Collaborating Across Realities: Analytical Lenses for Understanding Dyadic Collaboration in Transitional Interfaces

Jan-Henrik Schröder
University of Lübeck
Germany

Daniel Schacht
University of Lübeck
Germany

Niklas Peper
University of Lübeck
Germany

Anita Marie Hamurculu
University of Lübeck
Germany

Hans-Christian Jetter
University of Lübeck
Germany

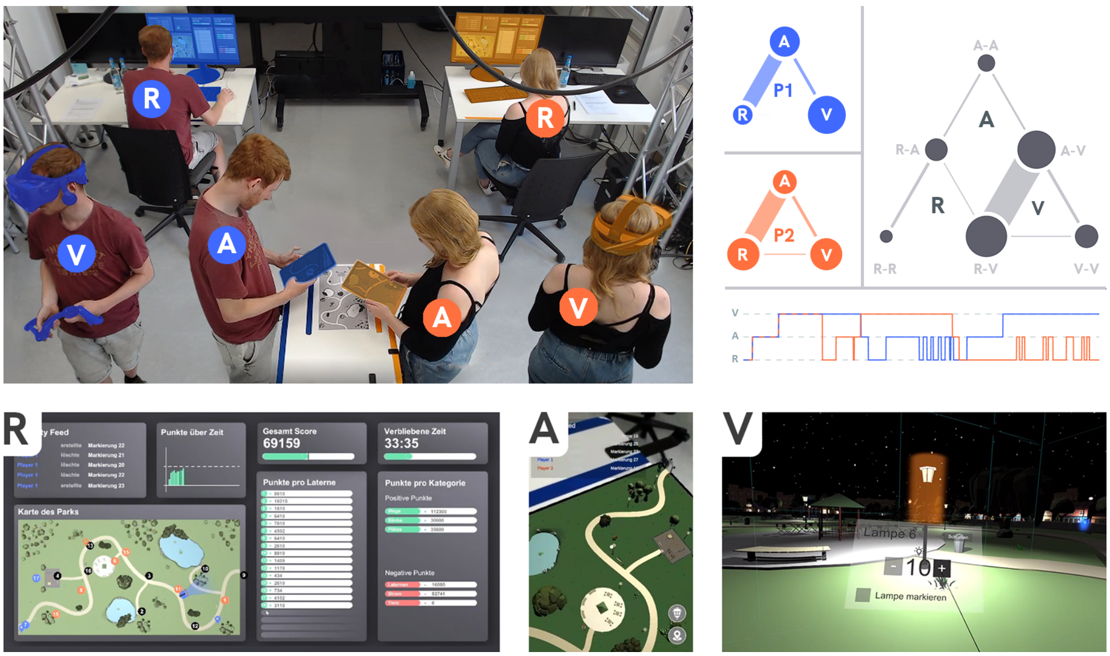

Figure 1: Top left: Collaborative, dyadic use of a TI enabling arbitrary transitions between combinations of the contexts *R* (Desktop PC), *A* (tablet-based AR), and *V* (VR HMD). Bottom: Users' views of *R*, *A*, and *V*. Top right: Analytical lenses for visual analysis of collaboration behavior showing individual, group, and temporal usage of transitions between contexts.

ABSTRACT

Transitional Interfaces are a yet underexplored, emerging class of cross-reality user interfaces that enable users to freely move along the reality-virtuality continuum during collaboration. To analyze and understand how such collaboration unfolds, we propose four *analytical lenses* derived from an exploratory study of transitional collaboration with 15 dyads. While solving a complex spatial optimization task, participants could freely switch between three contexts, each with different displays (desktop screens, tablet-based augmented reality, head-mounted virtual reality), input techniques (mouse, touch, handheld controllers), and visual representations (monoscopic and allocentric 2D/3D maps, stereoscopic egocentric views). Using the rich qualitative and quantitative data from our study, we evaluated participants' perceptions of transitional collaboration and identified commonalities and differences between dyads. We then derived four lenses including metrics and visualizations to analyze key aspects of transitional collaboration: (1) place and distance, (2) temporal patterns, (3) group use of contexts, (4) individual use of contexts.

CHI '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9421-5/23/04…$15.00
https://doi.org/10.1145/3544548.3580879

CCS CONCEPTS

• **Human-centered computing** → **Computer supported cooperative work**; **Collaborative and social computing theory, concepts and paradigms**; **Mixed / augmented reality**; **Virtual reality**; **Collaborative interaction**.

KEYWORDS

transitional interfaces, transitional collaboration, user study, analytical lenses

ACM Reference Format:

Jan-Henrik Schröder, Daniel Schacht, Niklas Peper, Anita Marie Hamurculu, and Hans-Christian Jetter. 2023. Collaborating Across Realities: Analytical Lenses for Understanding Dyadic Collaboration in Transitional Interfaces. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3544548.3580879

1 INTRODUCTION

Transitional Interfaces (TIs) [2] are a yet underexplored, emerging class of mixed reality (MR) or cross reality (XR) user interfaces [17] that enable users to freely move along the *reality-virtuality continuum* (RVC) of Milgram and Kishino [23] during their work. Thereby, TIs promise to bridge the gap between (1.) interacting with objects and computers that are situated *in reality*, e.g., using PCs or mobile/large touch screens, and (2.) interacting *inside of augmented and virtual realities* (AR/VR), e.g., using head-mounted displays (HMDs) [17]. At any time, TIs enable users to move freely along the RVC for individually choosing and switching between those displays, input/output modalities, and representations of data or functionality that are preferred for the users' task at hand [7, 17].

TIs predate and are foundational for more recent concepts such as cross reality [22, 36] or cross virtuality interfaces [7]. Notably, from their very beginning with Billinghurst et al.'s Magic Book [2] in the early 2000s, TIs have always emphasized the importance of collaboration between TI users (henceforth *transitional collaboration*) and the difficult challenge of maintaining awareness and coherence when collaborators transition to different locations on the RVC [10–12]. For example, Figure 1 shows our TI prototype that enables users to individually move between the three contexts *R* (Desktop PC), *A* (tablet-based AR), and *V* (VR HMD) during collaboration, allowing for 3×3 possible context combinations of which six are each located at two different locations on the RVC. Nonetheless, except the very first generation of TI research, there have been no further attempts to design, implement, and evaluate novel visualization or interaction techniques for collaborative TIs to address this challenge of frequent transitions between contexts. To this day, there exist no generalizable design guidelines or best practices for designing collaborative TIs. There is also not enough quantitative and qualitative empirical data available from user studies of actual collaboration in TIs. In particular, there are yet no analytical frameworks and methods to thoroughly analyze and interpret such data from transitional collaboration. Our paper addresses this with two *research contributions*.

Contribution 1: In an exploratory study in our lab, we collected rich qualitative and quantitative data from 15 dyads who collaborated to solve a complex spatial optimization task. Participants could freely switch between three contexts including displays (desktop screens, tablet-based augmented reality, and head-mounted virtual reality), input techniques (mouse, touch, handheld controllers), and visual representations (allocentric maps vs. egocentric views, monoscopic vs. stereoscopic 3D). Our first contribution is our TI prototype and study design as a blueprint for future research on transitional collaboration, and we share insights about participants' strategies and perceptions of collaboration as well as observations about characteristic commonalities and differences of collaboration between dyads.

Contribution 2: Based on that data, we derive four *analytical lenses* as our main contribution. These four lenses comprise metrics and analytical visualizations that TI researchers can apply for future user studies to thoroughly analyze, interpret, and understand different key aspects of transitional collaboration: (1.) place and distance, (2.) temporal patterns, (3.) group use of contexts, (4.) individual use of contexts. Figure 1 (top right) gives a first visual impression of how these lenses and their visualizations represent the observed transitional collaboration.

In the following, we first discuss previous and related work to identify gaps in TI research. We then describe our user study, prototype, and participants. This is followed by sharing details about our data analysis and the general nature and results of the observed collaboration. We then derive the above-mentioned four lenses as our main contribution and demonstrate their applicability by analyzing our collected data. We conclude by discussing limitations and future work.

2 RELATED WORK

Our discussion of previous and related work focuses on three categories of work: (1.) core TI research that is foundational for TIs, introduces important terminology, or has studied transitional collaboration; (2.) work that does not explicitly mention TIs but uses closely related concepts from MR or XR research for collaborative *and* single-user interfaces; (3.) descriptive and analytical frameworks from computer-supported cooperative work (CSCW) that have influenced our data analysis and proposed analytical lenses.

2.1 Core Research on Transitional Interfaces

TIs were first conceived of, defined, and demonstrated by Billinghurst et al. in 2001 [2]. Their system Magic Book already contained animated transitions and user representations across AR and VR to enable collaborative exploration using 3D storytelling content as an example. After Magic Book, however, research on TIs lay largely dormant until recently with notable exceptions by Grasset et al. from 2005 to 2008 [10–12] and Carvalho et al. in 2012 [4].

Grasset et al.'s concept and framework for TIs from 2006 introduces helpful definitions and clear terminology [12]: A *context* is an environment in which users collaborate and interact that is defined by its position on the RV continuum (e.g., AR, VR, reality) but also by properties such as scale (e.g., macro, micro, nano in relation to the data space or virtual environment), representation (e.g., photorealistic, non-photorealistic, symbolic), or any other user parameters such as navigation mode (e.g., natural walking, teleporting). When users are in the same context (i.e., *co-context*), but would not see each other because they wear HMDs, they see each other as virtual *proximal embodiments*. When they are separated by different contexts (i.e., *cross-context*), they become visible in the other context using virtual *distal embodiments*. Henceforth, we adopt Grasset et al.'s model and terminology.

Grasset et al. also conducted TI user studies [10, 11], but they were either single-user studies [10] or focused on comparably simple maze navigation tasks with static role assignments [11] that did not involve more complex collaboration or coordination, e.g., typical mixed-focus collaboration [37]. Similarly, Carvalho et al. proposed some initial design guidelines for TIs but only for single-user interfaces [4]. Only recently, the advances in MR technologies such as novel AR/VR HMDs have renewed research interest in collaborative TIs, for example (1.) in their potential for converging AR and VR technology into future MR collaboration tools [5], (2.) in their potential for collaborative cross-virtuality data visualization [7, 8, 32], and (3.) for better integration of AR/VR in existing workplaces [17]. This recent and renewed interest in TIs has initiated a wave of new and promising research activities related to TI, and we look forward to seeing their results published in the near future.

2.2 Related Work on Cross, Augmented, and Mixed Reality (XR, AR, MR)

A user's possibility to transition between different locations on the RVC is considered by some as the defining characteristic of an XR interface. For example, Simeone et al. define XR interfaces as systems that enable a smooth transition and/or collaboration of users between different degrees of virtuality [36] and Maurer et al. follow this definition [22]. There seems to be no clear line that separates these newer concepts such as XR [22, 36] or *cross-virtuality* [7] from the original concept of TIs [2]. Therefore, we consider these concepts closely related and also feel that these newer concepts could often benefit from adopting the elaborated conceptual frameworks and terminology for TIs (e.g., [12]). Also, some systems that are usually considered MR or XR behave very similarly to a single-user TI. For example, the discontinued AR/VR HMD by SULON was advertised with a demo experience including an impressive transition between AR and VR to increase players' immersion[1]. Similarly, AR systems such as Reipschläger et al.'s share many key qualities with TIs and XR by augmenting interactive surfaces (e.g., screens for touch and pen input) with AR content to combine the benefits of monoscopic and stereoscopic content and interactions [29–31].

MR research becomes particularly relevant for TIs in the context of asymmetric collaboration between desktop PCs, AR, or VR. In current MR research, users are typically assigned a fixed position in the RVC that is different from that of their collaboration partner (e.g., asymmetric PC-AR, PC-VR, AR-VR collaboration). Unlike in a TI, this fixed assignment cannot be changed by users during runtime. A review by Ens et al. [5] identified 63 papers about such asymmetric collaboration in MR and a literature survey by Fröhler et al. lists nine papers alone in the context of asymmetric MR collaboration for visual analytics [7]. Some of these have directly or indirectly informed the design of our prototype and study. For example, Piumsonboom et al. and Bai et al. combined AR, VR, and natural communication cues such as gestures, head, and eye gaze to create new types of cross-context AR-VR collaboration for spatial tasks [1, 25]. Grandi et al. used handheld tablet-based AR with VR HMDs and compared users' performance for symmetrical and asymmetrical setups (VR-AR, VR-VR, AR-AR) during basic 3D manipulation tasks [9]. Handheld displays were also used by Gugenheimer et al.'s ShareVR to demonstrate new asymmetric ways of sharing a VR experience between a user wearing an HMD and a non-HMD user next to them [13]. The system that is probably closest to our understanding of collaboration in TI —but without explicitly referring to TIs as its conceptual basis— is Roo and Hachet's One Reality [33]. It allows one or more users to use and transition between multiple mixed reality modalities while interacting with augmented artifacts, e.g., moving from spatial augmentation of a tangible physical model to AR see-through displays and VR HMDs. However, Roo and Hachet have not used the system to conduct single-user or collaborative user studies.

[1]SULON Magic Beans Demo (2016). https://www.youtube.com/watch?v=pp90zGjydwI (last accessed Sep 2022)

2.3 Related Work on Descriptive Frameworks and Models in CSCW

There is a wealth of relevant research on descriptive frameworks and models for co-located, remote, and hybrid collaboration in computer-supported cooperative work (CSCW) that can support TI research. After the pioneering formalization of groupware by Johansen's *Time-Space Matrix* [18], many different and more detailed frameworks were formulated to characterize new collaborative technologies such as interactive tabletops for around-the-table collaboration [16, 34, 37], co-located cross-device collaboration [3, 26], networked interactive surfaces for remote collaboration [38], or hybrid collaboration in partially-distributed teams [24]. Therefore, it seems consequential to build upon this work and extend these models and frameworks to create new analytical lenses that can describe and analyze the specifics of TIs and transitional collaboration.

Given the aforementioned challenges of TI design, we considered a variety of key concepts in our study design and data analysis. Our priority was to provide sufficient *workspace awareness* [14] as well as *verbal communication* and *shared audio territories* that are crucial for efficient hybrid collaboration [24]. Furthermore, *natural communication cues* such as gaze and gestures can support this [1] and were integrated. The overall goal of the collaboration is usually achieved by dividing it into a series of phases of *mixed-focus collaboration* with more *tightly-coupled* or *loosely-coupled* collaboration styles [16, 37]. Our analytical lenses refer to these concepts but, unlike Isenberg et al. [16], Brudy et al. [3], and Neumayr et al. [24], we have not analyzed the data down to the level of single *coupling styles.* Here, also *territoriality* [34] and *proxemics* with *F-formations* [21] could play a key role in better understanding why and how users share their space and position themselves in relation to other users or objects. Therefore, coupling styles and proxemics are both promising directions for future data analysis of transitional collaboration and will be part of our future work.

3 USER STUDY

In our exploratory study, we observed and analyzed emergent participant behaviors and collaborative activities of 15 dyads (henceforth D1-D15) whom we invited into our lab to work on a time-capped spatial optimization task. During the task, we recorded, tracked, and logged rich data from different sources such as audio/video/screen recordings, virtual and physical positions of active devices and users, application logs, and the state of the virtual workspace. Our goal for the study was to observe user behavior during a demanding collaborative activity and identify recurring patterns and typical characteristics of the transitional collaboration that would unfold. We then used these observations to (1.) identify *characteristic commonalities and differences* between the observed collaborations (e.g., differences in the patterns of using contexts over time, the spatial distribution of participants in the virtual space, their preferences for cross-context or co-context work) and, from this, (2.) derive and propose meaningful analytical lenses as methodological contributions and tools for researchers through which they can view transitional collaboration in future.

We would like to emphasize that our study did not aim at generalizable findings about the "true nature" of transitional collaboration since that nature is strongly dependent on a multitude of factors such as the domain, tasks, users' training, users' relationships, users' spatial abilities, and the specifics of the system's design. Hence it is impossible to establish from a single study. Instead, our goal was to provoke and observe an ideally wide enough range of behaviors, usage patterns, and different forms of collaboration in TIs to derive and propose new analytical methods and visualization techniques (i.e., our "analytical lenses") that can capture the characteristic commonalities and differences in dyadic collaboration in TIs and can be applied for their detailed analysis in future studies.

3.1 Task

To provide our participants with a sufficiently complex and motivating task, we defined a collaborative spatial optimization problem and implemented it into a prototype in a game-like environment. The participants had to design the illumination of a fictitious, urban park. The park consists of various winding paths, different plazas (e.g., a playground and a pavilion), and several benches in addition to ponds, trees, and meadows (Figure 2).

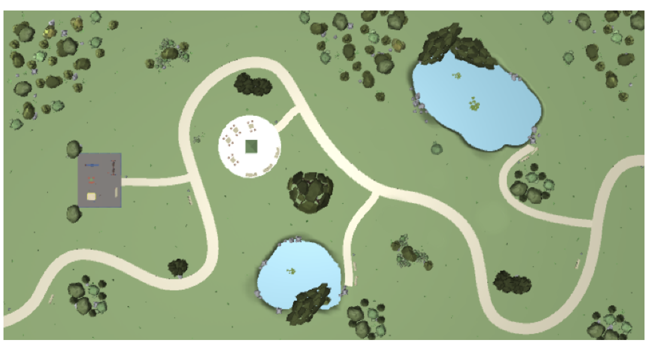

Figure 2: A top view of the park with paths, a playground (gray, left), a seating area around a pavilion (white, center), and two ponds.

Participants could place a limited number of street lamps inside the park and adjust each lamp's luminance level. Points were awarded for increasing the safety and security of each path, plaza, or park bench at night by sufficiently illuminating them with a lamp. Points were deducted for each new lamp, each lamp's power consumption, and for light pollution disturbing nocturnal animals. Participants could discover these animals when walking through the park so that they did not only have to repeatedly optimize the positions and brightness of lamps but also to navigate the park. This increased the spatial complexity of the collaboration and thus also the suitability of TIs for the task.

Based on informal experimentation, we limited the maximum number of lamps to 20 to make users optimize them more carefully instead of letting them randomly place large numbers all over the park. This felt like a good compromise between user freedom and the complexity of optimizing many lamps. Furthermore, a time limit ensured better comparability of the dyads' usage patterns, behaviors, and overall task performance. In pre-studies, we established that 45 minutes seemed sufficient for getting familiar with the task, initial experimentation, and solving the task with an iterative refinement of lamps' positions and luminance without being too repetitive.

3.2 Prototype

For our task, we created a TI with three different contexts between which participants could freely switch at any time (Figure 3).

3.2.1 *Context "Reality" (R).*

This context was provided by a desktop PC and a screen showing a dashboard with an overview of the current state of the park and scores. Participants could view a top-down mini-map of the park which contained the lamps they placed and the location of the other team member. In addition to the mini-map, participants were presented with score figures, timeline plots, and bar charts containing their score over time, score per category (e.g., losses due to light pollution), and score per lamp. Furthermore, they could see their remaining time. By using a mouse on the map, participants could also set pins to mark locations, highlight lamps, or hover over them to show how many points they contributed.

3.2.2 *Context "Augmented Reality" (A).*

This context contained a hologram-like 3D model of the current state of the park that participants could access using a tablet as a see-through AR device. The miniature park appeared as a 3D augmentation above and around a central AR marker on a tabletop. As an AR marker, we used the minimap from *R* and printed it in grayscale on a sheet of paper (size A3, 297 × 420 mm). Participants could explore the park at different scales and from different directions by moving the tablet closer or further away from the marker and by moving it around the marker to approach from different sides. Within this freely viewable 3D model of the park, participants could add and move lamps by using a gaze cursor in the center of the tablet and touching buttons on the tablet's screen.

3.2.3 *Context "Virtual Reality" (V).*

This context used a VR HMD to let participants immerse themselves into a stereoscopic VR visualization of the park at night. It provided a first-person experience of the illumination of the park. Participants could use handheld controllers to teleport inside the full-size representation of the park and use actual walking for locomotion. Buttons on the controllers could be used to adjust the brightness of lamps with a floating 3D pop-up menu. Only in this context, the luminous radius around each lamp became visible. Similarly, nocturnal animals became visible in this context while they were barely visible in context *A* and entirely invisible in context *R*.

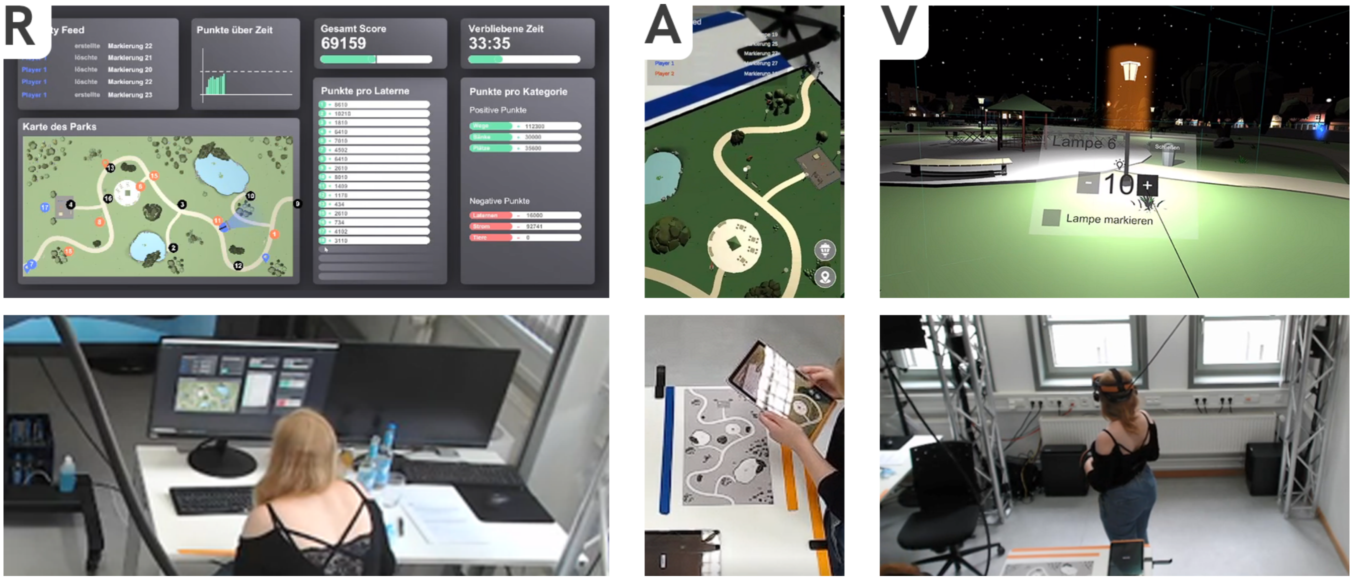

Figure 3: Prototype shown as screenshots (top) and physical setups (bottom) for each of the three contexts *R*, *A*, *V*.

3.2.4 Distribution of Information and Functionalities. It is noteworthy that our distribution of information and functionality between contexts was not chosen for the best possible usability or support of the task. Since the focus of our study was on observing and understanding how participants collaborate across contexts in a TI, we chose a design that motivated participants to use transitions between different contexts by distributing information and tools across contexts. While this at times might have created more cumbersome interactions, we could observe a much greater variety of collaborative and transitional behaviors this way.

3.2.5 Awareness Cues. For enabling participants to establish sufficient group awareness, we referred to Gutwin and Greenberg's descriptive framework for workspace awareness [14] to design consistent awareness cues. An overview of how we integrated these awareness cues into the prototype is presented in Table 1.

First, to address workspace awareness elements in the category *Where* [14], each context always contained an embodiment of the other participant at that participant's current position (Figure 4). How the other participant appeared depended not only on the current context of the observing participant but also on the context of the other participant. If the other participant was in *V*, they were represented to their observer as a ludo game figure with an HMD and a visualization of their gaze (Figure 4). If the other participant was in *A*, they were represented at the position of their gaze cursor using a glowing semi-sphere on the ground (Figure 6). If the other participant was in *R*, they were only indirectly represented by their interactions with the map and scores, i.e., by highlighting the lamps they brushed or selected using their mouse as well as showing the pins that they placed on the map to mark locations.

To support awareness of *What* [14], all edited or brushed artifacts (i.e., lamps) were also visually highlighted. In *V*, additional spatial audio was used to indicate specific actions by playing earcons from the location where they happened. In addition, there was an Activity Feed in each context that provided information about the team's latest actions. To enable some way of communicating locations, participants were also able to set an arbitrary number of pins that appeared in all contexts to mark either a location of their choice or the lamps themselves.

The category *Who* of workspace awareness [14] was irrelevant because participants worked in dyads and thus could easily differentiate themselves from the other team member. However, to address *Authorship* and *Identity*, we assigned participants a personal color (either blue or orange) for differentiating between their own artifacts or actions and those of others (e.g., highlights, pins, embodiments, etc.).

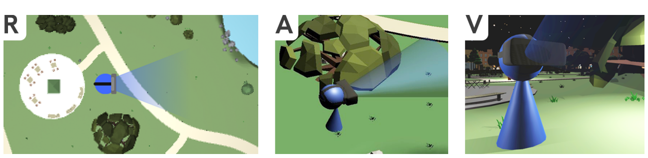

Figure 4: Screenshots of how the other team member was represented in *R*, *A*, and *V*. Here the other team member is in *V*.

3.3 Apparatus

To ensure that the setup of our study did not introduce differences in how easily either team member could access and use a device or context, our setup consisted of two exactly mirrored subsetups, one for each participant (Figure 5).

Each subsetup consisted of a tracking space for the VR HMD (*V*) (size 3×2m), a desk for the desktop PC and screen (*R*), and a shared table with the AR marker in the middle (*A*). For easier access, the VR HMDs were hung on a hook attached to this table. We chose a

Table 1: Overview of our awareness cues corresponding to the elements of workspace awareness by Gutwin and Greenberg [14].

Category	Element	Awareness cues in our prototype
Who	Presence	-
	Identity	Personal color of the embodiment
	Authorship	Colors of pins corresponding to the personal color of its creator
What	Action	Activity feed & spatial sounds in *V*
	Intention	Marking locations and lamps & highlighting hovered or brushed artifacts
	Artifact	Highlighting of currently edited lamps
Where	Location	Appearance (i.e., which context) and position (i.e., virtual location) of the embodiment
	Gaze	Visualized gaze cone of the embodiment of users in *V*
	View	Visualized gaze cursor of the embodiment of users in *A*
	Reach	-

desktop 27-inch monitor with a mouse for *R*, a Samsung 10.5-inch tablet for *A*, and Valve Index VR HMDs and controllers for *V*.

Since participants were co-located in a single room, they were able to talk and listen to each other in a natural manner. However, apart from this shared auditory space, we strictly ensured that there were no collaborative activities that would have been impossible in a non-co-located, remote scenario, i.e., sharing of devices (e.g., dyads looking at the same physical screen in *R*), deictic gestures (e.g., dyads discussing and pointing with their fingers at the physical map in *A*), or handing over devices (e.g., passing a VR-HMD to the other participant after using *V*). Also, participants were always restricted in their movement to their own subsetup and were not allowed to leave their space during the session. To clearly distinguish and communicate ownership of devices and spaces, we used colored duct tape with the personal color from the prototype.

Figure 5: Setup of the study in the lab.

3.4 Participants

We recruited 34 participants by sending invitations to mailing lists and online forums linked to a German university. Each participant was compensated for their two hours of participation with 25 Euro (above German legal minimum wage). Of the participants, 13 identified as female, 20 as male, and 1 as diverse. The mean age of our participants was 23.2 (*SD*=2.9, *min*=18, *max*=34). Our participants had various professional and educational backgrounds. Twelve participants had received or were receiving higher education in computer science, eight in psychology, three in medicine, two in mathematics, and one each in cultural studies, engineering, or legal studies. As their highest degree, three participants had a postgraduate degree, 11 had an undergraduate degree, and 18 were enrolled as undergraduate students. Two participants were attending high school.

To assess if our participants had an unusually positive or negative relation to technology, we asked each participant to complete the Affinity for Technology Interaction (ATI) scale developed by Franke et al. [6]. With M=3.93 on a scale from 1.0 to 6.0 (SD=1.06, min=2, max=5.9), our sample had a higher mean compared to the quota sample provided by Franke et al. [6] (M_q=3.61) that approximates the distribution of age, gender, and education level of the general population of Germany. However, based on a one-sample t-test, there was no significant difference to that quota sample ($t(33)$=1.746, p=.09) confirming the validity of the sample for a German (and thus only WEIRD [20]) user population.

Dyads D1-D17 were formed by the participants themselves, i.e., they signed up in pairs during the recruitment process and thus knew each other from before the study. The degree of familiarity among dyads ranged from work colleagues or fellow students to a pair of twin brothers (D12) living together.

Due to technical problems during the study, D16 and D17 had to restart the task and were unable to work on it for 45 minutes. Accordingly, their data were excluded from further analysis.

3.5 Procedure

All procedures were in line with European Union, federal, state, and institutional regulations for COVID-19 social distancing and data protection. After arriving in the lab, all participants received a brief introduction to the study, hygiene regulations, and gave informed consent. After that, the recording of the session started.

Both participants were given a standardized introduction to explain the task. Following a short demonstration of the prototype's functionality (up to 10 minutes), all participants were asked to sit down at *R* as the common starting point for their task. As discussed earlier, participants then had 45 minutes to solve the task. During the task, the two experimenters did not provide any hints about solutions to the task itself and participants could only ask for technical support, for example in those cases in which devices or client apps needed to be restarted.

After completion of the task, each participant was asked to sit at the desktop PC to complete a questionnaire about their session. While this questionnaire was designed to capture the individual

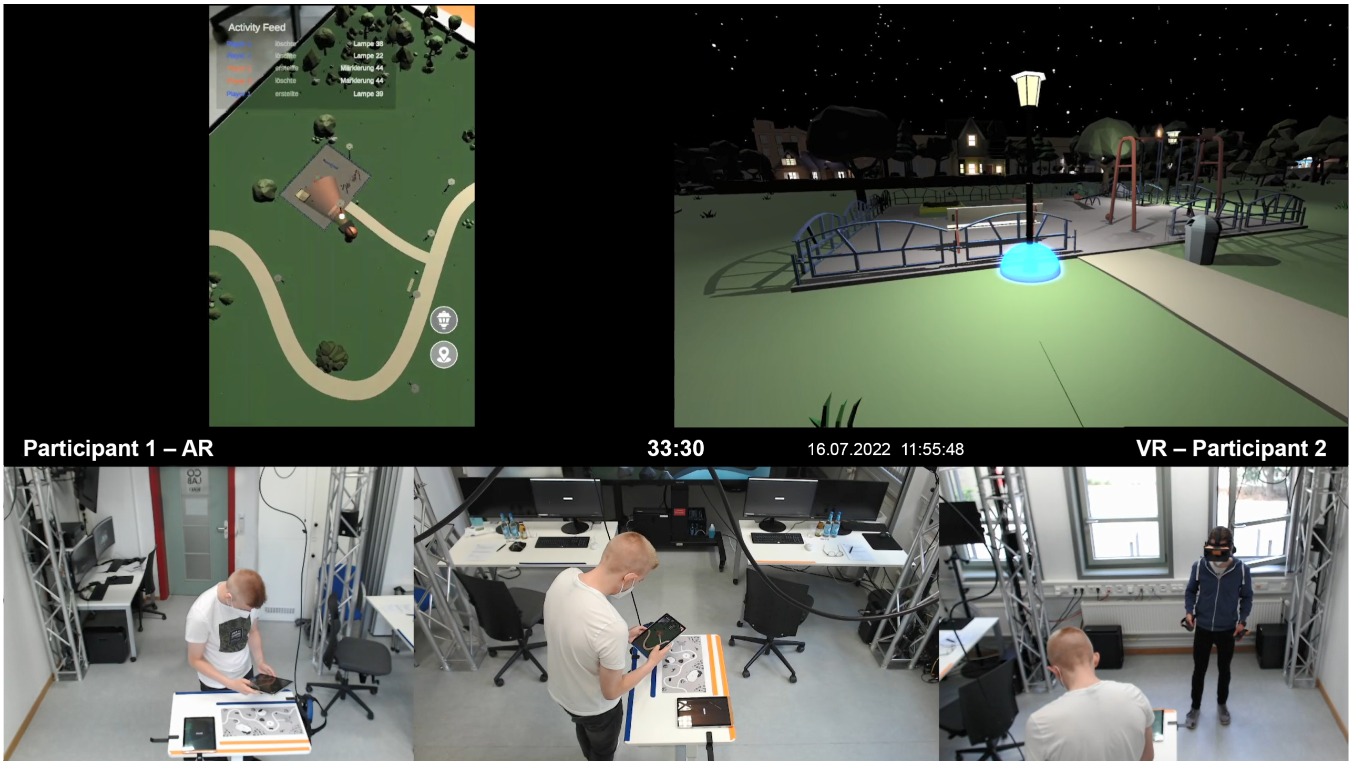

Figure 6: Video recording from D8 with screen captures of both current contexts (top) and three cameras recording participants in the physical lab (bottom). (This also shows how the gaze cursor of P1 in A is represented as a glowing semi-sphere in the view of P2 in V).

experience of each participant, we also captured the shared experience of the team during a subsequent semi-structured interview. At the beginning of these interviews, the lead experimenter asked a set of pre-defined, team-independent questions. These questions were also mixed with team-specific questions that were based on the lead experimenter's own observations during the session and on team-specific suggestions for questions from the other experimenter. This semi-structured interview was aimed at clarifying and consolidating our first impressions and interpretations during the observation by asking participants about their perceptions and intentions. To conclude the session, the lead experimenter debriefed the team.

4 DATA COLLECTION, VALIDATION, AND ANALYSIS

Our results are based on a three-step process of data collection, data validation, and data analysis. To ensure the credibility and validity of our findings during the process, we used multiple parallel data sources for triangulation.

4.1 Data Collection

Our goal was to identify as many relevant patterns or behaviors as possible, e.g., collaborative usage of contexts, the spatial distribution of the participants in the virtual space, and strategies for creating awareness. To truly understand the causes and users' motivations behind them, we used rich data sets collected from multiple sources including both qualitative and quantitative data.

4.1.1 Qualitative Data. All sessions and interviews were recorded by three cameras and a microphone to capture users' movements, actions, and utterances in the physical lab. In addition, for each participant, the screen content of the currently used device was streamed to a centralized capture PC that could be remotely controlled using network messages. The resulting synchronized video recording used a split screen with two sections (Figure 6). In the upper half, the screen content of each participant's currently used device was shown in adjacent streams. The bottom half contained the three adjacent camera images showing participants' actions in the physical lab. The sections were separated with an information strip that displayed the participants' current contexts, as well as the date and elapsed time of the session.

Additionally, the lead experimenter kept a handwritten log of their observations about the collaboration and user behavior of the current dyad. This was later used in the semi-structured interview to formulate team-specific questions related to the dyad's reasons for how they approached their task and collaboration.

4.1.2 Quantitative Data. As a basis for quantitative analysis, we continuously logged participants' positions in the virtual park, their current contexts, and their scores on a second-by-second basis. Furthermore, relevant events such as participants marking locations, changing contexts, placing lamps, or changing brightness were recorded with their timestamps in an extensive event log.

The post-study questionnaires contained several scales to measure different constructs such as individual workload through the *NASA-TLX* [15], the workload in the team through the *Team Workload Questionnaire* [35], workspace awareness through a custom questionnaire, and some information on demographics to collect further data on their perceived collaboration. The custom questionnaire about workspace awareness used the descriptive questions from Gutwin and Greenberg's framework [14] as input for formulating six-point Likert-scaled items, e.g., their question on *location* (*'Where are they working?'*) was turned into *'I knew at all times where my partner was working.'*. While this custom questionnaire is of course not validated, we used it as a rough instrument to detect if there were major breakdowns in awareness.

4.2 Data Validation

Before starting with an in-depth analysis to derive our analytical lenses, we first used objective user performance and subjective user ratings to validate the plausibility of our data. More specifically, we used that data to assess how successful, demanding, and thus also realistic the transitional collaboration of the 15 observed dyads was. We here report descriptive statistics about resulting scores, frequency of transitions, perceived workload, and perceived workspace awareness. By this, we also establish a general understanding of the nature of collaboration to confirm that the observed collaborative practices were meaningful enough to be used to later derive our analytical lenses.

4.2.1 Scores and Transitions. Based on informal experimentation with our prototype before our study, we identified a score of 75,000 points as "adequate" and 175,000 points as "excellent". All dyads achieved "adequate" scores or above, including an almost "excellent" score by D11, confirming their sufficient understanding of the task and their commitment. To verify that our prototype provoked transitions as intended (see 3.2), we plotted D1-D15's resulting scores (M=121,793.9, SD=25,284.67, min=88,524, max=174,119) and their number of transitions (M=59.6, SD=46.38, min=14, max=170) in Figure 7. Spearman's rank correlation showed a positive correlation between the total score and the number of transitions ($r(13)$=.74, p=.002). Therefore, it can be concluded that our task and prototype created a setting in which transitions were beneficial and thus met expectations towards a sensible and realistic task for a TI.

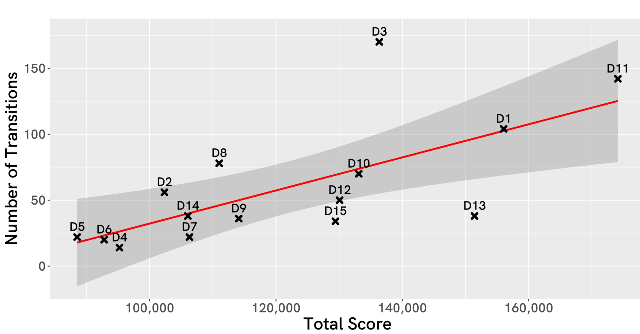

Figure 7: Correlation of total score and number of transitions.

4.2.2 Individual & Team Workload. To check if our data suffers from floor or ceiling effects in terms of perceived task difficulty and task load, we analyzed the data from the NASA-TLX. Figure 8 shows the results of our raw NASA-TLX with six 11-point Likert scales (coded as 0 to 10) where 0 means a generally desirable result (e.g., very low demand, perfect performance) and 10 a generally undesirable result (e.g., very high demand, failure performance).

Generally, participants reported moderate individual workloads with mean values outside the top and bottom quarters of the scales. From this, we conclude that the task and prototype were not clearly over- or under-challenging. However, *mental demand* and *effort* stood out with comparably high means of M=6.4 (SD=1.5) and M=5.6 (SD=1.8). This may be due to the inherent complexity of the task, but maybe also due to the nature of TIs in general, as participants need to familiarize themselves with different devices and visual representations and switch between them. Potentially, mental demand and effort could be reduced with a more elaborate design of animated transitions between contexts to increase visual and spatial coherence as already suggested by Grasset et al. [10] and recently explored by Pointecker et al. [27, 28].

To our own surprise, the *physical demand* scored comparably low (M=3.1, SD=2.3), even though our apparatus required participants to move to different devices at different physical locations including (un)wearing VR HMDs. Contrary to our expectations, this physical activity was not perceived negatively. Instead, our interviews revealed that participants perceived walking around in the lab and changing devices as desirable to counter fatigue and hence felt aroused by these changes (D2, D6, D7, D11). Also, the *temporal demand* was moderate (M=4.4, SD=2.5), given the obvious temporal costs for switching between different devices (especially HMDs) and our imposed time limit. In summary, this indicates that even without elaborate designs for minimizing physical and temporal costs of device switching (e.g., using a combined AR/VR HMD instead of a separate tablet and HMD), our TI was not perceived as particularly cumbersome and demanding by individual users. On the contrary, there are some cases of ergonomic, motivational, and attentional advantages associated with device switching.

We also used the team workload questionnaire (*TWLQ*) to collect each participant's assessment of the demands towards the team. The TWLQ consists of three aspects: (1.) the *task workload* using NASA-TLX (Figure 8), (2.) the *team workload* concerned with the

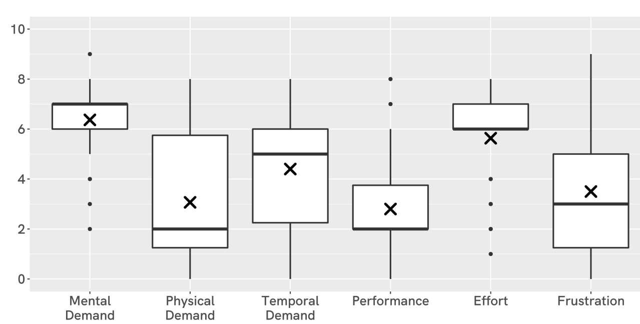

Figure 8: Results of each item of the NASA-TLX.

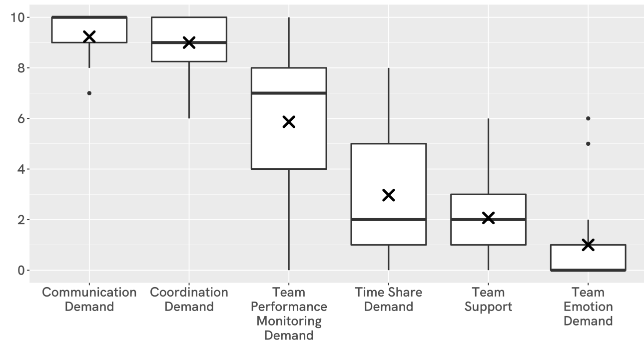

Figure 9: Results of each item of the TWLQ.

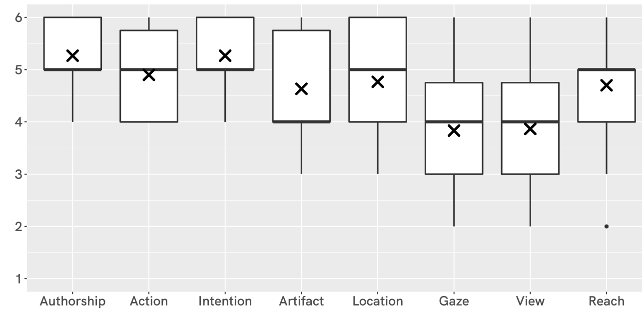

Figure 10: Results of our custom workspace awareness questionnaire.

demands of team interaction (Figure 9, first 3 items), and (3.) *task-team balancing* concerned with the management of taskwork within the team (Figure 9, last 3 items). The scale of each item was the same as NASA-TLX, therefore coding *"low"* as 0 and *"high"* as 10.

The results of the aspect team workload showed high to very high demands for team interaction. Especially the *communication* (*M*=9.2, *SD*=1.0) and *coordination demand* (*M*=9.0, *SD*=1.0) was very high. This confirms our observation that all dyads (D1-D15) had to actively work to establish sufficient group awareness for a coordinated collaboration, primarily relying on intense verbal communication. In comparison, task-team balancing played only a subordinate role since teams consisted of only two participants and the task was limited to 45 minutes. This made questions of the fair division of labor and mutual support appear less important. In summary, our results show that the unfolding teamwork put a substantial load on participants, similar to how we would expect it for a real-world collaboration.

4.2.3 Awareness. Generally, participants reported they could establish sufficient group awareness for meaningful and successful collaboration. As discussed above, they relied heavily on verbal communication to create awareness and exchange information about their intentions, actions, goals, or locations. Participants also used verbal communication to enrich the given awareness cues within our prototype. For example, some dyads used the two user-specific colors of markers purposefully to let them symbolize different meanings (e.g., encountered animals, lamps needing optimization, locations for new lamps, etc.) (D1, D2, D6, D7, D10, D11, D14, D15). Beyond that, participants also used other system features in unanticipated ways to create awareness. For example, the pins on the map were used not only to mark a static location (D1-D15) but also to draw temporary visual paths on the map in *R* that appeared as visible waypoints in the view of the other participant in *V* to support their egocentric navigation to a specific destination (D8, D14). Some dyads also used the pins for getting their partner's attention to a specific location by continuously marking and unmarking a location to achieve a "blinking" pin (D6, D9, D10, D15). One dyad used the pins to mark the size of the luminous radius of a lamp which was only visible in *V* to show it to their partner in *R* (D11).

Thanks to that creativity in using the integrated awareness cues, there were no reports of major breakdowns for any element of workspace awareness according to our custom workspace awareness questionnaire (Figure 10). The lowest scores were *gaze* (*M*=3.8, *SD*=1.2) and *view* (*M*=3.9, *SD*=1.3). Although participants stated in the interview, that the visualized gaze of a user in *V* was very useful (D2, D4-D6, D9-D12, D13, D14), we believe that *gaze* and *view* scores suffered from not being consistently available for all contexts. For example, they could not be detected by users in *R* since there was no integrated head or eye tracking in front of the screen. Nonetheless, these items still scored in the upper half of the scale. As discussed above, elements like *action* (*M*=4.9, *SD*=0.8), *intention* (*M*=5.3, *SD*=0.7), and *location* (*M*=4.8, *SD*=1.0) had relatively high scores, very likely also benefiting from additional verbal communication.

In conclusion, participants could establish sufficient group awareness to collaborate successfully. This is noteworthy since our implemented awareness cues lacked some of the more sophisticated natural communication cues from recent research (e.g., gestures and eye gaze [1]). Still, our comparably simple awareness cues apparently succeeded in addressing the core elements of workspace awareness.

Finding #1: Our results show that our study prototype and tasks afforded sufficiently complex, meaningful, and realistic collaboration in TIs. Participants collaborated successfully and reported sufficient group awareness.
We also found that (1.) workspace awareness can be established in a TI, even in the absence of more sophisticated awareness cues, (2.) comparably simple designs of transitions and awareness cues can be sufficient but potentially increase mental demand, communication demand, and coordination demand, (3.) verbal communication plays a critical role and appears essential for transitional collaboration, (4.) in some cases, users' arousal by and enjoyment of changing contexts could outweigh the physical and temporal demands of switching devices in TIs.

4.3 Data Analysis

We conducted a two-step analysis in order to derive our analytical lenses. In step 1, we identified characteristic commonalities and differences in transitional collaboration between the dyads. In step 2, we used different metrics and visualizations to quantitatively substantiate a selection of four of the observations from step 1. The resulting four analytical lenses are thus thematically defined by step 1 and enriched with analytic tools from step 2.

4.3.1 Analysis Step 1: Identifying Characteristic Commonalities and Differences of Collaboration. In step 1, we used the experimenter's log and reviewed our recordings to identify characteristic commonalities and differences such as recurring behaviors, patterns, or themes of collaboration that we found relevant for further in-depth analysis. We convened for an analysis session at the end of each day of the study to collect and cluster the day's observations on a large interactive whiteboard in affinity diagrams. Thereby, we used handwritten notes, recordings, and interview answers. Each observation was represented by a sticky note containing the observation and dyads. Colors were used to categorize notes. The process of (re)clustering and (re)coloring of notes was repeated each day until all researchers felt that their observations were faithfully represented. A saturation of the clustering scheme was reached after the 11th dyad and further dyads only confirmed existing ones. After the last day of the study, the whiteboard contained 33 sticky notes with four different colors in 11 clusters. From these, four were chosen for further examination in step 2, based on our estimation of their relevance and generalizability for the HCI community: (1.) place and distance, (2.) temporal patterns, (3.) group use of contexts, (4.) individual use of contexts.

4.3.2 Analysis Step 2: Derivation of Analytical Lenses. In step 2, these four clusters became the starting points for deriving our analytical lenses. For each analytical lens, we started with various metrics and visualizations that we believed to represent the observations from the clusters in terms of qualitative and quantitative data. After implementing scripts for data processing, analysis, and visualization, we iteratively refined them in continuous discussion between both experimenters and senior researchers. We stopped iterations as soon as we found that the metrics and/or visualizations quantitatively supported or confirmed our qualitative observations and enabled us to report them concisely.

5 RESULTING ANALYTICAL LENSES

In the following, we present our four lenses along with their descriptive metrics and analytic visualizations. Thereby, we first motivate each lens with qualitative observations from our study and, where applicable, previous and related work. We then exemplify how each lens can be used for analysis by viewing our data through this lens to discuss our particular observations.

5.1 Lens #1: Understanding Place and Distance in Transitional Collaboration

5.1.1 Qualitative Observations: Seemingly basic concepts such as place and distance become non-trivial and ambiguous for transitional collaboration. Early work in CSCW such as Johansen's time-space matrix [18] introduced simple dichotomies between *co-located* vs. *remote collaboration* (i.e., "same place" vs. "different place"). With the more widespread use of AR/VR for collaboration, "same place" and "different place" have become more complex and ambiguous [7]. For example, two VR users can be perceptually separated by being immersed at opposite sides of a large virtual environment while still being physically co-located within a range of a few meters. On the other hand, users can perceive each other as if they were virtually co-located in the same room while being at different geographical locations with a physical distance of thousands of kilometers between them [7]. In a TI, the notions of "same place" and "different place" can become even more complex as place is now defined simultaneously by users' physical location, users' location in the virtual workspace, and the contexts in which the users are.

We repeatedly found evidence for this non-trivial relation between physical distance (in the lab), virtual distance (in the park), context (same or different?), and how participants collaborated. For example, our initial assumption was that *co-context collaboration* (i.e., participants working symmetrically in the same contexts, e.g., *R-R*, *A-A*, *V-V*) would generally be used to stand close to each other in the park, view artifacts together, and generally work more "tightly coupled" [37], i.e., working on a shared task building upon each other's results within close range of each other and the relevant artifact(s). However, in the study, the context combinations *A-A* and *V-V* were used to divide work and established a much more "loosely coupled" [37] collaboration. Participants worked on the same general problem but as individuals and used divide-and-conquer strategies to carve up the workspace, e.g., each participant editing lamps (D8, D9), placing lamps (D3, D4, D6-D10, D12, D13), or searching for animals (D1) only in one half of the park. Especially in *R-R*, participants moved independently through the virtual space without regard to their mutual distance (D1, D3-D10, D12, D14), as this context combination was primarily used for planning purposes and to discuss distant locations that could be rapidly reached by simply moving the mouse or clicking on the map.

On the contrary, during *cross-context collaboration*, participants seemed to work together more "tightly coupled" [37]. In the context combination *R-A*, for example, one participant often optimized the position of individual lamps in *A* while the other participant checked and commented on the resulting scores in *R* (D4-D9, D11-D13, D15). Also, all dyads D1-D15 used a similar strategy in *R-V* for optimizing the brightness level of lamps. Participants in *V* also frequently asked their partners with an allocentric map in *R* or *A* to give them directions to help with egocentric navigation (all except D1, D11). Cross-context combinations, especially those involving *V*, generally afforded tighter collaborative coupling between participants, so that they could together overcome the asymmetry of these contexts in terms of content and functionality.

5.1.2 Analytical Metrics and Visualizations: Virtual Euclidean Distance and Context-specific Distance Histograms. To substantiate our interpretation of these qualitative observations with quantitative data, we conducted an in-depth analysis of the collaboration distance. In a first step, we used the logging data from all contexts and all sessions to determine the position of both participants P_1 and P_2 inside the virtual park for each second. We then calculated the Euclidean distance d_{eucl} between participants and normalized it using the park's diagonal as the maximum distance d_{max}.

$$d = \frac{d_{eucl}(P_1, P_2)}{d_{max}}$$

In a second step, we plotted the distribution of all measurements of d in a 3×3 matrix of histograms containing each possible context combination (Figure 11).

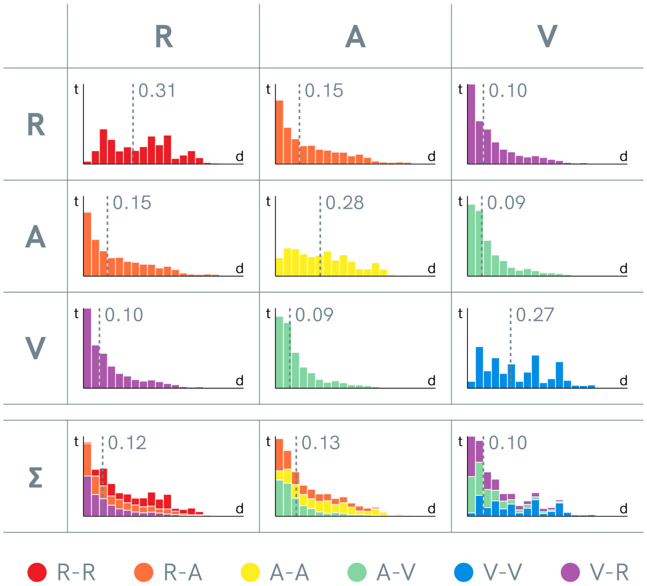

Figure 11: Visual Lens #1: Distribution of the virtual Euclidean distance d between $P1$ and $P2$ for each context combination. The x-axis of each histogram shows d normalized to the maximum possible distance with a resulting range of 0.0 to 1.0 and a bin size of 0.05. The y-axis shows t as the percentage of time that $P1$ and $P2$ spent within the bin for d. t ranges between 0.0 and 0.32 for rows *R*, *A*, and *V*, and 0.0 to 0.63 for row Σ). Histograms R-A and A-R, R-V and V-R, A-V and V-A are identical.

The x-axis of each histogram shows 20 bins (b_0 to b_{19}) for d with a bin size of 0.05. For easier comparisons between histograms, $median(d)$ of each histogram is indicated as a dashed line. The y-axis of each histogram shows the number of measurements of d falling into b_n divided by the total number of all measurements for that context combination. Since measurements were taken every second, it therefore also shows the percentage t of the total time that was spent in that combination. For better comparison, the y-axes always range from 0.0 to 0.32. With 15 dyads working on the task for 45 minutes, a total of $15 \times 45 \times 60 = 40,500$ measured values for d are visualized in the 3×3 matrix. A fourth row in Figure 11 shows stacked histograms to give an aggregated overview for *R*, *A*, and *V*. As each histogram in the fourth row is the sum of the three histograms above, their y-axes range from 0 to 0.63.

A visual comparison in Figure 11 between the histograms of the co-context diagonal (*R-R*, *A-A*, *V-V*) and their cross-context neighbors (*A-R*, *A-V*, *R-V*) reveals a great difference both in their distributions and medians. For cross-context collaboration, distributions are similar to exponential distributions with the maximum inside the first bin and medians of 0.15, 0.09, and 0.10 hinting at smaller distances, virtual co-location, and tight coupling. For co-context collaboration, the distances seem to be much more equally distributed without clear maxima and generally larger medians of 0.31, 0.28, and 0.27 hinting at loose coupling. This visually confirms our qualitative observations about differences in the closeness of collaboration between co-context and cross-context work.

Because of our large total sample size of 40,500 values, common statistical methods to test our observations for exponential distribution would always reveal trivial significant differences. Therefore, to quantitatively analyze our data, we fitted exponential probability distributions to our six histograms containing the observed probability distributions (Table 2). We then used the Kullback-Leibler Divergence [19] to calculate the information loss through those approximations. λ was established by determining the minimum resulting divergence for each distribution. The mean divergence of the co-context distributions (0.302) is eight times higher than that of cross-context distributions (0.038), indicating a much better fit of the exponential distribution for cross-context. Also, the mean of λ for cross-context distributions (0.299) is 2.5 times higher than that of co-context (0.119), meaning that the steepness of the exponential distribution is much higher and shows a better fit of the exponential distribution for cross-context than for co-context. In summary, this confirms our visual interpretation and qualitative observations.

Table 2: Kullback-Leibler Divergence D_{KL} between our observed histogram H and exponential distribution Exp defined by a given λ with $Exp = \lambda e^{-\lambda x}$.

Combination	Histogram H	Estimated λ	$D_{KL}(H\|\|Exp)$
Co-Context	*R-R*	0.091	0.397
	A-A	0.142	0.182
	V-V	0.125	0.388
	Mean	**0.119**	**0.302**
Cross-Context	*R-A*	0.224	0.057
	R-V	0.314	0.035
	A-V	0.358	0.028
	Mean	**0.299**	**0.053**

Finding #2: Transitional collaboration introduces new non-trivial notions of "place" and "distance". In addition to the geographical or physical location of users, they are now also defined by users' contexts and the virtual Euclidean distance between users in their shared virtual workspace or environment. For example, in our study, co-context collaboration afforded greater virtual Euclidean distances and loosely-coupled collaboration. On the contrary, cross-context collaboration afforded smaller distances and tightly-coupled collaboration.
To analyze the mutual effects between distance, context combinations, and collaborative coupling in the future, we propose (and have demonstrated) the use of matrices with context-specific histograms showing distributions of the virtual Euclidean distance d as analytical tools.

5.2 Lens #2: Temporal Patterns in Transitional Collaboration

5.2.1 Qualitative Observations: During our study, we observed that all sessions (D1-D15) seemed to follow the same general pattern:

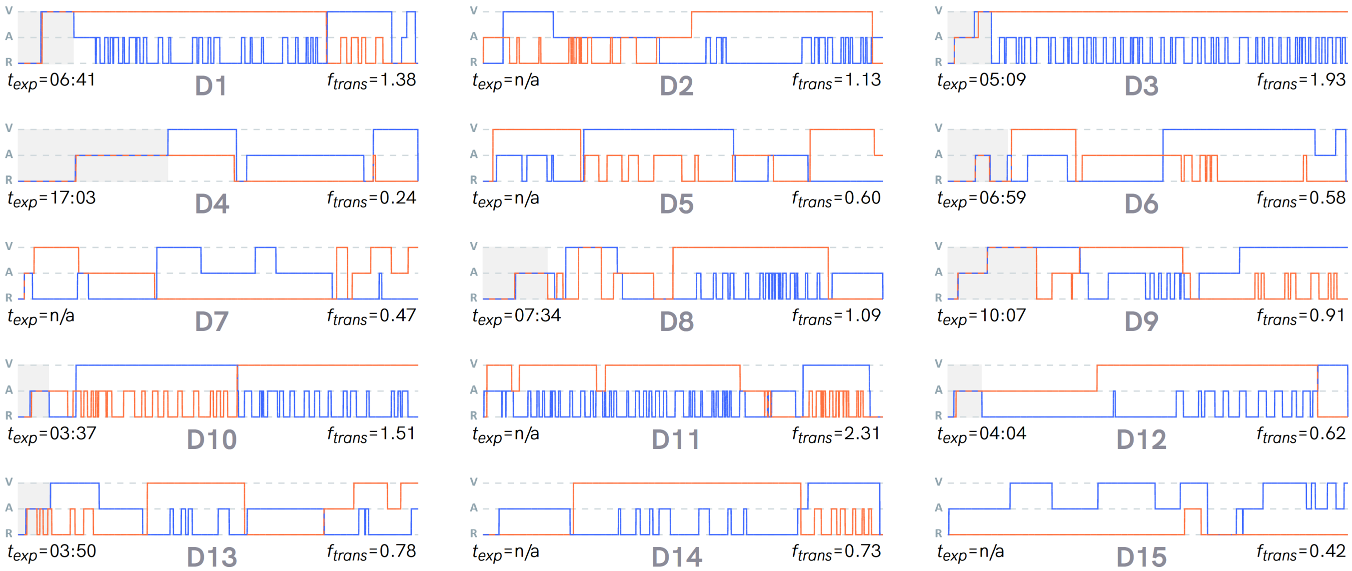

Figure 12: Visual Lens #2: Contexts over time of all dyads together with the transition frequency f_{trans} and the duration of the initial co-context exploration phase t_{exp} (where applicable). The duration of t_{exp} is also highlighted in light gray.

an initial phase of exploration of the system and task followed by a phase of coordinated work. Strategies, functionalities, and responsibilities were collaboratively explored and defined during the first phase. Thereby, it seemed as if the majority of dyads preferred co-context collaboration and moved in parallel through the different contexts. During the subsequent phase of coordinated work, however, dyads seemed to expose very different frequencies of transitions and individual working styles.

5.2.2 Analytical Metrics and Visualizations: Context-Time-Diagram. For more in-depth analysis, we use *small multiples* [39] of a *context-time-diagram* representing the usage of contexts over time of each dyad in Figure 12. Each small multiple contains two line graphs representing one participant of the dyad, colored in either blue (P_1) or orange (P_2). The x-axis represents the 45 minutes of session time. The y-axis represents one of the three contexts in the order of their position on the RVC (i.e., R bottom, A center, V top).

The majority of nine dyads (D1, D3, D4, D6, D8-D10, D12, D13) began their session with an extended co-context exploration phase. The duration t_{exp} of this phase lasted on average 07:14 minutes (*SD*=02:14, *Min*=03:37, *Max*=17:03 minutes), therefore taking up about 16.1% of the total time on average. After this exploration, participants began to engage more in cross-context collaboration.

In addition, some dyads strongly differed in how frequently they transitioned between contexts. To quantify that observation, we calculated the *frequency of transitions* f_{trans} for each dyad by dividing their total number of transitions by the 45 minutes of session time (Figure 12). This frequency f_{trans} was on average 0.98 (*SD*=0.59, *Min*=0.24, *Max*=2.31) transitions per minute. We then classified dyads into four high-frequency dyads (f_{trans}>1.3: D1, D3, D10, D11) and four low-frequency dyads (f_{trans}<0.6: D4, D6, D7, D15). The other seven dyads alternated very strongly between periods of frequent and infrequent transitions (D2, D5, D8, D9, D12-D14). As we discussed in 4.2.1, the number of transitions by a dyad correlated positively with the final scores. We found no other predictor in our data for the transition frequency of a dyad.

Finding #3: For understanding temporal patterns in transitional collaboration, we propose using context-time-diagrams and the transition frequency f_{trans} as analytical tools. Small multiples of these diagrams can visually reveal commonalities and differences between teams, in overall temporal patterns (e.g., phases of initial co-context exploration vs. subsequent cross-context work), average transition frequencies, and also the temporal development of the transition frequency over time.

5.3 Lens #3: Group Use of Contexts

5.3.1 Qualitative Observations: We observed that dyads seemed to have different preferences for certain context combinations. For example, some dyads seemed to prefer to have one member staying mostly in V while the second member was switching between the two other contexts. From our interviews and observations, the reasons were to avoid frequent (un)wearing of the HMD because adjusting display sharpness (D3, D5, D10, D13), taking off glasses (D8), or long hair (D3) proved to be time-consuming. However, other dyads seemed to have a strong preference for staying in R and switching between A and V. They commented that they enjoyed having a member in R for its permanent overview of what was going on (D6-D8, D10, D11, D13-D15).

5.3.2 Analytical Metrics and Visualizations: Context Triforce. We propose the *context triforce* as a new notation and formalization to analyze and fingerprint the dyadic usage of contexts in a TI. Since the context triforce is formally based on a graph, it can also be used as a basis for more advanced analysis or classification based on graph theory, pattern recognition, or probabilistic models (see section 6).

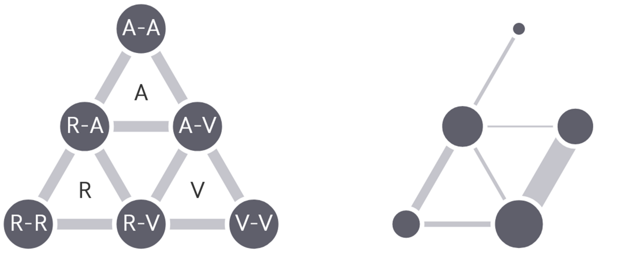

Figure 13: The general structure of the context triforce including labels for its context combinations (left). A specific instance of a context triforce visualizing collaboration data from D6 (right).

Figure 13 (left) shows the general structure of the notation. The name "triforce" originates from the use of this symbol in Nintendo's popular The Legend of Zelda series of video games. It consists of three equilateral triangles, which are joined to form a large equilateral triangle. We use its configuration as a graph in which the nodes represent all possible context combinations and the edges represent the observed transitions between them. The three outer subtriangles each represent one context (*R*, *A*, *V*) and their three corners cover all their possible context combinations. Since this visualization focuses on context combinations and the transitions between them, it does not differentiate between users P_1 or P_2. For example, a context combination with P_1 in *R* and P_2 in *V* equals a context combination with P_1 in *V* and P_2 in *R*. We consider this aggregation as useful as it reduces complexity and shows the true popularity of a context combination, even if participants took turns using a context, e.g., for reducing the physical demand of wearing an HMD (D5, D10, D13).

This structure allows us to visualize and compare the actual group use of contexts by different dyads (Figure 13 right). The total time spent by a dyad in a context combination is mapped linearly to the area of its corresponding node. The thickness of the connecting edge between two nodes represents the number of transitions between them. Thereby, it shows how many transitions happened between these edges in relation to the total number of all transitions between all edges. Thus, it shows no absolute number but a percentage.

We use 15 small multiples [39], one for each dyad, to visually represent the characteristics of that dyad's groupwise context usage patterns in a kind of visual fingerprint (Figure 14):

Preferences for contexts — In dyads D1, D3, and D10, the size of the nodes in the *V* subtriangle compared to the other nodes reveal the dyads' preference for almost constantly having a team member in *V*. Similarly, the dyads D7, D13, and D15 show a preference for having a team member in *R*. In the cases of D2, D4-D6, D8, D9, and D11-D13 node sizes were more equally distributed and therefore showed no clear preference for any context subtriangle.

Preferences for cross- vs. co-context collaboration — In terms of preferences for cross-context vs. co-context collaboration, the sizes of the nodes in the inner cross-context subtriangle (i.e., nodes *R-A*, *A-V*, and *R-V*) reveal the preference of 12 dyads for such collaboration (all except D4, D8, D9).

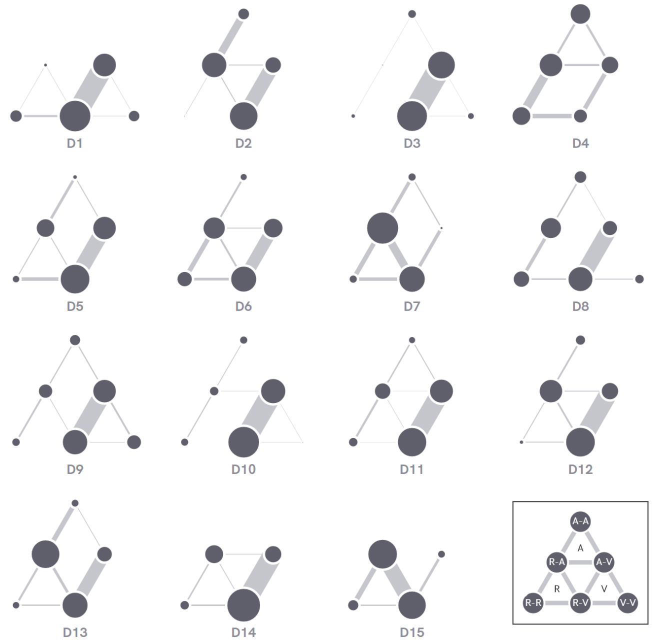

Figure 14: Visual Lens #3: All context triforces showing group usages of our TI. See legend in the bottom right corner.

Frequent transitions — The thickness of the edges of the graph also provides information about the way dyads interacted with the TI. The edge *R-V* to *A-V* was the most frequently used transition in 12 dyads (all except D4, D7, D15). In these cases, a participant always stayed in *V*, while the other one was switching between *R* and *A*. One reason for this could be the high cost of switching in and out of *V* as mentioned above. However, a thick edge does not only mean that many transitions happened there, but also that the time spent in the adjacent nodes was interrupted by many transitions. A counterexample for little interruptions is visible for D7. The nodes *R-A* and *R-V* are large, but the connecting edges are thin, showing that only a few transitions interrupted the time in these context.

Finding #4: We propose the context triforce which formalizes transitional collaboration as a graph and introduces a visual notation that fingerprints the dyadic use of a TI at a glance. This fingerprint can be used to get a fast yet detailed overview of preferred context combinations, co-context vs. cross-context collaboration, and frequently used transitions for single or multiple dyads. In future studies, this lens can be useful in evaluating specific transitions or awareness cues between contexts.

5.4 Lens #4: Individual Use of Contexts

5.4.1 Qualitative Observations: The previous lens did not differentiate between individual team members (i.e., between P_1 and P_2) in order to get an overview of the entire team's preference for different contexts and context combinations. However, looking into an individual's preferences during transitional collaboration in greater detail and enabling side-by-side comparisons of users as opposed to teams can reveal further insights.

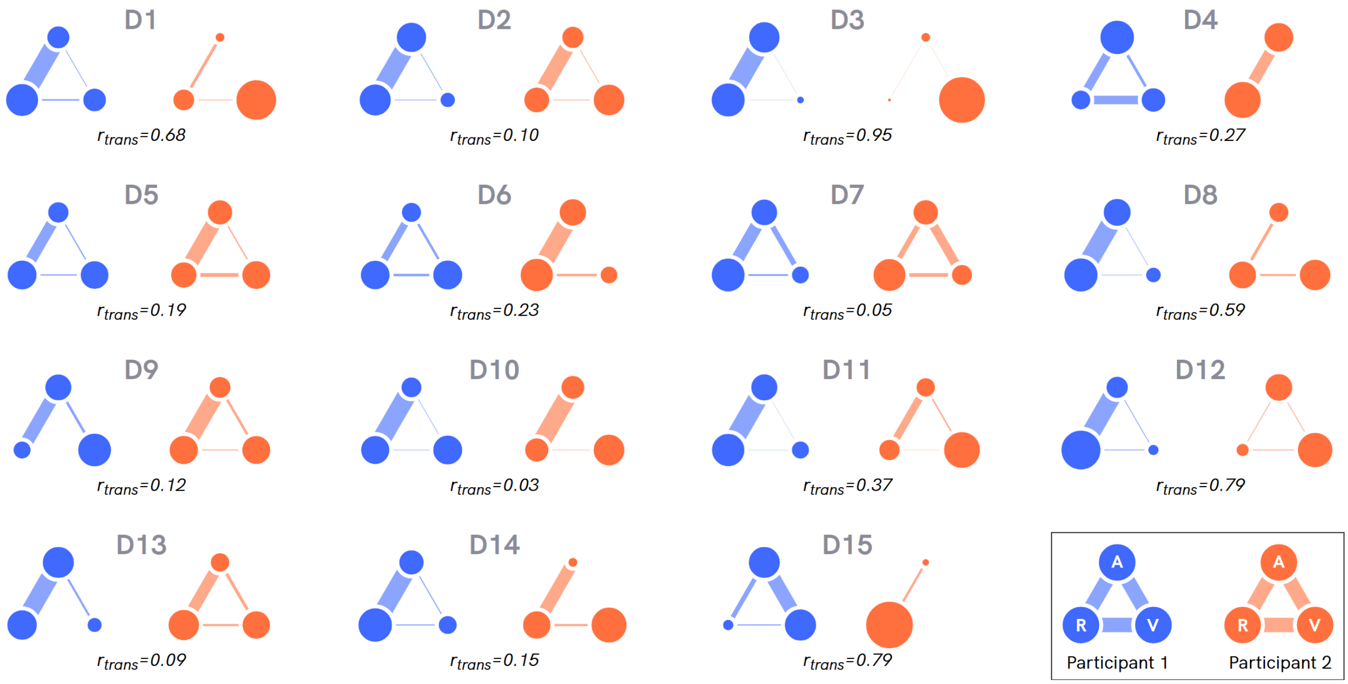

Figure 15: Visual Lens #4: Individual usages of our TI represented by our personal context graph and the ratio of r_{trans} of each dyad. See legend in the bottom right corner.

For example, in our study, some dyads seemed to divide up contexts among themselves, essentially turning one member into the person in charge or expert for a certain context. From our interviews, we learned three reasons for this: (1.) individual preferences for specific devices, (2.) gained expertise over time, and (3.) avoidance of transition costs (D1, D3, D12, D14, D15). On the contrary, dyads D2, D6, D7, and D11 stated that they did not want to work monotonously inside a single context and found frequent transitions desirable and stimulating.

Such behavior can also result in involuntary roles. If a team member decides to stay only in one context, the other team member must take responsibility for the other remaining contexts and therefore transition between them frequently. This results in asymmetry between P_1 and P_2 concerning not only the time spent within a context but also in the number of necessary transitions.

5.4.2 Analytical Metrics and Visualizations: Personal Context Graphs. To measure such asymmetry, we calculated the ratio r_{trans} as the delta of participants' individual transitions $n_{trans}(P_x)$ divided by the total number of transitions N_{trans}. This ratio ranges between 0 (i.e., both participants had the same number of transitions) to 1 (i.e., if only one participant made all transitions).

$$r_{trans} = \frac{|n_{trans}(P_1) - n_{trans}(P_2)|}{N_{trans}}$$

This ratio can be used to describe individuals' usage patterns and how they affect the group strategy. The median of all ratios r_{trans} was 0.23 (M=0.36, SD=0.31, min=0.03, max=0.95) indicating that the dyads tended towards symmetrical numbers of transitions.

To visualize individual usage behavior, we developed the *personal context graph* notation (Figure 15) which is similar to the context triforce. Each dyad is visualized as two small colored graphs, representing P_1 in blue and P_2 in orange. While in the previous lens, the nodes of the graphs represented context combinations, they are here representing only a single context. The nodes were sorted horizontally by their position on the RVC [23] (i.e., R bottom left, A top, and V bottom right). The size of the nodes and the thickness of the edges are mapped exactly like in the context triforce.

Within these small multiples, some extreme cases become immediately visible. For example, in D3 and D15, P_2 was almost permanently in V respectively R. As can be seen in the graphs, P_1 had to compensate for this behavior by having their time equally distributed between the other two contexts and using many transitions between them. Unsurprisingly, these two dyads were also the ones with the highest r_{trans} (r_{trans}(D3)=0.95, r_{trans}(D15)=0.79). On the contrary, D5, D7, and D10, for example, had very similar-looking personal context graphs and low r_{trans} (r_{trans}(D7)=0.05, r_{trans}(D10)=0.03, r_{trans}(D13)=0.09).

While D3 was the minimum and D10 the maximum of all r_{trans}, the corresponding context triforce appeared to be very similar (Figure 14). This is noteworthy as it means that on a team level, both dyads looked very similar, but on a personal level they applied completely different approaches to using the TI. This would have stayed hidden without the personal context graphs. A similar example is D5 & D12. Conversely, similar-looking individual context graphs do not necessarily lead to similar context triforces (e.g., D2 & D5) confirming the necessity for a team-level and person-level analysis using both lenses #3 & #4.

Finding #5: It became evident that participants influenced each other's usage patterns. From the moment a participant developed a strong preference for one context, their partner had to compensate for this behavior. This resulted in an asymmetry not only in the individual times that were spent in each context but also in the total number of transitions made by a user. Such asymmetries can be visualized with our proposed personal context graphs and our metric r_{trans}. Based on several examples, we were also able to show that it is necessary to consider both individual behavior and the behavior of the group with their respective lenses, because using only one of them is prone to leaving such patterns and mutual effects undiscovered.

6 LIMITATIONS AND FUTURE RESEARCH

We consider our findings #1 to #5 important steps towards an in-depth understanding of transitional collaboration with (1.) the analytical lenses #1 to #4 as *methodological contributions* for formalizing, analyzing, and visualizing transitional collaboration in future TI research and (2.) *observational contributions* about if and how dyadic transitional collaboration unfolds during a complex task under realistic conditions. Nonetheless, there are limitations to our work that we discuss in the following together with proposals on how the HCI community could address them in future research.

(1.) Generalizability of Observational and Methodological Contributions: The generalizability of our observations on TIs in finding #1 is limited by the specific task and design of our prototype. For example, we can only hypothesize about the role of verbal communication or the benefits of user enjoyment vs. costs for device/context switching for very different tasks or application domains. Therefore, we encourage HCI researchers to conduct in-depth user studies of transitional collaboration with realistic or real-world tasks. We believe that studies of visual-spatial tasks including visualization and sensemaking of spatio-temporal data (e.g., cross-virtuality analytics of complex data sets, 4D tomography, or also air traffic management) are a particularly promising domain for this research.

Some parts of our methodological contributions in the form of analytical lenses are currently limited to dyadic collaboration within three contexts. While future work on extending this to larger teams and more contexts is needed, we also believe that an initial focus on dyads and few contexts can already reveal the fundamental insights and give clear indications for how transitional collaboration would unfold for such larger teams and more contexts.

(2.) Strengthening Quantitative Analysis: In section 5.3, we introduced the context triforce and the personal context graphs, which formalize transitional collaboration as graphs and fingerprint the dyadic use of a TI. At this stage, we assessed similarities and differences between dyads using our own quantitative metrics, descriptive statistics, and visual comparisons. However, our formalizations as graphs invite to explore further avenues for data analysis. First, interpreting transitional collaboration as weighted and/or symmetric directed graphs could enable researchers to identify nodes of special relevance (e.g., using different centrality measures). Second, interpreting the graphs as Markov chains could help to create probabilistic models for sequences of transitions to inform design decisions or to integrate run-time recommendations. Third, interpreting the graph structure as feature vectors containing nodes and weighted edges could be used to establish similarity measures for quantitative comparisons and clustering of transitional behavior.

(3.) Strengthening Qualitative Analysis: In section 5.1, we used the Euclidean distance in the virtual workspace as an approximation for how tightly or loosely collaboration happened. This does not necessarily reflect the nuances of actual collaborative behavior. In future work, we will apply the concept of coupling styles to provide a more precise, qualitative description of the unfolding collaboration similar to their application for co-located [3, 16, 37], remote [38], and hybrid collaboration [24]. Such analyses are currently based on laborious human coding of user behaviors in audiovisual recordings. Our lenses could support them with a more quantitative and visual exploration of data.

7 CONCLUSION

In this paper, we presented four analytical lenses as methodological contributions through which transitional collaboration can be viewed to support a meaningful analysis and interpretation of the emerging collaborative behavior. They provide TI researchers with new tools for analyzing transitional collaboration or evaluating TIs based on new metrics and visualizations. These lenses were derived from rich qualitative and quantitative data taken from our user study of transitional collaboration with 15 dyads. Based on this data, we showed that our task was realistic in terms of the task load for individuals and teams and that a high level of collaboration and awareness was necessary to complete it. Using several examples, we demonstrated how our analytical lenses can be applied to identify and analyze various commonalities and differences between each of our observed dyads. Thereby, we also reported various findings about how TI design could affect users' performance and users' perceptions of transitional collaboration as observational contributions, including the role of awareness cues, verbal communication, task loads, and the costs of device switching.

ACKNOWLEDGMENTS

We thank André Calero Valdez for very helpful discussions and input on clustering techniques and statistical methods. This work is part of the X-PRO project. The project X-PRO is financed by research subsidies granted by the government of Upper Austria.

REFERENCES

[1] Huidong Bai, Prasanth Sasikumar, Jing Yang, and Mark Billinghurst. 2020. A User Study on Mixed Reality Remote Collaboration with Eye Gaze and Hand Gesture Sharing. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376550

[2] Mark Billinghurst, Hirokazu Kato, and Ivan Poupyrev. 2001. The magic book: A transitional AR interface. *Computers & Graphics* 25, 5 (2001), 745–753. https://doi.org/10.1016/S0097-8493(01)00117-0

[3] Frederik Brudy, Joshua Kevin Budiman, Steven Houben, and Nicolai Marquardt. 2018. Investigating the Role of an Overview Device in Multi-Device Collaboration. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3173874

[4] 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

[5] Barrett Ens, Joel Lanir, Anthony Tang, Scott Bateman, Gun Lee, Thammathip Piumsomboon, and Mark Billinghurst. 2019. Revisiting collaboration through

mixed reality: The evolution of groupware. *International Journal of Human-Computer Studies* 131 (Nov. 2019), 81–98. https://doi.org/10.1016/j.ijhcs.2019.05.011

[6] Thomas Franke, Christiane Attig, and Daniel Wessel. 2019. A Personal Resource for Technology Interaction: Development and Validation of the Affinity for Technology Interaction (ATI) Scale. *International Journal of Human–Computer Interaction* 35, 6 (April 2019), 456–467. https://doi.org/10.1080/10447318.2018.1456150

[7] Bernhard Fröhler, C. Anthes, Christoph, Fabian Pointecker, Judith Friedl, D. Schwajda, A. Riegler, Andreas, Shailesh Tripathi, Clemens Holzmann, Manuel Brunner, Herbert Jodlbauer, Hans-Christian Jetter, and Christoph Heinzl. 2022. A Survey on Cross-Virtuality Analytics. *Computer Graphics Forum* (2022), 465–494. https://doi.org/10.1111/cgf.14447

[8] Alexander Gall, Bernhard Fröhler, Julia Maurer, Johann Kastner, and Christoph Heinzl. 2022. Cross-virtuality analysis of rich X-ray computed tomography data for materials science applications. *Nondestructive Testing and Evaluation* (May 2022), 1–16. https://doi.org/10.1080/10589759.2022.2075864

[9] Jeronimo Gustavo Grandi, Henrique Galvan Debarba, and Anderson Maciel. 2019. Characterizing Asymmetric Collaborative Interactions in Virtual and Augmented Realities. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, Osaka, Japan, 127–135. https://doi.org/10.1109/VR.2019.8798080

[10] Raphael Grasset, Andreas Dunser, and Mark Billinghurst. 2008. Moving Between Contexts - A User Evaluation of a Transitional Interface. In *18th International Conference on Artificial Reality and Telexistence (ICAT 2008)*. Yokohama Japan, 137–143.

[11] Raphael Grasset, Philip Lamb, and Mark Billinghurst. 2005. Evaluation of Mixed-Space Collaboration. In *Proceedings of the 4th IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR '05)*. IEEE Computer Society, USA, 90–99. https://doi.org/10.1109/ISMAR.2005.30

[12] Raphael Grasset, Julian Looser, and Mark Billinghurst. 2006. Transitional interface: concept, issues and framework. In *Proceedings of the 5th IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR '06)*. IEEE Computer Society, USA, 231–232. https://doi.org/10.1109/ismar.2006.297819

[13] Jan Gugenheimer, Evgeny Stemasov, Julian Frommel, and Enrico Rukzio. 2017. ShareVR: Enabling Co-Located Experiences for Virtual Reality between HMD and Non-HMD Users. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 4021–4033. https://doi.org/10.1145/3025453.3025683

[14] Carl Gutwin and Saul Greenberg. 2002. A Descriptive Framework of Workspace Awareness for Real-Time Groupware. *Computer Supported Cooperative Work (CSCW)* 11, 3-4 (Sept. 2002), 411–446. https://doi.org/10.1023/A:1021271517844

[15] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in Psychology*. Vol. 52. Elsevier, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

[16] Petra Isenberg, Danyel Fisher, Sharoda A. Paul, Meredith R. Morris, Kori Inkpen, and Mary Czerwinski. 2012. Co-Located Collaborative Visual Analytics around a Tabletop Display. *IEEE Transactions on Visualization and Computer Graphics* 18, 5 (May 2012), 689–702. https://doi.org/10.1109/TVCG.2011.287

[17] Hans-Christian Jetter, Jan-Henrik Schröder, Jan Gugenheimer, Mark Billinghurst, Christoph Anthes, Mohamed Khamis, and Tiare Feuchtner. 2021. Transitional Interfaces in Mixed and Cross-Reality: A new frontier?. In *Interactive Surfaces and Spaces (ISS '21)*. Association for Computing Machinery, New York, NY, USA, 46–49. https://doi.org/10.1145/3447932.3487940

[18] Robert Johansen. 1988. *Groupware: computer support for business teams.* Free Press ; Collier Macmillan, New York : London.

[19] S. Kullback and R. A. Leibler. 1951. On Information and Sufficiency. *The Annals of Mathematical Statistics* 22, 1 (March 1951), 79–86. https://doi.org/10.1214/aoms/1177729694

[20] Sebastian Linxen, Christian Sturm, Florian Brühlmann, Vincent Cassau, Klaus Opwis, and Katharina Reinecke. 2021. How WEIRD is CHI?. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–14. https://doi.org/10.1145/3411764.3445488

[21] Nicolai Marquardt, Ken Hinckley, and Saul Greenberg. 2012. Cross-device interaction via micro-mobility and f-formations. In *Proceedings of the 25th annual ACM symposium on User interface software and technology - UIST '12*. ACM Press, Cambridge, Massachusetts, USA, 13–22. https://doi.org/10.1145/2380116.2380121

[22] Frank Maurer, Craig Anslow, Joaquim Jorge, and Mauricio Sousa. 2022. Enhancing cross-reality applications and user experiences. In *Proceedings of the 2022 International Conference on Advanced Visual Interfaces (AVI 2022)*. Association for Computing Machinery, New York, NY, USA, 1–3. https://doi.org/10.1145/3531073.3535256

[23] Paul Milgram and Fumio Kishino. 1994. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems* 77, 12 (1994), 1321–1329. Publisher: The Institute of Electronics, Information and Communication Engineers.

[24] Thomas Neumayr, Hans-Christian Jetter, Mirjam Augstein, Judith Friedl, and Thomas Luger. 2018. Domino: A Descriptive Framework for Hybrid Collaboration and Coupling Styles in Partially Distributed Teams. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW (Nov. 2018). https://doi.org/10.1145/3274397

[25] Thammathip Piumsomboon, Youngho Lee, Gun Lee, and Mark Billinghurst. 2017. CoVAR: A Collaborative Virtual and Augmented Reality System for Remote Collaboration. In *SIGGRAPH Asia 2017 Emerging Technologies (SA '17)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3132818.3132822

[26] Thomas Plank, Hans-Christian Jetter, Roman Rädle, Clemens N. Klokmose, Thomas Luger, and Harald Reiterer. 2017. Is Two Enough?!: Studying Benefits, Barriers, and Biases of Multi-Tablet Use for Collaborative Visualization. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*. ACM Press, Denver, Colorado, USA, 4548–4560. https://doi.org/10.1145/3025453.3025537

[27] Fabian Pointecker, Judith Friedl, Daniel Schwajda, Hans-Christian Jetter, and Christoph Anthes. 2022. Bridging the Gap Across Realities: Visual Transitions Between Virtual and Augmented Reality. In *2022 IEEE Interational Symposium on Mixed and Augmented Reality (ISMAR)*. Singapore. (to appear).

[28] Fabian Pointecker, Hans-Christian Jetter, and Christoph Anthes. 2020. Exploration of Visual Transitions Between Virtual and Augmented Reality. In *4th Workshop on Immersive Analytics: Envisioning Future Productivity for Immersive Analytics // @CHI 2020 Honolulu*. https://drive.google.com/file/d/1V1lXIk-DPjcfuJg2LgaS_UfYXQq-7e2k/view

[29] Patrick Reipschläger and Raimund Dachselt. 2019. DesignAR: Immersive 3D-Modeling Combining Augmented Reality with Interactive Displays. In *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces (ISS '19)*. ACM, New York, NY, USA, 29–42. https://doi.org/10.1145/3343055.3359718

[30] Patrick Reipschläger, Severin Engert, and Raimund Dachselt. 2020. Augmented Displays: Seamlessly Extending Interactive Surfaces with Head-Mounted Augmented Reality. In *Proceedings of the 2020 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, New York. https://doi.org/10.1145/3334480.3383138

[31] Patrick Reipschläger, Tamara Flemisch, and Raimund Dachselt. 2021. Personal Augmented Reality for Information Visualization on Large Interactive Displays. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (Feb. 2021), 1182–1192. https://doi.org/10.1109/TVCG.2020.3030460

[32] Andreas Riegler, Christoph Anthes, Hans-Christian Jetter, Christoph Heinzl, Clemens Holzmann, Jodlbauer Herbert, Manuel Brunner, Stefan Auer, Judith Friedl, and Bernhard Fröhler. 2020. Cross-Virtuality Visualization, Interaction and Collaboration. In *International Workshop on Cross-Reality (XR) Interaction co-located with 14th ACM International Conference on Interactive Surfaces and Spaces*. http://ceur-ws.org/Vol-2779/paper1.pdf

[33] Joan Sol Roo and Martin Hachet. 2017. One Reality: Augmenting How the Physical World is Experienced by combining Multiple Mixed Reality Modalities. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*. ACM, Québec City QC Canada, 787–795. https://doi.org/10.1145/3126594.3126638

[34] Stacey D. Scott, Carpendale Sheelagh, and Kori M. Inkpen. 2004. Territoriality in collaborative tabletop workspaces. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work - CSCW '04*. ACM Press, Chicago, Illinois, USA, 294. https://doi.org/10.1145/1031607.1031655

[35] James Sellers, William S. Helton, Katharina Näswall, Gregory J. Funke, and Benjamin A. Knott. 2014. Development of the Team Workload Questionnaire (TWLQ). *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 58, 1 (Sept. 2014), 989–993. https://doi.org/10.1177/1541931214581207

[36] Adalberto L. Simeone, Mohamed Khamis, Augusto Esteves, Florian Daiber, Matjaž Kljun, Klen Čopič Pucihar, Poika Isokoski, and Jan Gugenheimer. 2020. International Workshop on Cross-Reality (XR) Interaction. In *Companion Proceedings of the 2020 Conference on Interactive Surfaces and Spaces (ISS '20)*. Association for Computing Machinery, New York, NY, USA, 111–114. https://doi.org/10.1145/3380867.3424551

[37] Anthony Tang, Melanie Tory, Barry Po, Petra Neumann, and Sheelagh Carpendale. 2006. Collaborative Coupling over Tabletop Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06)*. Association for Computing Machinery, New York, NY, USA, 1181–1190. https://doi.org/10.1145/1124772.1124950

[38] Philip Tuddenham and Peter Robinson. 2009. Territorial Coordination and Workspace Awareness in Remote Tabletop Collaboration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. Association for Computing Machinery, New York, NY, USA, 2139–2148. https://doi.org/10.1145/1518701.1519026

[39] Edward R. Tufte. 2001. *The visual display of quantitative information* (2nd ed.). Graphics Press, Cheshire, Conn.