

# RELIVE: Bridging In-Situ and Ex-Situ Visual Analytics for Analyzing Mixed Reality User Studies

Sebastian Hubenschmid\*  
University of Konstanz  
sebastian.hubenschmid@uni.kn

Andrea Batch  
University of Maryland  
ajulca@umd.edu

Jonathan Wieland\*  
University of Konstanz  
jonathan.wieland@uni.kn

Johannes Zagermann  
University of Konstanz  
johannes.zagermann@uni.kn

Daniel Immanuel Fink\*  
University of Konstanz  
daniel.immanuel.fink@uni.kn

Niklas Elmqvist  
University of Maryland  
elm@umd.edu

Harald Reiterer  
University of Konstanz  
harald.reiterer@uni.kn



**Figure 1:** The RELIVE mixed-immersion tool. RELIVE combines an immersive analytics virtual reality view (left) with a synchronized non-immersive visual analytics desktop view (right) for analyzing mixed reality studies. The virtual reality view allows users to relive and analyze prior studies in-situ, while the desktop facilitates an ex-situ analysis of aggregated data.

## ABSTRACT

The nascent field of mixed reality is seeing an ever-increasing need for user studies and field evaluation, which are particularly challenging given device heterogeneity, diversity of use, and mobile deployment. Immersive analytics tools have recently emerged to support such analysis *in situ*, yet the complexity of the data also warrants an *ex-situ* analysis using more traditional non-immersive visual analytics setups. To bridge the gap between both approaches, we introduce RELIVE: a mixed-immersion visual analytics framework for exploring and analyzing mixed reality user studies. RELIVE combines an in-situ virtual reality view with a complementary ex-situ desktop view. While the virtual reality view allows users to

relive interactive spatial recordings replicating the original study, the synchronized desktop view provides a familiar interface for analyzing aggregated data. We validated our concepts in a two-step evaluation consisting of a design walkthrough and an empirical expert user study.

## CCS CONCEPTS

• **Human-centered computing** → **Visual analytics; Virtual reality; User studies.**

## KEYWORDS

Immersive analytics, visual analytics, data visualization, virtual reality.

## ACM Reference Format:

Sebastian Hubenschmid, Jonathan Wieland, Daniel Immanuel Fink, Andrea Batch, Johannes Zagermann, Niklas Elmqvist, and Harald Reiterer. 2022. RELIVE: Bridging In-Situ and Ex-Situ Visual Analytics for Analyzing Mixed Reality User Studies. In *CHI Conference on Human Factors in Computing Systems (CHI '22)*, April 29-May 5, 2022, New Orleans, LA, USA. ACM, New York, NY, USA, 21 pages. <https://doi.org/10.1145/3491102.3517550>

\*The first three authors contributed equally to this research.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

CHI '22, April 29-May 5, 2022, New Orleans, LA, USA

© 2022 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-9157-3/22/04.

<https://doi.org/10.1145/3491102.3517550>

## 1 INTRODUCTION

Given its rapid growth, the nascent field of mixed reality<sup>1</sup> (MR) is seeing an ever-increasing need for user studies and field evaluation, analyzing not only MR-specific metrics such as use of space [4], movement patterns [21, 51], or interaction ergonomics [31], but also more traditional HCI performance metrics such as time and error. For this purpose, researchers gather an abundance of quantitative and qualitative data such as interaction or movement logs (e.g., spatiotemporal data), audio/video recordings, and questionnaire responses in mixed-method evaluation approaches that allow for data triangulation. However, MR tools and techniques are particularly challenging to evaluate given their device heterogeneity, diversity of use, and mobile deployment. Additionally, many metrics are affected by the original environmental context [19, 51] or may be hard to analyze in a 2D context [21, 31]. Thus, Ens et al. [30] recently identified the establishment of a general evaluation framework as one of the current grand challenges in immersive analytics (IA) [62]. Several IA tools have already emerged that aim to streamline this process (e.g., [19, 51, 69]), offering increased immersion and flow [70]. This can be especially beneficial when analyzing spatiotemporal data (e.g., simulations of the actual study scene [19]), analyzing data within their environmental context [19, 51], or viewing 3D visualizations [19, 55, 69] **in situ**. However, “*being ‘in the data’ at times prevents an outside-in view that may be needed to get an overview of the data.*” [19]

In contrast to these novel immersive approaches, non-immersive visual analytics (VA) tools such as Tableau [81] and Spotfire [46] are widely used to analyze data through a more traditional **ex-situ** approach—where the analyst is detached from the original study environment—that can outperform immersive counterparts for overview tasks [54]. The familiar input methods in these non-immersive tools allow for precise interaction (e.g., via mouse pointer) and facilitate the creation of specialized analysis workflows (e.g., via computational notebooks [2]). In addition, the 2D environment is suited towards a range of relevant study analysis tasks, such as analyzing video data, transcribing audio, pre-processing data, or exporting results for use within other applications (e.g., for statistical tests).

We argue that a holistic analysis of MR study data can therefore benefit from both immersive and non-immersive VA tools, as they complement each other well [30]: On the one hand, IA tools excel in an in-situ analysis, allowing researchers to reconstruct the context of the original setting (akin to a crime scene investigation) and facilitating the analysis of inherently 3D data; on the other hand, non-immersive VA tools excel in an ex-situ analysis, providing a holistic overview of the data and inter-compatibility with other tools, and allowing users to define their own specialized analysis pipelines to compare data across multiple participants or conditions. Yet, there is a missing link between both approaches: Researchers have to reconstruct their (mental) workspaces when switching from one tool to the next, thus barring any kind of serendipitous findings that might occur if this transition was seamless. In addition, there has been little research on which tasks are best suited for immersive or non-immersive settings as well as on how to best transition

between these environments [30]. More specifically, which aspects of the analysis process, if any, benefit from immersive representation, and which aspects are instead better served by the use of a non-immersive visual analytics interface? How can we transfer context information and support users when switching from, e.g., an immersive VR device to a non-immersive desktop environment and back?

To investigate these challenges, we introduce **RELIVE**: a *mixed-immersion* visual analytics framework [80] that combines both immersive and non-immersive views to enable the holistic exploration and malleable analysis of MR user studies. **RELIVE** offers an **in-situ VR view** suited for immersing the user in an interactive spatial recording replicating the original study setting (cf. [59]). Similar to prior work (e.g., [19, 51, 69]), users can walk through the scene, create visualizations based on entities and events within the scene, and view the study data within its original environmental context. Unlike prior work, which uses AR to visualize spatiotemporal data (e.g., [19, 69]), we argue for the use of VR, which does not require access to the original study setting (cf. [19, 69]), can simulate studies across the whole virtuality continuum [47, 64], and allows for higher immersion, which, in turn, may provide deeper insights into the study participant’s environmental context. To cover both exploratory and analytical procedures of the analysis process and identify which aspects may be better served by an IA view or a 2D view, we complement our VR view with an **ex-situ desktop view**. Users can use this desktop view as a visual analysis workbook, taking advantage of a toolkit of visual analytics techniques for summarizing, linking, and exploring details of spatiotemporal, event, and nominal data to make comparisons between study sessions. The desktop view also allows for the playback of audio and video media, and offers a 2D window into the current VR view. Both the desktop and VR view are synchronized in real time, facilitating the switch between the different views—thus representing an *asynchronous* hybrid user interface [45]. Although our focus in this work is solely on a single-user system, such cross-platform environments can also open up the design space for asymmetric collaboration [26, 34, 82].

Prior work has already shown the viability of analyzing MR studies both on the desktop [93] and in MR [19, 51, 69, 73]. Closely related to our own work is the Mixed Reality Analytics Toolkit (MRAT) [69], both in terms of data capturing and analysis approach. MRAT allows researchers to define specific tasks and metrics in advance, which can then be analyzed in a tightly-integrated session inspecting tool, using a hybrid user interface to display both ex-situ 2D visualizations and in-situ 3D visualizations. In contrast, our open data logging approach aims to capture all data needed for a replication of the original study—thus enabling a holistic data exploration—and allows for both additional data capturing of non-MR devices as well as the conversion of prior user studies. **RELIVE** enables users to explore and evaluate this data based on customizable analysis component, similar to a computational notebook. Instead of focusing on novel visualizations, our goal is to investigate how a research platform could support both a flexible analysis on the desktop and an immersive analysis in VR at the same time, thereby combining the benefits of both. Consequently, **RELIVE** demonstrates a novel interplay between these different approaches, which lets us examine the unique opportunities and challenges (e.g., transitioning between environments).

<sup>1</sup>The term “mixed reality” includes both virtual reality (VR) and augmented reality (AR) [87].

We evaluated a prototype of ReLIVE in a two-step evaluation process to identify the benefits and challenges of combining immersive and non-immersive views for the analysis of MR user studies. In the first step, a guided design walkthrough was conducted with the authors of this paper to analytically evaluate and validate the concepts. In the second step, we invited 5 MR experts in an empirical user study to see how they use the interplay between immersive and non-immersive VA for the analysis of study data. Here, we were particularly interested in which tasks are best suited for which level of immersion, the interplay between in-situ and ex-situ analysis, and the actual applicability of our holistic analysis concept. Our contributions are thus twofold: (1) We contribute ReLIVE, a mixed-immersion visual analytics tool for the holistic analysis of MR user studies; and (2) design insights and research implications for the implementation of mixed-immersion systems extracted from a two-step evaluation process of ReLIVE.

## 2 RELATED WORK

Here we look into existing *metrics for evaluating mixed reality studies* and review prior work in terms of the necessary *degree of immersion for analysis of mixed reality studies*. Büschel et al. [19] already provide an extensive analysis of available systems for analyzing MR studies and their visualizations; therefore, we focus on the distinction between non-immersive and immersive approaches. To bridge these two fields, we also investigate *transitional interfaces* for the combination and transition between non-immersive and immersive environments.

### 2.1 Metrics for Evaluating Mixed Reality Studies

When evaluating mixed reality studies, researchers employ a variety of different metrics, the choice of which is heavily influenced by the study type and research objectives [30, 69]. Performance metrics such as task completion time, accuracy, or error rate are often used for MR studies and offer a well-understood point of reference [7]. However, established metrics for 2D interaction may not be easily transferred to these immersive environments: For example, applying ergonomics metrics created for interacting with vertical displays [42] to mid-air interaction may result in misleading results, especially when visualized in situ [31]. In addition, most MR studies “*need to cover many more factors than studies of non-immersive surroundings*” [7], such as place illusion and world awareness [85], environmental constraints [19, 31], or novelty bias [30]. Although an extensive analysis of MR study metrics exceeds the scope of this work, recent works [7, 30, 63] point towards a lack of standardization for MR study metrics.

An interview with domain experts conducted by Nebeling et al. confirms that “[*researchers*] mentioned many types of data specific to their projects but relatively few concrete metrics” [69]. Rather, many MR systems are evaluated qualitatively based on interview data, observations, or bespoke visualizations of available study data, such as movement data [21, 66]. To aid in these observations, some studies employed a passive observation client [44, 92, 98], which can provide more insights into the digital environment than the user’s point of view.

Another aspect of evaluating MR studies is capturing and calculating specific metrics: Here, Nebeling et al. [69] automated the calculation of a set of global metrics (e.g., task completion time, distance moved, area coverage of user movement) in a Unity framework; Kloiber et al. [51] integrated a clustering algorithm to automatically detect keyframes in a recording with high spatial activity; and Liliya et al. [59] enable the user to step through notable changes of a selected object. Prior work also supports creating annotations and tags [19], recording of gestures and voice commands [69], or defining tasks for calculating metrics [69]. However, some MR systems may not rely on evaluating user movement, but instead investigate olfactory (e.g., [5]) or taste (e.g., [35, 96]) feedback.

In summary, well-established performance metrics work especially well for simple interactions such as pointing and dragging, but are usually insufficient to completely characterize more complex activities in MR. The possibilities of multi-modal interaction, multi-user scenarios, and multi-device environments therefore demand an analysis environment that not only allows for the calculation of classical measures, but can also offer richer ways to capture, visualize, and analyze these complex activities (e.g., using a 3D environment [31]).

### 2.2 Degree of Immersion for Analysis of Mixed Reality Studies

When analyzing user studies of any kind, researchers usually turn towards (non-immersive) desktop analysis tools, offering a wide range of different well-established software suitable for analysis, the choice of which depends on the data set, the goal of the analysis, and the expertise of the user group [2]. Existing approaches offer a fully-featured graphical user interface (e.g., Tableau [81] or Spotfire [46]) which may require data preprocessing, or provide a powerful development environment (e.g., R [36] or Python) that allows users to both calculate their metrics and output visualizations. These non-immersive approaches are highly configurable, as simple visualizations may be expressed using a common specification (e.g., Vega-Lite [84]) to replicate the visualization in a wide range of analysis tools, or use powerful toolkits (e.g., D3.js [13]) to generate bespoke visualizations. To give more context to these visualizations, computational notebooks (e.g., [2, 38, 72, 74]) are often used, employing the concept of literate computing [65] to narrate the analysis process by combining explanations, code, and resulting visualizations as a visual analysis workbook.

Specific to the analysis of MR user studies, past efforts have also investigated non-immersive visualizations of spatiotemporal data. For visualizing user movement and orientation, such prior work has augmented top-down views [17, 27, 28, 91, 93] and 3D scene views [18, 28] with different visualizations, including trajectory plots [17, 18, 27, 28, 91], heatmaps [17, 27, 93], or field of view frustums [17, 27, 73]. Often (e.g., [17, 18, 27, 28, 61, 93]), visualizations can be controlled by a timeline in combination with playback controls. These visualizations of movement data are also often complemented by the playback of one or more video recordings (e.g., [17, 18, 28, 61, 93]), synchronizing the visualization of user movement with the actual video recordings. Furthermore, several works [17, 18, 61, 93] visualize calculated or manually annotated events (time points and time periods) as part of a timeline.

In contrast to these (non-immersive) desktop-focused approaches, recent work has also introduced various immersive prototypes and toolkits that facilitate the analysis of spatiotemporal data from AR [19, 69, 73] and VR studies [51, 59]. The capability of these prototypes depends on the individual research focus, which affects the choice of visualization: Here, 3D trajectory plots and 3D trails are often used to visualize the position and speed of objects of tracked devices [21, 59], participant’s head and hands [51, 59], or gaze cues [73]. In addition, Nebeling et al. [69] use 3D point plots to visualize events such as the position and direction of users, tracked objects, and physical markers, while a tablet shows an overview of the events in a 2D visualization. Büschel et al. [19] also enrich spatial 3D trajectories with additional videos and 2D visualizations (e.g., heatmaps, scatterplots) that can be placed in the AR environment, while Liliya et al. [59] use these 3D trajectories as an interactive, non-linear time slider.

In summary, both non-immersive and immersive analysis tools represent viable choices for analyzing data from MR user studies: Non-immersive tools offer flexibility and reproducibility (e.g., via computational notebooks) and are well-integrated in a rich ecosystem of established applications, for example allowing users to export their results to their research paper. In contrast, immersive tools can reveal the environmental context [19], increase spatial understanding [55], help in understanding physical measurements [57], and aid in the decision-making [76], but still suffer from novelty factors and discomfort issues [25, 71] such as HMD weight [99], temperature [99], or simulator sickness [49] which can make these immersive approaches unattractive for some users over longer periods of time. We therefore argue that a holistic analysis of MR user studies requires both immersive and non-immersive approaches—allowing users to choose and transition between different levels of immersion based on their current analysis task. Yet, most existing tools only support one or the other; users have to therefore either choose one, or spend significant effort in migrating their current analysis workflow to a different reality.

### 2.3 Transitional Interfaces

The combination of non-immersive and immersive devices is often used in the context of *hybrid user interfaces* [33] to offset the disadvantages of mixed reality head-mounted displays (HMDs), for example by offering text input on a keyboard. In the context of IA, recent work has demonstrated the use of a wide variety of devices with AR HMDs, such as tablets [44, 56, 83, 90], interactive surfaces [22, 77], smartphones [20, 53, 58, 60, 92, 100], or tangible user interfaces [88]. Often, the visualizations act as an augmentation of the surface [79] (e.g., extending tablet [56], desktop [77], or large display walls [78]). In contrast to these AR systems, the full immersion of VR systems makes the use of hybrid user interfaces more challenging. Here, the interaction device (e.g., tablet [29, 89], touch surface [86], or mobile devices [9]) is either tracked and replicated in VR, or the user’s immersion is reduced with a non-occlusive VR HMD [94, 95].

The use of a hybrid user interface often implies a *synchronous* usage of devices, e.g., interacting with a tablet *while* wearing a HMD (cf. [44]). A recent taxonomy of the broader field of *cross-device interaction* [16] also considers the *asynchronous* use of different

devices: In the context of VR, the HybridDesk [1, 23] allows users to transition between a 2D desktop interface and immersive 3D VR. Berns et al. [8] recently combined a live code editor on a desktop with a synchronized immersive 3D view for teaching. Similarly, the Unreal Engine SDK [37] complements their desktop editor with a dedicated VR interface, while Klein et al. [50] use multiple synchronized views across different immersive environments. Lastly, Cavallo et al. [24, 25] created an immersive environment covering desktop, cave automatic virtual environment (CAVE), and mixed reality systems. We thus consider the use case of *asynchronous* hybrid user interfaces [45], where heterogeneous devices are used in *sequence*.

Yet, switching between displays has been shown to incur significant overhead [40, 75] leading to high transaction costs [43]—a problem that is likely further exacerbated if users not only have to switch displays, but device form factors (e.g., switching from a desktop system to a VR HMD). To study these transition, the term *transitional interfaces* [10, 39] is used when transitioning between different realities (e.g., real world to virtual reality). In this context, Carvalho et al. [23] provide guidelines for the explicit transition between such environments. Yet, to the best of our knowledge, research on transitional interfaces is sparse, and there is little research on the actual transition between different device types (e.g., from a desktop to a VR HMD).

In summary, there is a growing interest in hybrid user interfaces that often complements mixed reality interfaces with more traditional input devices, where both devices are used in simultaneously. However, wearing a HMD over a longer period of time may cause significant discomfort [25, 71]. We therefore use our device combination asynchronously [45] (i.e., as a migratory interface [16]), which requires a transition between the VR and desktop environment. Although a CAVE environment [24, 32] could alleviate this transition, our focus is on VR HMDs, which are more affordable for end-users, provide higher immersion, and are thus more suited for use in personal workstations. Given the sparse amount of research in transitional interfaces, our work explores the underrepresented use case of transitioning between a VR HMD and a desktop environment.

## 3 RELIVE

We propose RELIVE: a visual analytics framework [80] that combines immersive with non-immersive views to enable the holistic exploration and malleable analysis of MR user studies. Based on current trends in immersive and visual analytics study analysis (see Section 2), RELIVE provides an immersive VR view for in-situ analysis with a synchronized, non-immersive desktop view for ex-situ analysis (see Figure 1): The VR view allows users to relive an interactive spatial recording replicating the original study (cf. [19, 59, 69, 73]), while the complementary desktop view facilitates the malleable analysis of aggregated study data (cf. [2, 11, 72, 74]). In this section, we describe RELIVE, which was informed by prior work and extended based on findings from a two-step evaluation process (see Section 4).

To ground RELIVE in authentic data and realistic evaluation scenarios, we refer to five reference studies (RS; four of these published [4, 44, 68, 97] and one of them as-of-yet unpublished) that

served as testbeds for the conceptual design, prototypical development, and evaluations of RELIVE:

- RS1 In this study [4], domain experts were tasked with individually exploring a domain-specific data set, configure their visualizations in a 3D space, and present their findings, which allowed to study e.g., the use of space or the resulting complexity of visualizations, focusing on quantitative measurements in a within-subjects design.
- RS2 In this study [97], two co-located participants were asked to collaboratively position and rotate virtual furniture to pre-defined target positions using handheld AR tablets. This allowed to study the influence of different interaction techniques on task performance (e.g., accuracy or task completion time), focusing on quantitative measurements in a within-subjects design.
- RS3 In this study [68], two participants (either co-located or remote) were instructed to remember the position of task objects using different display configurations: The study compared a handheld AR condition (tablets superimposing digital content on their real-world camera) with handheld VR condition (tablets showing a completely digital world), both controllable through egocentric navigation. Both display conditions featured visual landmarks in the form of digital furniture. The display configuration represented the within-subjects factor and the spatial dispersion (e.g., co-located or remote) represented the between-groups factor. The authors studied the influence of the display configuration on measures such as the perceived social presence (i.e., for remote collaboration) and task completion times, while combining other qualitative and quantitative measurements in a mixed design (between-groups and within-subjects factors).
- RS4 In this study [44], participants were tasked to individually explore and analyze a visualization of 3D parallel coordinates using the various input modalities provided by the handheld tablet (e.g., touch or tablet orientation) and AR HMD (e.g., head-gaze or voice input) to investigate the applicability of this multimodal interaction approach, focusing on qualitative measurements in a within-subjects design.
- RS5 In this study (as-of-yet unpublished), two remote participants were asked to jointly decide on a travel destination and plan a trip. In this realistic negotiation task, the influence of the location and representation of the remote peer allowed to study user experience and subjective perception of presence, focusing on qualitative measurements in a within-subjects design.

The selection of these reference user studies was guided by the aim to cover a wide range of MR user study situations to ensure the general applicability of RELIVE. It ranges from single-user studies using an AR/VR HMD (RS1, RS4), over co-located collaborative user studies using handheld AR/VR (RS2, RS3), to remote collaborative user studies using AR HMDs (RS5), each focusing on various types of measurements and applying different study designs.

To facilitate replication in RELIVE, we created a data specification that unifies the recorded data from these different studies. Existing data was converted to this specification using bespoke preprocessing pipelines and imported to RELIVE. To support and validate our

concepts in real world scenarios, we also created a data logging toolkit to easily recreate and analyze future studies in RELIVE.

The following sections<sup>2</sup> describe the *data specification*, the *component templates and instances* that allow for a malleable analysis, the *non-immersive desktop view* and the *immersive VR view*, and the *transition between in-situ and ex-situ analysis*. Lastly, we describe our *data logging toolkit* to support RELIVE for future user studies.

Throughout this section, we will refer to the following example scenario to showcase the possible benefits of using RELIVE, which is further illustrated in the figures in this section<sup>3</sup>. The scenario is based on findings that were unveiled when analyzing data from RS3 with RELIVE during development.

**Scenario** *The HCI researcher Sarah finished the mixed reality mixed-method experiment RS3 with 32 participants and two conditions. Before investigating the gathered data with statistical tools such as IBM SPSS, she starts exploring the qualitative and quantitative data with RELIVE to get a first impression and to see how the display configurations were adopted by the participants. For this, she opens up a new analysis notebook in the desktop view and selects RS3, which then loads all data based on the pre-defined dependent variables on the desktop and on the complementary VR view.*

### 3.1 Data Specification

Our data specification was designed to holistically reflect the data from a user study, allowing RELIVE to reconstruct the study as accurately as possible. The specification differentiates between three data types: *sessions*, *entities*, and *events*. Additional data, such as audio or video, may be included as attachment, but is stored as an ordinary file to facilitate external access.

**Sessions.** A session represents a self-contained subset of the study data pertaining to one study session containing multiple entities and events. Depending on the study design, a session can represent a trial, an experimental condition, or an entire study session.

**Entities.** Each entity describes an actor, input device, or object with a visual appearance, which should be replicated. Entities may change over time (e.g., their position or rotation) and are usually represented by a 3D model of the original object. Entities may also contain media (e.g., screen capture of a tablet) that can be used for a more holistic replication (e.g., displaying a tablet's screen recording directly on its virtual replica).

**Events.** Events represent ephemeral actions that typically do not have a physical representation (e.g., task start, touch event), but occur at a specific point or interval in time. Similar to entities, events may occur at a specific position within a room (cf. [69]) and may contain media data, such as screenshots.

**Scenario** *In RS3, each pair of participants is divided into two sessions, one for each condition: handheld AR and handheld VR. In this study, entities are represented by the participant's tablets, the virtual furniture acting as landmarks, and the cubes of the memory task. New events are created at the position of a participant's tablet whenever a cube is unveiled by the participant.*

<sup>2</sup>Please also refer to our supplemental video.

<sup>3</sup>An additional scenario showcasing the investigation of outliers in RS2 is illustrated in the supplemental video.



## 3.2 Component Templates and Instances

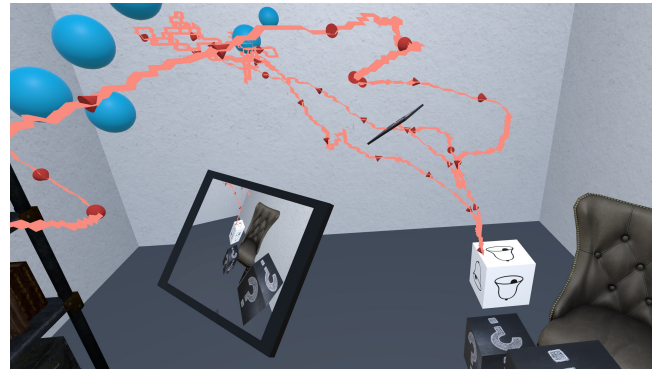
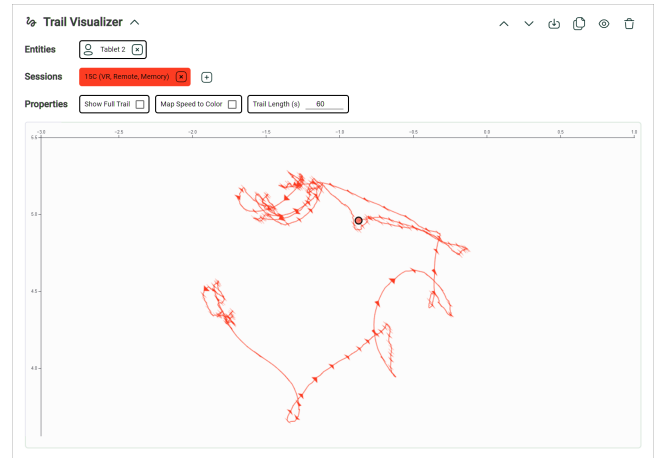
To establish a malleable analysis workflow that works in both VR and on a desktop, we adapt the concept of *components* from computational notebooks (cf. [2, 11, 74]). Instead of having to program each component separately, our concept differentiates between *component templates* that are programmable and *component instances* which execute the template's code with data provided by the analyst. RELIVE thus combines the benefits of a desktop environment for programming, while still offering a flexible analysis workflow in VR.

*Component templates* can be programmed by analysts to calculate and visualize metrics, or to add custom behavior to the interactive 3D study replication—thus allowing for the creation of custom analysis tools. By using the data specification, templates are generic and only need to specify how many entities or events are required for successful code execution. In addition, analysts can specify optional parameters to further customize the calculation. Component templates facilitate the distribution of metrics across different analysis workflows, as these templates can be easily packaged (cf. [11]) and shared with others. For example, consider a component template that visualizes the user's movement as a trail (cf. [19, 59], see Figure 2): Here, the analysts can specify that the component requires one entity, add code to convert the entity's movement data to a visualization, and provide additional parameters (e.g., length of visible trail in seconds) that can customize the resulting visualization. In future work, we aim to make these templates more easily programmable by the user, for example by offering an editor similar to existing computational notebook, allowing components to output a Vega-Lite specification [84] depending on their specified input parameters.

In contrast, *component instances* (see Figure 2) can be created from existing templates. Instances can be customized by the analyst by adding study data or adjusting parameters. Once configured, components automatically execute the code defined in their corresponding template. Component instances either add custom behavior to the 3D study replication, or visualize the calculated metrics. Visualizations can differ between the non-immersive and immersive environment akin to multiple coordinated views, making the best use of each environment. For example, when instantiating the trail component template, analysts have to specify a entity for this instance and can adjust the defined properties, such as trail length. Once one or more sessions (i.e., data subsets) are added to the instance, the template's code is executed, and a visualization appears in the component (see Figure 2 (top)). In addition, a corresponding 3D visualization appears in the 3D study replication if the affected entity is visible (see Figure 2 (bottom)).

## 3.3 Non-Immersive Desktop View

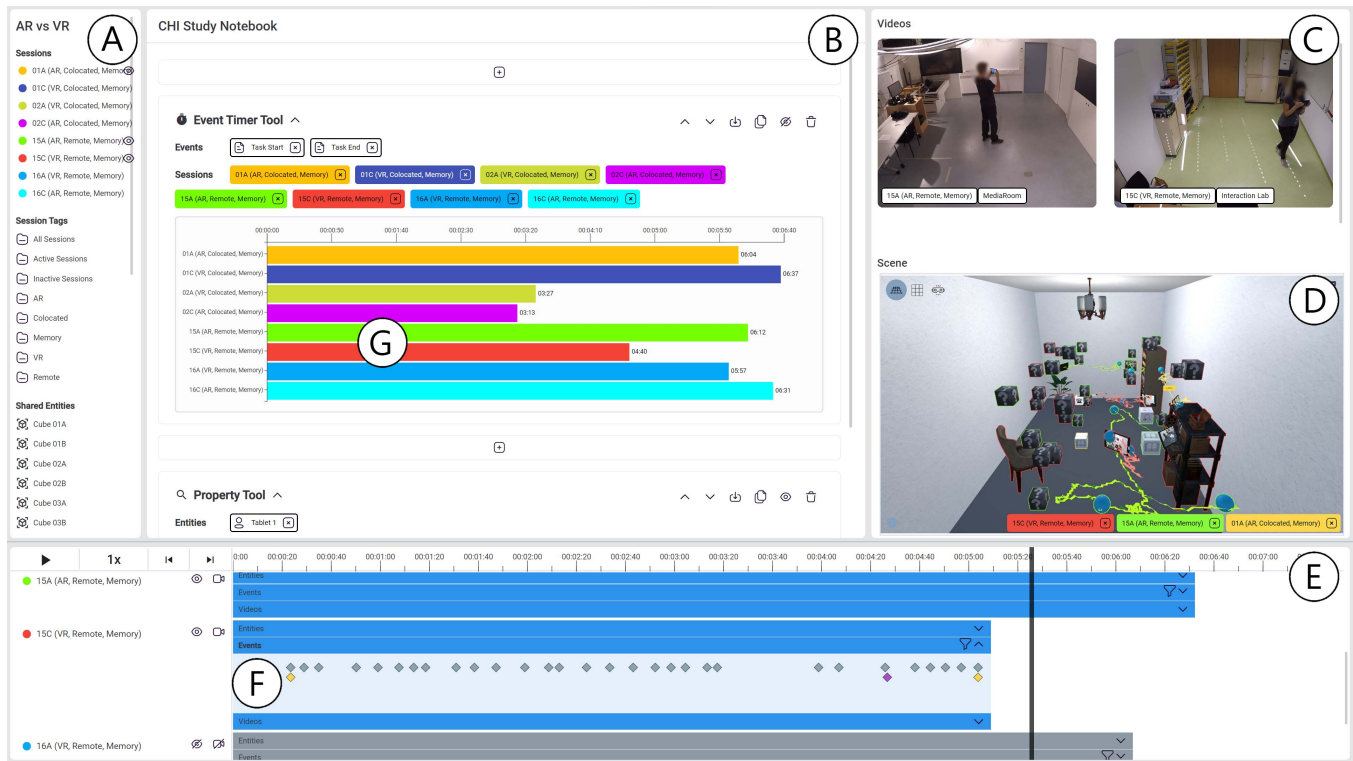
The non-immersive desktop interface of RELIVE is designed to provide analysts with a holistic ex-situ overview of the study's available data and is suited for an analytical evaluation of aggregated study data using a components-based approach (see Figure 3). Our concept takes inspiration from computational notebooks (e.g., [2, 11, 72, 74]), which use an interactive programming environment to display results such as visualizations inline (cf. literate computing [65]). Here, analysts can create component templates



**Figure 2: Component instance.** Used on the desktop (top) and in the 3D study replication (bottom), created from a *trail visualizer* component template with data from RS 3. On the desktop (top), the component requires one entity, as well as at least one data set (session) to create a top-down 2D visualization of the tablet's movement. The component exposes three properties to further customize the current visualization. Additional buttons in the top right corner allow for general control (e.g., delete component, export data). In VR (bottom), the same component instance also adds a 3D visualization in the 3D study replication, which can reveal additional information such as user behavior and interactions with entities.

and instances, allowing analysts to, for example, narrate their analysis process, calculate and visualize metrics (see Figure 3 (B)), or easily export data from these components for use within other applications.

This notebook approach is combined with a video playback interface akin to state-of-the-art video editors to better support the complementary in-situ analysis. Analysts can display available 2D media data (see Figure 3 (C)), or see into a MR study session via a reconstruction of the study session (see Figure 3 (D)). The current playback status is synchronized across all available media data and can be controlled with a timeline.



**Figure 3: ReLive desktop view.** This view combines a computational notebook approach with a video editor interface, specializing in an ex-situ analysis of aggregated study data. Access to additional media data such as videos and 3D study replication allows for a holistic insight into the user study.

The desktop interface is divided into five panels to merge the computational notebook approach with a video editor (see Figure 3): a *data panel* (A), a *component panel* (B), a *timeline panel* (E), a *video panel* (C), and a *3D scene panel* (D).

*Data Panel.* The data panel shows an overview of all available sessions, entities, and events within a study (see Figure 3 (A)), allowing analysts to assign this data to other panels via drag and drop (e.g., to configure component instances). To facilitate comparisons across study sessions, data with similar attributes (e.g., identical name) is automatically grouped together: *Session tags* act as a smart folder of sessions with similar attributes (e.g., the same experimental condition), allowing analysts to drag and drop multiple sessions at once; *shared entities* and *shared events* group together similar entities or events, respectively, across sessions and thus allow for an easy comparison of data across different sessions.

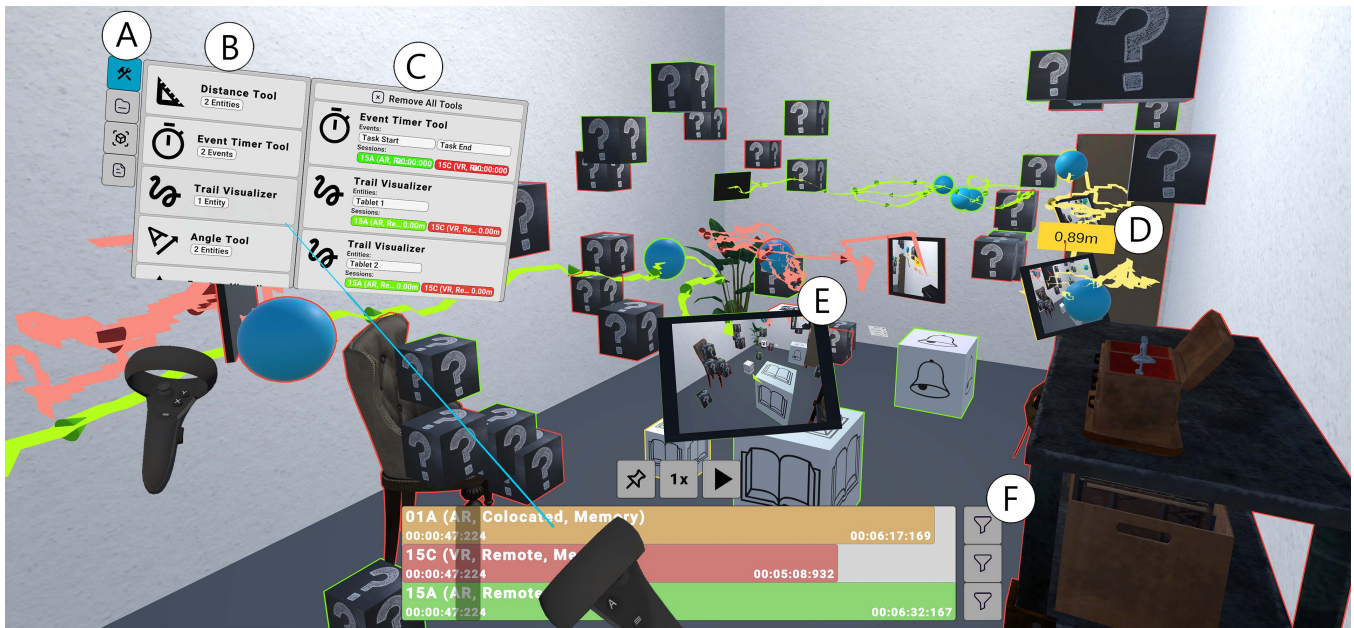
*Component Panel.* A component panel allows analysts to visualize study-specific metrics as visualizations via customizable components (see Figure 3 (B)). Here, analysts can create new component instances from existing component templates. Created component instances are organized linearly like a computational notebook, allowing analysts to narrate their analysis process. In addition, the visualizations can provide contextual information about the current playback time (e.g., as interactive playhead in a time series

line chart, see Figure 3 (G)), linking the ex-situ analysis with the available videos and 3D study replication.

*Timeline Panel.* Similar to a video editor, a timeline panel at the bottom controls the current playback time, speed, and status (see Figure 3 (E)). To facilitate the comparison across study sessions, the current playback time is synchronized across all sessions—each session is therefore akin to a single video track in a video editor. In addition, each session timeline can be expanded to reveal an overview over the lifetime of all entities and a timeline visualization of events (see Figure 3 (F)), which can be filtered (cf. [69]). Static data (e.g., screenshots contained within events) can be inspected by hovering over the data, while videos can be displayed in the *video panel*.

*Video Panel.* The video panel provides additional insights into the real-world context of the study (see Figure 3 (C)). Here, videos enabled in the timeline panel are displayed as separate videos and synchronized with the playback time displayed in the timeline panel, akin to a video editor.

*3D Scene Panel.* The 3D scene panel provides additional insights into the virtual context of the study (see Figure 3 (D)). Based on the available data, the original study is reconstructed as an interactive 3D scene suited for in-situ analysis (cf. [73]). This 3D scene panel provides controls similar to common 3D editing programs such as free camera movement and isometric (e.g., top-down) camera



**Figure 4: RE LIVE VR view.** This view immerses analysts in an interactive reconstruction of the original study (here: **RS3**). A user interface is affixed to the analyst’s left controller and a timeline at the bottom allows analysts to control the simulation. In addition, metrics and utilities from component instances are visualized in situ. Once multiple sessions are visible simultaneously, a colored outline helps in differentiating entities and events from different sessions.

perspectives, but can also provide additional insights, for example by replaying the scene from a user’s point of view. In addition, analysts can drag and drop data to instantly locate the corresponding 3D object. Lastly, visualizations from corresponding component instances are displayed in the 3D scene window, allowing for a quick glance into the VR view.

**Scenario** As Sarah is interested in task completion times, she creates a new component instance of an already available “Event Timer” component template in the component panel. To get an initial overview, Sarah adds the data from all participants to this component instance by dragging and dropping the “All Sessions” tag from the data panel to the newly created component instance. In addition, Sarah specifies a start and end point by dragging in the representative events from one of the sessions in the timeline panel into the component instance. Once the data has been defined, a 2D barchart automatically appears—revealing that one participant pair was considerably slower in their first trial (see Figure 3 (G)). To investigate this outlier, Sarah looks at the video data by dragging in the outlier sessions into the the video panel to show the real world video, and then dragging the same session into the 3D scene panel to show a digital reconstruction, controlling their playback with the timeline panel. However, viewing the study replay on a desktop makes it difficult to understand how the participants moved through the digital environment, as Sarah has to constantly adjust the camera to understand the spatial relations between objects.

### 3.4 Immersive Virtual Reality View

In contrast to the desktop view which focuses on ex-situ visual analytics, the VR view focuses on in-situ immersive analytics (see Figure 4). The VR view is inspired by prior work (e.g., [19, 59, 69, 73]), which enriches the in-situ analysis with environmental context. To that end, analysts can relive an interactive replication of the original study, which is enabled by our data specification: Session backgrounds (e.g., 3D model of the room) can be added to the data set and displayed in VR; entities are reproduced using, for example, their 3D model and move around the scene based on their captured movement data; and events are visualized in situ as colored spheres (cf. [69]).

For better cohesion with the desktop view, the user interface is structured similarly to its desktop counterpart (cf. [23]), yet also geared towards an in-situ analysis. Analysts can interact by clicking and pointing on user interface elements or objects within the scene, mirroring the features of the desktop where possible. In the future, we aim to provide an ex-situ view in VR similar to the desktop view, allowing for a more detailed control over the analysis. The interface of the VR view provides analysts with access to the *study data*, the created *components*, a *timeline*, and displays available *media* data within the scene.

**Study Data.** Analysts can reveal additional information about entities and events as a floating tooltip by pointing at a corresponding 3D object within the scene. Analysts can also use a user interface anchored to their left controller to get a basic overview of all entities and events within a scene (see Figure 4 (A)). This menu also allows analysts to focus on specific entities (e.g., by hiding all irrelevant



entities). Lastly, analysts can browse through all available sessions and load the data into the scene. When multiple sessions are active at the same time, a colored outline is added to each object in the scene, allowing analysts to map entities to their corresponding data set.

**Components.** To allow for an analysis of the data, analysts can create new component instances through a 2D menu anchored to the left controller (see Figure 4 (B)). However, as the input modalities for VR are unsuited for text-based programming, creation of new component templates is restricted to the desktop view. Thus, component instances can be created by selecting a template, then clicking on the corresponding entities or events within the 3D scene.

Existing component instances are shown in the user interface (see Figure 4 (C)), but can also be placed anywhere in the 3D space. The resulting visualizations may appear in situ (movement trail, see Figure 4 (D)), as a freely placeable visualization, or as a single value depicting the current value (see Figure 4 (D)).

**Timeline.** Similar to the desktop view, a timeline allows analysts to control the playback state, speed, and time (see Figure 4 (F)). In addition, analysts can wind forward or backwards by using the VR controller’s joysticks. The timeline can also provide an overview of all events, which can be filtered based on event properties.

**Media.** To simulate a device’s screen during the actual study, each device shows their current view on the virtual scene on their screen (see Figure 4 (E)). We aim to visualize more recorded media data in future work, for example by replaying screen recordings from a tablet on the corresponding replica’s screen, or adding another video playback panel to the *components* menu.

**Scenario** To better investigate the outlier, Sarah immerses herself in VR by putting on a VR HMD—putting her into the same environment as visible in the desktop’s 3D scene panel. For comparison, she loads in another participant pair by opening the study data interface on their left controller and activating another session. Thus, another set of tablets appears, which are now highlighted with a yellow outline to distinguish them from the tablets of the other session (outlined in red and green). To better visualize the participant’s movement, Sarah selects a “Trail Visualizer” component instance from a predefined component template and attaches this to each tablet by pointing and clicking on each tablet with her right controller. As a result, each tablet now shows a trail of its movement over the last few seconds (see Figure 4). Using the timeline interface, Sarah quickly scrubs through the session, allowing them to relive the study. By observing the tablet’s position and screen, Sarah notices that one participant’s tablet was occluded by the digital furniture placed in the room, which could be a potential cause for the longer task completion time. She also notices that the participants communicated by holding their tablets into the cubes they wanted to select (see Figure 2 (bottom)). For both cases, Sarah selects the “Camera” from the components panel to take a photo, thus saving these incidents within RELIVE for further analysis.

### 3.5 Transitioning Between In-Situ and Ex-Situ Analysis

Both the non-immersive desktop view and the immersive VR view were designed to suit different analysis workflows. Depending on

the task and context, analysts may prefer to work in situ (e.g., in the VR view) or ex situ (e.g., on a desktop). RELIVE therefore aims to facilitate the transition between in-situ and ex-situ analysis and follows the guidelines proposed by Carvalho et al. [23].

Most importantly, application state (e.g., components, playback time) is synchronized across both views in real time, thus “[making] users aware of the system state” [23]: Component instances created in the desktop view are instantly visible in the VR view and vice versa. This is further supported as analysts can explore the 3D study replication on the desktop view (see Figure 3 (D)). In addition, the desktop view allows analysts to drag data (e.g., entities, events) from the ex-situ overview to the 3D scene replication, causing the camera to focus on the related object in the 3D scene. Analysts can therefore instantly explore the environmental context from within the ex-situ overview.

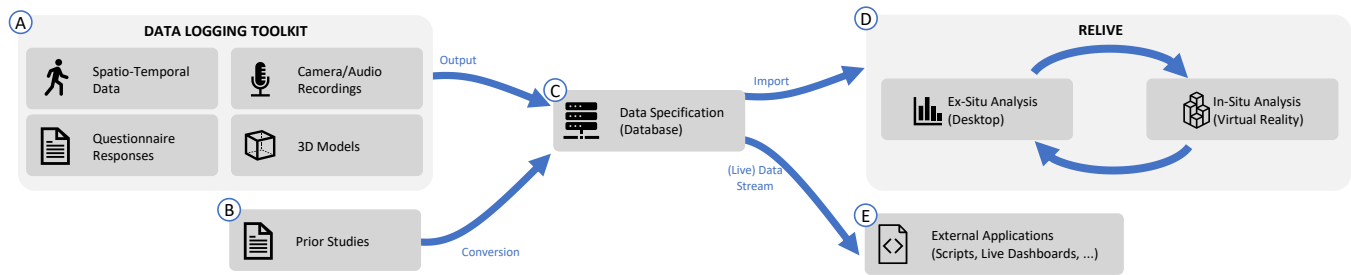
In future work, we also want to implement and examine *cross-reality linking and brushing*: Here, users could mark outliers in the component instance of the desktop view, which is then instantly highlighted in the VR view and vice versa—akin to linking and brushing [48] *across realities*.

**Scenario** While investigating the outlier in VR, Sarah also notices that the “cube selection” events of one pair of participants are clustered within one corner of the room—while the events of other pairs are spread throughout the entire room. To quantify this, Sarah attaches a “Property” component instance to the tablets of each session and configures the component instance in the component menu to show the “total distance moved”—thus displaying the current distance above each tablet. By scrubbing to the end of the session, Sarah can now compare their values between the currently loaded sessions, revealing a substantial difference between the different sessions. To better generalize this across all 32 participants, Sarah switches back to the desktop view. Here, Sarah can seamlessly resume her workflow with the previously created “Property” component instance, which shows a line chart of the “distance moved” metric. By dragging and dropping the “All Sessions” tag to this component instance, Sarah can easily generalize her findings for all participants and continue their analysis, for example by correlating the movement with task completion time or exporting her findings and visualizations to her research paper.

### 3.6 Data Logging Toolkit

To utilize RELIVE in a real world scenario, we developed a holistic logging and malleable evaluation environment, supporting researchers throughout all stages of a study (see Figure 5). At the center of this logging framework is a data specification that aims to capture the abundance of qualitative and quantitative data collected during a study (e.g., spatiotemporal data, interaction events, video/audio recordings, 3D models), allowing researchers to relive the original study as accurately as possible. With the increased availability of 3D scanning hardware (e.g., LIDAR sensors in Apple iPad), scans of the real world environment can be added to the resulting data set [73]. In comparison to prior work that offers similar workflows (e.g., [19, 69]), our concept follows two main design principles:

- (1) **Openness:** Our concept was designed to run on all platforms (e.g., Unity, web) and integrates well with external applications by relying on established standards (e.g., JSON



**Figure 5: Logging and evaluation overview.** This environment supports researchers throughout the all stages of a study: (A) An open data logging toolkit can facilitate data capturing, including spatiotemporal data (e.g., movement data from AR HMD), video data (e.g., imported or automatically captured from network cameras), questionnaire responses, and 3D models. (B) Data from prior studies can be converted to the data specification for use within RELIVE. (C) Our data specification contains holistic data about a study, and can be easily accessed by external applications. (D) RELIVE uses the data specification to enable a holistic study data analysis and reconstruct the original study setting. (E) Data from an ongoing study can be streamed to (and parsed by) external applications, enabling an analysis using bespoke scripts or live dashboards.

for interaction data, standardized file formats where possible). Our aim is to encourage open science by making it easy to share and reuse data sets.

- (2) **Extensibility:** To offer support beyond the status quo and accommodate a wide range of MR studies, the toolkit is designed to be easily extensible without requiring changes to the underlying code.

## 4 EVALUATION PROCESS

We evaluated RELIVE in a two-step evaluation process: First, a guided design walkthrough (Section 5) allowed us to analytically investigate RELIVE in a formative evaluation; second, an expert user study (Section 6) provides deeper insights into the real world applicability of RELIVE in an empirical evaluation. The overall goal of this evaluation process was to better understand the interplay between an in-situ and ex-situ analysis as exemplified by the VR and desktop view of RELIVE, respectively, and how this combination can support the analysis of MR user study data. As part of this goal, we defined three *research objectives* to narrow our analysis focus. In line with these objectives, we created an *evaluation prototype* that implements the core features of RELIVE.

### 4.1 Research Objectives

To investigate the interplay between an in-situ and ex-situ analysis, we focus on three research objectives:

- RO1 Task Allocation:** Which analysis tasks benefit from immersive analytics, which tasks are better suited for non-immersive visual analytics?
- RO2 Interplay Between Ex-Situ and In-Situ Analysis:** How does in-situ and ex-situ analysis complement each other?
- RO3 Applicability:** Can RELIVE match the diverse requirements of MR researchers and their use cases?

### 4.2 Evaluation Prototype

To address our research objectives, we created an evaluation prototype of RELIVE (see Figures 1–4), which was iteratively improved. First, we focused on implementing core concepts for the guided








design walkthrough. This first version already supported both a limited desktop view and a limited VR view, which were synchronized: The desktop view offered predefined component templates (see Table 1) that produced 2D visualizations, a timeline panel, and a 3D scene view; the VR view also supported the same predefined component templates that produced in-situ visualizations and a timeline. Based on the feedback from the guided design walkthrough, we refined our evaluation prototype to support an authentic evaluation of our reference studies. This study prototype includes all features described in Section 3, with the exception of user-programmable component templates, which were intentionally replaced with predefined *tools*. This reduced the complexity of our study and allowed us to focus on uncovering initial challenges relating to the task allocation **RO1** and interplay **RO2**, especially as potential participants would not necessarily be familiar with the used reference studies.

We decided on seven predefined tools listed in Table 1, based on the evaluation requirements of our reference studies **RS1–RS5** and our analysis of related work (Section 2). The tools are aimed to address a wide range of analysis tasks and allow for an easily understandable, yet authentic analysis of our five reference studies. Consequently, we intentionally decided against implementing other ideas for very specific and single use case tailored components that we discussed during our design walkthrough. To align with our research objectives **RO1** and **RO2**, each tool either provides a visualization on both the desktop and in VR, or utility within the 3D scene (Frustum Visualizer, Camera).

### 4.3 Technical Implementation

For the RELIVE evaluation prototype, we use a client/server structure with two separate clients (immersive VR and non-immersive desktop). Our server uses a multithreaded Node.js v12 runtime and is backed by MongoDB. The server is responsible for serving media files as well as preprocessing (e.g., data compression) and calculating metrics to ensure that the clients remain responsive. For this, we employ different transport protocols, such as HTTP for fetching static data, TCP/WebSockets for real-time data synchronization, and WebRTC for video transmission. For the desktop interface, we utilize web technologies for rapid prototyping (e.g., Angular), while

**Table 1: Overview of predefined tools. We defined seven component templates as predefined tools for our final prototype. Tools were derived based on the evaluation requirements of our reference studies and the discussed related works. The chosen tools showcase either visualizations available in both desktop and VR view, or utilities for analyzing the study in VR. The *event timer tool* was not part of the prototype for the guided walkthrough.**

Tool	Visualizes	Desktop view	VR view	Use Cases / References
 Distance Tool	Distance between two entities	Line chart (distance vs. time)	Virtual measuring tape connecting the entities with label showing the current distance	<ul style="list-style-type: none"> <li>Position error during docking task, as in <a href="#">RS2</a> and [7]</li> <li>Finding F-Formations, as in [17]</li> </ul>
 Angle Tool	Angle between two entities	Line chart (angle vs. time)	Angle visualization connecting the entities, projected to the floor	<ul style="list-style-type: none"> <li>Rotation error during docking task, as in <a href="#">RS2</a></li> <li>Finding F-Formations, as in [17]</li> </ul>
 Trail Visualizer	Entity movement over time	Top-down view on 2D trail (xz graph)	3D trail visualization in the VR scene	<ul style="list-style-type: none"> <li>Analyzing movement behavior / spatial activity, as in <a href="#">RS1</a> <a href="#">RS4</a> <a href="#">RS5</a> and [17, 18, 27, 28, 90]</li> <li>Detecting tracking issues</li> </ul>
 Event Timer Tool	Time between two events	Bar chart	Line connecting the events with label showing the time between them	<ul style="list-style-type: none"> <li>Task completion time as in <a href="#">RS1</a> <a href="#">RS2</a> <a href="#">RS3</a></li> <li>Action duration as in <a href="#">RS4</a></li> </ul>
 Property Tool	An entity's inherent properties and derived metrics	Line chart (property value vs. time)	Label with the properties' values hovering over the entities	<ul style="list-style-type: none"> <li>Displaying raw data and metrics during analysis, e.g., distance moved per minute, as in <a href="#">RS1</a> and [69]</li> </ul>
 Frustum Visualizer	An entity's field of view within the 3D scene	Configuration only	Lines visualizing the frustum directly at the entity in the VR scene	<ul style="list-style-type: none"> <li>Visualize what participant has seen, as in <a href="#">RS5</a> and [17, 27, 73]</li> <li>Attention grouping, as in [17]</li> </ul>
 Camera	Screenshots of the 3D study replication	Gallery of taken screenshots and possibility to take screenshots from scene view	Virtual camera for taking screenshots in the VR scene	<ul style="list-style-type: none"> <li>Documenting insights (e.g., outliers)</li> </ul>

the immersive VR client was written in Unity. The 3D scene in the desktop interface is streamed as a native HTML video from a rendertexture within Unity. The ReLIVE evaluation prototype, data specification, logging toolkit, and sample data is available as open-source project on GitHub<sup>4</sup>.

## 5 DESIGN WALKTHROUGH

We analytically evaluated the initial prototype of ReLIVE to validate our concepts, develop ideas for additional features that are essential for the analysis of MR user studies, and verify the general applicability of ReLIVE. We conducted design walkthroughs [41] using each of our five reference user studies ([RS1](#)–[RS5](#)) as testbeds. Each session lasted about 1.5 hours and 4–7 authors participated in the design walkthroughs. We prepared both request and evaluation sheets for task-related features for systematic feedback during the sessions, and used screen sharing in a video conferencing tool to share the same point of view, while following all ethical and sanitary guidelines provided by our universities.

*Roles.* During each session, attendees were given specific roles: A *moderator* that moderated and guided the design walkthrough and kept track of time; a *note-taker* responsible for filling out requests

and evaluation sheets for task-related features; a *presenter* who was involved in the reference user study (i.e., one of the authors) and could provide insights into the prior study analysis; and an *analyst* who was not involved in the reference user study and tried to replicate the study analysis within the ReLIVE prototype—mimicking a potential user. Other participants of the design walkthroughs without a dedicated role participated by discussing ideas and noting down possible issues by observing the analyst. We switched roles for each reference user study to avoid e.g., that a person who was involved in the prior study analysis takes on the role of the analyst as this might have influenced the workflow.

*Procedure.* First, the presenter briefly introduced their reference user study by presenting the study's goal, research questions, and other user study related aspects (e.g., procedure, data gathering methods, apparatus). This introduction was concluded with a list of research objectives and metrics that were investigated during prior analysis. This list was then briefly discussed and extended with additional analysis ideas of all participants. Next, the analyst started the ReLIVE application and progressed through each of the research objectives. All participants discussed possible solutions to each research objective and evaluated the use of existing features, which were noted down by the notetaker.

<sup>4</sup><https://github.com/hcigroupkonstanz/ReLIVE>

**Results.** We collected a total of 59 requests and 16 evaluation sheets for task-related features from all walkthroughs to prepare the prototype for an expert user evaluation. We first estimated the priority and implementation effort for each feature request and tagged it using the following tags: (1) Web, (2) VR, (3) Bridge, (4) Component, and (5) Visualization. We then ranked each feature request based on its generalizability, combined possible duplicates, and filtered out requests that did not fit within the scope of the following expert user study. Remaining features were further organized based on common topics, using a semantic clustering approach. Lastly, we discussed and sketched out possible ideas, before implementing 18 requests that were in line with our research objectives. This extended prototype was then used for our expert user study. The outcome of our walkthrough is merged with the insights from our expert user study and are discussed in Section 7.

Aside from revealing opportunities to improve usability, our requests were mainly concerned with *new visualizations and tools for study analysis* and *interaction techniques to facilitate the transition between environments*. In terms of *new visualizations and tools for study analysis*, many ideas were linked to visualizing the available data types: For example, transcribed audio data could be presented as a wordcloud in 2D for overview, while showing the exact position where the words were spoken in 3D. Similarly, a waveform visualization could be combined with a 3D trajectory trail in the 3D scene replication, highlighting areas where participants talked.

Concerning *interaction techniques to facilitate the transition between environments*, we gathered requests to increase the cohesion between the in-situ and ex-situ analysis: For example, users can drag and drop entities directly into the desktop’s 3D scene reconstruction to zoom in on the relevant object; show the point of view of the VR user or different entities within the 3D scene (e.g., showing the exact point of view of an AR tablet); or change the position of the VR view directly in the desktop view. To further increase cohesion, RELIVE could support linking and brushing on both the desktop and VR (i.e., *across realities*): Here, areas of interest can be marked in an ex-situ 2D visualization, which would automatically highlight relevant areas within the corresponding in-situ visualization. However, the implementation and evaluation of such a *cross-reality linking and brushing* can be different depending on the visualizations and therefore exceeds the scope of this work.

## 6 EXPERT USER STUDY

To empirically evaluate RELIVE, we conducted an expert user study, focusing on participants that had prior experience in conducting and analyzing MR studies. Our goal was to evaluate RELIVE guided by our three research objectives. Our tasks were based on real data from two of our five reference user studies. We collected both qualitative and quantitative data to gain insights into participants’ workflows.

### 6.1 Participants

We recruited 5 male MR experts between 25 and 38 years ( $M = 29.80$ ,  $SD = 5.17$ ) from different research labs as participants. We intentionally looked for researchers with prior experience in analyzing or conducting MR studies and who were not affiliated with any of the current works of the authors. Thus, we invited researchers

from a data analysis, an immersive analytics, and a virtual reality lab from the University of Konstanz. We also invited a colleague from the human-computer interaction lab at the University of Konstanz who was not involved with the design or implementation of RELIVE. All participants had degrees in computer science, 4 with master’s degrees and 1 with a doctorate. All of them were working in academia (e.g., as lecturers or research assistants) and had conducted MR user studies and analyzed their results before. All participants had normal or corrected to normal eyesight and did not suffer from color blindness; consequently, they had no problems with text sizes and the color-coding used in the different visualizations. We asked them to rank their experience in conducting AR/VR studies ( $M = 4.20$ ,  $SD = 0.45$ ) and analyzing the results ( $M = 3.60$ ,  $SD = 0.55$ ) on a scale from 1 (very inexperienced) to 5 (very experienced). On the same scale, they also ranked their experience with computational notebooks ( $M = 2.40$ ,  $SD = 0.55$ ) and virtual reality applications ( $M = 4.60$ ,  $SD = 0.55$ ). Four of them already developed a VR application on their own. We also asked participants if they prefer to analyze MR study data on their own ( $n=3$ ) or together with others ( $n=2$ ). All participants interacted with RELIVE for the first time during the expert user study—this means, none of them were involved in the design nor the implementation.

### 6.2 Apparatus

All studies took place in one of our labs in which we allotted a walkable area of approximately  $2 \times 2$  meters where participants could freely move. At one side of this area, we set up two tables (each  $1.40 \times 0.80$ m) with the long sides aligned. At one table, the participant was seated. The experimenter sat down at the opposite table. We equipped the participant’s table with a 27’’ 4k display, a mouse, and a keyboard connected to a desktop PC (simulating a desktop workspace, similar to Figure 1) that they used to work with the non-immersive desktop view of RELIVE and to fill out questionnaires. Additionally, we provided participants with a tethered VR HMD (Oculus Quest 2) and the accompanying controllers to work with the immersive VR view of RELIVE. The connection cables of the HMD were mounted via ceiling trusses to provide participants freedom in their movement while avoiding tripping hazards. On the experimenter’s table, we placed two displays (both 27’’ 4k), a mouse, and a keyboard connected to the participant’s desktop PC. One display was mirrored with the participant’s display and the other display showed the VR scene. This allowed us to observe and support the participant in both views.

### 6.3 Procedure

Participants were welcomed and provided with introductory documents explaining the purpose and procedure of our study. They signed a consent form and filled out a demographic questionnaire. Using a slide show, the experimenter then introduced participants to RELIVE and the data set of **RS 3**. After that, participants started with the guided phase (see tasks below) to familiarize themselves with RELIVE. Then, participants received a short introduction to the data set of **RS 4** before starting with the *free phase* (not guided by the experimenter). At the end of the free phase, participants filled out the System Usability Scale [15] and each session was concluded



with a semi-structured interview, which included a subjective rating of the interplay of both views. Sessions took approximately 1.5 h and participants received compensation for their time. We followed all ethical and sanitary guidelines provided by the local institution at the time of the study.

## 6.4 Tasks

Participants started the tasks sitting in front of the desktop view. They were allowed to stand up and move through the allotted walkable area at any time. For the guided phase, we carefully ensured that all tasks were balanced between the desktop view and VR view. This means that the guided phase was not favoring one of the views. For the free phase, participants were given the opportunity to use both views as they suited them for their analysis workflow.

*Guided phase.* Participants were guided by the experimenter to solve the following six tasks using step-by-step instructions. For each task, we summarize in italics which instructions the participants received. The tasks resembled authentic analysis scenarios based on the actual analysis of [RS3](#). Participants started with the desktop view and over the course of these tasks, they were required to switch between the desktop and VR view 5 times and therefore had 3 phases in the desktop view alternating with 3 phases in VR.

- (1) **Task completion time** (one session). Visualize the task completion time of a given session by measuring the time between the start and end event. *Participants solved this task using the desktop view. They created an event timer tool and added the given session, its start event, and its stop event to compute and visualize the session's task completion time.*
- (2) **Task completion time** (all sessions of a condition). Visualize the task completion time of all sessions of a given condition. *Participants solved this task using the desktop view. They added the given condition's session tag to the event timer tool they created in the previous task.*
- (3) **Accuracy** (one session). Measure how accurately the participants of a given session placed a given entity in comparison to a given target position. *Participants started with the desktop view to select the session and find the entities' position within the scene using the 3D scene panel. They were then instructed to switch to the VR view and create a distance tool by visually connecting the two entities. This also involved winding through time and teleporting through the 3D scene.*
- (4) **Accuracy** (all sessions of a condition). Measure how accurately all participants of a given condition placed a given entity in comparison to a given target position. *Participants switched back to the desktop view for this task. Here, they added the session tag for the given condition to the distance tool they created in the VR view during the previous task.*
- (5) **Events.** Investigate where the participants of a given session were located in the room when a given event happened. *Participants switched back to the VR view, filter the events, and use the VR controller's joystick to wind forward and backward to investigate the events' position and occurrence.*
- (6) **Tracking Issues.** Investigate if a given tablet (entity) in a given session had tracking issues by visualizing the speed (unrealistic high speed is an indicator for tracking issues). *Participants switched back to the desktop view and used the*

*property tool to investigate the time series line chart for the speed of the given tablet. In the chart, they identified segments where the tablet moved with an unrealistic high speed. They then switched to the VR view and created a trail visualizer to investigate the direction of the jumps in context.*

Participants needed between 15 min 8 s and 25 min 36 s ( $M = 20.25$  min,  $SD = 3.53$  min) for the guided phase—excluding the times for switching between desktop and VR, as they were not representative due to hygiene requirements. In total, they spent between 6 min 38 s and 13 min 44 s ( $M = 9.25$  min,  $SD = 2.52$  min) in VR. While E1 and E2 always stood up when in VR, E4 remained seated during all tasks. E3 stood only briefly during the last phase in VR. E5 started the first use of VR seated but then decided to stand up for the remaining time in VR. In total, participants were standing between 0 min and 10 min 6 s ( $M = 6.01$  min,  $SD = 4.16$  min).

*Free phase.* Participants explored and analyzed the data set of [RS4](#) on their own and were free to use either environment on their own accord. As a starting point, the experimenter suggested the following four analysis goals. However, participants were free to follow their own analysis approaches.

- (1) Investigate the distance between HoloLens and the Interaction Tablet over time. Did the distance increase while the user was holding the tablet?
- (2) Which participant moved the most?
- (3) Were there any tracking issues?
- (4) When holding the Interaction Tablet vertically, at which height did the participants roughly hold their Interaction Tablet: At eye-level, or at shoulder level?

The free phase had a soft limit of 10 minutes. However, all participants decided to continue their individual analysis at the end of these 10 minutes and spent between 11 min 20 s and 17 min 14 s ( $M = 13.37$  min,  $SD = 2.19$  min) in the free phase—again excluding the times for switching between the views. All participants decided to start their analysis using the desktop view. E3 switched 3 times between both views, while the others switched only once. Participants spent between 2 min and 10 min 21 s ( $M = 5.70$  min,  $SD = 2.89$  min) in VR. Only E3 remained seated during the whole free phase. All other participants stood up for using the VR view. In total, they were standing between 0 min and 10 min 50 s ( $M = 4.93$  min,  $SD = 386$  min).

## 6.5 Data Collection and Analysis

We used two ceiling-mounted cameras with opposing views on the scene to capture video data. We recorded the content of the participant's display and HMD. For audio recordings (e.g., interviews), we placed a dedicated microphone in the center of the tables. Further, we transcribed all interviews and analyzed them following an inductive thematic analysis approach [14]: After familiarization with the data, one author identified data extracts that are relevant to our research objectives and generated descriptive codes to label them. These descriptive codes were then counterchecked and validated by two other authors. We then thematically clustered the codes to identify potential themes. In the further process, we iteratively revised and refined the themes to ensure that they were in line with the dataset and our research objectives.

## 6.6 Findings

Here we present the findings of the expert user study. We use the five themes that resulted from the thematic analysis of the data collected in our semi-structured interviews as structure to report our findings—this means that we present the scores of the subjective rating and system usability scale thematically aligned.

**Theme #1: Desktop view for overview, and VR view for reasoning in the environmental context.** All experts generally expressed that the desktop view is better suited for getting an overview of the study data compared to the VR view. For example, they said that “*the advantage of the [desktop] view is overview*” [E1] or that “*the 2D charts provide a quick overview*” [E3]. In addition, E1 noted that the “*VR view lacks [this] overview over the data*” [E1] and that the “*VR tools are not powerful enough to quickly get the big picture.*” [E1] Two experts (E1, E4) emphasized that this enabled them to find points of interest (e.g., outliers) in the data that they then could analyze more in detail using the VR view. Besides, E3 explained that the desktop view is better suited for performing statistical analyses and to “*correlate different factors.*” [E3] Also, experts (E1, E3, E4) mentioned that the desktop view was faster to use when configuring components.

Concerning the VR view, experts noted that it allows you to immerse yourself in the data (E3) and to relive the study in its spatial context (E3, E5). Our experts mentioned various situations and tasks for which this can be beneficial. For example, E3 noted that reliving the study can be especially helpful when analyzing the study data weeks after you conducted it: “*[Then the VR view] offers the possibility to dive back into the study. It’s like being inside the study again [...] and maybe seeing aspects that were no longer present.*” [E3] Two experts (E2, E4) additionally mentioned that they found the VR view helpful to get a better overview of the course of the sessions. According to our experts, this spatial-visual impression (E1, E2) helps to assess where entities and events were located in the scene (E2) and to see what the participants of the study have seen (E1, E2). Although E5 stated that he did not use the VR view during the free phase as the desktop view already provided an interactive representation of the VR scene, he described the potential utility of the VR view for tasks that benefit from an increased depth perception. In this context, experts (E1, E2, E4, E5) also positively highlighted the stereoscopic 3D view in VR as an advantage over the scene view in the desktop view. They perceived it beneficial for spatial measurements (E2, E4, E5) and assessing the entities’ and events’ depth and spatial constellations (E1, E5). “*The [desktop scene view] helps a lot. [...] I can explore it like a game. However, only as 2D representation of a 3D world with which I don’t get the spatial impression. I only get that when I put on the headset. For example, when I look at [these objects] in the 2D scene view it is hard to assess their depth. I get that much, much better in VR.*” [E1]

With that, the VR view helps to explain the data (E2, E3, E4, E5), e.g., to find reasons for outliers and also allows for exploring the data to discover points of interest that you would not see in the 2D diagrams presented in the desktop view (E1, E3). Participants’ descriptions of their workflows during the free phase reflected these different strengths of the two views: Three experts (E3, E4, E5) stated that they started in the desktop view to get an overview of the data and then—if needed—switched to VR for reasoning (E4,

E5) or to explore the data further (E3): “*For me, the [desktop] view is overview, and if I need details, I’ll switch to VR, and that’s a coherent workflow for me.*” [E5] In contrast, E2 stated that he would start in VR to better understand the scene and switch to the desktop afterwards. E1 could imagine to use both workflows and E1, E3, and E5 agreed that the workflow “*depends on the research questions and the data*” [E3] to be analyzed. We also discovered three different patterns of participant behavior when using VR: (1) Remaining seated, (2) standing up, and (3) switching between standing and sitting. Interestingly, participants’ behavior partly changed between the guided and the free phase, indicating that their behavior not only relates to personal preferences but also to the type of task (cf. [6]).

All in all, experts stated that both views complement each other well (E1, E3), that their combination has no disadvantages (E1, E2, E4), and that the VR view is necessary to solve the tasks (E4).

**Theme #2: Interplay between in-situ and ex-situ visual analytics for analysis workflow.** A common theme during the evaluation was the interplay between the different views (i.e., desktop for ex-situ, VR for in-situ) and, thus, analysis types (i.e., in situ or ex situ). Three experts (E2, E4, E5) appreciated the instantaneous synchronization between the desktop and VR environment, with one expert noting that: “*If I start in 2D and then see something [...] interesting within the data, I’ll jump there, put the headset on and look at it*” [E1]. Another expert (E4) considered the synchronization the most important feature for the interplay. Three experts (E1, E2, E4) noted that, while the interplay was “*not perfect*” [E1], it worked “*flawlessly*” [E1] and there were “*no problems*” [E2, E4], removing the need to configure the VR view (E1). One expert (E1) highlighted the consistency of both views in terms of their icons and interaction with the components.

When transitioning between devices, problems occurred for 2 experts (E3, E4) due to loss of orientation: “*You somehow have a cut, and then you have to reorient yourself.*” [E3] Similarly, one expert (E4) also lost touch with the desktop’s position in the real world. Here, E3 suggested showing the real desktop device in VR, which could act as an anchor point and could enable the use of the keyboard in VR. In addition, one expert (E4) noted that the transition could be more fluid, while two experts noted that the transition “*worked well*” [E1, E2] and was “*quite fast*” [E2]. Here, two experts (E2, E4) mentioned that the 3D scene panel in the desktop view helped them when switching between devices, noting that it was “*really helpful*” [E2]. Another expert (E1) appreciated the combination of a “*familiar*” [E1] desktop view with its link to a native 3D application.

The availability of a desktop view in VR for quick access to the ex-situ analysis view was appreciated by four experts (E1, E2, E3, E4): “*I thought the possibility to jump directly to the desktop view was very cool. I can imagine that it helps in ensuring a quick transition.*” [E3]. However, three experts (E1, E2, E3) requested a deeper integration of this desktop view into the VR environment. Similarly, four experts (E1, E3, E4, E5) missed the availability of their component’s 2D charts in VR, for example for finding the exact point in time when an outlier occurred: “*I would’ve liked to see [the 2D charts] in VR! Because then I get the impression in 2D: okay, here the curve rising, then I can jump to this point and view it in detail.*” [E1]. During the concluding interview, we additionally

asked experts to rank the interplay between both views on a scale from 1 (“not very useful”) to 10 (“very useful”). Here, experts ranked the prototype with an average score of 8.00 ( $SD = 1.73$ ).

In terms of using both in-situ and ex-situ view for their analysis workflow, opinions were divided: Two experts (E1, E5) expressed a preference for working in the desktop view, especially as the desktop still offers a higher resolution (E1) and as HMDs can be uncomfortable (E1, E5). E5 therefore preferred working solely on the desktop view, treating both views separately and commenting that “*the [desktop view] can do everything, why should I still switch to VR?*” [E5]. In contrast, E1 could also imagine staying in VR if a well-integrated desktop view was available. Nevertheless, all experts used both views during the free phase of our expert study and, on average, spent almost half their time in VR. In addition, two experts (E1, E2) were positive about integrating the switch between desktop and VR into their analysis workflow, noting that each view has its distinct advantages and disadvantages: “*You can completely use the [desktop view] [...], and then for parts where there is an advantage for VR, you can use VR, and vice versa*” [E2]. E3 highlighted the potential for future hardware, which could offer better ergonomics and enable more fluid transitions.

**Theme #3: Modularity to suit individual analysis requirements.** Throughout the interview, experts mentioned several features that are required for analyzing their specific study data, “[*because*] everyone has different requirements for what is analyzed” [E1]. Here, experts highlighted the variety of different analysis use cases from their own work, such as crime scene investigation (E2, E5), collective behavior (E2), ergonomics (E3), balance issues (E1), or general asynchronous remote tasks (E4). Two experts (E1, E4) therefore emphasized the system’s extensibility, with one expert (E1) appreciating that ReLIVE uses a similar approach to a computational notebook. Experts also mentioned integrating additional visualizations, such as aggregated trails (E5), 3D heatmaps (E1, E5), or eye-tracking data (E4). To support this, experts noted that the data specification should support eyetracking data (E4), questionnaire data (E4), interaction data from non-immersive systems (E5), and audio recordings (E5).

Furthermore, experts suggested different metrics that could provide further insights, such as metrics to calculate a user’s balance (E1), speed (E2), room coverage (E2), task completion time (E2), metrics to analyze ergonomics (E3), and the user’s height as an indicator if users have bent down (E4, E5). Here, one expert highlighted the need for an export functionality for interoperability with other tools: “*If I really want to calculate something with the data, I would want to do that in a different tool*” [E1]. Experts also suggested extending the timeline to support aligning different sessions (E2, E3) or dividing sessions (E3) to facilitate comparisons. In addition, experts (E1, E4) suggested different forms of video playback in VR, including 360° videos (E1), displaying screen recordings on their corresponding devices (E1, E4), or even adopting different points of view (e.g., from individual participants) (E5). Lastly, four experts (E1, E2, E4, E5) emphasized that, to use ReLIVE, importing the data must be as easy as possible, highlighting the need for a data logging toolkit: “*If it’s easily applicable, then I would definitely like to use it. I even would have liked to use it for my last study!*” [E2].

**Theme #4: Despite the limited extent of the prototype, the concept was appreciated.** Generally, all experts (E1–5) were positive about the ReLIVE framework, noting that ReLIVE was “*intuitive*” [E1, E2], “*very positive*” [E4], “*very useful*” [E3], “*cool*” [E2, E3, E5], “*easy to use*” [E1, E2], and well-suited for the analysis of lab studies (E3). One expert cited ReLIVE’s adaptiveness as an advantage, “[*enabling*] the analysis of all kinds of studies” [E5]. Furthermore, E5 appreciated having the data in one place, while E4 mentioned that the data specification helped to create a mental map of the data. E4 also highlighted the potential for open science several times, mentioning ReLIVE’s potential for whiteboxing analysis, providing data provenance, and helping with replication and reproducibility of research results: “[...] *and I think that is indeed a step into this direction [of replication and, generally, reproducible research]. Because for VR, it’s often the case that [...] it’s flexible, but not traceable.*” [E4].

Since our evaluation prototype did not cover our entire concept, experts also mentioned current limitations. Experts expressed that the system is “*unfinished*” [E4] and that several details should be improved (E2, E3, E5), noting various minor usability issues (E1, E3, E4, E5). Specifically, E3 mentioned that the mapping between sessions and participants was unclear and the terminology should be clarified, while three experts (E1, E3, E5) felt that it was unclear when a session was visible in the 3D scene reconstruction. One expert also noted that pointing at objects to interact with them can be imprecise (E1). Experts also addressed the complexity of ReLIVE: four experts (E1, E2, E3, E5) mentioned that the preceding tutorial or training is mandatory, especially as the system was “*overwhelming*” [E5] at first. Here, experts provided the “*feature-richness*” [E5] as a reason, and ranked ReLIVE as an expert system with many different use cases (E2).

The experts’ impression of our prototype was also reflected in the scores of the System Usability Scale [15]. With a mean Usability Score of 74.50, the usability of our prototype was rated as good [3]. Additionally, all experts expressed, that they would like to use ReLIVE for the analysis of their next MR study.

**Theme #5: Potential for collaborative analysis.** A recurring topic across all experts was ReLIVE’s potential for a collaborative analysis, which was unanimously seen as advantage. Experts provided several possible collaboration opportunities, such as looking together at the desktop (as opposed to sharing VR glasses) (E2), symmetric collaboration in VR (E2), or asymmetric collaboration with one user at the desktop and another user in VR (E2, E4, E5). Here, user could take on different roles (E1, E2, E5): for example, the desktop user could act as a director, guiding the VR user to look at the data in context (E5). This asymmetric collaboration could also remove the need to transition between a desktop and VR (E5), allowing users to maximize the advantages of both VR and desktop (E2). Experts also pointed at additional opportunities during collaboration, such as the use of a private space (E3) or handing over control (E1).

## 7 INSIGHTS AND IMPLICATIONS

This section provides *design insights* (D1–D5) and *research implications* (I1–I5) based on our design of ReLIVE (Section 3), which combines recent trends in analyzing MR user studies (Section 2),

as well as findings from our design walkthrough (Section 5) and expert user study (Section 6). The structure of the insights and implications follows our research objectives of **RO1** *task allocation*, **RO2** *interplay between in-situ and ex-situ analysis*, and **RO3** *applicability*.

## 7.1 Task Allocation

All experts agreed with our intended distribution of using a desktop view for an ex-situ overview of the data and using VR for an in-situ analysis of the details in its environmental context, akin to an overview + detail visualization: “For me, the [desktop] view is overview, and if I need details, I’ll switch to VR, and that’s a coherent workflow for me.” [E5] The desktop view therefore allows experts to quickly compare data across different study sessions, find outliers, and create an analysis pipeline (**D1**). In contrast, the in-situ perspective (i.e., both 3D scene reconstruction on the desktop view and the VR environment) was used for sensemaking and reasoning tasks, focusing on *why* something happened—thus providing additional (environmental) context that cannot be found in the overview (**D2**). This is in line with prior research, which suggests that 2D outperforms immersive environments for overview tasks (cf. [54]). Here, further research could be conducted in directly comparing in-situ or ex-situ tasks in either environment (cf. [92]) (**I1**). However, the use of VR for an in-situ view on the data (as opposed to using the 3D study replication on the desktop) depends on both the task complexity (cf. [55]) as well as the user’s own preference.

### Design Insights – Task Allocation

- D1** Non-immersive visual analytics desktop view allows for analysis, ex-situ overview, finding outliers.
- D2** Immersive analytics view suited in-situ analysis, looking at data in context, reasoning, but also exploration.

### Research Implications – Task Allocation

- I1** Comparison of tasks in in-situ and ex-situ scenarios (cf. [54]).

## 7.2 Interplay Between In-Situ and Ex-Situ Analysis

Although ReLIVE’s desktop interface focused on an ex-situ analysis, adding an in-situ view (i.e., 3D scene panel) proved to be useful, providing an at-a-glance window into the environmental context. Depending on the task complexity, this window can render the full transition to VR obsolete, as the cost of switching displays may outweigh the benefits of full immersion and stereoscopy provided by VR (**D3**).

While the desktop’s 3D scene panel allowed for an instant glimpse into the VR scene, the transition between the desktop and the VR environment can still be disorienting. Here, future research could investigate a more explicit switch, requiring the user to specify a position in the VR scene before allowing the user to transition into VR (**I2**). In addition, anchors in VR (e.g., displaying a virtual desktop in the actual position of the physical desktop) can provide spatial context to the user, aiding in both transition into VR and

back to the desktop (cf. [67]) (**D4**). Future devices may further help this transition by gradually fading between VR and real world (**I3**), or by allowing the user to work exclusively in VR.

Furthermore, a task queue can be helpful in further reducing the need to transition between devices. Here, a *cross-reality linking and brushing* technique could be investigated, that could, for example, allow users to mark areas of interest (e.g., outliers) in the ex-situ view which are then highlighted in the in-situ view (**I4**). While the ex-situ view in VR should be easily accessible for an at-a-glance overview (cf. [19]), the 3D visualizations can be further improved with an on-demand 2D overview visualization.

In terms of video playback, both ex-situ and in-situ view complement each other: The ex-situ can act as a familiar 2D video player, allowing users to investigate video details (e.g., with a high-resolution desktop display). In contrast, the in-situ view helps in immersing the user in the original study setting, as videos can be played back in its original context (e.g., displaying screen recordings on a replica of the original device). The in-situ replication also enables analysts to view the study from different points of view, for example reliving the study from a participant’s perspective. Considering the limited field of view of current AR and VR devices, this emulation can provide additional insights into what was actually in the user’s view.

### Design Insights – Interplay Between In-Situ and Ex-Situ Analysis

- D3** Avoid needless transitions by offering basic in-situ view in desktop (and vice versa for VR).
- D4** Use anchors when transitioning between VR and desktop (cf. [67]).

### Research Implications – Interplay Between In-Situ and Ex-Situ Analysis

- I2** Investigate requiring an explicit switch before allowing transition into VR.
- I3** Investigate effect of gradual transition between the real world and VR.
- I4** Investigate benefits and challenges of cross-reality linking and brushing.

## 7.3 Applicability

Although we only evaluated our concepts as an initial prototype, our findings confirm the general applicability of our concept. The vast amount of different study goals, metrics, and visualizations provided by the experts highlights the need for a flexible analysis tool. Here, our computational notebook approach with its component instances proved useful in allowing users to easily employ analysis tools and create visualizations both on the desktop and in VR (**D5**). However, our prototype only provides a predefined set of component templates. Therefore more research is necessary in terms of creating these component templates (**I5**). Furthermore, although there is an increasing number of IA applications focusing on a set of fixed visualizations, converting study data for each tool can be time-consuming and error-prone. A standardized data



specification could help in both holistically recording the data and provide interoperability with other applications. We therefore see ReLIVE as an initial starting point that highlights the data variety of MR studies, but further work is necessary to reach a generally agreed-upon format.

#### Design Insights – Applicability

- D5 Computational notebook approach using components allows for flexible analysis of user study data both on the desktop and in VR (cf. [2]).

#### Research Implications – Applicability

- I5 Investigate benefits and limitations of programming computational components with in-situ and ex-situ representations.

## 8 LIMITATIONS AND FUTURE WORK

Our evaluation prototype of ReLIVE was tailored towards evaluating the interplay between in-situ and ex-situ analysis of MR user study data. As the analysis of unfamiliar MR user study data with a novel analysis environment can be challenging, we intentionally tried to reduce the complexity of the prototype, the data, and its visualizations. While this lowered the threshold of interacting with ReLIVE, it also left several aspects of the overarching concept untouched: Conceptually, the desktop view can be organized linearly like a computational notebook, allowing analysts e.g., to configure or program components that suit their analysis. We did not implement programmable component templates to reduce complexity. Further studies are necessary to study the general feasibility and applicability of such components for a mixed-immersion analysis workflow (cf. [2, 72]). In this regard, prior work by Borowski et al. [11] has already demonstrated the benefits of providing packages (cf. components) for computational notebooks and how they can even support shared activities [12]. Future work could investigate this concept, allowing analysts to access, share, or contribute to a public collection of available metrics, visualizations, or scripts. This would in turn facilitate the idea of open science, allowing for a whiteboxing of the analysis (supporting reproducibility and data provenance) by, for example sharing the analysis notebook along with the available study data, in addition to preregistering a user study.

To further reduce complexity, our evaluation prototype only provides simple, yet easily understandable predefined visualization templates (see Table 1). Given that participants were not familiar with the used reference studies, we intended to not overwhelm participants further. This allowed us to study the suitability of in-situ and ex-situ analysis for different analysis tasks. While our insights already allude to the necessity of 2D overview visualizations inside VR, additional studies are necessary to investigate the role of 2D and 3D visualizations in the different views (e.g., desktop, VR). Furthermore, future work could investigate the mapping between 2D visualizations and their 3D equivalents and how established techniques such as linking and brushing could facilitate the interplay between in-situ and ex-situ analysis.

The reduced evaluation prototype and relatively small amount of expert participants also limits the extent of our results and discussion. Due to the predefined tasks and missing comparison to existing tools, our evaluation provides only limited insights into the real world efficacy of ReLIVE. Thus, we concentrated on investigating the transition between different environments instead of quantitatively comparing if ReLIVE provides any ecologically valid benefits over existing tools. However, given that participants were experts in their respective fields, they provided valuable insights into initial opportunities and challenges. In addition, although the tasks artificially forced a transition between desktop and VR, they were grounded on our own experience when analyzing MR studies. We aim to extend ReLIVE with the missing described concepts, including component templates and richer media integration in VR, allowing us to study possible benefits in more detail.

Although ReLIVE is designed to provide access to a wide range of data, gathered using an abundance of available quantitative and qualitative methods, we intentionally limited the scope of available data for the evaluation. This reduced complexity facilitated especially the free phase in the expert user study and reduced the overall duration of the user study. However, we see a great potential for the interplay between in-situ and ex-situ analysis of e.g., qualitative data such as annotating and coding of video material: Traditionally, the coding of video material is a part of the user study analysis workflow done in a desktop environment. The analyst carefully watches recordings of user study sessions to understand participants' activities, annotates remarkable events, and assigns codes to them to quantify the material for further analysis. However, having a pre-defined and static perspective of the camera can be a limitation (due to occlusion or limited resolution). Here, the in-situ analysis approach can complement the ex-situ analysis by investigating the activities from any desired point of view. Future research could investigate to what extent sophisticated video analysis tools (e.g., [52]) in combination with in-situ approaches can support MR user study analyses either *post-hoc* after completing all study sessions or even with the help of additional experimenters *ad hoc* during the runtime of a study session (e.g., adopting the participant's field of view).

Lastly, we intentionally focused on a single-user scenario to investigate the interplay between in-situ and ex-situ analysis. However, the findings of our expert user study highlighted the potential for ReLIVE as a collaborative system, supporting different constellations (e.g., collaborating on a desktop, or assuming different roles such as for overview on desktop and detail in VR), with one participant expressing that “[...] *you need another person, and, I think, then it'll be really great*” [E5]. Further studies are necessary to investigate the potential benefits and challenges of such a collaborative system for the analysis of MR user studies (cf. heterogeneous remote assistance systems [26, 34]).

## 9 CONCLUSION

We have presented ReLIVE: a *mixed-immersion* visual analytics framework that combines an immersive VR environment with a non-immersive desktop environment to enable the holistic exploration and malleable analysis of MR user studies. The immersive VR environment allows users to relive an interactive recording

of a replica of the original study, providing the possibility for in-situ analysis of the data. In contrast, the non-immersive desktop view facilitates the analysis of aggregated study data and provides a holistic overview over the available study data. In our concept, users can program components to calculate metrics and create visualizations in an interface akin to a computational notebook. RE LIVE also supports the user's transition between the VR and desktop environment, for example by synchronizing both environments in real-time and offering a glimpse of the VR environment in the desktop view and vice versa—thus representing an asynchronous hybrid user interface.

We provided design insights and research implications, based on our design of RE LIVE as well as a two-step evaluation process consisting of a design walkthrough and an expert user study. Here, we demonstrate the general applicability of our concept, showed the benefits of bridging in-situ and ex-situ analysis, and provided insight into transitioning between immersive (VR) and non-immersive (desktop) visual analytics. With our work, we contribute to the Grand Challenges of Immersive Analytics [30] by offering a flexible analysis workspace and investigating the role of IA when analyzing mixed reality study data. In addition, our work contributes toward the underrepresented field of transitional user interfaces by investigating the transition between a VR HMD and a desktop interface.

## ACKNOWLEDGMENTS

We thank Dieter Schmalstieg for constructive discussions on early ideas of the RE LIVE framework and Tiare Feuchtner for her invaluable feedback. This research was partially funded by the *Deutsche Forschungsgemeinschaft* (DFG, German Research Foundation) – Project-ID 251654672 – TRR 161 (Project C01), SMARTACT (BMBF, Grant 01EL1820A), and U.S. National Science Foundation grant IIS-1908605. Any opinions, findings, and conclusions or recommendations expressed here are those of the authors and do not necessarily reflect the views of the funding agencies.

## REFERENCES

- [1] Marcus F. C. Alencar, Alberto B. Raposo, and Simone D. J. Barbosa. 2011. Composition of HCI Evaluation Methods for Hybrid Virtual Environments. In *Proceedings of the ACM Symposium on Applied Computing*. ACM, New York, NY, USA, 1237–1244. <https://doi.org/10.1145/1982185.1982457>
- [2] Sriram Karthik Badam, Andreas Mathisen, Roman Rädle, Clemens N. Klokmoose, and Niklas Elmqvist. 2019. Vistrates: A Component Model for Ubiquitous Analytics. *IEEE Transactions on Visualization and Computer Graphics* 25, 1 (Jan. 2019), 586–596. <https://doi.org/10.1109/TVCG.2018.2865144>
- [3] Aaron Bangor, Philip Kortum, and James Miller. 2009. Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. *J. Usability Studies* 4, 3 (May 2009), 114–123.
- [4] Andrea Batch, Andrew Cunningham, Maxime Cordeil, Niklas Elmqvist, Tim Dwyer, Bruce H. Thomas, and Kim Marriott. 2020. 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 (Jan. 2020), 536–546. <https://doi.org/10.1109/TVCG.2019.2934803>
- [5] Andrea Batch, Biswaksen Patnaik, Moses Akazue, and Niklas Elmqvist. 2020. Scents and Sensibility: Evaluating Information Olfaction. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–14. <https://doi.org/10.1145/3313831.3376733>
- [6] Martin Bellgardt, Sebastian Pick, Daniel Zielasko, Tom Vierjahn, Benjamin Weyers, and Torsten W. Kühlen. 2017. Utilizing Immersive Virtual Reality in Everydaywork. In *2017 IEEE 3rd Workshop on Everyday Virtual Reality (WEVR)*. 1–4. <https://doi.org/10.1109/WEVR.2017.7957708>
- [7] Joanna Bergström, Tor-Salve Dalsgaard, Jason Alexander, and Kasper Hornbæk. 2021. How to Evaluate Object Selection and Manipulation in VR? Guidelines from 20 Years of Studies. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–20. <https://doi.org/10.1145/3411764.3445193>
- [8] Christopher Berns, Grace Chin, Joel Savitz, Jason Kiesling, and Fred Martin. 2019. MYR: A Web-Based Platform for Teaching Coding Using VR. In *Proceedings of the ACM Symposium on Computer Science Education*. ACM, New York, NY, USA, 77–83. <https://doi.org/10.1145/3287324.3287482>
- [9] Verena Biener, Daniel Schneider, Travis Gesslein, Alexander Otte, Bastian Kuth, Per Ola Kristensson, Eyal Ofek, Michel Pahud, and Jens Grubert. 2020. Breaking the Screen: Interaction Across Touchscreen Boundaries in Virtual Reality for Mobile Knowledge Workers. *IEEE Transactions on Visualization and Computer Graphics* PP (Aug. 2020), 14. <https://doi.org/10.1109/TVCG.2020.3023567>
- [10] M. Billinghurst, H. Kato, and I. Poupyrev. 2001. The MagicBook - Moving Seamlessly between Reality and Virtuality. *IEEE Computer Graphics and Applications* 21, 3 (May 2001), 6–8. <https://doi.org/10.1109/38.920621>
- [11] Marcel Borowski, Roman Rädle, and Clemens N. Klokmoose. 2018. Codestrate Packages: An Alternative to "One-Size-Fits-All" Software. In *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–6. <https://doi.org/10.1145/3170427.3188563>
- [12] Marcel Borowski, Johannes Zagermann, Clemens N. Klokmoose, Harald Reiterer, and Roman Rädle. 2020. Exploring the Benefits and Barriers of Using Computational Notebooks for Collaborative Programming Assignments. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*. ACM, Portland OR USA, 468–474. <https://doi.org/10.1145/3328778.3366887>
- [13] Michael Bostock, Vadim Ogievetsky, and Jeffrey Heer. 2011. D<sup>3</sup>: Data-Driven Documents. *IEEE Transactions on Visualization and Computer Graphics* 17, 12 (2011), 2301–2309. <https://doi.org/10.1109/TVCG.2011.185>
- [14] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- [15] John Brooke. 1995. SUS: A quick and dirty usability scale. *Usability Eval. Ind.* 189 (11 1995).
- [16] Frederik Brudy, Christian Holz, Roman Rädle, Chi-Jui Wu, Steven Houben, Clemens Nylandsted Klokmoose, and Nicolai Marquardt. 2019. Cross-Device Taxonomy: Survey, Opportunities and Challenges of Interactions Spanning Across Multiple Devices. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–28. <https://doi.org/10.1145/3290605.3300792>
- [17] Frederik Brudy, Suppachai Suwanwatcharachart, Wenyu Zhang, Steven Houben, and Nicolai Marquardt. 2018. EagleView: A Video Analysis Tool for Visualising and Querying Spatial Interactions of People and Devices. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 61–72. <https://doi.org/10.1145/3279778.3279795>
- [18] Bernhard Brüning, Christian Schnier, Karola Pitsch, and Sven Wachsmuth. 2012. Integrating PAMOCAT in the Research Cycle: Linking Motion Capturing and Conversation Analysis. In *Proceedings of the ACM Conference on Multimodal Interaction*. ACM, New York, NY, USA, 201–208. <https://doi.org/10.1145/2388676.2388716>
- [19] Wolfgang Büschel, Anke Lehmann, and Raimund Dachselt. 2021. MIRIA: A Mixed Reality Toolkit for the In-Situ Visualization and Analysis of Spatio-Temporal Interaction Data. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–15. <https://doi.org/10.1145/3411764.3445651>
- [20] Wolfgang Büschel, Annett Mitschick, Thomas Meyer, and Raimund Dachselt. 2019. Investigating Smartphone-Based Pan and Zoom in 3D Data Spaces in Augmented Reality. In *Proceedings of the ACM Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, New York, NY, USA, 1–13. <https://doi.org/10.1145/3338286.3340113>
- [21] Wolfgang Büschel, Patrick Reipschläger, Ricardo Langner, and Raimund Dachselt. 2017. Investigating the Use of Spatial Interaction for 3D Data Visualization on Mobile Devices. In *Proceedings of the ACM Conference on Interactive Surfaces and Spaces*. ACM, New York, NY, USA, 62–71. <https://doi.org/10.1145/3132272.3134125>
- [22] Simon Butscher, Sebastian Hubenschmid, Jens Müller, Johannes Fuchs, and Harald Reiterer. 2018. Clusters, Trends, and Outliers: How Immersive Technologies Can Facilitate the Collaborative Analysis of Multidimensional Data. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–12. <https://doi.org/10.1145/3173574.3173664>
- [23] Felipe G. Carvalho, Daniela G. Trevisan, and Alberto Raposo. 2012. Toward the Design of Transitional Interfaces: An Exploratory Study on a Semi-Immersive Hybrid User Interface. *Virtual Reality* 16, 4 (Nov. 2012), 271–288. <https://doi.org/10.1007/s10055-011-0205-y>
- [24] Marco Cavallo, Mishal Dholakia, Matous Havlena, Kenneth Ocheltree, and Mark Podlaseck. 2019. Dataspace: A Reconfigurable Hybrid Reality Environment for Collaborative Information Analysis. In *Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces*. IEEE, Piscataway, NJ, USA, 145–153. <https://doi.org/10.1109/VR.2019.8797733>

- [25] Marco Cavallo, Mishal Dolakia, Matous Havlena, Kenneth Ocheltree, and Mark Podlaseck. 2019. Immersive Insights: A Hybrid Analytics System for Collaborative Exploratory Data Analysis. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*. ACM, New York, NY, USA, 1–12. <https://doi.org/10.1145/3359996.3364242>
- [26] Henry Chen, Austin S. Lee, Mark Swift, and John C. Tang. 2015. 3D Collaboration Method over HoloLens™ and Skype™ End Points. In *Proceedings of the International Workshop on Immersive Media Experiences*. ACM, New York, NY, USA, 27–30. <https://doi.org/10.1145/2814347.2814350>
- [27] L. Chittaro, R. Ranon, and L. Ieronutti. 2006. VU-Flow: A Visualization Tool for Analyzing Navigation in Virtual Environments. *IEEE Transactions on Visualization and Computer Graphics* 12, 6 (Nov 2006), 1475–1485. <https://doi.org/10.1109/TVCG.2006.109>
- [28] Philip DeCamp, George Shaw, Rony Kubat, and Deb Roy. 2010. An Immersive System for Browsing and Visualizing Surveillance Video. In *Proceedings of the ACM Conference on Multimedia*. ACM, New York, NY, USA, 371–380. <https://doi.org/10.1145/1873951.1874002>
- [29] Tobias Drey, Jan Gugenheimer, Julian Karlbauer, Maximilian Milo, and Enrico Rukzio. 2020. VRSketchIn: Exploring the Design Space of Pen and Tablet Interaction for 3D Sketching in Virtual Reality. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–14. <https://doi.org/10.1145/3313831.3376628>
- [30] Barrett Ens, Benjamin Bach, Maxime Cordeil, Ulrich Engelke, Marcos Serrano, Wesley Willett, Arnaud Prouzeau, Christoph Anthes, Wolfgang Büschel, Cody Dunne, Tim Dwyer, Jens Grubert, Jason H. Haga, Nurit Kirshenbaum, Dylan Kobayashi, Tica Lin, Monsurat Olaosebikan, Fabian Pointecker, David Saffo, Nazmus Saquib, Dieter Schmalstieg, Danielle Albers Szafir, Matt Whitlock, and Yalong Yang. 2021. Grand Challenges in Immersive Analytics. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–17. <https://doi.org/10.1145/3411764.3446866>
- [31] João Marcelo Evangelista Belo, Anna Maria Feit, Tiare Feuchtnr, and Kaj Grønbaek. 2021. XRgonomics: Facilitating the Creation of Ergonomic 3D Interfaces. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–11. <https://doi.org/10.1145/3411764.3445349>
- [32] Alessandro Febretti, Arthur Nishimoto, Terrance Thigpen, Jonas Talandis, Lance Long, J. D. Pirtle, Tom Peterka, Alan Verlo, Maxine Brown, Dana Plepys, Dan Sandin, Luc Renambot, Andrew Johnson, and Jason Leigh. 2013. {CAVE}2: A Hybrid Reality Environment for Immersive Simulation and Information Analysis. In *The Engineering Reality of Virtual Reality 2013*. SPIE, Bellingham, WA, USA, 864903. <https://doi.org/10.1117/12.2005484>
- [33] Steven Feiner and Ari Shamash. 1991. Hybrid User Interfaces: Breeding Virtually Bigger Interfaces for Physically Smaller Computers. In *Proceedings of the ACM Symposium on User Interface Software and Technology*. ACM, New York, NY, USA, 9–17. <https://doi.org/10.1145/120782.120783>
- [34] Philipp Fleck, Fernando Reyes Aviles, Christian Pirchheim, Clemens Arth, and Dieter Schmalstieg. 2020. MAUI: Tele-Assistance for Maintenance of Cyber-Physical Systems. In *Proceedings of the International Conference on Computer Vision Theory and Applications*. SciTePress, Setúbal, Portugal. <https://doi.org/10.5220/0009093708000812>
- [35] Florian "Floyd" Mueller, Tim Dwyer, Sarah Goodwin, Kim Marriott, Jialin Deng, Han D. Phan, Jionghao Lin, Kun-Ting Chen, Yan Wang, and Rohit Ashok Khot. 2021. Data as Delight: Eating Data. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, Article 621, 14 pages. <https://doi.org/10.1145/3411764.3445218>
- [36] R Foundation. 2021. *R Project for Statistical Computing*. Retrieved August 31, 2021 from <https://www.r-project.org/>
- [37] Epic Games. 2021. Unreal Engine – The Most Powerful Real-Time 3D Creation Platform. <https://www.unrealengine.com/en-US/>
- [38] Google. 2021. *Colaboratory*. Retrieved August 31, 2021 from <https://colab.research.google.com>
- [39] Raphael Grasset, Julian Looser, and Mark Billinghurst. 2006. Transitional Interface: Concept, Issues and Framework. In *2006 IEEE/ACM International Symposium on Mixed and Augmented Reality*. 231–232. <https://doi.org/10.1109/ISMAR.2006.297819>
- [40] Jens Grubert, Matthias Kranz, and Aaron Quigley. 2016. Challenges in Mobile Multi-Device Ecosystems. *mUX: The Journal of Mobile User Experience* 5, 1 (Dec 2016). <https://doi.org/10.1186/s13678-016-0007-y>
- [41] R. Hartson and P. Pyla. 2018. *The UX Book: Agile UX Design for a Quality User Experience*. Elsevier Science. <https://books.google.de/books?id=Pp5RjgEACAAJ>
- [42] Juan David Hincapié-Ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani. 2014. Consumed Endurance: A Metric to Quantify Arm Fatigue of Mid-Air Interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. Association for Computing Machinery, Toronto, Ontario, Canada, 1063–1072. <https://doi.org/10.1145/2556288.2557130>
- [43] Steven Houben, Nicolai Marquardt, Jo Vermeulen, Clemens Klokmoose, Johannes Schöning, Harald Reiterer, and Christian Holz. 2017. Opportunities and Challenges for Cross-Device Interactions in the Wild. *Interactions* 24, 5 (Aug. 2017), 58–63. <https://doi.org/10.1145/3121348>
- [44] Sebastian Hubenschmid, Johannes Zagermann, Simon Butscher, and Harald Reiterer. 2021. STREAM: Exploring the Combination of Spatially-Aware Tablets with Augmented Reality Head-Mounted Displays for Immersive Analytics. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–14. <https://doi.org/10.1145/3411764.3445298>
- [45] Sebastian Hubenschmid, Johannes Zagermann, Daniel Fink, Jonathan Wieland, Tiare Feuchtnr, and Harald Reiterer. 2021. Towards Asynchronous Hybrid User Interfaces for Cross-Reality Interaction. In *ISS'21 Workshop Proceedings: "Transitional Interfaces in Mixed and Cross-Reality: A New Frontier?"*, Hans-Christian Jetter, Jan-Henrik Schröder, Jan Gugenheimer, Mark Billinghurst, Christoph Anthes, Mohamed Khamis, and Tiare Feuchtnr (Eds.). <https://doi.org/10.18148/kops/352-2-84mm0sggcqz02>
- [46] TIBCO Software Inc. 2021. TIBCO Spotfire®. <https://www.tibco.com/products/tibco-spotfire>
- [47] Hans-Christian Jetter, Roman Rädle, Tiare Feuchtnr, Christoph Anthes, Judith Friedl, and Clemens Nylandstedt Klokmoose. 2020. "In VR, Everything Is Possible!": Sketching and Simulating Spatially-Aware Interactive Spaces in Virtual Reality. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–16. <https://doi.org/10.1145/3313831.3376652>
- [48] Daniel A. Keim, Jan.-March/2002. Information Visualization and Visual Data Mining. *IEEE Transactions on Visualization and Computer Graphics* 8, 1 (Jan.-March/2002), 1–8. <https://doi.org/10.1109/2945.981847>
- [49] Robert S. Kennedy, Norman E. Lane, Kevin S. Berbaum, and Michael G. Lilienthal. 1993. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology* 3, 3 (July 1993), 203–220. [https://doi.org/10.1207/s15327108ijap0303\\_3](https://doi.org/10.1207/s15327108ijap0303_3)
- [50] Karsten Klein, Michael Aichem, Ying Zhang, Stefan Erk, Björn Sommer, and Falk Schreiber. 2021. TEAMWISE: Synchronised Immersive Environments for Exploration and Analysis of Animal Behaviour. *Journal of Visualization* 24, 4 (Aug. 2021), 845–859. <https://doi.org/10.1007/s12650-021-00746-2>
- [51] Simon Kloiber, Volker Settgest, Christoph Schinko, Martin Weinzerl, Johannes Fritz, Tobias Schreck, and Reinhold Preiner. 2020. Immersive Analysis of User Motion in VR Applications. *The Visual Computer* 36, 10–12 (Oct. 2020), 1937–1949. <https://doi.org/10.1007/s00371-020-01942-1>
- [52] Clemens N. Klokmoose, Christian Remy, Janus Bager Kristensen, Rolf Bagge, Michel Beaudouin-Lafon, and Wendy Mackay. 2019. Videotrates: Collaborative, Distributed and Programmable Video Manipulation. In *Proceedings of the ACM Symposium on User Interface Software and Technology*. ACM, New York, NY, USA, 233–247. <https://doi.org/10.1145/3332165.3347912>
- [53] Pascal Knierim, Dimitri Hein, Albrecht Schmidt, and Thomas Kosch. 2021. The SmARtPhone Controller. *i-com* 20, 1 (2021), 49–61. <https://doi.org/doi:10.1515/i-com-2021-0003>
- [54] Matthias Kraus, Katrin Angerbauer, Juri Buchmüller, Daniel Schweitzer, Daniel A. Keim, Michael Sedlmair, and Johannes Fuchs. 2020. Assessing 2D and 3D Heatmaps for Comparative Analysis: An Empirical Study. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–14. <https://doi.org/10.1145/3313831.3376675>
- [55] M. Kraus, N. Weiler, D. Oelke, J. Kehr, D. A. Keim, and J. Fuchs. 2019. The Impact of Immersion on Cluster Identification Tasks. *IEEE Transactions on Visualization and Computer Graphics* (2019), 1–1. <https://doi.org/10.1109/TVCG.2019.2934395>
- [56] Ricardo Langner, Marc Satkowski, Wolfgang Büschel, and Raimund Dachselt. 2021. MARVIS: Combining Mobile Devices and Augmented Reality for Visual Data Analysis. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–17. <https://doi.org/10.1145/3411764.3445593>
- [57] B. Lee, D. Brown, B. Lee, C. Hurter, S. Drucker, and T. Dwyer. 2020. Data Visceralization: Enabling Deeper Understanding of Data Using Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics* (2020), 1–1. <https://doi.org/10.1109/TVCG.2020.3030435>
- [58] Chi-Jung Lee and Hung-Kuo Chu. 2018. Dual-MR: Interaction with Mixed Reality Using Smartphones. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*. ACM, New York, NY, USA, 1–2. <https://doi.org/10.1145/3281505.3281618>
- [59] Klemen Lilija, Henning Pohl, and Kasper Hornbæk. 2020. Who Put That There? Temporal Navigation of Spatial Recordings by Direct Manipulation. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–11. <https://doi.org/10.1145/3313831.3376604>
- [60] Weizhou Luo, Eva Goebel, Patrick Reipschläger, Mats Ole Ellenberg, and Raimund Dachselt. 2021. Exploring and Slicing Volumetric Medical Data in Augmented Reality Using a Spatially-Aware Mobile Device. In *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality Adjunct*. IEEE, Piscataway, NJ, USA.
- [61] Nicolai Marquardt, Frederico Schardong, and Anthony Tang. 2015. EXCITE: Exploring Collaborative Interaction in Tracked Environments. In *Proceedings of IFIP Human-Computer Interaction*. Springer International Publishing, Cham, Germany, 89–97.

- [62] Kim Marriott, Falk Schreiber, Tim Dwyer, Karsten Klein, Nathalie Henry Riche, Takayuki Itoh, Wolfgang Stuerzlinger, and Bruce H. Thomas (Eds.). 2018. *Immersive Analytics*. Number 11190 in Lecture Notes in Computer Science. Springer, Cham.
- [63] L. Merino, M. Schwarzl, M. Kraus, M. Sedlmair, D. Schmalstieg, and D. Weiskopf. 2020. Evaluating Mixed and Augmented Reality: A Systematic Literature Review (2009–2019). In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 438–451. <https://doi.org/10.1109/ISMAR50242.2020.00069>
- [64] Paul Milgram and Fumio Kishino. 1994. A Taxonomy of Mixed Reality Visual Displays. *IEICE Transactions on Information and Systems* 77, 12 (1994), 1321–1329.
- [65] K.Jarrold Millman and Fernando Pérez. 2018. Developing Open-Source Scientific Practice. In *Implementing Reproducible Research* (first ed.), Victoria Stodden, Friedrich Leisch, and Roger D. Peng (Eds.). Chapman and Hall/CRC, 149–183. <https://doi.org/10.1201/9781315373461-6>
- [66] Florian Müller, Martin Schmitz, Daniel Schmitt, Sebastian Günther, Markus Funk, and Max Mühlhäuser. 2020. Walk The Line: Leveraging Lateral Shifts of the Walking Path as an Input Modality for Head-Mounted Displays. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–15. <https://doi.org/10.1145/3313831.3376852>
- [67] Jens Müller, Roman Rädle, and Harald Reiterer. 2017. Remote Collaboration With Mixed Reality Displays: How Shared Virtual Landmarks Facilitate Spatial Referencing. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 6481–6486. <https://doi.org/10.1145/3025453.3025717>
- [68] Jens Müller, Johannes Zagermann, Jonathan Wieland, Ulrike Pfeil, and Harald Reiterer. 2019. A Qualitative Comparison Between Augmented and Virtual Reality Collaboration with Handheld Devices. In *Proceedings of Mensch und Computer 2019 (Hamburg, Germany) (MuC'19)*. Association for Computing Machinery, New York, NY, USA, 399–410. <https://doi.org/10.1145/3340764.3340773>
- [69] Michael Nebeling, Maximilian Speicher, Xizi Wang, Shwetha Rajaram, Brian D Hall, Zijian Xie, Alexander R E Raistrick, Michelle Aebbersold, Edward G Happ, Jiayin Wang, Yanan Sun, Lotus Zhang, Leah E Ramsier, and Rhea Kulkarni. 2020. MRAT: The Mixed Reality Analytics Toolkit. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 12. <https://doi.org/10.1145/3313831.3376330>
- [70] Federica Pallavicini and Alessandro Pepe. 2019. Comparing Player Experience in Video Games Played in Virtual Reality or on Desktop Displays: Immersion, Flow, and Positive Emotions. In *Extended Abstracts of the ACM Symposium on Computer-Human Interaction in Play*. ACM, New York, NY, USA, 195–210. <https://doi.org/10.1145/3341215.3355736>
- [71] Leonardo Pavanatto, Chris North, Doug A. Bowman, Carmen Badea, and Richard Stokley. 2021. Do We Still Need Physical Monitors? An Evaluation of the Usability of AR Virtual Monitors for Productivity Work. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. IEEE, Lisboa, Portugal, 759–767. <https://doi.org/10.1109/VR50410.2021.00103>
- [72] Fernando Perez and Brian E. Granger. 2007. IPython: A System for Interactive Scientific Computing. *Computing in Science Engineering* 9, 3 (2007), 21–29. <https://doi.org/10.1109/MCSE.2007.53>
- [73] Michael Prilla and Lisa M. Rühmann. 2018. An Analysis Tool for Cooperative Mixed Reality Scenarios. In *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. 31–35. <https://doi.org/10.1109/ISMAR-Adjunct.2018.00026>
- [74] Roman Rädle, Midas Nouwens, Kristian Antonsen, James R. Eagan, and Clemens N. Klokose. 2017. Codestrates: Literate Computing with Webstrates. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 715–725. <https://doi.org/10.1145/3126594.3126642>
- [75] Umar Rashid, Miguel A. Nacenta, and Aaron Quigley. 2012. The Cost of Display Switching: A Comparison of Mobile, Large Display and Hybrid UI Configurations. In *Proceedings of the ACM Conference on Advanced Visual Interfaces*. ACM, New York, NY, USA, 99–106. <https://doi.org/10.1145/2254556.2254577>
- [76] Carolin Reichherzer, Andrew Cunningham, Tracey Coleman, Ruochen Cao, Kurt McManus, Dion Sheppard, Mark Kohler, Mark Billinghurst, and Bruce H Thomas. 2021. 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*. ACM, Yokohama Japan, 1–12. <https://doi.org/10.1145/3411764.3445464>
- [77] Patrick Reipschläger and Raimund Dachsel. 2019. DesignAR: Immersive 3D-Modeling Combining Augmented Reality with Interactive Displays. In *Proceedings of the ACM Conference on Interactive Surfaces and Spaces*. ACM, New York, NY, USA, 29–41. <https://doi.org/10.1145/3343055.3359718>
- [78] Patrick Reipschläger, Tamara Flemisch, and Raimund Dachsel. 2021. Personal Augmented Reality for Information Visualization on Large Interactive Displays. *IEEE Transactions on Visualization and Computer Graphics* (Feb. 2021). <https://doi.org/10.1109/TVCG.2020.3030460> arXiv:2009.03237
- [79] Jun Rekimoto and Masanori Saitoh. 1999. Augmented Surfaces: A Spatially Continuous Work Space for Hybrid Computing Environments. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 378–385. <https://doi.org/10.1145/302979.303113>
- [80] Yvonne Rogers. 2012. HCI Theory: Classical, Modern, and Contemporary. *Synthesis Lectures on Human-Centered Informatics* 5, 2 (May 2012), 1–129. <https://doi.org/10.2200/S00418ED1V01Y201205HCI014>
- [81] Salesforce. 2021. Tableau: Business Intelligence and Analytics Software. <https://www.tableau.com/>.
- [82] Prasanth Sasikumar, Max Collins, Huidong Bai, and Mark Billinghurst. 2021. XRTB: A Cross Reality Teleconference Bridge to Incorporate 3D Interactivity to 2D Teleconferencing. In *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–4. <https://doi.org/10.1145/3411763.3451546>
- [83] Marc Satkowski, Weizhou Luo, and Raimund Dachsel. 2021. Towards In-situ Authoring of AR Visualizations with Mobile Devices. In *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality Adjunct*. IEEE, Piscataway, NJ, USA.
- [84] Arvind Satyanarayan, Dominik Moritz, Kanit Wongsuphasawat, and Jeffrey Heer. 2017. Vega-Lite: A Grammar of Interactive Graphics. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (Jan. 2017), 341–350. <https://doi.org/10.1109/TVCG.2016.2599030>
- [85] Richard Skarbez, Missie Smith, and Mary C. Whitton. 2021. Revisiting Milgram and Kishino's Reality-Virtuality Continuum. *Frontiers in Virtual Reality* 2 (March 2021), 647997. <https://doi.org/10.3389/frvir.2021.647997>
- [86] Mauricio Sousa, Daniel Mendes, Soraia Paulo, Nuno Matela, Joaquim Jorge, and Daniel Simões Lopes. 2017. VRRRRoom: Virtual Reality for Radiologists in the Reading Room. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 4057–4062. <https://doi.org/10.1145/3025453.3025566>
- [87] Maximilian Speicher, Brian D. Hall, and Michael Nebeling. 2019. What Is Mixed Reality?. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–15. <https://doi.org/10.1145/3290605.3300767>
- [88] Seung Youb Ssin, James A. Walsh, Ross T. Smith, Andrew Cunningham, and Bruce H. Thomas. 2019. GeoGate: Correlating Geo-Temporal Datasets Using an Augmented Reality Space-Time Cube and Tangible Interactions. In *Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces*. IEEE, Piscataway, NJ, USA, 210–219. <https://doi.org/10.1109/VR.2019.8797812>
- [89] Hemant Bhaskar Surale, Aakar Gupta, Mark Hancock, and Daniel Vogel. 2019. TabletInVR: Exploring the Design Space for Using a Multi-Touch Tablet in Virtual Reality. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–13. <https://doi.org/10.1145/3290605.3300243>
- [90] Zsolt Szalavári and Michael Gervautz. 2008. The Personal Interaction Panel – a Two-Handed Interface for Augmented Reality. *Computer Graphics Forum* 16 (June 2008), C335–C346. <https://doi.org/10.1111/1467-8659.16.3.conferenceissue.35>
- [91] Anthony Tang, Michel Pahud, Sheelagh Carpendale, and Bill Buxton. 2010. VisTACO: Visualizing Tabletop Collaboration. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. ACM, New York, NY, USA, 29–38. <https://doi.org/10.1145/1936652.1936659>
- [92] Katja Vock, Sebastian Hubenschmid, Johannes Zagermann, Simon Butscher, and Harald Reiterer. 2021. IDIAR: Augmented Reality Dashboards to Supervise Mobile Intervention Studies. In *Mensch und Computer*. ACM, New York, NY, USA. <https://doi.org/10.1145/3473856.3473876>
- [93] Ulrich von Zadow and Raimund Dachsel. 2017. GIANT: Visualizing Group Interaction at Large Wall Displays. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, Denver Colorado USA, 2639–2647. <https://doi.org/10.1145/3025453.3026006>
- [94] Jia Wang and Robert Lindeman. 2014. Coordinated 3D Interaction in Tablet- and HMD-Based Hybrid Virtual Environments. In *Proceedings of the ACM Symposium on Spatial User Interaction*. ACM, New York, NY, USA, 70–79. <https://doi.org/10.1145/2659766.2659777>
- [95] Jia Wang and Robert Lindeman. 2015. Coordinated Hybrid Virtual Environments: Seamless Interaction Contexts for Effective Virtual Reality. *Computers & Graphics* 48 (May 2015), 71–83. <https://doi.org/10.1016/j.cag.2015.02.007>
- [96] Yun Wang, Xiaojuan Ma, Qiong Luo, and Huamin Qu. 2016. Data Edification: Representing Data with Food. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 409–422. <https://doi.org/10.1145/2851581.2892570>
- [97] Jonathan Wieland, Johannes Zagermann, Jens Müller, and Harald Reiterer. 2021. Separation, Composition, or Hybrid? – Comparing Collaborative 3D Object Manipulation Techniques for Handheld Augmented Reality. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, Bari, Italy, 403–412. <https://doi.org/10.1109/ISMAR52148.2021.00057>
- [98] Wanze Xie, Yining Liang, Janet Johnson, Andrea Mower, Samuel Burns, Colleen Chelini, Paul D'Alessandro, Nadir Weibel, and Jürgen P. Schulze. 2020. Interactive Multi-User 3D Visual Analytics in Augmented Reality. *Electronic Imaging* 2020, 13 (Jan. 2020), 363–363. <https://doi.org/10.2352/ISSN.2470->



- 1173.2020.13.ERVR-363
- [99] Yan Yan, Ke Chen, Yu Xie, Yiming Song, and Yonghong Liu. 2019. The Effects of Weight on Comfort of Virtual Reality Devices. In *Advances in Ergonomics in Design*, Francisco Rebelo and Marcelo M. Soares (Eds.), Vol. 777. Springer International Publishing, Cham, 239–248. [https://doi.org/10.1007/978-3-319-94706-8\\_27](https://doi.org/10.1007/978-3-319-94706-8_27)
- [100] Fengyuan Zhu and Tovi Grossman. 2020. BISHARE: Exploring Bidirectional Interactions Between Smartphones and Head-Mounted Augmented Reality. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–14. <https://doi.org/10.1145/3313831.3376233>