



# Engagement vs. Understanding: Comparing Immersive Virtual Reality and Desktop Displays for Climate Data Visualization

A. Arunkumar<sup>1</sup> , J. Liu<sup>2</sup>, and C. Bryan<sup>2</sup> 

<sup>1</sup>Northeastern University, Boston, United States, <sup>2</sup>Arizona State University, Tempe, United States

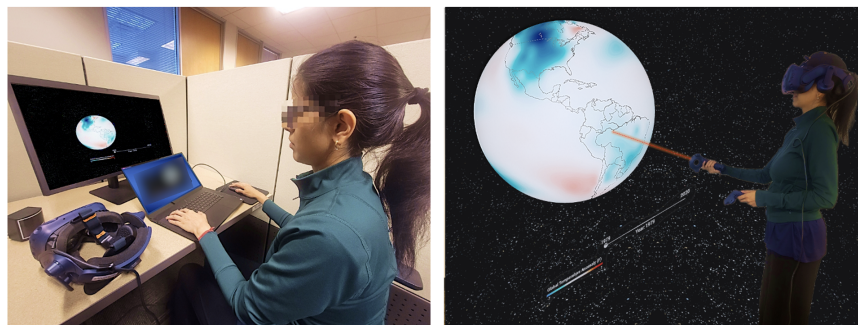


Figure 1: A user explores climate data projected onto a 3D sphere, with either a desktop display (left) or head-mounted Immersive VR (IVR; right). Desktop displays use a 24-inch monitor (3840 × 2160 resolution) with keyboard/mouse navigation, connected to a Razer laptop running Windows 11. IVR uses a VIVE Headset with VIVE VR Controllers.

## Abstract

*Immersive virtual reality (IVR) is increasingly used for scientific data visualization, with the expectation that greater immersion will enhance both engagement and understanding. However, prior research suggests a potential trade-off: IVR can heighten affective responses while impairing cognitive learning due to increased cognitive load. To examine this tension in abstract scientific visualization, we conducted a between-subjects study (N=84) comparing head-mounted display (HMD) VR with traditional desktop displays as participants explored three climate datasets on an interactive 3D globe. We measured cognitive and affective learning outcomes after each exploration trial and at a two-week follow-up. We additionally logged patterns of visualization exploration during trials, and measured user experience post-study and at two-week follow-up. Results reveal systematic divergences: desktop displays led to significantly higher recall accuracy and promoted more systematic exploration patterns, while HMD-based VR produced stronger short-term increases in climate concern and higher satisfaction ratings. Critically, affective changes largely reverted to baseline at follow-up regardless of modality. We further identify five distinct exploration strategies that emerge differentially across modalities and relate directly to learning outcomes. Overall, our findings highlight how modality choice should align with visualization goals, and offer actionable insights into designing effective scientific visualizations that balance cognitive and affective learning objectives. All data and materials are available at: <https://osf.io/24w7s/>.*

## CCS Concepts

• **Human-centered computing** → **Empirical studies in visualization; Virtual reality; Visual analytics;**

## 1. Introduction

Scientific visualization serves audiences with diverse goals, from exploratory data analysis requiring precision and systematic examination, to public communication emphasizing engagement and emotional resonance. Climate data visualization exemplifies this tension. As climate modeling and satellite observations generate

increasingly large and complex datasets [NSB\*08], effective visualizations must balance two competing demands: supporting accurate interpretation of spatial-temporal patterns while fostering the emotional engagement needed to motivate public concern and action [HLSC16, Mos14].

In response, news organizations, museums, and research institu-

tions have increasingly adopted immersive journalism, including data visualizations viewable via virtual reality (VR), augmented reality (AR), and 360° video for climate storytelling [BSK\*22, SL20]. Immersive VR (IVR) is often assumed to enhance both cognitive understanding and affective response through increased presence and immersion. While IVR reliably increases presence, enjoyment, and emotional response compared to desktop displays [KJD\*20, NEW\*20], it can also introduce additional cognitive load due to unfamiliar interactions and the demands of navigating stereoscopic 3D environments [MP05, Swe11, MOM19]. This can lead to declines in cognitive learning despite increased engagement [PM18, RT15], raising an unresolved question for scientific visualization: **does immersion promote or hinder comprehension during free-form exploration of abstract data?**

Despite growing interest in immersive visualization for science communication, existing work rarely examines cognitive and affective outcomes together [KJD\*20, PTS\*21], leaving unclear whether gains in one domain trade off with the other. Few studies include delayed measurements, even though retention and lasting attitude change are critical for climate communication [ODC\*20]. Additionally, while interaction patterns strongly shape insight formation in desktop visualization [DYL20], how these patterns differ across modalities (and how they relate to learning) remains largely unexamined. Finally, the role of multimodal cues such as spatialized audio in modulating cognitive and affective outcomes across display types has received little systematic attention.

To investigate these gaps, we conducted a between-subjects study (N=84) comparing an immersive head-mounted display (HMD) VR condition with a traditional desktop display condition (see Figure 1). Participants explored the same interactive visualization: a 3D globe displaying climate data on global temperature anomalies, sea surface temperature, and Arctic sea ice age. The desktop condition rendered the globe monoscopically on a standard monitor, while the HMD condition presented it stereoscopically with tracked controllers. Half of the participants in each modality experienced spatialized audio during exploration of one of the datasets. Our goal is not to test persuasive strategies, or advance climate-specific techniques, but to use climate data as a realistic and socially relevant case study to provide empirical evidence on how display modality shapes exploration behavior, cognitive learning, and affective response during free-form analysis of abstract scientific data. We assess outcomes immediately after the study and two weeks later across three dimensions: cognitive learning (recall accuracy and consistency), affective learning (climate change awareness and behavioral intentions), and interaction patterns (exploration strategies derived from logged behaviors). This multi-dimensional approach allows us to directly examine the relationship between engagement (hypothesized to be higher in IVR) and understanding (hypothesized to be higher in desktop displays) while accounting for individual differences in prior knowledge and baseline attitudes. Specifically, our contributions are:

- A systematic comparison of immersive and desktop modalities for scientific visualization (using climate data as an ecologically valid case study of abstract, globe-based data exploration), providing empirical evidence for the engagement–understanding

trade-off across cognitive, affective, and behavioral outcomes measured immediately and two weeks later (Sections 4, 5.2, 5.3).

- A characterization of exploration strategies in immersive and desktop environments, identifying five distinct interaction patterns and demonstrating how modality shapes exploration behavior and learning performance (Section 5.4).
- Design implications for modality selection in scientific data communication, outlining when desktop displays better support analytical understanding and long-term recall, when immersive VR is more effective for short-term engagement, and how multimodal cues (e.g., audio) provide primarily cognitive rather than affective benefits (Section 6).

As a preview, our findings reveal a consistent engagement–understanding trade-off. Desktop displays produced significantly higher short- and long-term recall and more systematic exploration patterns, while IVR generated stronger short-term affective shifts, particularly among participants initially unconcerned about climate change. These affective effects were transient: by two weeks, roughly 73% had reverted to baseline. Five distinct exploration strategies emerged across modalities, and were strongly predictive of cognitive performance. Additionally, spatialized audio modestly improved recall but had no meaningful effect on affective outcomes. Taken together, these findings demonstrate that no single modality optimizes both cognitive and affective learning, underscoring the need for deliberate modality choice when designing immersive scientific visualizations.

## 2. Related Work

Our work draws on research across immersive analytics, VR-based learning, immersive journalism, and interaction patterns in visualization. We organize this review around the central question motivating our study: when does immersion enhance and hinder learning during free-form exploration of abstract scientific data?

### 2.1. Immersive Technology for Data Visualization

Immersive analytics investigates how virtual and augmented reality (VR/AR) can support data exploration by offering large spatial workspaces, natural 3D interaction, and potential benefits for spatial memory [ECN\*22, LPED20, SEC\*20, RCC\*21]. The underlying premise is that physical space can act as external memory and reduce cognitive load during sensemaking [Kir95, AEN10]. However, evidence for these benefits is mixed, particularly for abstract or multidimensional scientific data. While 3D environments provide abundant display space [AEN10, LDC\*21], they may also introduce clutter or unnecessary complexity [CM02, EHR14]. Climate data, for instance, is often mapped onto 3D spheres [Tim19], yet it remains unclear whether stereoscopic rendering and head-tracked navigation actually improve learning outcomes relative to desktop displays. To address this gap, we directly compare desktop and HMD-based IVR presentations of the same 3D climate visualization, examining not only cognitive and affective learning but also the exploration strategies that emerge across modalities.

### 2.2. Virtual Reality and Learning Outcomes

The relationship between immersion and learning has been widely examined. Generative learning theory argues that deep understanding requires actively selecting, organizing, and integrating

information [FM16]. VR may support this process by increasing presence [SB18, SKO17] and flow [SB18], which can enhance attention and motivation [BHRRW16]. Yet empirical findings remain inconsistent: some studies report learning benefits in IVR [MBGM19, LAES18], while others find mixed or negative effects [LHP\*19, PM18, MTM19, MM02]. Directly relevant to our focus on recall, Hurter et al. [HRT\*24] measured memory for data visualizations across 2D, 3D, VR, and MR conditions, finding that VR did not significantly outperform 2D for simple recall despite increased presence, consistent with Vindenes et al. [VdGW18], who found conventional displays superior to IVR in a memory recall task. These results suggest that the assumed recall benefits of immersion are not guaranteed for abstract visual information. A central explanation is cognitive load. According to multimedia learning theory [MP05, May09], high extraneous load—from unfamiliar controllers or sensory complexity—can interfere with essential and generative processing [Swe11, MOM19], reducing learning despite increased engagement [PM18]. Critically, most prior work examines cognitive or affective outcomes, but science communication requires both; climate communication aims to inform, foster concern, and motivate action [Mos14]. VR has been shown to increase emotional engagement [ABP14], but whether these affective benefits trade off with cognitive learning, and whether either persists over time, is unclear. Our study addresses this by measuring cognitive and affective outcomes immediately and at delayed follow-up, and by testing whether spatialized audio modulates this balance.

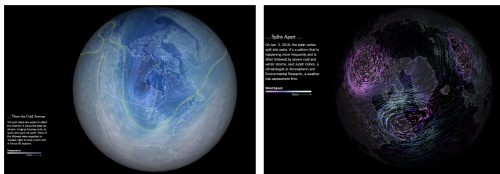


Figure 2: A Closer Look at the Polar Vortex's Dangerously Cold Winds, desktop immersive journalism, *New York Times*, 2019.

### 2.3. Immersive Journalism and Climate Communication

News organizations increasingly use “immersive journalism” to elicit emotional response and empathy [DIPWL\*10, Has20, SL20]. Advocates argue that immersive media can close knowledge gaps about scientific issues by elevating engagement [Jon17, SKO17, Sch18]. Prior work shows that IVR often produces stronger presence and enjoyment than traditional video [SB18, SKO17] and can improve story recall [SKO17], though learning gains are far from guaranteed [PM18, DAP\*15]. Most immersive journalism employs photorealistic simulations (e.g., war zones, coral reefs [BSK\*22, HDBdG22]). Climate data visualization, however, frequently requires abstract representations such as contour maps projected onto a globe (Figure 2) [Tim19]. In these contexts, viewers are not “inside” the scene; stereoscopic depth and head tracking may offer less benefit. It remains unclear whether immersion helps or hinders the recall and learning of abstract climate data. Our study addresses this gap using a skeuomorphic 3D globe visualization to test how presence, flow, and cognitive absorption operate in this underexplored scenario. Prior climate communication research highlights tailored visuals for diverse audiences [DYL20], landscape-based imagery

for emotional connection [She05], and tools for adaptation planning [SASD14]. Work on interactive graphics [HVC17, Vis17] and VR experiences [ABP14, Gre17, MD17] demonstrates IVR’s potential for engagement. However, to our knowledge, no prior study systematically examines how cognitive and affective learning intersect across modalities for abstract climate data, or whether observed effects persist beyond immediate exposure. Our study provides controlled comparison of cognitive, affective, and behavioral outcomes across Desktop and IVR.

### 2.4. Interaction Patterns and Exploration Strategies

In desktop visualization, interaction behaviors strongly shape insight formation. Users often develop distinct exploration strategies when analyzing data [BOZ\*14], and these patterns correlate with task performance. Yet it remains unclear how these strategies differ between desktop and immersive VR, or how modality shapes the link between interaction and learning. Immersive environments may shift exploration in several ways. Embodied navigation can encourage different viewing patterns than mouse-and-keyboard control [NSB\*21]. Head-tracked viewing requires physical effort, which may lead users to focus on salient features rather than systematic scanning. VR novelty and sensorimotor demands can also increase engagement while fragmenting attention [Sla09, DAP\*15]. Understanding these shifts is essential for designing effective immersive visualizations. To address this gap, we extend prior work on desktop interaction patterns to immersive contexts. By logging all rotations, zooms, and time-slider movements, we characterize exploration strategies across modalities and show how these differences help explain divergent learning outcomes.

## 3. Research Questions and Hypotheses

We structure our investigation around three research questions, each examining a different aspect of how display modality influences learning and user experience. Before presenting these, we clarify the conceptual distinctions between outcomes we measure. *Cognitive learning* refers to the formation of durable mental models of the data, operationalized through recall accuracy (accurate retrieval of factual information) and recall consistency (stability of representations over time). We use *understanding* as a broader informal term in framing and discussion, while recognizing that our empirical measure of cognitive learning is specifically recall accuracy and consistency. *Affective learning* refers to changes in attitudes, concern, and behavioral intentions, which we treat as conceptually and empirically distinct from cognitive outcomes.

**RQ1 (Cognitive Learning). How does display modality affect cognitive learning?** Desktop interaction may support more systematic visual analysis due to lower extraneous load and more familiar controls [MP05, Swe11, MOM19], while IVR may induce perceptual immersion without analytic benefit [KJD\*20, NEW\*20, PM18]. Cognitive learning is measured using recall tasks requiring identification of maxima/minima and broad spatial-temporal patterns in the data, assessed immediately and two weeks later. These items are adapted from prior climate-education studies [She05, HLSC16]. We hypothesize that **H1: Desktop displays will produce higher cognitive learning outcomes than IVR**, as measured by higher recall accuracy immediately post-study and at two-week follow-up and greater consistency in recall across time points.

**RQ2 (Affective Learning): How does display modality affect affective learning outcomes?** IVR often increases presence and emotional engagement [SB18, Mos14], but affective gains may not persist. We measure concern and behavioral intentions using standard climate-attitude scales [BWPMD20, ODC\*20], at pre-test, post-test, and two-week follow-up. This longitudinal design allows us to distinguish transient emotional responses from durable attitude change. We hypothesize that **H2: IVR will produce stronger short-term affective responses and long-term retention compared to desktop displays**, as measured by changes in level of concern and adoption of behaviors/intents addressing climate change.

**RQ3 (Exploration Strategies): How does display modality shape visualization exploration behavior, and how do exploration patterns relate to learning outcomes?** Interaction patterns strongly influence insight formation in desktop visualization [BOZ\*14, NSB\*21], but how these patterns differ in immersive environments remains poorly understood. Different exploration strategies may also mediate learning outcomes. For instance, systematic scanning may support better recall than novelty-seeking or passive observation. We measure exploration behavior through logged interactions across VR exploration trials. We hypothesize that **H3: Display modality will shape exploration strategies in ways that directly influence learning outcomes**.

**Exploratory Analysis: Audio Effects.** We additionally examine spatialized audio as a multimodal cue. Prior work in VR learning suggests audio can support attention allocation [RKS\*24], but its effects for abstract climate visualization remain unclear.

Together, these questions build a comprehensive view of how immersion influences learning and exploration, establishing a basis to guide modality choices in scientific visualization design.

## 4. Methodology

We conducted a between-subjects study to investigate how display modality (Desktop Display vs. IVR) affects cognitive and affective learning outcomes when exploring climate data visualizations.

### 4.1. Stimuli

We used three climate datasets from the National Oceanic and Atmospheric Administration (NOAA), each visualized as annual 2D data snapshots projected onto an interactive 3D sphere (Figure 3). All globe renderings were implemented in Unreal Engine [Epi].

**(1) Temperature Anomalies (1980–2020).** Derived from NOAA's Merged Land–Ocean Surface Temperature Analysis (MLOST), this dataset shows yearly deviations from a 1991–2020 baseline. Red regions indicate warmer-than-average temperatures; blue regions indicate cooler anomalies. **(2) Sea Surface Temperature (2000–2020).** From NOAA's Optimum Interpolated SST product, this dataset combines satellite and buoy measurements into yearly January 1 snapshots. Values range from near-freezing waters (dark blue) to tropical storm-forming temperatures (white/orange). **(3) Arctic Sea Ice Age (2000–2020).** Produced by the National Snow and Ice Data Center (NSIDC), this dataset estimates annual sea-ice age using satellite and drift-tracking data. Colors encode ice longevity and thickness, from first-year ice (dark blue) to 4+-year-old multi-year ice (white).

These datasets vary in analytical demand: (1) temperature trends require anomaly comparison against a baseline, (2) SST emphasizes absolute magnitude and gradients, and (3) sea-ice age emphasizes categorical spatial patterns, allowing us to examine how modality affects different forms of inference [LBI\*12].

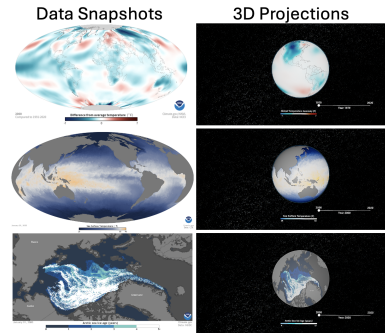


Figure 3: **LHS:** Data snapshots for the three datasets used in the study (Temperature Anomalies ([link](#)), Sea Surface Temperature([link](#)), Arctic Sea Ice Age([link](#)); all shown for the year 2000). **RHS:** Corresponding projections on the interactive 3D globe in IVR and Desktop Display conditions.

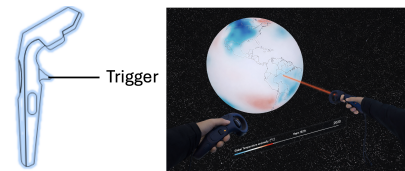


Figure 4: VIVE Controller interactions. Left: The trigger is held down for selection of the sphere before zoom/rotate. Right: The user looks down at their hand to track a red ray; this ray is cast at the globe and selected by holding down the trigger for movement.

**Navigation and Interaction.** Participants could rotate, zoom, and scrub through years in all conditions. In the Desktop condition, participants rotated the globe by clicking and dragging with the mouse, zoomed using the scroll wheel, and scrubbed through years using keyboard arrow keys [KJD\*20, NEW\*20]. In the IVR condition, participants pointed the VIVE controller toward the globe, held the trigger to select it, and physically moved the controller to rotate and zoom, and the time slider was manipulated by pointing the raycast at the slider toggle and holding the trigger [LJKM\*17] (Figure 4). Both offered identical functionality, differing only in interaction mechanics. The globe was rendered at a virtual diameter of 0.8 m in the IVR condition, positioned approximately 0.9 m from the participant, and rendered at approximately 25 cm diameter on the 24-inch desktop monitor.

To examine multimodal augmentation, 50% of participants experienced spatialized audio during the Sea Ice Age exploration. We used underwater recordings from Kongsfjorden ice-melt events [MLV\*20], chosen because they are semantically tied to this dataset, with louder ‘pops’ in audio indicating ice cracks during melting; we did not include audio in other trials to avoid introducing a confound between audio presence and audio-content align-

ment. Playback was scaled to the proportion of 4+-year-old ice each year, paralleling visible temporal change.

## 4.2. Apparatus and Participants

**Display Conditions:** We employed a between-subjects design with participants randomly assigned to one of two modalities: (i) *Desktop Display*: 24-inch monitor (3840×2160 resolution) with keyboard and mouse navigation, connected to a Razer Blade laptop running Windows 11, or (ii) *Immersive VR (IVR)*: VIVE head-mounted display with VIVE VR Controllers, providing stereoscopic 3D rendering and 6-degree-of-freedom tracking. The application ran at a stable 90 Hz on the VIVE HMD and 60 Hz on the desktop display (Intel i7 processor, NVIDIA RTX-series GPU, 16 GB RAM), with no noticeable dropped frames during the study.

Power analysis using G\*Power [Coh13] (ANOVA: repeated measures, between-within interaction;  $f = 0.25$ ,  $\alpha = 0.05$ , power = 0.8) indicated 35 participants per modality were needed post-exclusions. We recruited 84 undergraduate student participants from a local university, yielding 42 per modality. Audio was a secondary exploratory factor applied only to the Arctic Sea Ice Age trial; equal allocation within each modality group ensured the audio sub-analysis was adequately powered to detect medium-sized effects within that subset, with 21 participants per audio condition. Participants volunteered their time in the study, which lasted 38:33±5:21 minutes, with two 5 minute breaks permitted between data exploration trials. We collected visualization familiarity (3.28±1.34), virtual reality familiarity (3.04±1.09), age (27.05 ± 3.33 y/o), gender (46 males, 38 females), and climate change familiarity (2.98±0.87), all measured on a 5-point Likert scale (1 = not at all familiar, 5 = extremely familiar); no identifying information was recorded. We note that baseline climate familiarity did not significantly differ by the modality (Desktop vs. IVR) participants were assigned ( $\Delta 0.21$ ). All participants had normal color vision. Before the main study, we conducted a pilot study with four participants to validate the experimental design, survey instruments, and timing estimates. Minor adjustments to survey wording and training duration were made based on pilot feedback.

**Exclusion criteria.** We excluded participants who: (i) did not complete all three exploration trials, (ii) failed attention checks embedded in the recall tasks, or (iii) exhibited unusable interaction logs due to sensor dropout or software freezes (e.g.,  $\geq 20\%$  missing navigation data). No participants were excluded, yielding a final dataset of 84 participants used in the analyses.

## 4.3. Study Procedure

The study followed a four-phase sequence:

**(1) Pre-Study Survey.** Participants completed a 5-minute questionnaire capturing demographics, self-reported familiarity with data visualization and VR, and baseline attitudes toward climate change. They then completed a short training session using a blank sphere and time-slider interface, ensuring they were comfortable with the interaction controls in their assigned modality. **(2) Exploration Trials.** Each participant completed three 8-minute exploration trials, one per dataset, presented in a randomized order to minimize sequencing effects. At the start of each trial, the administrator provided a concise scripted description of the dataset and

color scale. Participants then freely explored the visualization using the interaction techniques supported by their assigned modality. Navigation events were logged continuously for later analysis. Following each trial, participants completed a 2-minute post-trial survey assessing the recall of and attitudes towards of the dataset they had just viewed. For the Arctic sea-ice dataset, half of the participants also experienced an accompanying auditory stimulus (see Section 3). **(3) Post-Study Survey.** After completing all trials, participants filled out a questionnaire measuring immersion, user experience, immediate recall of takeaways, and post-study climate attitudes. Survey items and scoring procedures are detailed in Section 4.4. **(4) Follow-Up.** Two weeks later, participants completed an online follow-up survey assessing longer-term recall and any sustained changes in climate attitudes.

## 4.4. Measures

We collected four classes of measures: user experience, cognitive outcomes, affective outcomes, and interaction behavior (see Supplemental Material for survey items and scoring rubrics).

**User Experience.** Although user experience (UX) is not one of our primary research questions, we include standard UX measures: sense of immersion using the Slater–Usoh–Steed Presence Scale [Sla99] and usability (enjoyment, engagement, satisfaction, and ease of use) [LHS08], following common practice in VR evaluation studies. Prior work shows that differences in presence and usability can shape how people approach immersive visualizations, even when they do not directly predict learning outcomes. Measuring UX therefore allows us to (a) confirm expected modality differences in presence and usability, (b) verify that participants interpreted each condition as intended, and (c) interpret RQ1–RQ3 results in light of general experience-level differences between Desktop and IVR (e.g., whether IVR’s higher presence coincides with higher cognitive load, or whether Desktop’s higher usability supports more systematic exploration). By reporting UX measures up front, but keeping them analytically separate from our hypothesis tests, we provide a transparent characterization of participant experience (confirming the expected immersion gradient across conditions) without overextending UX as an explanatory variable.

**Cognitive Learning Outcomes.** Cognitive outcomes were assessed through recognition and inference tasks adapted from climate-visualization education research (e.g., trend identification and spatial-temporal comparisons) [She05, HLSC16]. Each post-trial survey included a small set of dataset-specific questions (see Section 5.2), which participants answered again immediately post-study and at the two-week follow-up. Responses were coded for two metrics: (i) *recall accuracy*, scored as correct/incorrect using expert-validated answers (each question had one correct numeric/ordinal answer; partial credit was not used); and (ii) *recall consistency*, defined as the stability of each participant’s answers across the three time points (post-trial, post-study, and follow-up), including repeated misconceptions.

**Affective Learning Outcomes.** Attitudes toward climate change were measured using a 19-item instrument (16 likert ratings on a 5-point scale (1 = strongly disagree, 5 = strongly agree) + 3 free-response answers) adapted from prior climate research [YWZ\*13]. Items were administered pre-study, post-study, and follow-up. To

reduce dimensionality, we applied varimax factor analysis with a 0.40 loading threshold to likert-items [Kef10], yielding six components: climate change awareness, perceived community impact, human responsibility, self-reported expertise, discussion propensity, and behavioral intentions (see Section 5.3). Free-response answers were independently coded by two raters and converted to numeric scales (Cohen's  $\kappa = 0.83$ ).

**Interaction Patterns.** We measure exploration behavior through logged interactions (rotations, zoom levels, time-slider scrubbing) across three 8-minute exploration trials. We used these logs to characterize exploration strategies following prior work on interaction-pattern analysis in visualization [LPED20, DFP\* 18]. Patterns were classified into distinct exploration strategy categories (see Sec 5.4) by two independent raters, resolving disagreements (Cohen's  $\kappa = 0.89$ ) through discussion, characterized by: (i) *Temporal coverage*: Time spent viewing different years/periods, (ii) *Spatial coverage*: Geographic regions explored and order of exploration, (iii) *Interaction frequency*: Rate and type of navigation actions, and (iv) *Temporal sequencing*: Whether exploration proceeded chronologically, in reverse, or non-linearly.

## 5. Result Analysis

We organize our results around the three research questions: cognitive learning (RQ1), affective learning (RQ2), and exploration behavior (RQ3). Before presenting these primary outcomes, we briefly summarize participants' user experience (UX) ratings (immersion and usability) to contextualize how participants perceived each modality. For cognitive outcomes, we use repeated-measures ANOVA to examine how recall accuracy and consistency vary by modality, audio condition, and time. For affective outcomes, we analyze changes in climate-attitude components (derived via factor analysis) using linear mixed-effects models to assess both short-term and two-week shifts. For exploration behavior, we combine quantitative interaction-log metrics with a qualitative pattern-coding analysis to identify and compare exploration strategies across modalities. Together, these analyses allow us to evaluate the engagement–understanding trade-off from complementary cognitive, affective, and behavioral perspectives.

### 5.1. User Experience

We focus on reporting the results of standard UX measures: sense of immersion [Sla99] and usability [LHS08]. We find that participants report their perceived level of immersion in a uniform manner across all modalities in general, with IVR outstripping desktop displays. We also find that the presence of auditory stimuli increases the perceived level of immersion for both desktop displays and IVR. Users also report highest levels of enjoyment, engagement, and satisfaction for IVR+Audio (see Figure 5); however, ease of use is seen to be greatest for desktop displays.

### 5.2. RQ1: Cognitive Learning Outcomes

We evaluated cognitive learning using *recall accuracy* and *recall consistency*; participants completed identical recall questions immediately after each trial, again at the end of the study, and at the two-week follow-up (see Supplemental Material for full set of survey questions).

**Analysis Procedure.** We conducted a  $2 \times 2 \times 3$  mixed-design

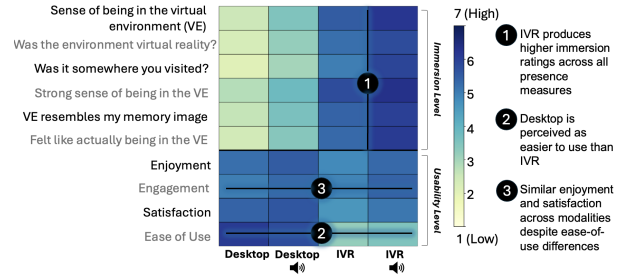


Figure 5: Results for User Experience from post-study survey; mean ratings per immersion and usability measure are visualized using heatmaps, based on a 1–7 Likert Scale. 🗣️: represents the presence of auditory stimuli. Top: reported level of immersion. Bottom: reported level of usability ratings.

ANOVA with *Modality* (Desktop vs. IVR) and *Audio* (present vs. absent) as between-subjects factors and *Time* (post-trial, post-study, follow-up) as a within-subjects factor. For trials on *Temperature Anomalies* and *Sea Surface Temperature*, audio was collapsed because no audio-related effects approached significance ( $F < 1.2$ ,  $p > .28$ ). For *Arctic Sea Ice Age*, we retained the audio factor because 50% of participants received spatialized audio. Responses were independently coded by two raters using expert-validated answer keys (Cohen's  $\kappa = 0.91$ ). We report partial  $\eta^2$  for effect sizes and apply Bonferroni correction for post-hoc comparisons.

**Recall Accuracy.** Figure 6(a) shows accuracy trends over time. We found a significant main effect of *Modality* ( $F(1,84) = 24.17$ ,  $p < .001$ , partial  $\eta^2 = 0.24$ ), with Desktop participants outperforming IVR participants overall (Desktop: 54.03%, IVR: 44.73%). A strong main effect of *Time* also emerged ( $F(2,168) = 47.93$ ,  $p < .001$ , partial  $\eta^2 = 0.38$ ), reflecting expected memory decay: accuracy dropped from post-trial to post-study and again at follow-up for both modalities. The *Modality*  $\times$  *Time* interaction was not significant ( $F(2,168) = 1.83$ ,  $p = 0.16$ ), indicating that although Desktop yielded higher accuracy at each measurement, the two modalities showed comparable forgetting rates. Bonferroni-corrected post-hocs showed significant Desktop advantages at post-trial ( $t(79) = 4.91$ ,  $p < .001$ ,  $d = 1.13$ ) and post-study ( $t(79) = 2.87$ ,  $p = 0.005$ ,  $d = 0.65$ ). A smaller, marginal advantage remained at two-week follow-up ( $t(79) = 1.42$ ,  $p = 0.08$ ). Trial-level analyses revealed especially pronounced Desktop benefits on *anomaly-identification* and *trend-change detection*.

**Recall Consistency.** Consistency, i.e., whether participants repeated the same answer across time-points is summarized in Figure 6(b). Desktop users showed higher overall consistency ( $F(1,84) = 6.83$ ,  $p = 0.011$ , partial  $\eta^2 = 0.08$ ); consistency decreased with longer retention intervals ( $F(1,84) = 32.41$ ,  $p < .001$ ), particularly between post-trial and follow-up.

**Audio Effects.** For the Arctic dataset, we found a modest main effect of *Audio* ( $F(1,42) = 4.21$ ,  $p = 0.047$ , partial  $\eta^2 = 0.10$ ): participants who heard spatialized sea-ice audio more accurately identified peak ice-volume years (61.2%) than those without audio (43.8%). No significant interactions with modality were observed, indicating that the audio cue provided a uniform but small benefit.

**Summary and Interpretation.** Across all datasets and mea-

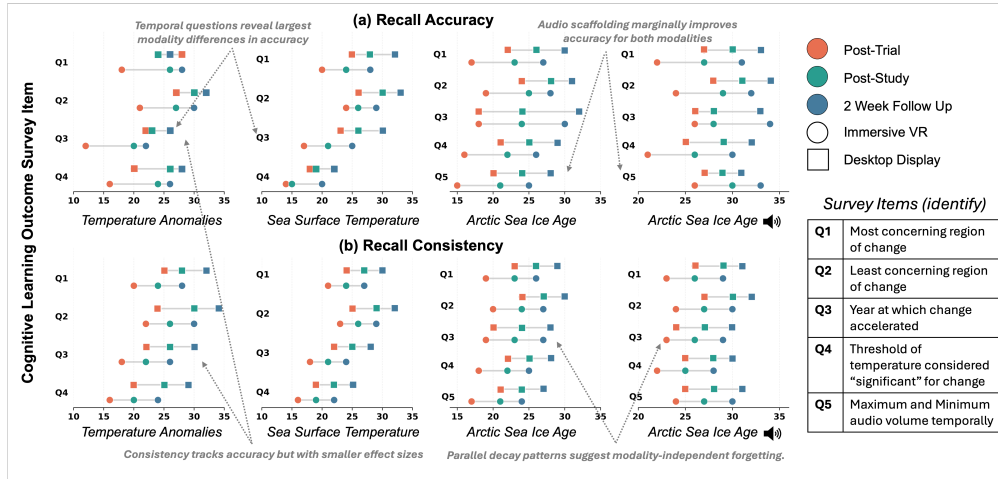


Figure 6: Cognitive learning outcomes across datasets and time. (a) Recall accuracy and (b) recall consistency shown for three climate datasets. Each row shows Desktop Display and Immersive VR (IVR) conditions across cognitive learning questions. We note that the audio-based question (Q5) is only asked for Arctic Sea Ice Age exploration. Three time points are shown: Post-Trial, Post-Study, Follow-Up Study. 🎧: indicates presence of auditory stimuli.

surement points, Desktop displays consistently supported stronger cognitive learning than IVR. Desktop users achieved higher immediate recall and maintained more stable representations after two weeks, whereas IVR participants showed lower accuracy and greater variability in their mental models. This aligns with prior work showing that immersive systems introduce additional sensorimotor and attentional demands that elevate extraneous cognitive load, which can hinder analytical encoding despite increasing engagement [BSK\*22, DAP\*15]. In contrast, Desktop environments provide predictable control dynamics and reduced perceptual load, supporting systematic temporal-spatial comparisons [DYLM20, LH14]. Forgetting rates were comparable across modalities, suggesting that differences arise primarily during initial encoding rather than during memory decay. The modest audio benefit in the Arctic dataset further indicates that tightly coupled multimodal cues can scaffold encoding—consistent with multimedia learning principles [May05]. Overall, Desktop displays proved more effective for analytic precision and long-term retention, compared to IVR. Hence, **H1 is fully supported**.

**5.3. RQ2: Affective Learning Outcomes**

We assessed affective learning using a 19-item climate-attitudes instrument (16 Likert items + 3 free-response questions; adapted from Yu et al. [YWZ\*13]), administered at pre-study, post-study, and follow-up. **Factor Analysis.** We conducted a varimax factor analysis (loading threshold = 0.40 [Kef10]) on the 16 Likert items. One item (“Who is responsible for responding to climate change?”) failed to load and was removed. The remaining 15 items yielded six components explaining 73.4% of variance (see Figure 7):

- **Climate Change Awareness** (4 items,  $\alpha = 0.84$ ): Belief in climate change, severity, willingness to take individual action
- **Community Impact** (3 items,  $\alpha = 0.79$ ): Perceived local/community-level effects, support for collective action
- **Human Responsibility** (2 items,  $\alpha = 0.71$ ): Attribution of climate change to human activities

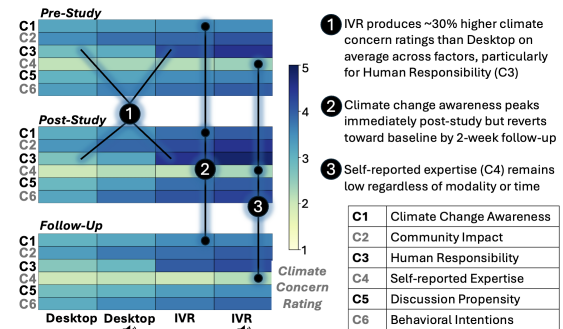


Figure 7: Heatmap with average values for each factor component, compared across different modalities and averaged across all 3 datasets. Standard deviation is less than 1 in all cases. Three time points are shown: Pre-Study, Post-Study, and Follow-Up. 🎧: indicates presence of auditory stimuli.

- **Self-Reported Expertise** (2 items,  $\alpha = 0.68$ ): Confidence in understanding climate science and trends
- **Discussion Propensity** (2 items,  $\alpha = 0.73$ ): Willingness to discuss climate views with others
- **Behavioral Intentions** (2 items,  $\alpha = 0.76$ ): Plans to adopt climate-mitigating behaviors

Free-response items (3) were analyzed using a structured thematic coding scheme. Two trained raters independently categorized each response along three dimensions: expressed concern, perceived personal impact, and behavioral intention; each rated on: 1 = minimal concern/impact/intent – 5 = strong concern/impact/intent. Disagreements were resolved via discussion (inter-rater reliability: Cohen’s  $\kappa = 0.83$ ). The resulting numeric scores were incorporated into the corresponding factor components (Climate Change Awareness, Community Impact, Behavioral Intentions).

**Analysis Procedure.** For each component, we computed change scores from pre-study to post-study ( $\Delta_1$ ) and from pre-study to

Pattern Name	Description	Avg. Exploration Time	Modality Per Pattern
<b>Geographic Sweepers</b>	Participants fixate on a particular geographical region, and observe temporal change for only that region.	6:41±1:12	IVR: 24%; Desktop: 42%
<b>Temporal Analysts</b>	Participants begin at the max/min year for analysis, and explore all regions for that year before moving linearly forward or backward in time.	6:01±1:35	IVR: 14%; Desktop: 21%
<b>Event-Driven Inspectors</b>	Participants choose an intermediate year to begin analysis, based on the occurrence of a striking visual change.	4:57±2:01	IVR: 6%; Desktop: 27%
<b>Constrained Focus Explorers</b>	Participants focus on only one side of the sphere throughout the exploration trial.	4:32±1:49	IVR: 24%; Desktop: 4%
<b>Opportunistic Browsers</b>	Participants show no clear patterns during exploration, and asynchronously switch either region or time period being observed.	4:01±1:43	IVR: 32%; Desktop: 6%

Table 1: Interaction patterns observed during exploration as well as the average interaction times (minutes) and modalities associated with that exploration pattern (percentage of participants exhibiting that pattern).

follow-up ( $\Delta_2$ ). We tested modality effects using linear mixed-effects models with *Modality* (Desktop vs. IVR) as a fixed effect and *Participant* as a random effect, complemented by independent-samples *t*-tests for  $\Delta_1$  and  $\Delta_2$  with Bonferroni correction ( $\alpha = 0.008$ ). Effect sizes are reported as Cohen's *d* [Coh13].

**Short-Term Shifts** (Pre-Study → Post-Trial). As shown in Figure 7, IVR produced significantly larger immediate increases than Desktop on Climate Change Awareness, Community Impact, and Discussion Propensity (all  $p < 0.001$ ,  $d = 0.66$ – $0.88$ ). Perceived Expertise increased similarly in both modalities, suggesting that exploration itself improved confidence regardless of display condition. Behavioral Intentions showed a modest IVR advantage that did not survive multiple-comparison correction.

**Long-Term Shifts** (Pre-Study → Follow-Up). By the two-week follow-up, most shifts had attenuated toward baseline for both modalities. None of the modality contrasts remained significant (all  $p > 0.09$ ). Component-level correlations showed that follow-up scores were more tightly aligned with pre-study attitudes than post-study, indicating limited long-term retention of affective change.

**Baseline Attitude Moderation.** When stratifying participants by their initial level of climate concern, IVR produced the strongest short-term gains among initially low-concern participants (1.0–2.5,  $n = 23$ ); however, these gains showed the steepest decay by follow-up (Desktop: +0.67, IVR: +1.24, 73% reversion). Moderate-concern participants (2.6–3.5,  $n = 31$ ) showed smaller shifts (Desktop: +0.41, IVR: +0.82, 39% reversion), while high-concern participants (3.6–5.0,  $n = 27$ ) exhibited minimal movement across modalities (Desktop: +0.18, IVR: +0.31, 21% reversion).

**Audio Effects.** Contrary to expectations, spatialized audio did not significantly enhance affective outcomes at either time point (all  $p > 0.16$ ), suggesting that auditory cues were more cognitively supportive (RQ1) than emotionally evocative.

**Summary and Interpretation.** IVR reliably elicited stronger immediate affective responses than Desktop, particularly among participants who began the study with low or moderate concern about climate change. These short-term “boosts” in concern, perceived community impact, and willingness to discuss climate issues mirror prior VR persuasion research showing that immersion can amplify emotional salience and perceived presence [BBC16]. However, these shifts proved transient: attitudes largely reverted toward baseline within two weeks, consistent with work demonstrating rapid decay of presence-driven effects without repeated reinforcement [SWHB25]. Taken together, Desktop produces smaller yet more stable affective profiles than IVR. Hence, **H2 is partially**

**supported:** immersion strengthens short-term affective responses but does not produce durable transformation in climate attitudes.

#### 5.4. RQ3: Interaction Patterns

To examine how modality shapes exploration behavior, we analyzed all logged navigation events (rotations, zooms, and time-slider movements) across the three 8-minute trials. Two coders independently classified each participant's strategy using criteria adapted from prior visualization-interaction work [BOZ\*14], achieving 89% agreement before resolving discrepancies.

**Pattern Identification.** We identified five distinct exploration patterns (Table 1), characterized by temporal coverage (viewing different years), spatial coverage (geographic regions explored), and temporal sequencing (chronological vs. non-linear navigation):

**Modality Differences.** A chi-square test revealed significant differences in pattern distribution across modalities ( $\chi^2(4) = 18.73$ ,  $p = 0.001$ , Cramér's  $V = 0.48$ ). Desktop users were far more likely to adopt systematic strategies (*geographic sweepers and temporal analysts*: 63% Desktop vs. 38% IVR), whereas IVR users more frequently exhibited fragmented behaviors (*constrained focus explorers and opportunistic browsers*: 56% IVR vs. 10% Desktop). *Event-driven inspection* was also notably more common in Desktop (27%) than IVR (6%).

**Exploration Time.** Strategy type was strongly associated with exploration duration ( $F(4,84) = 8.34$ ,  $p < 0.001$ ). Systematic patterns required significantly more time (6:01±1:35 minutes) than fragmented ones (4:32±1:49 minutes). This pattern aligns with prior research suggesting structured navigation is more effortful but supports deeper engagement [BJHA21].

**Patterns and Cognitive Performance.** Exploration strategy strongly predicted recall accuracy ( $F(4,84) = 14.28$ ,  $p < 0.001$ ). Systematic strategies achieved the highest accuracy, fragmented patterns the lowest (*geographic sweepers*: 72.1%, *temporal analysts*: 68.4% vs. *constrained focus explorers*: 48.3%, *opportunistic browsers*: 41.7%; all post-hoc  $p < .001$ ), and *event-driven inspection* fell in between (58.2%). No interaction with modality emerged, indicating that the performance gap was driven by the strategy itself rather than the display.

**Spatial and Temporal Coverage.** Desktop participants viewed significantly more of both the globe (78.4% vs. 54.2%;  $t(79) = 4.21$ ,  $p < 0.001$ ) and the available timeline (84.1% vs. 69.6%;  $t(79) = 4.73$ ,  $p < .001$ ). Coverage strongly predicted recall (spatial:  $r=0.61$ ; temporal:  $r=0.58$ ; both  $p < 0.001$ ). *Geographic sweepers* and *temporal analysts* achieved the highest coverage.

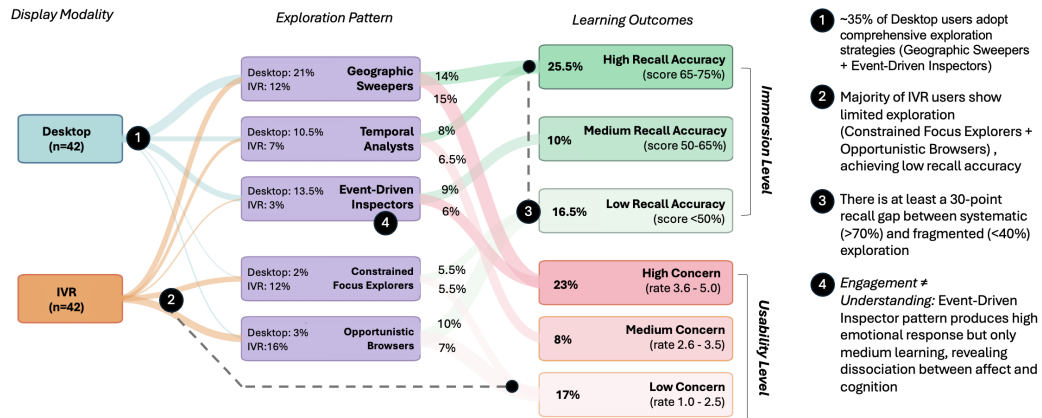


Figure 8: The Sankey diagram illustrates how display modality influences exploration patterns, which in turn predict dual learning outcomes: cognitive (recall accuracy) and affective (climate concern). Flow width represents the proportion of participants following each path. For visual clarity, flows representing less than 5% of total participants between patterns and outcomes are not shown, though all connections exist in the data. Audio condition showed no significant effect on pattern distribution or outcomes

**Audio Effects.** Audio presence did not significantly alter pattern distribution or coverage ( $p > .27$ ), though zoom activity increased during intense audio segments in the Arctic dataset ( $r = .34, p = .03$ ), suggesting attention capture without structural changes in strategy.

**Summary and Interpretation.** As shown in Figure 8, modality played a meaningful role in shaping how participants approached free exploration. Desktop users tended to adopt structured strategies, such as systematically sweeping regions or progressing chronologically, which led to broader data coverage and, in turn, stronger recall. IVR participants, by contrast, more frequently exhibited fragmented or spatially constrained behavior, consistent with prior work showing that immersive environments impose higher sensorimotor and attentional demands that can diffuse users' analytic focus. Across modalities, however, the clearest predictor of cognitive performance was the *pattern itself*: participants who engaged in systematic exploration achieved substantially higher recall accuracy, regardless of modality. This aligns with cognitive theories suggesting that deliberate, organized navigation supports the formation of stable mental models, whereas opportunistic or narrowly focused exploration limits exposure to key patterns. Hence, **H3 is fully supported**.

## 6. Discussion

Our findings reveal a systematic engagement–understanding trade-off in immersive scientific visualization: while IVR enhanced short-term emotional engagement and satisfaction, Desktop displays consistently supported superior cognitive learning and systematic exploration. This divergence challenges assumptions about immersion's benefits and has important implications for visualization design aligned with specific communication goals.

**The Engagement–Understanding Trade-Off:** Our central contribution is empirical evidence for a fundamental tension between affective engagement and cognitive learning in immersive climate visualization, manifesting across multiple dimensions. We note that while our affective outcomes are tied to climate-specific attitudes, the cognitive findings (recall advantage, exploration strate-

gies) and interaction burden effects are likely domain-general. Similar patterns would be expected in other contexts involving abstract, globe-based or spatial 3D data, such as geospatial analytics, medical imaging, or astronomical visualization, where the same tension between immersive presence and analytic precision applies. Future work should test this generalizability explicitly.

**Immediate vs. Sustained Effects.** IVR produced substantially stronger immediate affective responses across climate concern, perceived community impact, and discussion willingness; however, these gains proved ephemeral, with the majority of initially low-concern participants reverting to baseline attitudes within two weeks. In contrast, Desktop's moderate cognitive advantages persisted across the follow-up period with comparable retention rates. This suggests immersion creates strong emotional responses that rapidly decay without reinforcement [SWHB25, SKO17], whereas systematic exploration in Desktop conditions supports more durable mental model formation, as evidenced by higher recall accuracy and consistency.

**Baseline Attitude Moderation.** The differential impact on participants with varying initial concern levels reveals important boundary conditions. Low-concern participants showed the largest affective shifts but also the steepest reversion, suggesting IVR may effectively raise awareness among disengaged audiences but remains insufficient for sustained attitude change sans follow-up. High-concern participants, already motivated, showed minimal movement regardless of modality, indicating potential ceiling effects.

**Exploration Strategy Mediation.** The relationship between modality and learning was substantially driven by exploration patterns. Systematic strategies led to superior recall regardless of display; fragmented strategies yielded poor performance in both conditions. However, Desktop users were significantly more likely to adopt systematic approaches, explaining much of the modality effect; IVR's additional sensorimotor and navigational demands may consume working memory resources that would otherwise support deliberate, structured exploration [Swe11]. This aligns with Huang et al. [HPY23], who found that embodied interaction does not uni-

formly benefit visualization tasks such as structured analytical navigation, as physical interaction demands can redirect attention away from systematic exploration, consistent with our observation that IVR users more frequently exhibited fragmented strategies.

**Implications for Cognitive Load Theory:** Our results support Mayer's cognitive theory of multimedia learning [May05], which distinguishes extraneous (task-irrelevant), essential (core representational), and generative (deep integration) processing. Desktop displays appear to minimize extraneous load through familiar interactions, freeing resources for systematic temporal-spatial comparisons and stable mental model formation. IVR, despite potential advantages for spatial learning [RCC\*21], introduced extraneous load both from the medium itself (stereoscopic rendering, head-tracked navigation) and from interaction design (potentially unfamiliar controllers, raycast-based selection). While these features increased presence, they may have diverted attention from analytic processing, evidenced by lower spatial and temporal coverage. We note that Wagner Filho et al. [WFSN19] found higher usability and lower mental workload for an immersive space-time cube than desktop alternatives, contrasting with our desktop ease-of-use advantage. This likely reflects task differences: embodied spatial navigation suits immersive trajectory exploration, whereas analytic temporal comparison of abstract data favors familiar mouse-keyboard controls. However, our findings also suggest scaffolding opportunities. The modest benefit of spatialized audio for Arctic dataset recall demonstrates that carefully designed multimodal cues can support encoding without overwhelming learners, consistent with research on spatially integrated redundant cues [MLV\*20]. Future IVR designs might reduce interaction design induced extraneous load through simplified controllers, guided tours, or progressive disclosure of navigation options.

**Limitations.** Finally, several limitations warrant consideration. Our participants were university students from a single institution with moderate baseline climate familiarity, limiting generalizability to broader populations. We examined one visualization design (climate data on a 3D globe), which may not generalize to other encodings or less emotionally salient topics. Our two-week follow-up cannot capture truly long-term outcomes over months or years. We did not directly measure cognitive load, relying on behavioral and outcome measures to infer differences. Individual differences in spatial ability or working memory were not assessed and may moderate effects. Our 2019-era HTC Vive hardware may not reflect newer systems' capabilities. Further, since our comparison bundles display (HMD vs monitor), resolution, and input devices (controllers vs mouse-keyboard), our findings speak to realistic modality 'packages' rather than isolating display hardware. Our audio manipulation was also exploratory and applied to only one dataset, limiting conclusions about multimodal effects. Finally, free exploration without specific tasks may produce different patterns than goal-directed analysis. Despite these limitations, our multi-method approach combining cognitive, affective, and behavioral measures across immediate and delayed assessments provides robust evidence for the engagement-understanding trade-off.

## 6.1. Design Implications

Our findings also yield actionable guidance organized around three key decisions:

**Modality Selection Based on Communication Goals.** Desktop displays consistently supported superior cognitive learning through higher recall accuracy, greater consistency, and systematic exploration patterns. IVR generated stronger immediate affective responses (+43% concern for initially neutral participants) but with rapid decay. Designers should select modalities based on whether analytical precision (e.g.: policy making/analysis) or emotional engagement (e.g.: public outreach) is paramount.

**Exploration Strategy matters more than Modality.** The strongest predictor of cognitive performance was exploration strategy, not display type. Systematic strategies (geographic sweepers, temporal analysts) achieved 72% accuracy vs. 45% for fragmented strategies, regardless of modality. However, IVR users adopted systematic strategies less frequently (38% vs 63% for Desktop Displays), suggesting immersive environments may require additional scaffolding, for e.g.: simplified navigation through snap-to-region buttons and automated playback. Our findings on audio augmented viewing suggest multimodal cues can potentially support encoding when tightly coupled to visual patterns, though designers should avoid gratuitous stimulation increasing perceptual load.

**Affective Changes Require Reinforcement.** Visualization exposure alone produced transient affective shifts, with 73% of initially participants with low-concern about climate change reverting to baseline within two weeks across both modalities. This decay pattern suggests single-exposure visualizations are insufficient for sustained attitude change, highlighting the need for integrated reinforcement mechanisms that extend beyond the visualization itself.

## 7. Conclusion and Future Work

This study provides empirical evidence for a fundamental *engagement-understanding trade-off* in immersive scientific visualization. IVR enhanced short-term affective engagement and satisfaction, while Desktop Displays consistently supported superior cognitive learning through systematic exploration and lower extraneous load. Critically, affective gains proved transient, with most initially low-concern participants reverting to baseline within two weeks, whereas cognitive advantages persisted. Our characterization of five exploration patterns revealed that viewing strategy predicted learning outcomes more strongly than modality, suggesting that supporting systematic exploration represents a key design leverage point. These findings challenge assumptions that greater presence automatically yields deeper learning. Instead, cognitive and affective effects diverged, suggesting that visualization designers should align modality choices with communication goals.

Future work should explore a broader range of visualization encodings and scientific domains to establish generalizability and evaluate whether reinforcement can sustain affective gains—particularly in real-world deployment settings such as museums or classrooms. Additional directions include examining individual differences in spatial ability or working memory, testing collaborative immersive scenarios, and evaluating newer VR hardware with improved ergonomics and reduced cognitive overhead. Together, these directions can clarify when and for whom immersion supports meaningful learning, enabling visualization designers to better balance cognitive learning with emotional resonance.

## References

- [ABP14] AHN S. J. G., BAIENSON J. N., PARK D.: Short-and long-term effects of embodied experiences in immersive virtual environments on environmental locus of control and behavior. *Computers in Human Behavior* 39 (2014), 235–245. 3
- [AEN10] ANDREWS C., ENDERT A., NORTH C.: Space to think: large high-resolution displays for sensemaking. In *Proceedings of the SIGCHI conference on human factors in computing systems* (2010), pp. 55–64. 2
- [BBC16] BAILEY J. O., BAIENSON J. N., CASASANTO D.: When does virtual embodiment change our minds? *Presence* 25, 3 (2016), 222–233. 8
- [BHRRW16] BACHEN C. M., HERNÁNDEZ-RAMOS P., RAPHAEL C., WALDRON A.: How do presence, flow, and character identification affect players' empathy and interest in learning from a serious computer game? *Computers in Human Behavior* 64 (2016), 77–87. 3
- [BJHA21] BODZIN A., JUNIOR R. A., HAMMOND T., ANASTASIO D.: Investigating engagement and flow with a placed-based immersive virtual reality game. *Journal of science education and technology* 30 (2021), 347–360. 8
- [BOZ\*14] BROWN E. T., OTTLEY A., ZHAO H., LIN Q., SOUVENIR R., ENDERT A., CHANG R.: Finding waldo: Learning about users from their interactions. *IEEE Transactions on visualization and computer graphics* 20, 12 (2014), 1663–1672. 3, 4, 8
- [BSK\*22] BARNIDGE M., SHERRILL L. A., KIM B., COOKS E., DEAVOURS D., VIEHOUSER M., BROUSSARD R., ZHANG J.: The effects of virtual reality news on learning about climate change. *Mass Communication and Society* 25, 1 (2022), 1–24. 2, 3, 7
- [BWPM20] BROOMELL S. B., WONG-PARODI G., MORSS R. E., DEMUTH J. L.: Do we know our own tornado season? a psychological investigation of perceived tornado likelihood in the southeast united states. *Weather, Climate, and Society* 12, 4 (2020), 771–788. 4
- [CM02] COCKBURN A., MCKENZIE B.: Evaluating the effectiveness of spatial memory in 2d and 3d physical and virtual environments. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (2002), pp. 203–210. 2
- [Coh13] COHEN J.: *Statistical power analysis for the behavioral sciences*. Academic press, 2013. 5, 8
- [DAP\*15] DIEMER J., ALPERS G. W., PEPPERKORN H. M., SHIBAN Y., MÜHLBERGER A.: The impact of perception and presence on emotional reactions: a review of research in virtual reality. *Frontiers in psychology* 6 (2015), 26. 3, 7
- [DFP\*18] DIMARA E., FRANCONERI S., PLAISANT C., BEZERIANOS A., DRAGICEVIC P.: A task-based taxonomy of cognitive biases for information visualization. *IEEE transactions on visualization and computer graphics* 26, 2 (2018), 1413–1432. 6
- [DIPWL\*10] DE LA PEÑA N., WEIL P., LLOBERA J., SPANLANG B., FRIEDMAN D., SANCHEZ-VIVES M. V., SLATER M.: Immersive journalism: Immersive virtual reality for the first-person experience of news. *Presence* 19, 4 (2010), 291–301. 3
- [DYLM20] DONG W., YANG T., LIAO H., MENG L.: How does map use differ in virtual reality and desktop-based environments? *International Journal of Digital Earth* 13, 12 (2020), 1484–1503. 2, 3, 7
- [ECN\*22] ENS B., CORDEIL M., NORTH C., DWYER T., BESANÇON L., PROUZEAU A., LIU J., CUNNINGHAM A., DROGEMULLER A., SATRIADI K. A., ET AL.: Immersive analytics 2.0: Spatial and embodied sensemaking. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts* (2022), pp. 1–7. 2
- [EHR14] ENS B., HINCAPIÉ-RAMOS J. D., IRANI P.: Ethereal planes: a design framework for 2d information space in 3d mixed reality environments. In *Proceedings of the 2nd ACM symposium on Spatial user interaction* (2014), pp. 2–12. 2
- [Epi] EPIC GAMES: Unreal engine. URL: <https://www.unrealengine.com>. 4
- [FM16] FIORELLA L., MAYER R. E.: Eight ways to promote generative learning. *Educational Psychology Review* 28 (2016), 717–741. 3
- [Gre17] GREENPEACE: Adventurers needed for a virtual expedition, 2017. URL: <https://secure.greenpeace.org.uk/page/content/vr-splash>. 3
- [Has20] HASSAN R.: Digitality, virtual reality and the 'empathy machine'. *Digital journalism* 8, 2 (2020), 195–212. 3
- [HDBdG22] HERRERA DAMAS S., BENÍTEZ DE GRACIA M. J.: Immersive journalism: Advantages, disadvantages and challenges from the perspective of experts. *Journalism and Media* 3, 2 (2022), 330–347. 3
- [HLSC16] HAROLD J., LORENZONI I., SHIPLEY T. F., COVENTRY K. R.: Cognitive and psychological science insights to improve climate change data visualization. *Nature Climate Change* 6, 12 (2016), 1080–1089. 1, 3, 5
- [HPY23] HUANG H. H., PFISTER H., YANG Y.: Is embodied interaction beneficial? a study on navigating network visualizations. *Information Visualization* 22, 3 (2023), 169–185. 9
- [HRT\*24] HURTER C., ROGOWITZ B., TRUONG G., ANDRY T., ROMAT H., GARDY L., AMINI F., RICHE N. H.: Memory recall for data visualizations in mixed reality, virtual reality, 3d and 2d. *IEEE Transactions on Visualization and Computer Graphics* 30, 10 (2024), 6691–6706. doi:10.1109/TVCG.2023.3336588. 3
- [HVC17] HERRING J., VANDYKE M. S., CUMMINS R. G., MELTON F.: Communicating local climate risks online through an interactive data visualization. *Environmental Communication* 11, 1 (2017), 90–105. 3
- [Jon17] JONES S.: Disrupting the narrative: immersive journalism in virtual reality. *Journal of media practice* 18, 2-3 (2017), 171–185. 3
- [Kef10] KEFI H.: Using a systems thinking perspective to construct and apply an evaluation approach of technology-based information systems. In *Emerging Systems Approaches in Information Technologies: Concepts, Theories, and Applications*. IGI Global, 2010, pp. 294–309. 6, 7
- [Kir95] KIRSH D.: The intelligent use of space. *Artificial intelligence* 73, 1-2 (1995), 31–68. 2
- [KJD\*20] KLINGENBERG S., JØRGENSEN M. L., DANDANELL G., SKRIVER K., MOTTELSON A., MAKRANSKY G.: Investigating the effect of teaching as a generative learning strategy when learning through desktop and immersive vr: A media and methods experiment. *British Journal of Educational Technology* 51, 6 (2020), 2115–2138. 2, 3, 4
- [LAES18] LAMB R., ANTONENKO P., ETOPIO E., SECCIA A.: Comparison of virtual reality and hands on activities in science education via functional near infrared spectroscopy. *Computers & Education* 124 (2018), 14–26. 3
- [LBI\*12] LAM H., BERTINI E., ISENBERG P., PLAISANT C., CARPENDALE S.: Empirical studies in information visualization: Seven scenarios. *IEEE Transactions on Visualization and Computer Graphics* 18, 9 (2012), 1520–1536. doi:10.1109/TVCG.2011.279. 4
- [LDG\*21] LISLE L., DAVIDSON K., GITRE E. J., NORTH C., BOWMAN D. A.: Sensemaking strategies with immersive space to think. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)* (2021), IEEE, pp. 529–537. 2
- [LH14] LIU Z., HEER J.: The effects of interactive latency on exploratory visual analysis. *IEEE transactions on visualization and computer graphics* 20, 12 (2014), 2122–2131. 7
- [LHP\*19] LEDER J., HORLITZ T., PUSCHMANN P., WITTSTOCK V., SCHÜTZ A.: Comparing immersive virtual reality and powerpoint as methods for delivering safety training: Impacts on risk perception, learning, and decision making. *Safety science* 111 (2019), 271–286. 3
- [LHS08] LAUGWITZ B., HELD T., SCHREPP M.: Construction and evaluation of a user experience questionnaire. In *Symposium of the Austrian HCI and usability engineering group* (2008), Springer, pp. 63–76. 5, 6
- [LJKM\*17] LAVIOLA JR J. J., KRUIJFF E., MCMAHAN R. P., BOWMAN D., POUPYREV I. P.: *3D user interfaces: theory and practice*. Addison-Wesley Professional, 2017. 4

- [LPED20] LIU J., PROUZEAU A., ENS B., DWYER T.: Design and evaluation of interactive small multiples data visualisation in immersive spaces. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (2020), IEEE, pp. 588–597. 2, 6
- [May05] MAYER R. E.: Cognitive theory of multimedia learning. *The Cambridge handbook of multimedia learning* 41, 1 (2005), 31–48. 7, 10
- [May09] MAYER R. E.: *Contents*, 2 ed. Cambridge University Press, 2009, pp. vii–viii. 3
- [MBGM19] MAKRANSKY G., BORRE-GUDE S., MAYER R. E.: Motivational and cognitive benefits of training in immersive virtual reality based on multiple assessments. *Journal of Computer Assisted Learning* 35, 6 (2019), 691–707. 3
- [MD17] MOSER S., DANIELS C.: Look ahead: Results from visualization research conducted for san mateo county, 2017. 3
- [MLV\*20] MAHANTY M. M., LATHA G., VENKATESAN R., RAVICHANDRAN M., ATMANAND M., THIRUNAVUKARASU A., RAGURAMAN G.: Underwater sound to probe sea ice melting in the arctic during winter. *Scientific reports* 10, 1 (2020), 16047. 4, 10
- [MM02] MORENO R., MAYER R. E.: Learning science in virtual reality multimedia environments: Role of methods and media. *Journal of educational psychology* 94, 3 (2002), 598. 3
- [MOM19] MEYER O. A., OMDAHL M. K., MAKRANSKY G.: Investigating the effect of pre-training when learning through immersive virtual reality and video: A media and methods experiment. *Computers & Education* 140 (2019), 103603. 2, 3
- [Mos14] MOSER S. C.: Communicating adaptation to climate change: the art and science of public engagement when climate change comes home. *Wiley Interdisciplinary Reviews: Climate Change* 5, 3 (2014), 337–358. 1, 3, 4
- [MP05] MAYER R. E., PILEGARD C.: Principles for managing essential processing in multimedia learning: Segmenting, pretraining, and modality principles. *The Cambridge handbook of multimedia learning* (2005), 169–182. 2, 3
- [MTM19] MAKRANSKY G., TERKILDSEN T. S., MAYER R. E.: Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and instruction* 60 (2019), 225–236. 3
- [NEW\*20] NOWAK G. J., EVANS N. J., WOJDYNSKI B. W., AHN S. J. G., LEN-RIOS M. E., CARERA K., HALE S., MCFALLS D.: Using immersive virtual reality to improve the beliefs and intentions of influenza vaccine avoidant 18-to-49-year-olds: Considerations, effects, and lessons learned. *Vaccine* 38, 5 (2020), 1225–1233. 2, 3, 4
- [NSB\*08] NOCKE T., STERZEL T., BÖTTINGER M., WROBEL M., ET AL.: Visualization of climate and climate change data: An overview. *Digital earth summit on geoinformatics* (2008), 226–232. 1
- [NSB\*21] NEWBURY R., SATHIADI K. A., BOLTON J., LIU J., CORDEIL M., PROUZEAU A., JENNY B.: Embodied gesture interaction for immersive maps. *Cartography and Geographic Information Science* 48, 5 (2021), 417–431. 3, 4
- [ODC\*20] OTTO I. M., DONGES J. F., CREMADES R., BHOWMIK A., HEWITT R. J., LUCHT W., ROCKSTRÖM J., ALLERBERGER F., MCCAFFREY M., DOE S. S., ET AL.: Social tipping dynamics for stabilizing earth's climate by 2050. *Proceedings of the National Academy of Sciences* 117, 5 (2020), 2354–2365. 2, 4
- [PM18] PARONG J., MAYER R. E.: Learning science in immersive virtual reality. *Journal of Educational Psychology* 110, 6 (2018), 785. 2, 3
- [PTS\*21] PANDE P., THIT A., SØRENSEN A. E., MOJSOSKA B., MOELLER M. E., JEPSEN P. M.: Long-term effectiveness of immersive vr simulations in undergraduate science learning: Lessons from a media-comparison study. *Research in Learning Technology* 29 (2021). 2
- [RCC\*21] REICHERTZER C., CUNNINGHAM A., COLEMAN T., CAO R., MCMANUS K., SHEPPARD D., KOHLER M., BILLINGHURST M., THOMAS B. H.: Bringing the jury to the scene of the crime: Memory and decision-making in a simulated crime scene. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (2021), pp. 1–12. 2, 10
- [RKS\*24] ROSSKOPF S., KROCZEK L. O., STÄRZ F., BLAU M., VAN DE PAR S., MÜHLBERGER A.: The impact of binaural auralizations on sound source localization and social presence in audiovisual virtual reality: converging evidence from placement and eye-tracking paradigms. *Acta Acustica* 8 (2024), 72. 4
- [RT15] RICHARDS D., TAYLOR M.: A comparison of learning gains when using a 2d simulation tool versus a 3d virtual world: An experiment to find the right representation involving the marginal value theorem. *Computers & Education* 86 (2015), 157–171. 2
- [SASD14] SCHROTH O., ANGEL J., SHEPPARD S., DULIC A.: Visual climate change communication: From iconography to locally framed 3d visualization. *Environmental Communication* 8, 4 (2014), 413–432. 3
- [SB18] SHIN D., BIOCCA F.: Exploring immersive experience in journalism. *New media & society* 20, 8 (2018), 2800–2823. 3, 4
- [Sch18] SCHEUFELE D. A.: Beyond the choir? the need to understand multiple publics for science. *Environmental Communication* 12, 8 (2018), 1123–1126. 3
- [SEC\*20] SATHIADI K. A., ENS B., CORDEIL M., CZAUDERNA T., JENNY B.: Maps around me: 3d multiview layouts in immersive spaces. *Proceedings of the ACM on Human-Computer Interaction* 4, ISS (2020), 1–20. 2
- [She05] SHEPPARD S. R.: Landscape visualisation and climate change: the potential for influencing perceptions and behaviour. *Environmental science & policy* 8, 6 (2005), 637–654. 3, 5
- [SKO17] SUNDAR S. S., KANG J., OPREAN D.: Being there in the midst of the story: How immersive journalism affects our perceptions and cognitions. *Cyberpsychology, behavior, and social networking* 20, 11 (2017), 672–682. 3, 9
- [SL20] SÁNCHEZ LAWS A. L.: Can immersive journalism enhance empathy? *Digital journalism* 8, 2 (2020), 213–228. 2, 3
- [Sla99] SLATER M.: Measuring presence: A response to the witmer and singer presence questionnaire. *Presence* 8, 5 (1999), 560–565. doi: 10.1162/105474699566477. 5, 6
- [Sla09] SLATER M.: 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. 3
- [Swe11] SWELLER J.: Cognitive load theory. In *Psychology of learning and motivation*, vol. 55. Elsevier, 2011, pp. 37–76. 2, 3, 9
- [SWHB25] SANTOSO M., WANG P., HAN E., BAIENSON J.: Virtual reality reduces climate indifference by making distant locations feel psychologically close. *Scientific Reports* 15, 1 (2025), 37102. 8, 9
- [Tim19] TIMES N. Y.: A closer look at the polar vortex's dangerously cold winds, 2019. Jan 20, 2019. URL: <https://www.nytimes.com/interactive/2019/01/30/science/polar-vortex-extreme-cold.html>. 2, 3
- [VdGW18] VINDENES J., DE GORTARI A. O., WASSON B.: Mnemosyne: adapting the method of loci to immersive virtual reality. In *International conference on augmented reality, virtual reality and computer graphics* (2018), Springer, pp. 205–213. 3
- [Vis17] VISUALS R. W.: Real world visuals, 2017. 2017. URL: <http://www.realworldvisuals.com>. 3
- [WFSN19] WAGNER FILHO J. A., STUERZLINGER W., NEDEL L.: Evaluating an immersive space-time cube geovisualization for intuitive trajectory data exploration. *IEEE transactions on visualization and computer graphics* 26, 1 (2019), 514–524. 10
- [YWZ\*13] YU H., WANG B., ZHANG Y.-J., WANG S., WEI Y.-M.: Public perception of climate change in china: results from the questionnaire survey. *Natural hazards* 69 (2013), 459–472. 5, 7