Evaluating Replay Techniques for Asynchronous Task Handover in Immersive Analytics

Zhengtai Gou[*∗]
Georgia Tech

Junxiao Long[*†]
Georgia Tech

Tao Lu[‡]
Georgia Tech

Jian Zhao[§]
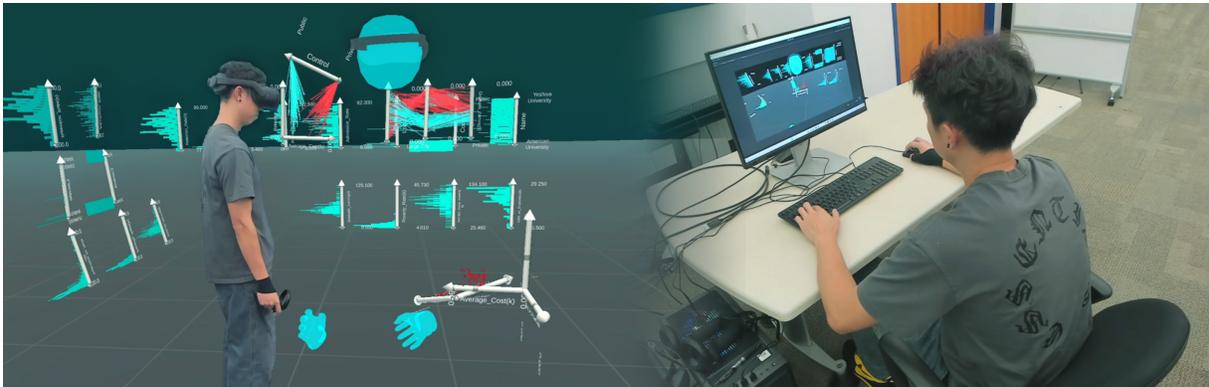University of Waterloo

Yalong Yang[¶]
Georgia Tech

Figure 1: A data analyst (i.e., the real human illustrated in the figure) replays an immersive analytics session, observing the actions of a previous analyst represented by a floating blue avatar. The image depicts the two optimized conditions evaluated in our Phase 2 comparison: the active immersive VR replay and the hybrid desktop PC replay.

ABSTRACT

Immersive analytics enables collaborative data analysis in shared virtual spaces. While synchronous collaboration in such environments is well-established, real-world analysis often requires an effective task handover—the transfer of knowledge and analytical context between analysts working asynchronously. Traditional handover methods often rely on static annotations that fail to capture the dynamic problem-solving process and spatial context inherent in immersive workflows. To address this handover challenge, we explore session replay as a comprehensive approach for analysts to re-experience a predecessor's work, facilitating a deeper understanding of both the visual details and the insight formation process. Two phases of studies were conducted to establish design guidelines for such replay systems by investigating the impact of viewing platform (PC vs. VR), perspective (first-person vs. third-person), and navigation control (active vs. passive). Phase 1 identified the optimal replay configurations within each viewing platform, revealing a platform-dependent divergence: PC users favored a guided, first-person perspective for its focused detail, while VR users benefited significantly from the agency afforded by a third-person perspective with active navigation. After refining each condition based on user feedback—including developing a novel hybrid 1PP/3PP format for PC—Phase 2 compared the two optimized systems (PC vs. VR). Our results show that the immersive VR replay led to significantly better task comprehension and workflow reconstruction accuracy, demonstrating the critical role of embodied agency in understanding complex analytical processes.

Index Terms: Immersive Analytics, Handover, Collaboration

[*]e-mail: zgou7@gatech.edu. [*]Authors contributed equally.
[†]e-mail: jlong343@gatech.edu
[‡]e-mail: tlu307@gatech.edu
[§]e-mail: jianzhao@uwaterloo.ca
[¶]e-mail: yalong.yang@gatech.edu

1 INTRODUCTION

Immersive analytics is an emerging field that shows significant promise in many applications by leveraging immersive technologies to improve data analysis and decision making [32, 40, 55]. In general, the primary benefits of this approach stem from the vast spatial canvases [23, 37, 66] and the embodied interactions [22, 25, 67, 73] that these technologies afford. Compared to standard desktop-based analytical workflows, immersive analytics allows analysts to render graphics in a large 3D space around them and interact with data visualizations using intuitive physical movements [2, 24, 49, 65].

Although these features offer significant advantages for the individual analyst, complex real-world data analysis is now usually conducted by teams in a collaborative manner [3, 28, 51, 58]. Asynchronous collaboration, in particular, has become prevalent due to its flexibility, especially for spatially distributed teams [21, 35]. In such asynchronous scenarios, a critical workflow is the *handover*, where an incoming analyst must first understand what the previous collaborator completed before they can begin their own work [61, 72].

The challenge for successors in a handover process is to understand insights, intents and analytical history from predecessors. While non-immersive visual analytic tasks often rely on oral communication or provenance tracking tools to facilitate this sensemaking process[72], these externalisations fail to capture the predecessor's spatial focus and embodied interactions, which limits their application in immersive tasks[16, 35, 71]. To address these challenges, replaying predecessor's session has emerged as a promising method for communicating the analytical process. It helps capture the spatial, temporal and interaction details needed for the successor's task comprehension, and saves the extra burden for the predecessor to explain their analytical history. Past studies have demonstrated the benefits of replay in immersive spatial manipulation tasks, including virtual furniture placement [62] and guided tours [19]. The design of these replay systems involves critical trade-offs concerning viewing perspective and user agency. Research by Iriye et al. [26], for example, has shown that a third-person perspective can enhance spatial memory, while recent work by Giovannelli et al. [19] indicates that active navigation control enhances engagement and spatial understanding over passive playback.

The novelty of our work lies in addressing the unique handover challenge posed by immersive analytics, a domain fundamentally different from prior studies. Previous replay systems have centered on spatial manipulation tasks, such as furniture placement, where the goal is to understand *what object moved where*. In contrast, the core of immersive analytics is *sensemaking*: identifying correlations, filtering datasets, and linking visualizations to form insights [40, 55, 57]. The goal of a replay system for immersive analytics is therefore not just to see spatial changes, but to reconstruct the analyst's thought process. This distinction introduces a critical and unaddressed research gap. The information vital to an analytical workflow often requires the detailed, focused view that a first-person perspective provides. Simultaneously, understanding the broader strategy demands the contextual overview of a third-person perspective. This creates a fundamental tension: a third-person view effective for spatial tasks may render crucial data illegible in an IA context, while a first-person view may obscure the overall analytical strategy. Thus, findings from simpler spatial tasks cannot be generalized, and the optimal replay configuration for this cognitively demanding domain remains unknown.

To resolve this critical tension, we designed a two-phase plan to systematically investigate three key design factors (see Figure 2): *viewing platform* (PC vs. VR), *perspective* (1PP vs. 3PP), and *navigation control* (active vs. passive). Crucially, our evaluation went beyond simple recall, employing measures such as a *Workflow Reconstruction* task to directly assess comprehension of the sensemaking involved. **Phase 1** aimed to identify the optimal configuration within each platform by comparing three conditions: 1PP-Passive, 3PP-Passive, and 3PP-Active. The results revealed a clear platform-dependent divergence. While VR users benefited significantly from the agency afforded by 3PP-Active, the findings for PC were more nuanced, with users valuing both the focused detail of 1PP-Passive and the contextual awareness of 3PP-Passive.

To address this, we developed a novel hybrid 1PP/3PP condition for the PC. Informed by these findings, **Phase 2** directly compared these optimized formats for PC and VR after enhancing them with user-requested features. Our final results show that the immersive VR replay led to significantly better task comprehension and workflow reconstruction accuracy, providing clear evidence for the benefits of an embodied, active approach for replaying complex analytical processes. Our contributions are:

- The design and implementation of replay prototype systems for immersive analytics task handover on both PC and in VR.

- Empirical knowledge of the trade-offs between different perspectives (1PP vs. 3PP), navigation control (passive vs. active) and platform (PC vs. VR) for replaying immersive analytic processes in task handover.

- A set of design guidelines for replay systems of immersive analytics tasks.

2 RELATED WORK

Replay in Extended Reality. Replay techniques are widely used in XR to facilitate asynchronous collaboration [62], tutoring [36], and socialization [63]. Unlike screen-based video, replaying immersive sessions requires systematically tracking user inputs, virtual objects, and avatars to reconstruct interactions [36, 62]. Prior systems have demonstrated replay's promise in collaborative contexts like virtual museum tours [11] and furniture placement tasks [62], showing improvements in teamwork and communication clarity. However, these systems are primarily designed for tasks involving simple object transformations (i.e., changes in position, rotation, and scale). This makes them poorly suited for the dynamic nature of immersive analytics, which involves frequent object instantiation, destruction, and complex data-centric interactions. Our work addresses this critical gap.

Design Factors in Replay Techniques. We identify three key design factors from prior work.
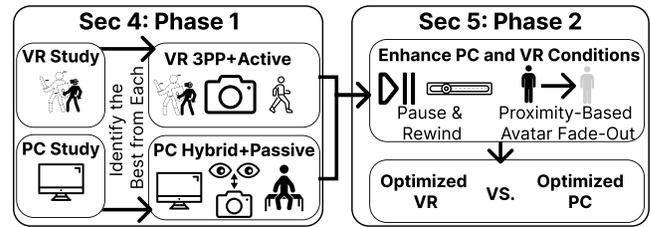

Figure 2: Project Procedure Overview. In Phase 1, we identified the best configuration separately from PC and VR. In Phase 2, we first enhanced the conditions by introducing most frequently mentioned features, and then compared the optimized VR and PC conditions.

Perspective: Replay can be presented from a first-person (1PP) or third-person (3PP) perspective. While 3PP is often used in VR to mitigate motion sickness [45], the choice involves trade-offs: 3PP can enhance spatial memory [26], while 1PP is often associated with a greater sense of immersion [12]. It remains unclear how these perspectives affect the comprehension of complex analytical workflows.

Navigation: Viewers can be granted active, user-controlled navigation or guided through a passive, pre-defined camera path [38]. Active navigation can improve spatial learning and engagement [7, 6, 19], but some studies suggest passive navigation can yield comparable results in spatial memory tasks [18]. These conflicting findings indicate that the effects of navigation are highly task-dependent, motivating our investigation.

Viewing Platform: The choice between immersive (VR) and non-immersive (PC) devices has yielded mixed results. Some studies report similar performance in spatial learning tasks, while others find VR can outperform PCs in engagement and retention [33]. Desktop systems may offer lower immersion but provide stable performance with minimal physiological stress [30], highlighting situational advantages of each platform.

Immersive Analytics and Task Handover. Immersive analytics (IA) leverages VR/AR to enhance data visualization and interaction [40], using embodied interaction to build complex visualizations like scatterplots and parallel coordinate plots in 3D space [9, 10]. An important step in collaborative IA tasks is handover, which is process of transferring knowledge and progress between collaborators working asynchronously [52]. For complex data tasks, a successful handover requires sharing analytical provenance—the historical record of the interactions and reasoning used to reach an insight [64]. In 2D environment, systems such as VisTrails[17] and KTGraph[72] have demonstrated that providing a structured investigation history helps subsequent analysts to recall the critical facts and the predecessor's insights. However, the spatial and embodied nature of immersive analytics makes these handovers more complex, as traditional externalizations like notes or static logs often fail to capture the full context of a predecessor's interactions and spatial focus[71]. While recent research into "embodied provenance" has begun to address how to track these 3D interactions [71], the effectiveness of using full-session replay as a primary communication tool remains established only for simpler, non-analytical tasks [11, 62], and the design guidelines to use replay for IA task handover remains underexplored. Our study addresses this gap by investigating how viewing platform, perspective, and navigation control in IA replay impact a user's ability to comprehend complex, data-rich workflows.

3 PROJECT PROCEDURE OVERVIEW

The overarching goal of this research is to systematically investigate the design trade-offs for replaying immersive analytics sessions. Informed by prior work in replaying generic VR sessions, we identified three key factors: viewing platform (PC vs. VR) [62], perspective (1PP vs. 3PP) [26], and navigation control (active vs. passive) [19]. A full factorial design would yield eight conditions; however, since a first-person perspective inherently implies pas-
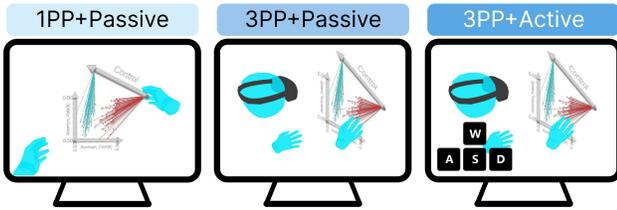
Figure 3: The three PC conditions evaluated in Phase 1. **1PP-Passive**: A first-person view from the original analyst's perspective with no user interaction. **3PP-Passive**: A third-person view that follows the analyst's avatar along a pre-defined spatial trajectory. **3PP-Active**: A third-person view that permits viewer-initiated navigation through mouse and keyboard, like some video games.
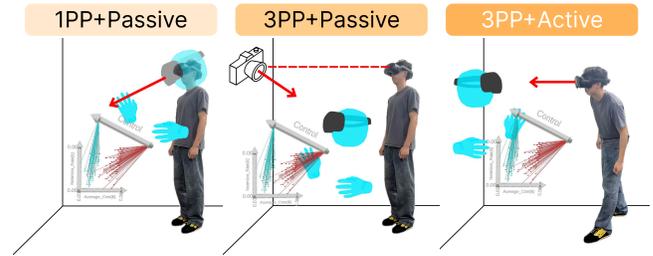

Figure 4: The three VR conditions evaluated in Phase 1. **1PP-Passive**: A first-person view from the original analyst's perspective with no user interaction. **3PP-Passive**: A third-person view that follows the analyst's avatar along a pre-defined spatial trajectory. **3PP-Active**: A third-person view that permits viewer-initiated navigation through body movements.

sive navigation, the number of valid conditions is reduced to six. This is still too many for a participant to experience in a single session without introducing significant fatigue. Furthermore, given the absence of prior empirical design guidelines for building immersive analytics replay systems, a single-phase study would risk comparing underdeveloped interfaces. We therefore adopted an iterative design process, beginning with basic prototypes, gathering feedback through evaluation, and refining the prototypes for a final comparative study. To ensure the feasibility of this approach and to generate more polished designs, we structured our systematic evaluation into a two-phase methodology (see Figure 2).

In **Phase 1**, our objective was to determine the optimal replay configuration for each platform (PC and VR) independently. To achieve this, we conducted two separate within-subject user studies, with each study focusing on one platform. Within each platform, we investigated the effects of perspective and navigation control on the replay experience across three study conditions: (1) 1PP with passive navigation, (2) 3PP with passive navigation, and (3) 3PP with active navigation.

In **Phase 2**, we first analyzed the results and qualitative feedback from Phase 1 to refine our prototypes. Based on this analysis, we implemented two additional features that were frequently requested by participants. We then conducted a comparative evaluation between the two improved designs. By incorporating the newly implemented user-centered features, this final study aimed not only to evaluate the trade-offs between viewing platforms (PC vs. VR) but also to provide more ecologically valid design guidelines and implications.

The proposed study procedure was approved by the Institutional Review Board (IRB) in our university before it was conducted.

4 PHASE 1: IDENTIFYING THE OPTIMAL REPLAY CONFIGURATIONS WITHIN PC AND VR

The first phase was designed to systematically evaluate the impact of perspective and navigation control on the effectiveness of replaying immersive analytics sessions. The primary goal of Study 1 was twofold: (1) to empirically determine the optimal replay configuration within both PC and VR platforms independently, and (2) to gather evidence-based insights that would inform the development of improved, user-centric replay systems for each medium.

4.1 Study Design and Conditions

Study 1 employed a within-subjects experimental design for each of the two viewing platforms. Each participant experienced three distinct replay conditions, see Figure 3 for illustrations of PC conditions and Figure 4 for VR conditions. The order of these conditions, along with the three experimental tasks, was fully counterbalanced using a Latin Square design to mitigate learning and fatigue effects. The three conditions experienced by each participant were: 1PP+Passive, 3PP+Passive, and 3PP+Active, as described below.

1PP+Passive: In this condition, the participant views the session from the original analyst's viewpoint. **On the PC**, this was pre-

sented as a standard, pre-recorded first-person video. **In VR**, early pilot testing revealed that directly mapping the raw, recorded head movements to the user's HMD induced significant motion-sickness due to the mismatch between the user's physical head motion and the virtual motion. To mitigate this, we implemented a system analogous to a 360-degree video. Participants were seated and could freely move their heads to look around, but their virtual position in the environment was determined by the analyst's recorded path. To reduce motion sickness from erratic head movements, we smoothed the recorded head position and rotation data using a one-second moving average filter, resulting in a more stable and comfortable viewing experience.

3PP+Passive: In this condition, the participant observes the original analyst, represented by a virtual avatar, from an external viewpoint. On both PC and in VR, the camera followed a carefully choreographed, pre-recorded path designed to capture the analyst's activities and their interactions with the data visualizations. As no mature techniques currently exist for automatically generating an optimal camera path for such tasks, we manually recorded the camera path for each task replay. This path was standardized across all participants and was designed to focus on the most salient information, including the specific visualizations being manipulated and the avatar's body language and gaze direction.

In the VR version, similar to the 1PP+Passive VR condition, the participant's viewpoint was fixed to the moving camera path, but they retained the freedom to move their head to observe the scene from that viewpoint, analogous to being on a guided tour track.

3PP+Active: This condition granted participants full control over their movement through the recorded scene while observing the analyst's avatar.

In VR, this was inherently supported, allowing participants to freely and physically walk around the virtual environment to explore the scene from any position or angle.

On PC, we used a control scheme common in video games. Participants used the W, A, S, and D keys for planar translation (forward, left, backward, and right) [15]. To accommodate vertical movement, which pilot users indicated was desirable for getting better views of 3D visualizations, two additional keys (Q and E) were mapped for moving up and down. Rotational control of the camera (yaw and pitch) was mapped to mouse movement.

The replay system was developed in the Unity3D game engine. For the immersive analytics environment, we adapted ImAxes [10], an open-source Unity3D-based library that supports the embodied creation and interaction with visualizations in VR. The replay functions were implemented through frame-based scene logging.

4.2 Datasets and Tasks

In each condition, the overall task for participants is first viewing a replay of an IA session, and then answer questions on task comprehension and reconstruct the wrokflow. To ensure our findings were grounded in realistic scenarios, we selected four widely-used

datasets to produce those replay sessions, including the Starbucks nutrition dataset, the US Colleges dataset, a wine quality dataset, and the Spotify song attributes dataset. The Spotify dataset was used for the initial training session, while the other three were used for the main experimental tasks. To control task difficulty and ensure comparability across the within-subject conditions, each dataset was manually curated to contain 10-12 variables and approximately 200 data points. All replay sessions were designed as structured exploratory analysis process. Each replay depicted an analyst investigating a central theme (e.g., identifying chemical properties that influence wine quality) while also exploring other emergent questions. The analyst's actions were adapted from a visual analytics task paradigm [53], including Construct, Discard, Categorize, Filter, Explore (Details-on-demand) and Identify. The visualizations created included 2D/3D scatterplots, linked visualizations, and parallel coordinate plots from the ImAxes system [10]. By keeping the number of analytical steps and visualizations constructed consistent across the three sessions, we maintained a comparable level of complexity.

The purpose of Phase 1 is to compare different perspective and navigation control settings. To avoid confounding factor of unnecessary interactions, we did not provide pausing or rewinding the replay. Participants watched the replay sessions uninterrupted from the beginning to the end. Each session was about six minutes.

4.3 Participants

A total of 24 participants were recruited for Study 1 from the local university community through mailing lists and campus announcements. Participants were required to be over 18 years of age with normal or corrected-to-normal vision and no prior history of conditions that could be exacerbated by VR use, such as severe motion sickness. The recruited participants were 13 males and 11 females, aged between 19 and 29 years ($M = 24.1$, $SD = 2.3$). We also collected their self-reported previous experience with VR ($M = 2.52$, $SD = 1.05$) and data visualisation($M = 3.68$, $SD = 0.90$).They were randomly and evenly assigned to one of two independent groups: the PC Group ($N = 12$) or the VR Group ($N = 12$).

4.4 Procedure

Each experimental session lasted approximately 85 minutes and followed a standardized procedure.

After providing informed consent and completing demographic and pre-study questionnaires, participants underwent a 10-minute training session, which was later repeated before each condition correspondingly. This session was designed to familiarize them with the replay interface for their assigned medium (PC or VR), the perspective and navigation control, and the nature of the immersive analytics tasks, using the pre-prepared Spotify music dataset.

Following the training, participants proceeded through the three counterbalanced experimental conditions. Before each condition, they were given a printed document providing the context for the specific analysis task (i.e., the College, Wine, or Starbucks dataset). They then watched the corresponding six-minute replay of the analyst's session. They were instructed to look for answers to the *Comprehension Questionnaire*. In VR, the questions were shown on a wrist-mounted UI panel. The user can look at the questions when they feel necessary. On the PC, the questions were printed on a paper for the same purpose. Immediately after each replay, participants completed the two primary task performance measures: the *Comprehension Questionnaire* and the *Workflow Reconstruction* task. Afterward, they filled out the NASA-TLX and a user satisfaction questionnaire. A mandatory three-minute break was provided between each condition to minimize fatigue.

Upon completing all three conditions, participants were asked to rank their preference for each format. The session concluded with a ten-minute semi-structured interview to gather in-depth qualitative feedback on their experience, the challenges they faced, and the perceived benefits of each replay format.

4.5 Measures

Our evaluation focused on holistically assessing each participant's understanding of both (a) the data-related insights and (b) the analytical workflow used to reach those insights. We employed a mix of quantitative and qualitative measures.

Task Performance: We apply *Comprehension Questionnaire:* A five-item questionnaire designed to test understanding of the session. The questions assessed recall of basic visual details, comprehension of simple data insights and understanding of the analyst's intent. *Workflow Reconstruction:* Participants were given a pen and paper and asked to reconstruct the entire analytical workflow they had just observed. They were instructed to use the pre-defined action types to create a flowchart representing the analyst's process.

Subjective Measures: After each condition, participants completed standard psychometric questionnaires: *Cognitive Load:* Measured using the NASA-TLX questionnaire. *Satisfaction:* Measured using a standard user satisfaction rating scale.

Qualitative Feedback: At the end of the study, a semi-structured interview was conducted to gather in-depth qualitative feedback on participants' experiences, preferences, and the reasoning behind their performance.

We did not compare the completion time, as the completion time was determined by the sessions rather than the user performance. All sessions also share similar durations.

4.6 Hypotheses

Our hypotheses for Phase 1 were formulated based on previous literature [38, 18, 26] to address how perspective and navigation control impact a user's task understanding, cognitive load, and user satisfaction in a replayed immersive analytics session. These hypotheses were tested separately within both the PC and VR groups.

$H(Pers \cdot Perf)$: We hypothesized that the 3PP-Passive condition would lead to better task performance (higher comprehension and workflow reconstruction scores) compared to the 1PP-Passive condition, as the exocentric view provides greater situational awareness.

$H(Pers \cdot Load)$: We hypothesized that the 3PP-Passive condition would result in a lower cognitive load and higher satisfaction compared to the 1PP-Passive condition, as it reduces the mental effort for spatial alignment and perspective-taking from a potentially disorienting first-person viewpoint.

$H(Nav \cdot Perf)$: We hypothesized that the 3PP-Active condition would result in better task performance compared to the 3PP-Passive condition. We predicted that user-directed navigation would allow participants to seek information more effectively and clarify spatial relationships on their own terms.

$H(Nav \cdot Load)$: We hypothesized that while active navigation might increase physical workload, it would result in lower cognitive load and higher user satisfaction, benefited from the agency to control one's viewpoint.

4.7 Results

This section presents the statistical analysis of the quantitative data from Phase 1. Our analysis aimed to identify the optimal replay configurations for both the PC and VR media platforms. For all tests, we report $p < 0.05$ as statistically significant, and $p < 0.1$ as marginally significant. Given that Shapiro-Wilk tests indicated that the data did not consistently meet the assumptions of normality, we employed non-parametric methods for our analysis. The Friedman test was used to detect overall differences among the three within-subjects conditions. When a significant effect was found, post-hoc pairwise comparisons were conducted using the Wilcoxon signed-rank test with a Bonferroni correction for multiple comparisons. Task performance results and subjective measure results are shown in Figure 5 and Figure 6 respectively.

4.7.1 VR Group

For the 12 participants in the VR group, analysis of task performance and preference revealed a clear advantage for the active navigation condition.

Task Performance: A Friedman test showed a statistically significant effect of the replay condition on the workflow reconstruction scores ($\chi^2(2) = 13.56, p = .001$). Post-hoc tests revealed that participants in the 3PP-Free Navigation condition achieved significantly higher reconstruction scores than those in both the 1PP-Passive ($Z = 0.0, p = .015$) and 3PP-Passive ($Z = 0.0, p = .010$) conditions.

Similarly, a significant effect was found for the comprehension question scores ($\chi^2(2) = 9.77, p = .008$). Pairwise comparisons again showed that the 3PP-Free Navigation condition significantly outperformed both the 1PP-Passive ($Z = 4.0,\ p = .038$) and 3PP-Passive ($Z = 0.0,\ p = .047$) conditions. These results support $H(Nav \cdot Perf)$ and $H(Per \cdot Perf)$, indicating that for VR-based replays, providing users with agency to freely navigate the scene leads to a significantly better understanding of both the analytical workflow and the data insights.

Subjective Ratings: User preference data strongly corroborated the performance results. A Friedman test on the preference rankings was statistically significant ($\chi^2(2) = 6.20, p = .045$), with 3PP-Free Navigation being the most preferred condition. Analysis of the NASA-TLX data also showed significant differences in Frustration ($\chi^2(2) = 6.00, p = .049$) and Satisfaction ($\chi^2(2) = 9.05, p = .011$).Post-hoc tests revealed that participants were significantly less frustrated ($Z = 12.5, p = .0065$) and more satisfied($Z = 9.0, p = .031$) in the 3PP-Free Navigation condition compared to the 1PP-Passive condition.

Feedbacks: Qualitative feedback from interviews strongly supported the quantitative preference for the 3PP-Active condition, revealing three key themes: ***Agency in Free Navigation Was Critical for Comprehension.*** The preference for 3PP-Active was driven by the sense of control it afforded. Eight participants highlighted that freedom of movement allowed them to investigate the scene from optimal angles and revisit points of interest. As P12 summarized, *"Free navigation felt most natural... it offered the most flexible angle."* This ability to *"go around and double check"* (P1) was directly linked to better performance and lower cognitive load, as it empowered users to resolve their own confusion. ***Passive Viewpoints Caused Disorientation and Frustration.*** Conversely, passive conditions were criticized for being restrictive and uncomfortable. Seven participants described the 1PP condition as dizzying and unnatural, with P2 noting, *"First person was more dizzy, the mental strain was higher."* A key issue was the lack of agency, which made p7 *"feel a loss of control"*. While the 3PP-Passive condition was an improvement, its pre-recorded camera path was still problematic; some found its movements disorienting (P10) or the viewpoint too distant from important details (P3). ***Visual Occlusion Was a Major Obstacle.*** A recurring complaint across both third-person views was visual occlusion by the analyst's avatar. Six participants mentioned that the avatar's head would occasionally block critical information. As P4 stated, the avatar "helped in understanding the intention, except for blocking the view."

4.7.2 PC Group

In contrast to the VR group, the results for the 12 participants in the PC group were less conclusive regarding task performance but still revealed clear trends in user experience and preference.

Task Performance: Friedman tests on the performance data for the PC group showed no statistically significant differences across the three conditions for either workflow reconstruction ($\chi^2(2) = 3.62, p = .016$) or comprehension question scores ($\chi^2(2) = 3.35, p = .019$). However, pairwise comparisons revealed marginal significance exists between 1PP Passive and 3PP Active conditions

($Z = 3.0, p = .09$) in workflow reconstruction, which does not support $H(Nav \cdot Perf)$ or $H(Nav \cdot Perf)$. As an extension, we conducted between-group analysis in corresponding PC and VR conditions. While there were no significant performance differences between passive viewing conditions, the VR group achieved significantly higher comprehension scores than the PC group ($U = 20.50$, $p = .0024$) in the 3PP-Active condition, highlighting the platform-dependent impact.

Subjective Ratings: While performance was comparable across conditions, subjective ratings revealed notable differences. A Friedman test on preference rankings was significant ($\chi^2(2) = 8.00, p = .018$), with the 1PP-Passive video format being the most preferred. The 3PP-Free Navigation condition was rated as the least preferred. The NASA-TLX results showed a borderline significant difference in Physical Demand ($\chi^2(2) = 5.84, p = .005$), with post-hoc tests indicating that the 3PP-Free Navigation condition was perceived as significantly more physically demanding than the 3PP-Passive condition ($p = .034$). There was also a borderline significant difference in Satisfaction ($\chi^2(2) = 4.75, p = .093$), with pairwise comparisons showing that both the 1PP-Passive and 3PP-Passive conditions were significantly more satisfying than the 3PP-Free Navigation condition ($p = .039$ and $p = .045$, respectively). This suggests that on a desktop, the additional effort required for manual navigation did not translate into a better user experience and was in fact detrimental to satisfaction. The results reject both $H(Nav \cdot Pref)$ and $H(Per \cdot Pref)$.

Qualitative Feedback: The interviews with the 12 participants in the PC group revealed a strong preference for passive, guided experiences. The key feedback includes: ***Preference for Guided, Focused Views.*** The most prevalent theme was a preference for the 1PP-Passive video condition, which 9 out of 12 participants ranked as their top choice. The primary reason was that it minimized distractions and cognitive overhead. In contrast, the third-person views had *"more visual contents that are not related to data insight."* Participants felt direct perspective made it easier to follow the analyst's focus. **Free Navigation on PC Increases Cognitive Load.** In opposition to the VR group, 3PP-Free Navigation was the least preferred. Eight participants ranked it last, citing the difficulty of control and the increased mental effort required. P23 found that having to manually control the camera *"consumes attention and makes you skip important information."* P22 elaborated on this, stating that *"I'd rather focus on data, instead of paying attention to control it on PC."* This difficulty was compounded by unfamiliar controls (P23) and the challenge of simultaneously tracking the avatar and the complex 3D visualizations (P18). **Avatar Occlusion and Distraction.** Similar to the VR group, visual occlusion by the avatar was a frequently mentioned issue in the 3PP and Free Navigation conditions. However, on the PC, some participants also perceived the avatar itself as a distraction(P18, P21, P24). This indicates that without a strong sense of co-presence, the on-screen avatar could be seen as unnecessary visual element.

4.8 Discussion

The results of Study 1 provide key insights into how the choice of perspective and navigation control impacts the user experience differently on PC and VR platforms. The most interesting finding is the clear divergence in user preference and performance between the two media groups, indicating that the optimal design for an immersive analytics replay system is highly platform-dependent.

The Effect of Navigation Control: Agency in VR and Cognitive Load on PC. Our study reveals a fundamental difference in how users perceive and benefit from navigation control in VR versus on a PC. For the **VR** group, providing active navigation (3PP-Free) led to a statistically significant improvement across all measures: task performance, user preference, and satisfaction. This finding aligns with prior work on spatial learning, which posits that active, embodied exploration leads to the formation of more robust

cognitive maps and enhances comprehension [8]. In our study, the embodied nature of VR made physical navigation feel intuitive. As participants noted in interviews, the freedom to reposition themselves to get a better vantage point or "go and double check" details was crucial for identifying data from different dimension, which were significant barriers in the passive conditions. This user agency appears to be a key factor in reducing frustration and improving overall task understanding.

Conversely, the **PC** group showed the opposite trend. The 3PP-Free Navigation condition was the least preferred, rated as significantly less satisfying, and was perceived as more physically demanding than the passive conditions. Crucially, this increased effort did not translate to better performance, as there was no significant improvement in task scores. This finding is consistent with Cognitive Load Theory [56], which suggests that mentally demanding tasks unrelated to the primary learning objective can hinder performance. On a 2D screen, navigating a 3D space with a keyboard and mouse is an abstract, learned skill that imposes significant cognitive load. As participant P22 mentioned, it is "better to focus on data, instead of paying attention to control it on PC." We also observed that some users in this condition would navigate to a single, preferred viewpoint and then remain stationary for the majority of the replay, effectively opting out of continuous navigation. For desktop users, the mental effort of navigation competed with the primary task of understanding the analysis, making guided, passive conditions a more efficient choice.

The Effect of Perspective: Situational Awareness vs. Guided Focus. The choice of perspective also yielded different results across the platforms, highlighting a trade-off between the situational awareness of a third-person view and the direct focus of a first-person view.

In **VR**, the **1PP** condition was consistently rated as the least preferred and was a significant source of discomfort for many participants, confirming that replaying raw, first-person head movements can be a disorienting experience. The **3PP** view, which provided a more stable, external perspective, was preferred, as it allowed users to understand the analyst's actions within the broader context of the virtual environment.

On the **PC**, however, the **1PP**-Passive video was the most preferred format. Participants found this view to be the most efficient and least distracting, as it directed their attention precisely where the original analyst was looking, with no avatar to cause occlusion. On a limited screen space, the 3PP view could place key visuals further away, and the avatar itself was occasionally occluding with visualisations. However, our qualitative feedback also indicated that while the 1PP view was better for seeing details, the 3PP view was still useful for gaining a contextual overview, especially when more than one visualisation were explored. This suggests that for PC users, both a clear, ego-centric focus for detail and a broader, exo-centric view for context are important.

Deriving the Optimal Configurations for Study 2

Our findings from Study 1 demonstrate that a universal approach to designing replay systems is suboptimal. Therefore, two different formats were chosen for Study 2 configurations.

For **VR**, the data unanimously points to 3PP-Free Navigation as the superior format. It yielded the best task performance, the highest user satisfaction, and was the most preferred condition. The benefits of embodied agency in VR are clear.

For the **PC**, the optimal choice is less straightforward. The 1PP-Passive condition was preferred for its focused view, while the 3PP-Passive view was valued for its contextual awareness. The two conditions did not produce a salient difference in task performance. Recognizing this trade-off, and in response to user feedback (P17) suggesting a desire for both, we chose to develop an improved, hybrid 1PP/3PP-Passive condition for Study 2. This design aims to synthesize the benefits of both perspectives, allowing users to switch freely between the focused view of 1PP for detailed analysis
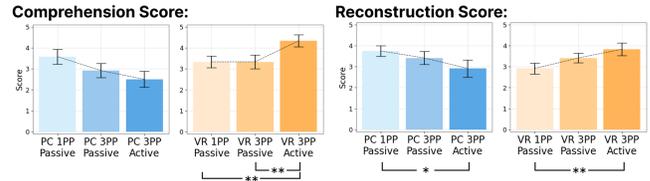

Figure 5: Average Comprehension and Reconstruction Scores in PC and VR Conditions. * indicates borderline significance ($.05 < p < .1$), ** indicate significance ($p < .05$)
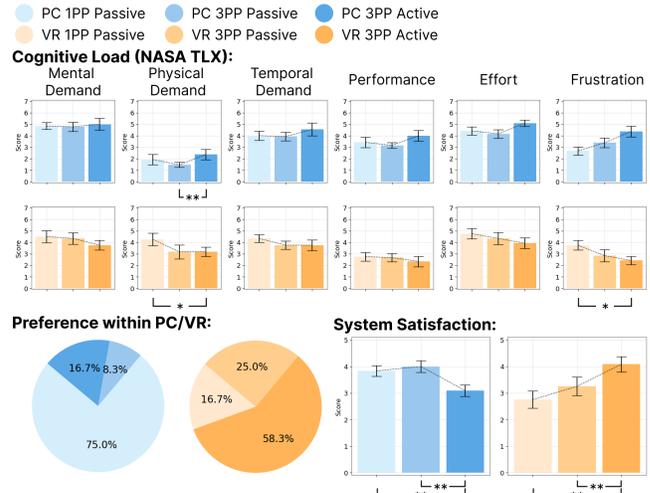

Figure 6: Subjective ratings for PC and VR conditions: Cognitive Load, Preference and Satisfaction

and the contextual overview of 3PP for understanding the broader workflow.

Furthermore, based on consistent feedback from both groups, we identified two additional features to be implemented in the systems for Study 2. First, to address the frequently mentioned issue of avatar occlusion, a **proximity-based fade-out** feature will be added, making the avatar more transparent as the user gets closer. Second, to address the high temporal demand reported in the NASA-TLX scores and make the settings more applicable for complex IA tasks, **pause and rewind** capabilities will be integrated.

5 PHASE 2: COMPARING OPTIMIZED PC AND VR REPLAY SYSTEMS

The second phase of the user study directly compared the optimized replay configurations for PC and VR that were derived from the findings of Study 1. By comparing these two formats, we aim to compare the optimized VR and PC replay system in terms of task performance, cognitive load, and user satisfaction, and produce empirical insights on the design of IA replay systems.

5.1 Conditions and System Improvements

Based on the results and qualitative feedback from Phase 1, we developed two optimized replay conditions for this second phase. **Optimized VR Condition** was 3PP+Active, while **Optimized PC Condition** featured a novel hybrid 1PP/3PP+Passive format. This design was created to synthesize the benefits of both passive perspectives from Study 1, addressing the trade-off between the focused, detailed view of 1PP and the contextual, situational awareness of 3PP. Users could instantly switch between the two perspectives at any time by pressing a key.

Additionally, two new features were added for both conditions:

Pause and Rewind: Participants were given full control over the playback timeline. On PC, a standard slider and pause button were provided. In VR, when the user presses the controller trigger, a timeline slider appears in front of them. By holding the trigger
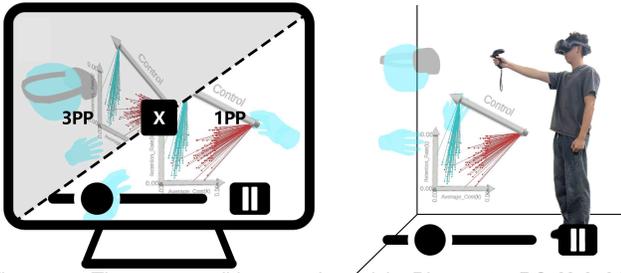
Figure 7: The two conditions evaluated in Phase 2. **PC Hybrid (1PP/3PP-Passive)**: A guided experience where the viewer cannot navigate freely but can instantly switch between 1PP and 3PP. **VR (3PP-Active)**: An embodied experience where the viewer observes from a 3PP and has full navigational control via physical movement. Both conditions were enhanced with user-requested features from Phase 1, including full playback controls and a proximity-based avatar fade-out to mitigate visual occlusion.

and moving the controller horizontally, they can drag the slider to the left or right and set a new time point. This feature was added to represent real-world use cases by giving users the agency to revisit specific moments and reduce the temporal demand of the task.

Proximity-Based Avatar Fade-Out: To address the prevalent issue of avatar occlusion, a dynamic fade-out feature was implemented. The analyst's avatar would automatically become more transparent as the user's viewpoint moved closer to it within 1.5 meters threshold, ensuring that the user's view of the data visualizations was clear.

5.2 Tasks and Measures

The tasks for Phase 2 used the Starbucks and College datasets from the first study, with the Spotify dataset again used for training. To better leverage the new playback controls and to increase the task's analytical depth, the replayed sessions were redesigned to include more analytical steps and visualizations. Each session last about 30 minutes. The measures were also refined to capture quantified task performance and adapt to the new tasks:

Comprehension Questionnaire: The questionnaire was expanded to seven items, including two questions on *visual details* (e.g. what did blue and red represent in the visualisations respectively?), two on the *analyst's intention* (e.g. how did the author group the variables in the first step?), and three on *data-related insights* (e.g. what was the relationship between retention rate and admission rate?), with one question designed to test a *deeper understanding across multiple analytical steps* (e.g. Compare the difference of data distribution in average cost and graduate earnings).

Spatial Recall: A new measure was introduced to specifically test spatial memory. After each replay, participants were asked to use a pen and paper to draw the relative position of a chosen visualization from the session.

Interaction Logging: We quantitatively logged user interactions, including total task completion time, the number and timing of pauses and rewinds, the number of perspective switches in the PC condition, and the frequency and duration of question panel views.

Other measures, including workflow reconstruction, NASA-TLX, and satisfaction ratings, remained the same as in Phase 1.

5.3 Participants and Study Design

A new group of 12 participants, recruited using the same criteria as Phase 1, took part in the second study. The recruited participants were 7 males and 5 females, aged between 18 and 25 years ($M = 22.5$, $SD = 1.1$). We also collected their self-reported previous experience with VR ($M = 3.08$, $SD = 1.08$) and data visualization ($M = 3.5$, $SD = 1.09$). The study employed a within-subjects design, where each participant experienced both of the refined replay conditions on PC and in VR. The order of the two conditions as well as the two experimental tasks were counterbalanced using a Latin
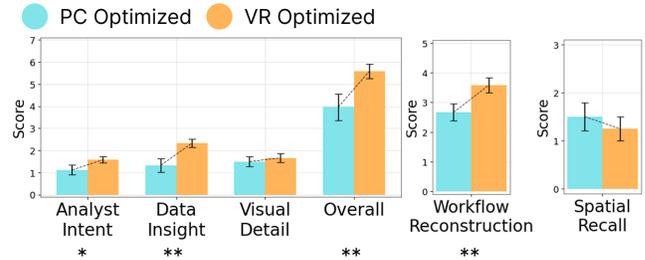

Figure 8: Performance Comparison in Optimized PC and VR Condition: Average Comprehension Sub-scores, Reconstruction Score and Spatial Recall
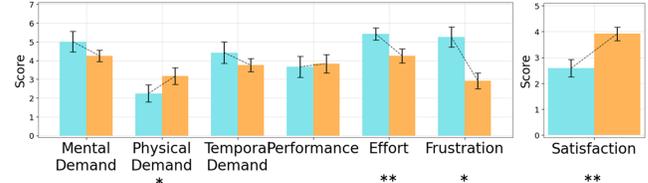

Figure 9: Subjective Ratings Comparison for Optimized PC and VR Condition: Workload and Satisfaction

Square design.

5.4 Procedure

Each session lasted approximately 70 minutes. The procedure was similar to that of Phase 1, with participants first completing a training session for one of the conditions. Before each replay, they were given the task context. The comprehension questions were made available to the participants during the replay. On the PC, the questions could be toggled on and off with a keypress. In VR, the questions were displayed on a wrist-mounted UI panel. After completing the first condition and its associated measures, participants took a break before proceeding to the second training and experimental session with the other condition. The study concluded with a final preference ranking and a semi-structured interview.

5.5 Hypotheses

Our hypotheses for Study 2 focused on the comparison between the two optimized media platforms.

$H(Perf)$**:** We hypothesized that the VR condition would lead to better task performance (shorter task completion time, higher comprehension, workflow, and spatial recall scores) than the PC condition, due to the benefits of embodied agency and spatial immersion.

$H(Load)$**:** We hypothesized that the VR condition would result in a higher physical workload and cognitive workload compared to the PC condition.

$H(Pref)$**:** We hypothesized that the VR condition would be rated as more satisfying and would be the preferred format due to its immersive and engaging nature.

5.6 Results

This section presents the statistical analysis comparing the optimized PC and VR replay formats from Phase 2. Our analysis aimed to evaluate differences in task performance, cognitive load, and user satisfaction. We used the same statistical methods as of Phase 1.

Task Performance: As shown in Figure 8 Our analysis revealed significant differences in task performance between the PC and VR conditions, as measured by comprehension, workflow reconstruction, and spatial recall scores. In support of $H(Perf)$, the VR condition led to significantly better task comprehension. A Wilcoxon signed-rank test showed that participants achieved a higher mean comprehension score in VR ($M = 5.58$, $SD = 1.12$) compared to PC ($M = 3.96$, $SD = 2.08$), with the difference being statistically sig-

nificant ($W = 6.5$, $p = .018$). Further supporting $H(Perf)$, the VR condition also resulted in superior workflow reconstruction accuracy. Participants' workflow scores were significantly higher in VR ($M = 3.58$, $SD = .90$) than in PC ($M = 2.67$, $SD = .98$), as indicated by a Wilcoxon signed-rank test ($W = 4.00$, $p = 0.047$).

However, we did not find a significant difference in spatial recall scores between the VR ($M = 1.25$, $SD = .87$) and PC ($M = 1.50$, $SD = 1.00$) conditions ($W = 9.50$, $p = .44$), partially rejecting $H(Perf)$. A follow-up analysis did reveal that within the VR condition, participants with greater data visualization experience tended to have higher spatial recall scores, showing a moderate positive correlation ($r = .56$), using Spearman's rank correlation.

Subjective Ratings: As shown in Figure 9 We assessed cognitive and physical load using the NASA-TLX survey. Our findings partially support $H(Load)$: There were no statistically significant differences in the mental, temporal, or performance demand subscales between the two conditions. However, regarding physical demand, VR ($M = 3.17$, $SD = 1.53$) yielded significantly higher scores than PC ($M = 2.25$, $SD = 1.60$) using Wilcoxon signed-rank test ($W = 8.0$, $p = .07$).

The PC condition was perceived as significantly more frustrating than the VR condition ($W = 11.5$, $p = .05$). In line with $H(Pref)$, a Wilcoxon signed-rank test on the satisfaction scores revealed that participants found the VR replay to be more satisfying than the PC replay ($W = 11.5$, $p = .03$). Furthermore, when asked to rank their preference, a vast majority of participants (10 out of 12) chose the VR condition as their preferred format.

Behavioral Data: We examined the difference of behavioral metrics, including task completion time, pause duration, pause counts and perspective switch counts. While no significant difference was detected, Figure 10 provided insights into the user interaction frequency. The darker areas signifies more frequent interactions, suggesting critical periods in the replay. Besides, on PC, higher proportion of 1PP duration was positively correlated with better comprehension score (Spearman's $r = 0.61$, $p = .04$).

Qualitative Feedback: A thematic analysis was adopted to synthesize feedback into three main topics: *1. perspective*, *2. new features* and *3. sensemaking strategy*.

Agency and visual clarity drove a preference for VR, which was consistent with $H(Pref)$. Eight out of 12 participants cited freedom of perspective as the primary reason, which not only allowed them to look from desired angles (P3, P5, P11) but also offered clear visual information (P3, P4, P5, P8, P9). In contrast, the PC condition was criticized for its limited perspective; participants noted that while the avatar no longer hindered the view, visual overlaps still occurred when multiple visualizations existed (P6) and the 3PP view made it difficult to investigate details in 3D scatterplots (P7).

New interactive features functioned as core cognitive tools for information processing. The ability to pause and rewind was acknowledged by ten participants as an essential feature. P4 and P7 also mentioned using it as a core part of their strategy to process information. The proximity-based avatar fade-out was a subtle but effective improvement that successfully prevented the avatar from occluding critical data (P5, P9). Regarding navigation, the perspective switch was actively used. Most particpipants found it smooth to use, but some described it as not ideal (P1, P6).

Participants adopted active sensemaking strategies to manage analytical complexity. Two distinct strategies emerged. The most prevalent approach used by six participants was to go over the entire replay first. Alternatively, some participants (P7, P11) adopted a layered strategy, watching the entire replay once for a general overview before using timeline controls to go back to specific moments for details. These behaviors demonstrate that the replay system supports personalized, active reconstruction of the analytical process rather than passive observation.
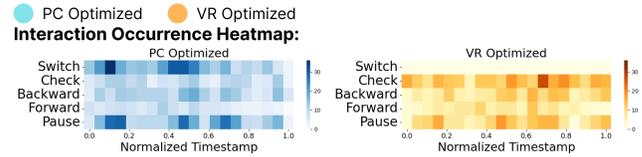

Figure 10: Heatmaps of interaction counts across the normalized task timeline for tested Conditions. Darker colors represent a higher frequency of user interactions within that time window.

5.7 Discussion

The results of Study 2, which compared the optimized replay formats for PC and VR, provide clear evidence supporting the immersive, embodied approach for re-experiencing complex analytical workflows. The discussion below interprets these findings, focusing on the fundamental differences between the two media and the role of the newly implemented interactive features.

Embodied Agency vs. Guided View: Phase 1 illustrated that guided view was an effective practice on PC, while embodied agency was critical for VR. The meaningful comparison in Phase 2, therefore, was to determine which of these optimized paradigms would better support complex IA replays. Our results show a clear advantage for the VR condition in task comprehension ($p = 0.018$) and workflow reconstruction ($p = 0.041$), suggesting that for understanding complex analytical processes, the benefits of embodied agency are more critical than those of a guided view.

The ability to navigate and control one's perspective in VR allows for the construction of a robust mental model of the complete data history. This aligns with findings from Chrastil & Warren[8], which suggest that self-directed navigation is superior for acquiring cognitive map knowledge. Moreover, the task of analyzing a 3D data visualization is intrinsically spatial. The depth dimension in VR provides a more direct representation of the data, reducing the cognitive effort needed to interpret 3D structures on a 2D screen.

Effectiveness of pause, rewind and perspective switch: Most participants believed pause and rewind was the most preferred feature. Participants explicitly described using the pause function to "process information" and "think" (P4), suggesting a direct mitigation of the high temporal demand. By giving users control over the pace of information flow, these features allowed them to manage their own cognitive load. On the other hand, for perspective switch, while most users do it multiple times and think of it as positive, we do not find correlation between perspective switch counts and task performance. Moreover, a Spearman's rank correlation identified the proportion of time spent in 1PP to be positively correlated to comprehension score ($r = 0.608$, $p = 0.0359$), which further implies that 3PP view in PC tend to miss details in visualised data, due to less guided focus and longer distance.

Why spatial recall indicated no signicant difference between PC and VR: This result partially rejects hypothesis $H(Perf)$, suggesting that an immersive environment does not automatically confer better spatial memory for all users. Our demographic analysis provides a critical insight here, as prior experience with data visualization showed a moderate positive correlation with spatial recall performance ($r = 0.56$) only in the VR condition. This implies that the spatial advantages of VR are not automatic; instead, they are most effectively leveraged by users who already possess well-developed spatial reasoning skills. This finding resonates with literature suggesting that individual differences can moderate performance in virtual environments [26].

6 DESIGN GUIDELINES AND IMPLICATIONS FOR IA REPLAY

With the insights from the two studies, we try to convert the practical lessons from our work into four concrete, implementable guidelines for designers of IA replay systems. Notably, while we concluded VR as the better performing platform, there are many scenarios where only traditional desktop settings are available, espe-

cially for asynchronous, remote collaboration. Therefore, we try to target both PC and VR in the design guidelines we present.

Match navigation to the medium: give VR agency, keep PC guided. A primary takeaway from our work is that the optimal user perspective and navigation control is fundamentally tied to the medium. In VR, the replay system should let viewers physically reposition themselves or use natural locomotion to inspect visual elements from different angles and distances. It is an overarching guideline to preserve the users' agency for navigation, so as to immerse users in the analytical environment for better performance and ensure usability. These observations align with previous work finding physical navigation outperform virtual navigation [1, 23]. On the contrary, for desktop platforms, where users view a 3D visualisation scene through a 2D window, the priority should be a guided and focused perspective. Navigation effort should be minimized, with precurated camera paths, as manual navigation increased cognitive load without measurable performance gains. Such navigation burden were also reported in previous studies [15, 58]. The perspective change feature needs to be implemented with caution, as it may lead to loss of focus. It should be considered for coarse-grained, procedural tasks instead of fine-grained, detail-rich analysis.

Reduce visual occlusion and increase clarity. Common obstacles across both studies were the avatar or other visuals blocking important information and users struggling to keep pace with replayed timelines. Implementations such as proximity-based avatar fading [14], pause/rewind, and timeline slider proved to be useful. These features are easily overlooked, but they could be the key point to user experience and task comprehension, as shown in the comparison of 1PP and 3PP Passive conditions. Research indicates that while 1PP enhances immersion and precision, 3PP significantly aids spatial awareness and context gathering during task review [50, 59]. Consequently, we recommend treating these not as optional add-ons, but as baseline requirements for IA replay systems and VR usability guidelines. To further reduce visual overlap, it is beneficial to keep track of the visualization being interacted with or investigated, based on proximity or interaction, a concept aligned with Focus+Context visualization techniques [29, 68]. The tracked visualization could then be highlighted or zoomed in to provide detail. Other visual elements could also reduce opacity to further emphasize the targeted visualization, ensuring that contextual information is preserved without creating visual clutter.

Support replay as an active sensemaking tool instead of a passive video. Our studies showed that participants rarely treated replay as a passive video. Instead, they actively managed the pace and sequence of playback to support their own reasoning, a behavior consistent with the "active viewing" strategies observed in interactive video learning [41]. In Study 2, the pause and rewind feature was described as "most important" and became a core strategy. Participants frequently paused after a key analyst action to consolidate their understanding or used rewind to revisit details they had missed. Some also reported deliberately watching the entire replay once to gain an overview, then returning to specific points to answer detailed questions. This specific workflow mirrors Shneiderman's Visual Information-Seeking Mantra: Overview first, zoom and filter, then details-on-demand [53], confirming that replay is already functioning as an analytic instrument rather than a static recording [47]. Building on this, future systems should strengthen replay's role in collaborative sensemaking by tying replay more directly with analytic actions, moving towards analytic provenance [47]. For example, segments could be linked to operations such as construct, filter, or compare [20], allowing collaborators to navigate by action sequence rather than just timestamps. Similarly, key frame function could make it easier for analysts to capture and share the key steps they pause to reflect on. In this way, replay evolves from a passive narrative into an interactive workspace for reconstructing and communicating analytic provenance.

7 LIMITATIONS AND FUTURE WORKS

Annotation and Note-Taking: Our investigation focused on the core parameters of the replay experience through the two-phase controlled study. For our purpose, we did not test all potentially beneficial features, most notably the integration of explicit insight communication. While our replay system captures the procedural workflow, the original analyst's specific "aha" moments and rationale remain implicit. Future immersive analytics replay systems could benefit from techniques such as audio/video annotation [27], text [13, 34, 60] and highlights [39, 48, 60]. Investigating the most effective methods for embedding these insights directly into the replay could significantly enhance a collaborator's understanding of the analyst's intentions.

Navigational and Temporal Interactions: Another limitation lies in the navigational guidance and temporal control mechanisms, especially for long-duration replay sessions. While active navigation in VR provides agency, users may struggle to identify where to look, increasing cognitive load and the risk of missing critical visual details. Future systems could address this by incorporating visual attention cues such as gaze rays [5] or head rays [44], in order to indicate avatar gaze and highlight visualizations currently being manipulated, and thereby guide viewers without sacrificing their sense of agency. To mitigate the physical demand and fatigue noted by participants in extended sessions, future systems can incorporate teleportation [4, 46] or "follow analyst" re-centering features [42] to provide proactive guidance while maintaining comfort.

Furthermore, to allow more effective timeline control, future work should explore novel temporal interfaces, such as allowing users to place their own keyframe marks [72] for efficient navigation to critical moments. Additionally, more precise slider interaction techniques [43, 69] could be integrated to provide finer temporal control when analyzing dense sequences of data interactions.

Semantic Filtering and Clipping: Inherently, raw replays of analytical processes are noisy, often containing idle time and locomotion that are irrelevant to the data analysis. Removing these segments would make the replay more focused and time-efficient, as watching lengthy replays imposes a significant cognitive burden on the viewer. Future systems could address this by implementing automated filtering and clipping based on key periods identified in provenance logs and user behavior [70]. Furthermore, the integration of LLM-assisted analytics frameworks [31] could help transform complex interaction sequences into a structured, concise narrative. Vision Language Models help remove noises through comprehending and clipping lengthy session recordings [54]. By automatically pruning idle periods and highlighting high-value analytical events, these techniques can lower the burden of session summarization, allowing incoming analysts to rapidly grasp a predecessor's rationale during a handover.

8 CONCLUSIONS

In this paper, we presented an empirical study evaluating session replay techniques for asynchronous task handover in immersive analytics. We evaluated the impact of three key design dimensions: media, perspective, and navigation control, through a two-phase within-subject comparison. Our results reveal that the optimal design for such systems is highly platform-dependent. For VR, a third-person perspective with free, active navigation significantly improved task comprehension and workflow reconstruction. Conversely, on PC, A passive, guided experience was more preferable. A direct comparison of the two optimized formats confirmed that the immersive VR replay led to significantly better task completion time and comprehension accuracy. Our findings can inform the design of future asynchronous collaborative systems in data-centric immersive environment.

REFERENCES

[1] R. Ball, C. North, and D. A. Bowman. Move to improve: promoting physical navigation to increase user performance with large displays. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 191–200, 2007. 9

[2] A. Batch, A. Cunningham, M. Cordeil, N. Elmqvist, T. Dwyer, B. H. Thomas, and K. Marriott. There is no spoon: Evaluating performance, space use, and presence with expert domain users in immersive analytics. *IEEE transactions on visualization and computer graphics*, 26(1):536–546, 2019. 1

[3] M. Billinghurst, M. Cordeil, A. Bezerianos, and T. Margolis. Collaborative immersive analytics. In *Immersive Analytics*, pages 221–257. Springer, 2018. 1

[4] F. Buttussi and L. Chittaro. Locomotion in place in virtual reality: A comparative evaluation of joystick, teleport, and leaning. *IEEE transactions on visualization and computer graphics*, 27(1):125–136, 2019. 9

[5] D. L. Chen, M. Giordano, H. Benko, T. Grossman, and S. Santosa. Gazeraycursor: Facilitating virtual reality target selection by blending gaze and controller raycasting. In *Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology*, pages 1–11, 2023. 9

[6] E. R. Chrastil and W. H. Warren. Active and passive contributions to spatial learning. *Psychonomic bulletin & review*, 19(1):1–23, 2012. 2

[7] E. R. Chrastil and W. H. Warren. Active and passive spatial learning in human navigation: acquisition of survey knowledge. *Journal of experimental psychology: learning, memory, and cognition*, 39(5):1520, 2013. 2

[8] E. R. Chrastil and W. H. Warren. Active and passive spatial learning in human navigation: acquisition of graph knowledge. *Journal of experimental psychology: learning, memory, and cognition*, 41(4):1162, 2015. 6, 8

[9] M. Cordeil, A. Cunningham, B. Bach, C. Hurter, B. H. Thomas, K. Marriott, and T. Dwyer. Iatk: An immersive analytics toolkit. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 200–209. IEEE, 2019. 2

[10] M. Cordeil, A. Cunningham, T. Dwyer, B. H. Thomas, and K. Marriott. Imaxes: Immersive axes as embodied affordances for interactive multivariate data visualisation. In *Proceedings of the 30th annual ACM symposium on user interface software and technology*, pages 71–83, 2017. 2, 3, 4

[11] N. Correia, L. Alves, R. Sa, J. Santiago, and L. Romero. Hypermem: a system to store and replay experiences in mixed reality worlds. In *2005 International Conference on Cyberworlds (CW'05)*, pages 8–pp. IEEE, 2005. 2

[12] A. Denisova and P. Cairns. First person vs. third person perspective in digital games: do player preferences affect immersion? In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, pages 145–148, 2015. 2

[13] J. Dominic and A. Robb. Exploring effects of screen-fixed and world-fixed annotation on navigation in virtual reality. In *2020 IEEE conference on virtual reality and 3d user interfaces (VR)*, pages 607–615. IEEE, 2020. 9

[14] N. Elmqvist and P. Tsigas. A taxonomy of 3d occlusion management for visualization. *IEEE transactions on visualization and computer graphics*, 14(5):1095–1109, 2008. 9

[15] D. Enriquez, W. Tong, C. North, H. Qu, and Y. Yang. Evaluating layout dimensionalities in pc+ vr asymmetric collaborative decision making. *Proceedings of the ACM on Human-Computer Interaction*, 8(ISS):112–132, 2024. 3, 9

[16] B. Ens, B. Bach, M. Cordeil, U. Engelke, M. Serrano, W. Willett, A. Prouzeau, C. Anthes, W. Büschel, C. Dunne, et al. Grand challenges in immersive analytics. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–17, 2021. 1

[17] J. Freire and C. T. Silva. Making computations and publications reproducible with vistrails. *Computing in Science & Engineering*, 14(4):18–25, 2012. 2

[18] F. Gaunet, M. Vidal, A. Kemeny, and A. Berthoz. Active, passive and snapshot exploration in a virtual environment: influence on scene memory, reorientation and path memory. *Cognitive Brain Research*, 11(3):409–420, 2001. 2, 4

[19] A. Giovannelli, L. Pavanatto, S. Davari, H. Miao, V. Chheang, B. Giera, T. Bremer, and D. A. Bowman. Investigating the influence of playback interactivity during guided tours for asynchronous collaboration in virtual reality. In *2025 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, pages 23–33. IEEE, 2025. 1, 2

[20] D. Gotz and M. X. Zhou. Characterizing users' visual analytic activity for insight provenance. *Information Visualization*, 8(1):42–55, 2009. 9

[21] J. Heer, F. B. Viégas, and M. Wattenberg. Voyagers and voyeurs: supporting asynchronous collaborative information visualization. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 1029–1038, 2007. 1

[22] H. H. Huang, H. Pfister, and Y. Yang. Is embodied interaction beneficial? a study on navigating network visualizations. *Information Visualization*, 22(3):169–185, 2023. 1

[23] S. In, E. Krokos, K. Whitley, C. North, and Y. Yang. Evaluating navigation and comparison performance of computational notebooks on desktop and in virtual reality. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, pages 1–15, 2024. 1, 9

[24] S. In, E. Krokos, K. Whitley, C. North, and Y. Yang. Investigating seamless transitions between immersive computational notebooks and embodied data interactions. In *Proceedings of the 2025 31st ACM Symposium on Virtual Reality Software and Technology*, pages 1–11, 2025. 1

[25] S. In, T. Lin, C. North, H. Pfister, and Y. Yang. This is the table i want! interactive data transformation on desktop and in virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, 30(8):5635–5650, 2023. 1

[26] H. Iriye and P. L. St. Jacques. Memories for third-person experiences in immersive virtual reality. *Scientific reports*, 11(1):4667, 2021. 1, 2, 4, 8

[27] A. Irlitti, S. Von Itzstein, L. Alem, and B. Thomas. Tangible interaction techniques to support asynchronous collaboration. In *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 1–6. IEEE, 2013. 9

[28] P. Isenberg, N. Elmqvist, J. Scholtz, D. Cernea, K.-L. Ma, and H. Hagen. Collaborative visualization: Definition, challenges, and research agenda. *Information Visualization*, 10(4):310–326, 2011. 1

[29] D. Kalkofen, E. Mendez, and D. Schmalstieg. Interactive focus and context visualization for augmented reality. In *2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, pages 191–201. IEEE, 2007. 9

[30] K. Kim, M. Z. Rosenthal, D. Zielinski, and R. Brady. Comparison of desktop, head mounted display, and six wall fully immersive systems using a stressful task. In *2012 IEEE Virtual Reality Workshops (VRW)*, pages 143–144. IEEE, 2012. 2

[31] Y. Kim, Z. Aamir, M. Singh, S. Boorboor, K. Mueller, and A. E. Kaufman. Explainable xr: Understanding user behaviors of xr environments using llm-assisted analytics framework. *IEEE Transactions on Visualization and Computer Graphics*, 2025. 9

[32] M. Kraus, J. Fuchs, B. Sommer, K. Klein, U. Engelke, D. Keim, and F. Schreiber. Immersive analytics with abstract 3d visualizations: A survey. In *Computer Graphics Forum*, volume 41, pages 201–229. Wiley Online Library, 2022. 1

[33] K.-W. K. Lai and H.-J. H. Chen. A comparative study on the effects of a vr and pc visual novel game on vocabulary learning. *Computer Assisted Language Learning*, 36(3):312–345, 2023. 2

[34] H. Lee, H. Kim, D. V. Monteiro, Y. Goh, D. Han, H.-N. Liang, H. S. Yang, and J. Jung. Annotation vs. virtual tutor: Comparative analysis on the effectiveness of visual instructions in immersive virtual reality. In *2019 IEEE International symposium on mixed and augmented reality (ISMAR)*, pages 318–327. IEEE, 2019. 9

[35] L. S. Lenz, A. R. Fender, J. Chatain, and C. Holz. Comparing synchronous and asynchronous task delivery in mixed reality environments. *IEEE Transactions on Visualization and Computer Graphics*, 30(5):2776–2784, 2024. 1

[36] Y. Li, Y. Kameda, and Y. Ohta. Ar replay in a small workspace. In *2013 23rd International Conference on Artificial Reality and Telexis-*

tence (ICAT), pages 97–101. IEEE, 2013. 2

[37] L. Lisle, K. Davidson, E. J. Gitre, C. North, and D. A. Bowman. Sensemaking strategies with immersive space to think. In *2021 IEEE virtual reality and 3D user interfaces (VR)*, pages 529–537. IEEE, 2021. 1

[38] T. Lu, Q. Zhu, T. Ma, W. Kam-Kwai, A. Xie, A. Endert, and Y. Yang. Ego vs. exo and active vs. passive: Investigating the individual and combined effects of viewpoint and navigation on spatial immersion and understanding in immersive storytelling. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, pages 1–19, 2025. 2, 4

[39] T. Mahmood, W. Fulmer, N. Mungoli, J. Huang, and A. Lu. Improving information sharing and collaborative analysis for remote geospatial visualization using mixed reality. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 236–247. IEEE, 2019. 9

[40] K. Marriott, F. Schreiber, T. Dwyer, K. Klein, N. H. Riche, T. Itoh, W. Stuerzlinger, and B. H. Thomas. *Immersive analytics*, volume 11190. Springer, 2018. 1, 2

[41] M. Merkt, S. Weigand, A. Heier, and S. Schwan. Learning with videos vs. learning with print: The role of interactive features. *Learning and Instruction*, 21(6):687–704, 2011. 9

[42] A. Nguyen, P. Wüest, and A. Kunz. Human following behavior in virtual reality. In *Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology*, pages 1–3, 2020. 9

[43] G. Perelman, M. Serrano, A. Marchal, and E. Dubois. 3d timelines in vr: Adaptive speed for 3d data navigation on a concave timeline. *International Journal of Human-Computer Studies*, page 103672, 2025. 9

[44] T. Piumsomboon, A. Dey, B. Ens, G. Lee, and M. Billinghurst. The effects of sharing awareness cues in collaborative mixed reality. *Frontiers in Robotics and AI*, 6:5, 2019. 9

[45] K. Ponto, J. Kohlmann, and M. Gleicher. Effective replays and summarization of virtual experiences. *IEEE Transactions on Visualization and Computer Graphics*, 18(4):607–616, 2012. 2

[46] A. Prithul, I. B. Adhanom, and E. Folmer. Teleportation in virtual reality; a mini-review. *Frontiers in Virtual Reality*, 2:730792, 2021. 9

[47] E. D. Ragan, A. Endert, J. Sanyal, and J. Chen. Characterizing provenance in visualization and data analysis: an organizational framework of provenance types and purposes. *IEEE transactions on visualization and computer graphics*, 22(1):31–40, 2015. 9

[48] N. Reski, A. Alissandrakis, J. Tyrkkö, and A. Kerren. "oh, that's where you are!"–towards a hybrid asymmetric collaborative immersive analytics system. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*, pages 1–12, 2020. 9

[49] D. Saffo, S. Di Bartolomeo, T. Crnovrsanin, L. South, J. Raynor, C. Yildirim, and C. Dunne. Unraveling the design space of immersive analytics: A systematic review. *IEEE Transactions on Visualization and Computer Graphics*, 30(1):495–506, 2023. 1

[50] P. Salamin, T. Tadi, O. Blanke, F. Vexo, and D. Thalmann. Quantifying effects of exposure to the third and first-person perspectives in virtual-reality-based training. *IEEE Transactions on Learning Technologies*, 3(3):272–276, 2010. 9

[51] M. Sereno, X. Wang, L. Besançon, M. J. Mcguffin, and T. Isenberg. Collaborative work in augmented reality: A survey. *IEEE transactions on visualization and computer graphics*, 28(6):2530–2549, 2020. 1

[52] N. Sharma. Sensemaking handoff: When and how? *Proceedings of the American Society for Information Science and Technology*, 45(1):1–12, 2008. 2

[53] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *The craft of information visualization*, pages 364–371. Elsevier, 2003. 4, 9

[54] Y. Shu, Z. Liu, P. Zhang, M. Qin, J. Zhou, Z. Liang, T. Huang, and B. Zhao. Video-xl: Extra-long vision language model for hour-scale video understanding. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 26160–26169, 2025. 9

[55] R. Skarbez, N. F. Polys, J. T. Ogle, C. North, and D. A. Bowman. Immersive analytics: Theory and research agenda. *Frontiers in Robotics and AI*, 6:82, 2019. 1, 2

[56] J. Sweller and P. Chandler. Evidence for cognitive load theory. *Cog-*

nition and instruction, 8(4):351–362, 1991. 6

[57] W. Tong, H. Li, M. Xia, W. Kam-Kwai, T.-C. Pong, H. Qu, and Y. Yang. Exploring spatial hybrid user interface for visual sensemaking. *IEEE Transactions on Visualization and Computer Graphics*, 2025. 2

[58] W. Tong, M. Xia, K. K. Wong, D. A. Bowman, T.-C. Pong, H. Qu, and Y. Yang. Towards an Understanding of Distributed Asymmetric Collaborative Visualization on Problem-solving, Feb. 2023. 1, 9

[59] Y. Ueyama and M. Harada. Effects of first-and third-person perspectives created using a head-mounted display on dart-throwing accuracy. *Virtual Reality*, 26(2):687–695, 2022. 9

[60] T. Ulusoy, K. T. Danyluk, and W. J. Willett. Beyond the physical: Examining scale and annotation in virtual reality visualizations. Technical report, Department of Computer Science, University of Calgary, 2018. 9

[61] J. Walny, C. Frisson, M. West, D. Kosminsky, S. Knudsen, S. Carpendale, and W. Willett. Data changes everything: Challenges and opportunities in data visualization design handoff. *IEEE transactions on visualization and computer graphics*, 26(1):12–22, 2019. 1

[62] C. Y. Wang, L. Drumm, C. Troup, Y. Ding, and A. S. Won. Vr-replay: Capturing and replaying avatars in vr for asynchronous 3d collaborative design. In *2019 IEEE conference on virtual reality and 3D user interfaces (VR)*, pages 1215–1216. IEEE, 2019. 1, 2

[63] C. Y. Wang, M. Sakashita, U. Ehsan, J. Li, and A. S. Won. Again, together: Socially reliving virtual reality experiences when separated. In *Proceedings of the 2020 chi conference on human factors in computing systems*, pages 1–12, 2020. 2

[64] K. Xu, S. Attfield, T. Jankun-Kelly, A. Wheat, P. H. Nguyen, and N. Selvaraj. Analytic provenance for sensemaking: A research agenda. *IEEE computer graphics and applications*, 35(3):56–64, 2015. 2

[65] H. Yang, E. H. Faa, W. Liu, S. Guo, D. H. Chau, and Y. Yang. Litforager: Exploring multimodal literature foraging strategies in immersive sensemaking. *IEEE Transactions on Visualization and Computer Graphics*, 2025. 1

[66] Y. Yang, M. Cordeil, J. Beyer, T. Dwyer, K. Marriott, and H. Pfister. 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):1214–1224, 2020. 1

[67] Y. Yang, T. Dwyer, K. Marriott, B. Jenny, and S. Goodwin. Tilt Map: Interactive Transitions Between Choropleth Map, Prism Map and Bar Chart in Immersive Environments. *IEEE Transactions on Visualization and Computer Graphics*, 27(12):4507–4519, Dec. 2021. 1

[68] Y. Yang, W. Xia, F. Lekschas, C. Nobre, R. Krüger, and H. Pfister. The pattern is in the details: An evaluation of interaction techniques for locating, searching, and contextualizing details in multivariate matrix visualizations. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pages 1–15, 2022. 9

[69] C. Zhang, T. S. Ma, J. Andrews, E. J. Gonzalez, M. Gonzalez-Franco, and Y. Yang. Forcepinch: Force-responsive spatial interaction for tracking speed control in xr. In *Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology*, pages 1–16, 2025. 9

[70] G. Zhang, M. A. N. Ahmed, Z. Hu, and A. Bulling. Summact: Uncovering user intentions through interactive behaviour summarisation. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, pages 1–17, 2025. 9

[71] Y. Zhang, B. Ens, K. A. Satriadi, Y. Yang, and S. Goodwin. Embodied provenance for immersive sensemaking. *Proceedings of the ACM on Human-Computer Interaction*, 7(ISS):198–216, 2023. 1, 2

[72] J. Zhao, M. Glueck, P. Isenberg, F. Chevalier, and A. Khan. Supporting handoff in asynchronous collaborative sensemaking using knowledge-transfer graphs. *IEEE transactions on visualization and computer graphics*, 24(1):340–350, 2017. 1, 2, 9

[73] Q. Zhu, T. Lu, S. Guo, X. Ma, and Y. Yang. Compositingvis: Exploring interactions for creating composite visualizations in immersive environments. *IEEE Transactions on Visualization and Computer Graphics*, 2024. 1