participants who were less sensitive to movements in VR, who mentioned this experience was similar to a Universal Studio tour on a club car (P4, P8). However, the overall current implementation of ego+passive is less acceptable to most participants.

Active navigation increases the difficulty to keep track of spatial information in exo viewpoint. As mentioned in feedback to exo+passive, active navigation introduced too much workload for VR non-experts in 2 exo stories. Another side effect we consistently received is the loss of object transform. This was particular the case for the 911 story (Case 1). P8, P9 and P23 reported that they could not tell WTC position as they rotated the map. This matches the lowest scores of exo+active in the spatial information depiction task.

7 Key Findings and Discussion

In this section, we summarize some key findings and compare our results with prior works to discuss the common and unique findings. We also discuss different expectations on immersive stories based on our participants feedback.

7.1 How does viewpoints influence story memorability?

Our study shows that the viewpoint is a more decisive factor than the navigation in story memorability. **The exocentric viewpoint is slightly better than the egocentric one in the overview and detail memorization.** This result matches the discovery of some prior works, in which the exocentric viewpoint benefits analytical tasks in immersive environments [35, 88, 89], but the advantage in our study is not as significant. Our anticipated reason is that the exocentric viewpoint provide comprehensive information within audiences' view, which alleviates the problem of targeting story elements in space. However, this also causes a larger information density and more occlusions than the egocentric viewpoint. The relatively smaller scale of visual representations in exo viewpoints increases the attention to the story textual content received and prolongs the retention time in the memory so that they can did a better content comprehension. However, at the same time, audiences could become less sensitive to content transitions and therefore they might miss some important details. Overall, our results indicate that the exocentric viewpoint can be a better design if the primary goal is to promote the understanding and retention of narratives, and aforementioned problems should be considered to further improve the story.

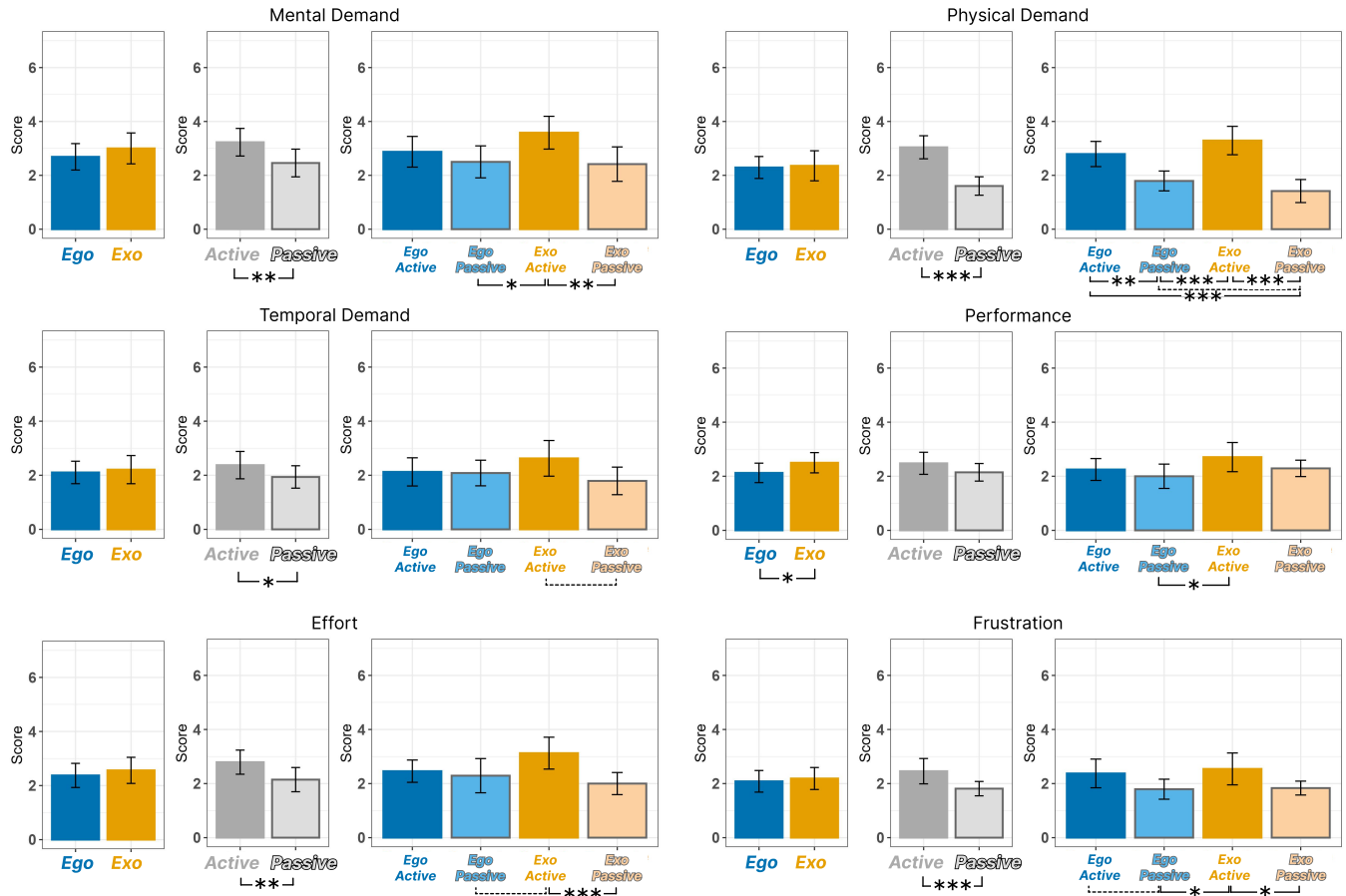


Figure 14: Average Sub-Scores of NASA-TLX Questionnaire. The dashed line indicates the marginal significance ($p < 0.1$) and the solid line indicates statistical significance ($p < 0.05$).

In contrast, **the egocentric viewpoint demonstrated better memorability on spatial information than exocentric viewpoint**. Similar result is also found in Krokos et al.'s [36] and Yang et al.'s [85] work. Objects in an egocentric viewpoint are often larger and more visible than the exocentric one. Resultedly, audiences can better recognize transforms (i.e. position, rotation and scale) and capture their changes. More importantly, audiences are involved as part of an egocentric story. This enable them to use themselves as a frame of reference to infer transforms. In exocentric stories, they have to refer to one central element (e.g. the base map in subsection 4.2) whose significant transform change can greatly distract the spatial information perception since audiences need to find a new frame of reference every time. Thus, the egocentric viewpoint is beneficial for those authors who want to convey understandings and perceptions related to the changes 3D spatial transforms (e.g. the renovation process of an office).

7.2 How does navigation influence spatial immersion?

As for the spatial immersion in our cases, navigation plays a more critical role than viewpoint. **Active navigation significantly enhances the sense of presence compared to passive, but their disparity of perceived workload varies between egocentric and exocentric conditions**. Specifically, we observed a substantial increase in workload for exocentric stories, while this effect was not as pronounced for egocentric ones.

This finding aligns with Lages et al.'s study [38], which showed that walking-based navigation improved performance for users who lacked proficiency in VR manipulation. Our results suggest that the need for higher precision in finding specific angles in exocentric scenarios (e.g., subsection 4.2, subsection 4.3, and the grab-and-move tasks in Lages et al.'s study) may explain the increased workload. The "changing frame of reference" issue discussed in subsection 7.1 adds to the demand for spatial awareness and motor skills to achieve the correct angle, resulting in higher effort and frustration, as reflected in the study.

In contrast, egocentric stories primarily use teleportation for navigation, where the most challenging task is locating the target

position. This task is not much more demanding than clicking a controller button for most participants. Consequently, we did not observe a significant increase in workload for active navigation in egocentric stories.

When designing navigation for VR stories, authors should consider the story’s viewpoint based on these findings. However, focusing solely on individual navigation steps is insufficient. Ferguson et al. [18], in their investigation of VR educational games, recommend that active navigation should seamlessly integrate with the story content and involve minimal learning curves. We argue that natural active navigation in VR storytelling should combine story progression with user exploration.

7.3 What are audiences’ expectations on VR stories?

From our study, we identified a significant difference in audience priorities, which can be classified as either **information-oriented** or **experience-oriented**. This distinction greatly influences their expectations for VR stories.

Information-oriented audiences focus on discovering and understanding detailed content. They prefer full control over the information they consume and favor clear, structured layouts. This group tends to prefer author-driven experiences, where the narrative is pre-designed to ensure they do not miss key details. For these users, an exocentric viewpoint and passive navigation are ideal, as they offer greater information clarity and a well-organized, linear progression with minimal interaction.

On the other hand, experience-oriented audiences value a more user-driven experience. They seek control over VR elements and prefer embodied interactions, even if it means potentially missing parts of the story. These users often desire a personalized experience that supports multiple storylines for free exploration. An egocentric viewpoint and active navigation are better suited to this group, as they encourage deeper engagement. These audiences can be motivated to “unlock” different storylines through interactions, maximizing their freedom to explore the narrative.

7.4 Summary and Generalizability

Our study underscores the significant impact of viewpoint and navigation on enhancing the memorability and spatial immersion of immersive stories. Among the tested conditions, Ego+Active and Exo+Passive were the most effective. The Ego+Active condition enhanced user immersion and creativity by allowing participants to actively explore and interact within the story, providing a strong sense of presence without significantly increasing the perceived workload. In contrast, the Exo+Passive condition improved content comprehension and recall by concentrating information and reducing cognitive load, making it ideal for users seeking a focused and efficient storytelling experience.

Immersive storytelling varies significantly across fields such as journalism, education, and entertainment, each with distinct objectives and audience expectations. In journalism, where clarity and factual accuracy are paramount, employing an ego viewpoint can enhance spatial understanding and engagement, making complex stories more accessible and memorable. In educational contexts, balancing ego and exo viewpoints can promote active learning and

critical thinking by aligning storytelling approaches with educational goals. In entertainment, where immersion and emotional engagement are crucial, optimizing the combination of viewpoints and navigation styles can deepen the audience’s connection to the narrative, enhancing their overall experience. Our preliminary exploration offers practical implications and guidance for making design decisions regarding viewpoint and navigation in immersive storytelling for these applications.

Additionally, future immersive stories could benefit from offering flexible viewpoints and navigation options to meet diverse audience needs for information and experiences, providing a personalized experience. For example, audiences might use an exocentric viewpoint to gain an overview and then “zoom in” on areas of interest for more detailed information from an egocentric view. Similarly, active navigation can be advantageous for audiences proficient in VR who seek more engagement and immersion, while passive navigation provides a suitable way for those less familiar with VR or who prefer to focus solely on story content to control story progression.

8 Limitations and Future Work

We adapted four selected web-based stories into immersive environments. To guide this adaptation, we conducted a formative user study to identify critical design considerations and explored various design options. Our exploration was not exhaustive; it aimed to facilitate the adaptation process by focusing on standard, off-the-shelf VR solutions. This project concentrated on specific aspects of viewpoint and navigation: ego vs. exo and active vs. passive. While we believe these are unique and essential areas to investigate, we acknowledge a much broader design space for immersive storytelling beyond these aspects. For instance, more sophisticated VR locomotion techniques [16], such as redirected walking, could enhance the active navigation experience. Similarly, advanced visual guidance methods [39], such as those for out-of-view scenarios [44, 59] and multisensory approaches [50], offer additional potential for enhancing user experience. We did not incorporate these techniques, as our primary focus was the user study and providing easily accessible techniques available on standard platforms. Building a comprehensive design space to inform future immersive story design is crucial, and we consider this an area for future work.

Our study materials were adapted from web-based stories that were initially designed for web browsers and contained design patterns specifically tailored for web environments [5]. It is unclear whether these design patterns are still applicable in immersive environments or whether we need to develop unique and more appropriate design patterns for immersive stories. Studying the adaptation or innovation of these design patterns for immersive stories has the potential to bring about a more native immersive experience. Furthermore, the web-based stories we adapted primarily follow a linear, single-scene structure. However, complex stories with multi-path or branching narratives also exist. While our studied scenario can serve as a foundational unit for these more complex stories, and our findings should be partially applicable, especially within their individual scenes. There is a need to explore transitions between scenes and more intricate active navigation, such as manipulating and selecting narrative branches.

Through our study, we found that exocentric and egocentric viewpoints each have their own strengths and limitations. We considered viewpoint as an exclusive design element, testing it in a controlled manner. There is a great opportunity to design immersive stories that leverage both viewpoints to complement each other. In movies and data videos, cinematic techniques like camera movements and zooms have been used to switch focuses and direct audiences' attention [3, 42, 84]. Similar incorporation such as viewing some story pieces from an exocentric view and others from an egocentric view is also promising to investigate. Our study results can provide guidelines on which parts of the story to use which viewpoint. To achieve this, smooth transition techniques are necessary to allow seamless and meaningful switching between viewpoints, which we aim to explore in our next research phase.

Finally, one important aspect of active vs. passive interaction is the agency of the readers. In our study, readers were given the freedom to navigate by themselves with guidance. However, there are other types of freedom that could be provided to readers, such as allowing them to freely decide their focus or even rearrange the storyline. This may empower readers but also risks them getting lost, particularly due to the more flexible interactivity in immersive environments. Balancing exploration and narrative control is a fundamental question in storytelling [73]. Studying the appropriate threshold or mechanism for this balance is an important research thread in immersive storytelling.

9 Conclusion

In this paper, we presented a user study investigating the effect of viewpoint and navigation on spatial immersion and understanding in immersive stories. To prepare the study materials, we first elicited design considerations from collected 3D spatial web stories. We then adapted four selected web stories to immersive environments based on these design considerations. Finally, we conducted a user study to empirically investigate the effect of viewpoints and navigation on spatial immersion and understanding. Our results showed a marginal significance of viewpoints on story understanding and a strong significance of navigation on spatial immersion, with preferences for the Ego+Active and Exo+Passive cases. Our study indicates the associations between viewpoints and navigation and provides a preliminary exploration of their individual and combined effects on VR storytelling, which could benefit future immersive storytelling designs.

References

- [1] Parastoo Abtahi, Mar Gonzalez-Franco, Eyal Ofek, and Anthony Steed. 2019. I'm a giant: Walking in large virtual environments at high speed gains. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13.
- [2] Fereshteh Amini, Matthew Brehmer, Gordon Bolduan, Christina Elmer, and Benjamin Wiederkehr. 2018. Evaluating data-driven stories and storytelling tools. In *Data-driven storytelling*. AK Peters/CRC Press, 249–286.
- [3] Fereshteh Amini, Nathalie Henry Riche, Bongshin Lee, Christophe Hurter, and Pourang Irani. 2015. Understanding Data Videos: Looking at Narrative Visualization through the Cinematography Lens. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) (CHI '15). Association for Computing Machinery, New York, NY, USA, 1459–1468. <https://doi.org/10.1145/2702123.2702431>
- [4] Ambre Assor, Arnaud Prouzeau, Pierre Dragicevic, and Martin Hachet. 2024. Augmented Reality Waste Accumulation Visualizations. *ACM Journal on Computing and Sustainable Societies* 2, 2 (2024), 1–29.
- [5] Benjamin Bach, Moritz Stefaner, Jeremy Boy, Steven Drucker, Lyn Bartram, Jo Wood, Paolo Ciuccarelli, Yuri Engelhardt, Ulrike Koeppen, and Barbara Tversky. 2018. Narrative design patterns for data-driven storytelling. In *Data-driven storytelling*. AK Peters/CRC Press, 107–133.
- [6] N. Baharuddin, H. Rosli, and Mohd. Suhaimi Juhan. 2023. The Creative Process: Developing Visual Storytelling through Design Thinking. *International Journal of Academic Research in Business and Social Sciences* (2023). <https://doi.org/10.6007/ijarbss/v13-i10/19113>
- [7] Sojung Bahng, Ryan M Kelly, and Jon McCormack. 2020. Reflexive VR storytelling design beyond immersion: facilitating self-reflection on death and loneliness. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. Association for Computing Machinery, New York, NY, USA, 1–13.
- [8] Nick Bartzokas, Mika Gröndahl, Karthik Patanjali, Miles Peyton, Bedel Saget, and Umi Syam. 2021. Why Opening Windows Is a Key to Reopening Schools. <https://www.nytimes.com/interactive/2021/02/26/science/reopen-schools-safety-ventilation.html>
- [9] Elliot Bentley. 2021. How the 9/11 Attacks Remade New York City's Financial District. <https://www.wsj.com/articles/how-the-9-11-attacks-remade-new-york-citys-financial-district-11630501200>
- [10] Mehmet Ilker Berkman and Güven Çatak. 2021. I-group presence questionnaire: Psychometrically revised English version. *Mugla Journal of Science and Technology* 7 (2021), 1–10.
- [11] Bloomberg. 2024. Graphics - Bloomberg. <https://www.bloomberg.com/graphics>
- [12] Matthew Conlen, Jeffrey Heer, Hillary Mushkin, and Scott Davidoff. 2023. Cinematic techniques in narrative visualization. *arXiv preprint arXiv:2301.03109* (2023).
- [13] James J Cummings and Jeremy N Bailenson. 2016. How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media psychology* 19, 2 (2016), 272–309.
- [14] Keshav Dasu, Yun-Hsin Kuo, and Kwan-Liu Ma. 2023. Character-Oriented Design for Visual Data Storytelling. *IEEE Transactions on Visualization and Computer Graphics* (2023).
- [15] Kylie Davidson, Lee Lisle, Kirsten Whitley, Doug A Bowman, and Chris North. 2022. Exploring the evolution of sensemaking strategies in immersive space to think. *IEEE transactions on visualization and computer graphics* 29, 12 (2022), 5294–5307.
- [16] Massimiliano Di Luca, Hasti Seifi, Simon Egan, and Mar Gonzalez-Franco. 2021. Locomotion vault: the extra mile in analyzing vr locomotion techniques. In *Proceedings of the 2021 CHI conference on human factors in computing systems*. 1–10.
- [17] Zhi-Chao Dong, Wenming Wu, Zenghao Xu, Qi Sun, Guanjie Yuan, Ligang Liu, and Xiao-Ming Fu. 2021. Tailored reality: Perception-aware scene restructuring for adaptive vr navigation. *ACM Transactions on Graphics (TOG)* 40, 5 (2021), 1–15.
- [18] Chris Ferguson, Egon L Van den Broek, and Herre Van Oostendorp. 2020. On the role of interaction mode and story structure in virtual reality serious games. *Computers & Education* 143 (2020), 103671.
- [19] National Geographic. 2024. National Geographic. <https://www.nationalgeographic.com>
- [20] Google. 2016. Beyond the Map, Rio De Janeiro - Introduction. <https://www.youtube.com/watch?v=gN9-ZdNC6s0>
- [21] Esther Greussing, S. Kessler, and H. Boomgaarden. 2020. Learning From Science News via Interactive and Animated Data Visualizations: An Investigation Combining Eye Tracking, Online Survey, and Cued Retrospective Reporting. *Science Communication* 42 (2020), 803 – 828. <https://doi.org/10.1177/1075547020962100>
- [22] The Guardian. 2024. Graphics | The Guardian. <https://theguardian.com/tone/graphics>
- [23] Brian D Hall, Lyn Bartram, and Matthew Brehmer. 2022. Augmented chironomia for presenting data to remote audiences. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*. Association for Computing Machinery, New York, NY, USA, 1–14.
- [24] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, 139–183.
- [25] Matthias Hoppe, Andrea Baumann, Patrick Chofor Tamunjoh, Tonja-Katrin Machulla, Pawel W Woźniak, Albrecht Schmidt, and Robin Welsch. 2022. There is no first- or third-person view in virtual reality: Understanding the perspective continuum. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–13.
- [26] Helen H Huang, Hanspeter Pfister, and Yalong Yang. 2023. Is embodied interaction beneficial? A study on navigating network visualizations. *Information Visualization* 22, 3 (2023), 169–185.
- [27] Jessica Hullman and Nick Diakopoulos. 2011. Visualization rhetoric: Framing effects in narrative visualization. *IEEE transactions on visualization and computer graphics* 17, 12 (2011), 2231–2240.
- [28] Sports Illustrated. 2017. Chapter 1: Preparing to Hike to the Top of Mount Everest in 4KVR | 360 Video | Sports Illustrated. https://www.youtube.com/watch?v=5_

- YNjD2HWVI&t=242s
- [29] Sungwon In, Eric Krokos, Kirsten Whitley, Chris North, and Yalong Yang. 2024. Evaluating Navigation and Comparison Performance of Computational Notebooks on Desktop and in Virtual Reality. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–15.
 - [30] Sungwon In, Tica Lin, Chris North, Hanspeter Pfister, and Yalong Yang. 2023. This is the table i want! interactive data transformation on desktop and in virtual reality. *IEEE Transactions on Visualization and Computer Graphics* (2023).
 - [31] Petra Isenberg, Bongshin Lee, Huamin Qu, and Maxime Cordeil. 2018. Immersive visual data stories. *Immersive analytics* (2018), 165–184.
 - [32] Maria Kandaurova and Seung Hwan Mark Lee. 2019. The effects of Virtual Reality (VR) on charitable giving: The role of empathy, guilt, responsibility, and social exclusion. *Journal of Business Research* 100 (2019), 571–580.
 - [33] Robert Kosara and Jock Mackinlay. 2013. Storytelling: The next step for visualization. *Computer* 46, 5 (2013), 44–50.
 - [34] Matthias Kraus, Karsten Klein, Johannes Fuchs, Daniel A Keim, Falk Schreiber, and Michael Sedlmair. 2021. The value of immersive visualization. *IEEE computer graphics and applications* 41, 4 (2021), 125–132.
 - [35] Matthias Kraus, Niklas Weiler, Daniela Oelke, Johannes Kehrer, Daniel A Keim, and Johannes Fuchs. 2019. The impact of immersion on cluster identification tasks. *IEEE Transactions on Visualization and Computer Graphics* 26, 1 (2019), 525–535.
 - [36] Eric Krokos, Catherine Plaisant, and Amitabh Varshney. 2019. Virtual memory palaces: immersion aids recall. *Virtual reality* 23, 1 (2019), 1–15.
 - [37] Shirish Kulkarni, Richard Thomas, Marlen Komorowski, and Justin Lewis. 2023. Innovating online journalism: New ways of storytelling. *Journalism Practice* 17, 9 (2023), 1845–1863.
 - [38] Wallace S Lages and Doug A Bowman. 2018. Move the object or move myself? walking vs. manipulation for the examination of 3D scientific data. *Frontiers in ICT* 5 (2018), 15.
 - [39] Daniel Lange, Tim Claudius Stratmann, Uwe Gruenefeld, and Susanne Boll. 2020. Hivefive: Immersion preserving attention guidance in virtual reality. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. 1–13.
 - [40] Benjamin Lee, Dave Brown, Bongshin Lee, Christophe Hurter, Steven Drucker, and Tim Dwyer. 2020. Data visualization: Enabling deeper understanding of data using virtual reality. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (2020), 1095–1105.
 - [41] Eric LG Legge, Christopher R Madan, Enoch T Ng, and Jeremy B Caplan. 2012. Building a memory palace in minutes: Equivalent memory performance using virtual versus conventional environments with the Method of Loci. *Acta psychologica* 141, 3 (2012), 380–390.
 - [42] Wenchao Li, Zhan Wang, Yun Wang, Di Weng, Liwenhan Xie, Siming Chen, Haidong Zhang, and Huamin Qu. 2023. GeoCamera: Telling Stories in Geographic Visualizations with Camera Movements. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (Hamburg, Germany) (CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 170, 15 pages. <https://doi.org/10.1145/3544548.3581470>
 - [43] E. S. D. Lima, Felipe João Gheno, and A. Viseu. 2020. Sketch-Based Interaction for Planning-Based Interactive Storytelling. *2020 19th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)* (2020), 154–162. <https://doi.org/10.1109/SBGames51465.2020.00029>
 - [44] Tica Lin, Yalong Yang, Johanna Beyer, and Hanspeter Pfister. 2021. Labeling out-of-view objects in immersive analytics to support situated visual searching. *IEEE transactions on visualization and computer graphics* 29, 3 (2021), 1831–1844.
 - [45] Weizhou Luo, Anke Lehmann, Hjalmar Widengren, and Raimund Dachselt. 2022. Where should we put it? layout and placement strategies of documents in augmented reality for collaborative sensemaking. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–16.
 - [46] K. Ma, Isaac Liao, Jennifer Frazier, H. Hauser, and H. Kostis. 2012. Scientific Storytelling Using Visualization. *IEEE Computer Graphics and Applications* 32 (2012), 12–19. <https://doi.org/10.1109/MCG.2012.24>
 - [47] Guido Makransky, Thomas S Terkildsen, and Richard E Mayer. 2019. Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and instruction* 60 (2019), 225–236.
 - [48] Mapbox. 2017. Mapbox. <https://www.mapbox.com/>
 - [49] Kim Marriott, Jian Chen, Marcel Hlawatsch, Takayuki Itoh, Miguel A Nacenta, Guido Reina, and Wolfgang Stuerzlinger. 2018. Immersive analytics: Time to reconsider the value of 3d for information visualisation. *Immersive analytics* (2018), 25–55.
 - [50] Miguel Melo, Guilherme Gonçalves, Pedro Monteiro, Hugo Coelho, José Vasconcelos-Raposo, and Maximino Bessa. 2020. Do multisensory stimuli benefit the virtual reality experience? A systematic review. *IEEE transactions on visualization and computer graphics* 28, 2 (2020), 1428–1442.
 - [51] Julián Méndez, Weizhou Luo, Rufat Rzayev, Wolfgang Büschel, and Raimund Dachselt. 2025. Immersive Data-Driven Storytelling: Scoping an Emerging Field Through the Lenses of Research, Journalism, and Games. *IEEE Transactions on Visualization and Computer Graphics* (2025).
 - [52] Seyedkoosha Mirhosseini, Parmida Ghahremani, Sushant Ojal, Joseph Marino, and Arie Kaufman. 2019. Exploration of large omnidirectional images in immersive environments. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 413–422.
 - [53] S M A Moin, Sameer Hosany, and J. O'Brien. 2020. Storytelling in destination brands' promotional videos. *Tourism Management Perspectives* 34 (2020), 100639. <https://doi.org/10.1016/j.tmp.2020.100639>
 - [54] Luiz Morais, Yvonne Jansen, Nazareno Andrade, and Pierre Dragicevic. 2020. Showing data about people: A design space of anthropographics. *IEEE Transactions on Visualization and Computer Graphics* 28, 3 (2020), 1661–1679.
 - [55] J. Motejlek and E. Alpay. 2023. The retention of information in virtual reality based engineering simulations. *European Journal of Engineering Education* 48 (2023), 929 – 948. <https://doi.org/10.1080/03043797.2022.2160968>
 - [56] K. Nash. 2018. Virtually real: exploring VR documentary. *Studies in Documentary Film* 12 (2018), 100 – 97. <https://doi.org/10.1080/17503280.2018.1484992>
 - [57] Matthew J Page, Joanne E McKenzie, Patrick M Bossuyt, Isabelle Boutron, Tammy C Hoffmann, Cynthia D Mulrow, Larissa Shamsseer, Jennifer M Tetzlaff, Elie A Akl, Sue E Brennan, et al. 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *bmj* 372 (2021).
 - [58] Yuliya Parshina-Kottas, Anjali Singhvi, Audra D. S. Burch, Troy Griggs, Mika Gröndahl, Lingdong Huang, Tim Wallace, Jeremy White, and Josh Williams. 2021. What the Tulsa Race Massacre Destroyed. <https://www.nytimes.com/interactive/2021/05/24/us/tulsa-race-massacre.html>
 - [59] Julian Petford, Iain Carson, Miguel A Nacenta, and Carl Gutwin. 2019. A comparison of notification techniques for out-of-view objects in full-coverage displays. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–13.
 - [60] The Washington Post. 2024. Visual Stories - The Washington Post. <https://www.washingtonpost.com/visual-stories/>
 - [61] The Pudding. 2024. The Pudding. <https://pudding.cool>
 - [62] Reuters. 2024. Reuters Graphics - Charts, Maps, Interactive Graphics and Videos. <https://www.reuters.com/graphics>
 - [63] Hugo Romat, Nathalie Henry Riche, Christophe Hurter, Steven Drucker, Fereshteh Amini, and Ken Hinckley. 2020. Dear pictograph: Investigating the role of personalization and immersion for consuming and enjoying visualizations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13.
 - [64] Kadek Ananta Satriadi, Barrett Ens, Maxime Cordeil, Tobias Czauderna, and Bernhard Jenny. 2020. Maps around me: 3d multiview layouts in immersive spaces. *Proceedings of the ACM on Human-Computer Interaction* 4, ISS (2020), 1–20.
 - [65] Edward Segel and Jeffrey Heer. 2010. Narrative visualization: Telling stories with data. *IEEE transactions on visualization and computer graphics* 16, 6 (2010), 1139–1148.
 - [66] J. Servotte, Manon Goosse, S. Campbell, N. Dardenne, Bruno Pilote, Ivan L. Simoneau, M. Guillaume, I. Bragard, and A. Ghuysen. 2020. Virtual Reality Experience: Immersion, Sense of Presence, and Cybersickness. *Clinical Simulation in Nursing* 38 (2020), 35–43. <https://doi.org/10.1016/j.ecns.2019.09.006>
 - [67] Anna Sheremetieva, Ihor Romanv, Sam Frish, M. Maksymenko, and Orestis Georgiou. 2022. Touch the Story: An immersive mid-air haptic experience. *2022 International Conference on Interactive Media, Smart Systems and Emerging Technologies (IMET)* (2022), 1–3. <https://doi.org/10.1109/IMET54801.2022.9929479>
 - [68] Jae-Eun Shin, Hayun Kim, Hyerim Park, and Woontack Woo. 2024. Investigating the Design of Augmented Narrative Spaces Through Virtual-Real Connections: A Systematic Literature Review. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–18.
 - [69] Jae-eun Shin, Hayun Kim, Callum Parker, Hyung-il Kim, Seoyoung Oh, and Woontack Woo. 2019. Is any room really ok? the effect of room size and furniture on presence, narrative engagement, and usability during a space-adaptive augmented reality game. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 135–144.
 - [70] Mel Slater. 2009. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364, 1535 (2009), 3549–3557.
 - [71] Charles D Stolper, Bongshin Lee, Nathalie Henry Riche, and John Stasko. 2018. Data-driven storytelling techniques: Analysis of a curated collection of visual stories. In *Data-driven storytelling*. AK Peters/CRC Press, 85–105.
 - [72] Justyna Świdrak, Orestis Georgiou, Domna Banakou, Marta Matamala-Gomez, Grzegorz Pochwatko, and Sofia Seinfeld. 2024. Beyond audiovisual: novel multi-sensory stimulation techniques and their applications. , 1491854 pages.
 - [73] Alice Thudt, Jagoda Walny, Theresia Gschwandtner, Jason Dykes, and John Stasko. 2018. Exploration and explanation in data-driven storytelling. In *Data-Driven Storytelling*. AK Peters/CRC Press, 59–83.
 - [74] Financial Times. 2024. Visual and data journalism. <https://www.ft.com/visual-and-data-journalism>
 - [75] Los Angeles Times. 2024. News from California, the nation and world. <https://www.latimes.com/>

- [76] The New York Times. 2022. Showcasing 31 Published Experiments in AR Storytelling. <https://rd.nytimes.com/projects/augmented-reality-storytelling/>
- [77] The New York Times. 2024. Graphics. <https://www.nytimes.com/spotlight/graphics>
- [78] Javlon Tursunov, Gregor Rozinaj, Ivan Minarik, and Marek Vanco. 2024. Creating a Metaverse-Based Meeting Room with Live Avatar Control and Web Communication. In *2024 International Symposium ELMAR*. IEEE, 57–60.
- [79] Austin Van Loon, Jeremy Bailenson, Jamil Zaki, Joshua Bostick, and Robb Willer. 2018. Virtual reality perspective-taking increases cognitive empathy for specific others. *PLoS one* 13, 8 (2018), e0202442.
- [80] Moises Velasquez-Manoff and Jeremy White. 2021. In the Atlantic Ocean, Subtle Shifts Hint at Dramatic Dangers. <https://www.nytimes.com/interactive/2021/03/02/climate/atlantic-ocean-climate-change.html>
- [81] W. Williams. 2019. Attending to the visual aspects of visual storytelling: using art and design concepts to interpret and compose narratives with images. *Journal of Visual Literacy* 38 (2019), 66 – 82. <https://doi.org/10.1080/1051144X.2019.1569832>
- [82] Margaret Wilson. 2002. Six views of embodied cognition. *Psychonomic bulletin & review* 9 (2002), 625–636.
- [83] Kaifeng Wu and Dana Karwas. 2024. The Metamorphosis of Storytelling: Time-based Interactivity in Virtual Reality Filmmaking. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–5.
- [84] Xian Xu, Leni Yang, David Yip, Mingming Fan, Zheng Wei, and Huamin Qu. 2022. From ‘Wow’ to ‘Why’: Guidelines for Creating the Opening of a Data Video with Cinematic Styles. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) (CHI '22). Association for Computing Machinery, New York, NY, USA, Article 599, 20 pages. <https://doi.org/10.1145/3491102.3501896>
- [85] Fumeng Yang, Jing Qian, Johannes Novotny, David Badre, Cullen D Jackson, and David H Laidlaw. 2020. A virtual reality memory palace variant aids knowledge retrieval from scholarly articles. *IEEE Transactions on Visualization and Computer Graphics* 27, 12 (2020), 4359–4373.
- [86] Leni Yang, Aoyu Wu, Wai Tong, Xian Xu, Zheng Wei, and Huamin Qu. 2023. Understanding 3d data videos: From screens to virtual reality. In *2023 IEEE 16th Pacific Visualization Symposium (PacificVis)*. IEEE, 197–206.
- [87] Leni Yang, Xian Xu, XingYu Lan, Ziyang Liu, Shunan Guo, Yang Shi, Huamin Qu, and Nan Cao. 2021. A design space for applying the freytag’s pyramid structure to data stories. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (2021), 922–932.
- [88] Yalong Yang, Maxime Cordeil, Johanna Beyer, Tim Dwyer, Kim Marriott, and Hanspeter Pfister. 2020. Embodied navigation in immersive abstract data visualization: Is overview+ detail or zooming better for 3d scatterplots? *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (2020), 1214–1224.
- [89] Yalong Yang, Bernhard Jenny, Tim Dwyer, Kim Marriott, Haohui Chen, and Maxime Cordeil. 2018. Maps and globes in virtual reality. In *Computer Graphics Forum*, Vol. 37. Wiley Online Library, 427–438.
- [90] Gürkan Yildirim, Serkan Yildirim, and Emrah Dolgunsöz. 2019. The effect of VR and traditional videos on learner retention and decision making. *World Journal on Educational Technology: Current Issues* (2019). <https://doi.org/10.18844/WJET.V11I1.4005>
- [91] Xinyi Yu and Lei Shi. 2018. A Comparative Research on Cognitive Effect of Infographic and Animation Presentation in News. *2018 International Joint Conference on Information, Media and Engineering (ICIME)* (2018), 73–76. <https://doi.org/10.1109/ICIME.2018.00024>
- [92] Zhenpeng Zhao and Niklas Elmqvist. 2023. The stories we tell about data: Surveying data-driven storytelling using visualization. *IEEE Computer Graphics and Applications* 43, 4 (2023), 97–110.
- [93] Ziyue Zhao, Yue Li, and Hai-Ning Liang. 2023. LeanOn: Simulating Balance Vehicle Locomotion in Virtual Reality. In *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 415–424.
- [94] Shuhua Zhou, Fei Xue, and Peiqin Zhou. 2005. VISUAL DIFFERENCES IN U.S. AND CHINESE TELEVISION COMMERCIALS. *Journal of Advertising* 34 (2005), 112 – 119. <https://doi.org/10.1080/00913367.2005.10639182>
- [95] Xiaoyan Zhou, Yalong Yang, Francisco Ortega, Anil Ufuk Batmaz, and Benjamin Lee. 2023. Data-driven Storytelling in Hybrid Immersive Display Environments. In *2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, 242–246.
- [96] Qian Zhu, Tao Lu, Shunan Guo, Xiaojuan Ma, and Yalong Yang. 2024. CompositingVis: Exploring Interactions for Creating Composite Visualizations in Immersive Environments. *IEEE Transactions on Visualization and Computer Graphics* (2024).
- [97] Qian Zhu, Linping Yuan, Zian Xu, Leni Yang, Meng Xia, Zhuo Wang, Hai-Ning Liang, and Xiaojuan Ma. 2024. From reader to experimenter: Design and evaluation of a VR data story for promoting the situation awareness of public health threats. *International Journal of Human-Computer Studies* 181 (2024), 103137.
- [98] S. Zollmann, Anthony Dickson, and Jonathan Ventura. 2020. CasualVRVideos: VR videos from casual stationary videos. *Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology* (2020). <https://doi.org/10.1145/3385956.3422119>