Mapping out Work in a Mixed Reality Project Room

Derek Reilly[1], Andy Echenique[2], Andy Wu[3], Anthony Tang[4], W. Keith Edwards[5]

[1]Dalhousie University
Computer Science
Halifax, NS, Canada
reilly@cs.dal.ca

[2] U. C. Irvine
ICS
Irvine, CA, USA
echeniqa@uci.edu

[3]GE Global Research
UX Innovation Lab
San Ramon, CA, USA
andycswu@gmail.com

[4]University of Calgary
Computer Science
Calgary, AB, Canada
tonyt@ucalgary.ca

[4]Georgia Tech
Interactive Computing
Atlanta, GA, USA
keith@cc.gatech.edu

ABSTRACT

We present results from a study examining how the physical layout of a project room and task affect the cognitive maps acquired of a connected virtual environment during mixed-presence collaboration. Results indicate that a combination of physical layout and task impacts cognitive maps of the virtual space. Participants did not form a strong model of how different physical work regions were situated relative to each other in the virtual world when the tasks performed in each region differed. Egocentric perspectives of multiple displays enforced by different furniture arrangements encouraged cognitive maps of the virtual world that reflected these perspectives, when the displays were used for the same task. These influences competed or coincided with document-based, audiovisual and interface cues, influencing collaboration. We consider the implications of our findings on WYSIWIS mappings between real and virtual for mixed-presence collaboration.

Author Keywords

Mixed reality; cross reality; mixed presence; display ecology; cognitive map; spatial cognition

ACM Classification Keywords

H5.3. Information interfaces and presentation (e.g., HCI): Group and Organization Interfaces----Computer Supported Cooperative Work.

INTRODUCTION

Teams in dedicated project rooms actively configure artifacts and space into meaningful regions for work. Over time, space in such rooms takes on special meaning [8,17]. For example, Teasley et al. [24] found that knowledge workers dedicated regions to specific activities, clustered task-relevant artifacts together, and used flipcharts and whiteboards as persistent at-a-glance records of team activity. Architects and designers respond by designing flexible and reconfigurable collaborative workspaces [20,22]. Space provides common ground for collaborators: it comes to take on meaning [8], and bodily action and deictic references become meaningful during collaboration.

When remote collaborators connect to a project room through audio/video conferencing, the affordances of the space are lost: the remote collaborator cannot easily refer to, nor understand references to, artifacts or regions of the space. Hindmarsh et al. name this problem *fragmented interaction*, referring both to the fragmentation of spaces (between the project room and the remote collaborator's space), as well as the fragmented interaction between collaborators as they encounter referential problems and the repair that needs to take place [9].

One approach to addressing this problem is to use *media spaces*, fusing two or more physical spaces with high-fidelity audio/visual connections, for example, as if collaborators worked on opposite sides of a pane of glass (e.g. Luff et al. [14]). Yet, such setups tend to connect only a single region of a workspace, ignoring other areas that may be relevant, and emphasize face-to-face communication over artifact-centric collaboration.

In contrast to the media space work, our approach has been to consider an artifact-centric fusion that links physical project rooms and virtual spaces, into what we call *mixed reality collaborative spaces*. Remote collaborators access digital documents from within the virtual space, while collocated collaborators work on the same documents through displays embedded in the physical space. In collaboration with industrial design and workspace experts at Steelcase, we have considered ways that physical and virtual environments can be designed together to benefit collaboration, and demonstrated these in a project room called the inSpace lab.

Figure 1. inSpace lab, "strict" WYSIWIS. Visitors to the project room understood the correspondence immediately, but the virtual world was difficult to navigate and use.

In the most literal variation of our approach (Figure 1), the physical project room is reproduced in a virtual world, allowing remote ("in-world") collaborators to join project room collaborators (the "in-room" collaborators) via audio

CHI 2015, April 18 - 23 2015, Seoul, Republic of Korea

ACM 978-1-4503-3145-6/15/04…$15.00
http://dx.doi.org/10.1145/2702123.2702506

links. Here, collaborators' communicative references to space are meaningful: "front of the room" can be easily understood, and with spatialized audio, "come over here" can be meaningfully understood by all.

This approach begs the question: does a strict What-You-See-Is-What-I-See (WYSIWIS) approach need to be followed, or when and in what ways can a relaxed or non-WYSIWIS approach be used [21]? A relaxed approach might be more effective for getting work done: we can tailor the virtual world to the remote collaborators' user interface (e.g., remove furniture and increase the relative size of documents and workspaces, Figure 2), we no longer need a 1-1 correspondence between displays and in-world content, and we can focus on key connection points for specific collaborative tasks rather than mapping the entire room. We may also want to support opportunistic, *ad hoc* and non-WYSIWIS connections (e.g., opening a virtual world client on a collaborator's laptop in the room).

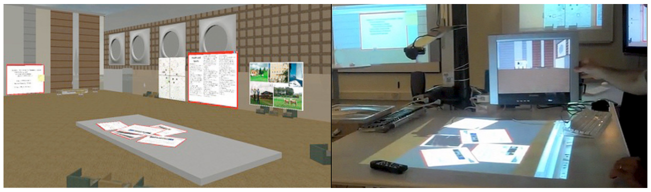

Figure 2. inSpace lab, "relaxed" WYSIWIS. The correspondence needed to be explained to visitors to the project room, but the virtual world was easier to use.

The layout of physical collaborative spaces influences the way people collaborate within that space, with or without digital technology, and sometimes in unanticipated ways [11]. One of the key intuitions guiding our design work was that the physical space may also guide collaboration with remote collaborators [18]. At the same time, our design and prototyping work [18-20] has generated a range of approaches for combining physical and virtual, some requiring strict correspondence, others relaxed correspondence, and others requiring no correspondence.

In this paper we present a study set in the inSpace lab that considers how people build mental models (*cognitive maps*) of a connected virtual space when collaborating. Our results show that the project room's physical layout combined with the collaborative activity impacted the cognitive maps formed by in-room collaborators, and that these effects competed with content and interface cues. Collaborators need ways to reference work materials, maintain awareness of each other's actions, and coordinate next steps. Accurate cognitive maps of the virtual world are only useful if they make it easier to work with remote collaborators. In our study, some of our participant groups worked effectively despite having inaccurate and/or divergent cognitive maps. Some groups were extremely resourceful when figuring out how to coordinate work, testing interface limitations and using both spatial and non-spatial attributes (such as document content) when coordinating. For others, conflicting or inaccurate models led to confusion and breakdowns in collaboration. We consider the implications for WYSIWIS, relaxed WYSIWIS, and non-WYSIWIS approaches to mixed reality collaborative spaces.

BACKGROUND AND MOTIVATION

The mixed reality collaborative spaces concept builds on considerable work addressing the "remote collaborator" problem, in particular work involving *mixed presence* (collocated and remote) collaboration [23], and most directly on three lines of research: media spaces, cooperative buildings, and mixed reality. We briefly touch on this research to illustrate how it has influenced our thinking on collaborative mixed reality workspaces. We then review theories from the psychology literature on how cognitive maps are constructed, and consider how these theories might apply in the mixed reality context.

Connecting Spaces

In the early media space work, which connected physical spaces together using audio-visual links, researchers were interested in preserving the function of the spaces. Bly et al. [3] emphasized connecting functionally similar spaces, for example, placing portals to remote offices in physical hallways, fusing "common areas", and creating shared offices. This early work used the location of a portal as an expression of its function, among other cues (placing coffee pots in view, for example). Buxton describes experiences with a similar system [4], finding that the placement of a media portal relative to the physical work environment impacts expectations and social protocol. These explorations were highly context-specific, however, and general design principles based on room layout are not offered. More recent research has incorporated vision-based tracking with video links to promote collaboration and play (see Hunter et al. [13] for example), yet these emphasize single connection points rather than work environments, and don't permit multiple remote connections.

Architectural & Digital Co-Design

The semantic and functional characteristics of space have been a central part of the cooperative buildings research. For example, Streitz et al. [22] illustrated how technology and architecture could be designed together—the insight being that display surfaces could be designed and configured depending on the function of the spaces they inhabit. These structures provided means for collaborators to engage one another, while also allowing collaborators to reassemble the spaces to support different styles of engagement (just as physical furniture does). Streitz et al. discuss but do not demonstrate an extension to mixed presence collaboration [22]. The swisshouse project [12] applied digital-physical co-design to connect houses in ways that reflect the functional needs of spaces in the home.

Mixed Reality and Cognitive Maps

Dix et al. argue that mixed reality relies on the cognitive process of mapping (drawing connections between) multiple spaces [6]. They identify three types of spaces present in mixed reality systems: the physical environment (*real space*), maps and/or sensed positioning (*measured space*), and the connected digital information (*virtual space*). Interacting with mixed reality systems requires moving from a spatial model of the physical space to a model of the sensed or mapped physical-virtual boundary to interact with the virtual space, which again one must build a model of through its manifestation in the real space. The success of a mixed reality environment will depend on appropriate mappings between these spaces.

Following Dix et al. [6], cognitive maps of the virtual space would impact how shared resources are referenced, how users across the physical-virtual divide maintain awareness of each other, and how they coordinate actions. We now outline research in spatial cognition that considers how such maps are formed.

The Landmark-Route-Survey (LRS) [15] model of spatial knowledge acquisition and its variants emphasize moving through space as the primary way to build cognitive maps. It was unclear to us how moving around our project room could impact a cognitive map of the connected virtual space, since the distances were short and much collaboration occurs while stationary. Colle and Read [5] propose a "gaze viewing mode" of spatial knowledge acquisition, where a spatial model is constructed from visual information, supplementing LRS-mode spatial knowledge. They provide evidence that cognitive maps formed from gaze viewing tend to be structured hierarchically according to distinct physical regions (they dub this the "room effect"), and so it is more difficult to recall the relative location of objects in different subsections of the hierarchy than objects in the same subsection. This effect may also apply to work regions within a project room, and possibly extend to a connected virtual world: one's perception of the relative location in the virtual world of content shown on displays would be influenced by the relative placement of the displays in the physical room if the displays are in the same work region, but not when they are in different work regions.

Physical room layout is not likely the sole determinant of cognitive map formation in a mixed reality collaborative space. As Tversky et al. have noted, cognitive maps are often *collages* of different levels of detail and focus, reflecting the fact that these maps are derived from qualitatively different cues [25]. In particular, non-spatial attributes of shared work documents may have an impact: Hommel and Knuf [10] have shown that grouping in cognitive maps can occur due to non-spatial factors including performing common actions with objects, while Aliakseyeu et al. [1] demonstrate user preference for semantic vs. spatial metaphors on control interfaces for multiple display environments. We anticipated that these factors would also influence cognitive map formation.

EXPERIMENT

We conducted an experiment using a subset of our project room consisting of two regions: a meeting area with a conference table and three large displays, and a brainstorming region comprised of a small table with high seating and a single large landscape display (see Figure 3, left). We expected that our participants would perceive these two regions as distinct. We contrast this with a second, "1-table" room configuration (see Figure 3, right). Display positions were held constant across layouts to avoid perceptual grouping effects based on changes in relative display positions. Each display presented a perspective on a connected virtual world, focused on one shared document plus some of the surrounding virtual world foreground and background (see Figures 4,5).

The experiment was conducted with groups of three participants. A collocated (*in-room*) pair participated in the inSpace project room, while a remote (*in-world*) participant was in an office (Figure 4, top), and used a standard desktop virtual world client to navigate the virtual world connected to the project room. Participants performed two kinds of activities: an itinerary activity and an image search activity. The study employed a counterbalanced 2×2 (Layout × Activity) within-subjects design.

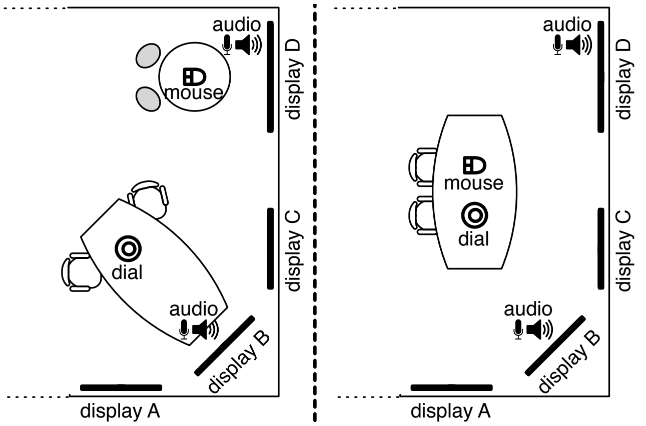

Figure 3. The two room layouts used in the study. 2-table (left) has two work regions, one using displays A-C and one display D. 1-table (right) has a single work area involving all displays.

Inspired by the "room effect" findings of Colle and Read [5], we expected that in the 2-table layout in-room participants would see displays A-C as a single left-to-right virtual world region, and would be less certain about how display D relates to A-C, while in the 1-table layout in-room participants would see displays A-D as a single left-to-right virtual world region.

We anticipated an influence of task on cognitive maps [10], expecting that the itinerary task (which uses display D differently than displays A-C) would amplify the 2-table effect, while the image search task (using displays A-D in the same way) would dampen the 2-table effect.

Following Aliakseyeu et al. [1] and building on our experiences demoing the inSpace lab, we were also interested in seeing what cues collaborators used to work together, and how these cues impacted cognitive maps.

Population and Recruitment

A total of 36 participants (12 groups of 3) were recruited, following two pilot sessions. Participants were drawn from the wider student populations of Georgia Tech and Emory University, through flyers and mailing lists. No specific expertise was required, and familiarity with collaborators was not controlled for. The entire experiment lasted 1½-2 hours, and each participant received $15.

Procedure

Participants first met in the remote participant's room and gave informed consent, after which the collocated pair was escorted to the project room (the remote collaborator did not see the project room). All participants then received training on their respective interfaces, lasting 5-10 minutes.

Participants conducted four group activities: an image search and an itinerary activity (see next section for details) under one layout condition, followed by an image search and an itinerary activity under the other layout condition. Each activity lasted 10-20 minutes. The activities were captured on video using two cameras in the project room, and one in the remote room.

After each activity, the in-room participants were given five minutes to independently sketch, annotate and describe on paper their conception of the virtual environment and the organization of the content within it. They could not see the displays at this time. This was our primary measure of in-room participants' cognitive maps of the virtual world. In order to avoid a learning effect, different virtual world backgrounds were used for each activity.

After the sketching task, in-room participants worked together to place miniature printouts of the documents they used in the activity where they believed they were located on a large top-down printout of the virtual world. This gave an opportunity for each in-room participant to vocalize details of their cognitive map as they relate to the top-down view, and to work out contradictions and disagreements with their partner. The task was captured on video, and a snapshot was taken of the result.

At the end of the experiment, all three participants were brought together for a videotaped 5-10 minute semi-structured interview, to discuss their strategies for collaborating in the activities, and to reflect on their cognitive maps as expressed in the sketching and item placement tasks.

Activities

As discussed, participants engaged in two kinds of group activity: *itinerary* (a simplified holiday itinerary planning exercise), and *image search* (finding a target image among 200 colour-matched photos). Each is detailed here.

Itinerary

In *itinerary*, groups were presented with a set of ~30 attractions, organized into three interest categories (nature, culture, and health and sport). For each category, items were numbered and displayed on a map broken into quadrants, descriptions and hours of operation were provided in a text listing, and labeled photos of points of interest and activities were displayed in a montage (Figure 4). Groups completed one itinerary activity under each layout condition, so materials were prepared for two destinations (Linz and Bologna); the destination-layout pairings were counterbalanced.

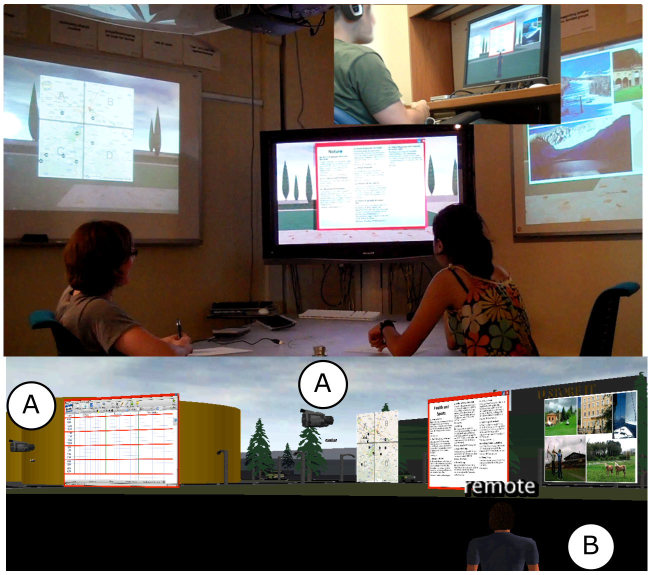

Figure 4 The Itinerary activity. A: cameras show what the in-room participants can see. B: the in-world participant's avatar

Participants selected six items together, choosing at least one from each category, to build a three-day holiday itinerary. Items were scheduled within hours of operation, and a fixed travel time was added if travelling between quadrants in the same day. Participants were asked to set "reasonable durations" for each item on their itinerary. Itinerary planning was split into two subtasks: item selection and scheduling. Participants were told to agree upon six itinerary items during item selection, and then to schedule them on a calendar. Participants were warned if still in the item selection phase after 12 minutes. The calendar displayed the three-day itinerary in an hourly grid (Figure 4, bottom left); items selected to be part of the itinerary could be dragged onto the calendar using a mouse by any collaborator. Participants could return to item selection as needed to review details or make a change to their selected items.

Figure 5 details the collaboration mechanics for the itinerary activity. The calendar was shown on display D, and was the leftmost document in the virtual world. Maps, item text, and photo montages were grouped together by category in-world, and these document groupings were placed in a line. Displays A-C showed map, text and image documents for one category. A dial controller was used to switch between categories: to the right if the dial is rotated clockwise or left if counter-clockwise. If already at the left or right-most category, turning the dial in that direction had no effect. Two audio links (mic and speakers, marked as au1 and au2 in the figure) were set up in the room, one at display D and the other in the center of displays A-C. If the remote collaborator was near the calendar they would be heard at display D, and if they were viewing the same category as the in-room collaborators, they would be heard at displays A-C. Otherwise, audio connection between local and remote collaborators was lost. The remote collaborator could see which category was currently visited by the in-room collaborators via a camera avatar placed in the air above and in front of the category being viewed. Another stationary camera avatar was situated in front of the calendar, indicating that the calendar was always visible to in-room collaborators. The remote collaborator could move their avatar manually using WASD keys, or use 1-4 as hotkeys to jump to the calendar and each of the categories.

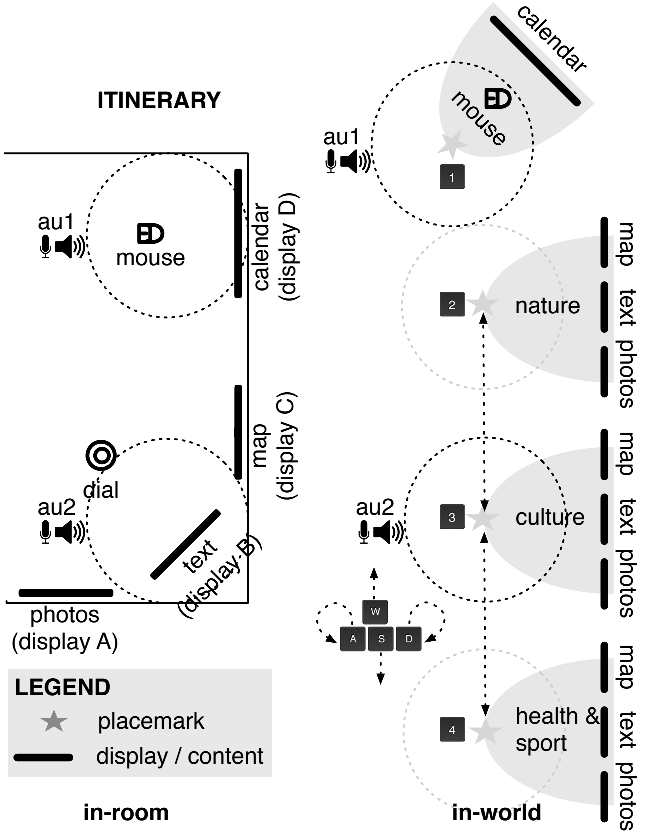

Figure 5 Real-virtual mappings for the Itinerary activity.

Image Search

In *image search* groups located a target image from among 200 colour-matched images presented on four image sets (each set presented images in a 5x10 grid, see Figure **6**). Each image set was visible from a separate in-room display. Participants completed 10 rounds, each involving a different set of 200 images. Groups completed an image search activity under each room layout, so two sets of 10x200 images were prepared.

The facilitator described the target image, and then the group worked together to locate the image (no constraints were imposed on strategy). Once a group member located the image, the other group members had to point to it (the remote participant did this using an in-world telepointer). Any round taking more than 2 minutes was aborted by the facilitator.

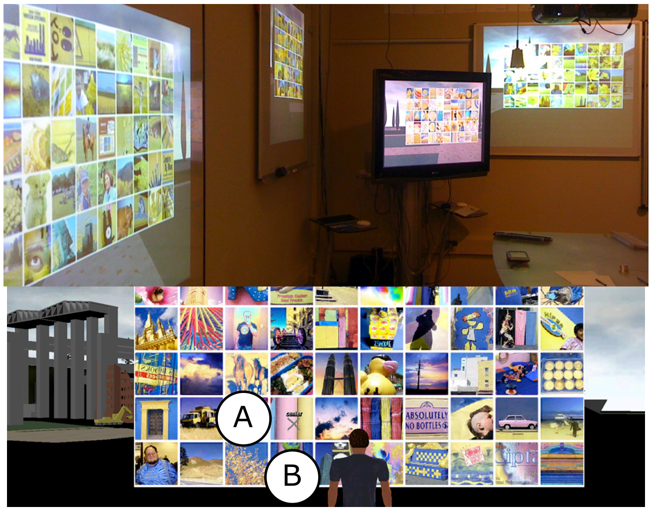

Figure 6 The Image Search activity. A: a telepointer is used to point to images. B: in-world participant's avatar.

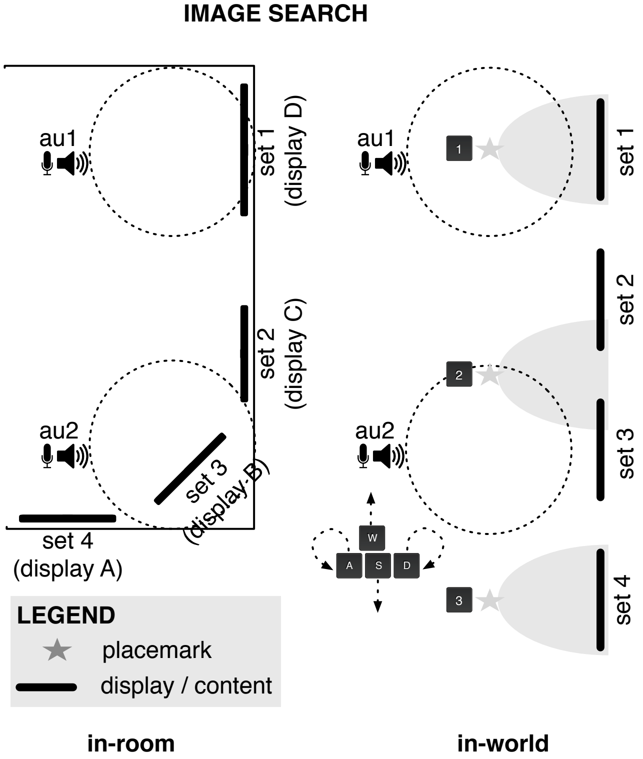

Figure 7 Real-virtual mappings for the Image Search activity.

Figure 7 details the collaboration mechanics for the image search activity. In Image Search, displays A-D each displayed one image set. The sets were presented in the same order in-room as in-world. Audio links (au1, au2) connected displays D and B to the virtual world regions in front of image sets labeled 1 and 3, respectively. Image sets were placed in a straight line in-world, far enough away from each other so that regions surrounding sets 2 and 4 were out of audio range. The remote collaborator could navigate manually using the WASD keys, or jump to a placemark in front of an image set using the 1-4 keys. Video camera avatars in front of image sets 1 and 3 indicated where the audio links were. In this activity, the virtual cameras did not move and there was no indication in-world that sets 2 and 4 were visible in-room.

Measures

Activity Video Coding

We were interested in events that might reflect the impact of cognitive maps, in particular:

Grounding work: local and remote participants talk about how they will collaborate, what each can see, etc.

Synchronization/Coordination: local and remote participants figure out together where to "move to" next.

Spatial reference: local and/or remote participants use spatial language during the activity.

The three video sources were combined, the video was transcribed and interactions were annotated. Following this, an iterative open coding process was followed using the generated text records. Common events were identified during axial coding and these were related back to the higher-level events listed above.

Sketching

Sketches were coded by two individuals who were unaware of their room layout conditions. Sketches were classified according to a set of nominal characteristics reflecting the relative location and organization of all task content, the correspondence of content placement in-world to the layout of displays in the physical project room, the depiction of virtual navigation and/or in-room navigation, and the depiction of the virtual world itself.

Sketching accuracy has previously been shown to correlate strongly with navigation efficacy in virtual environments [1], but using sketching as a means of expressing cognitive maps in general has been both challenged [7] and supported [16] in the spatial cognition literature. We employed several measures in addition to sketching, described below.

Group Content Placement Task and Interview

This data was reviewed for interesting points of disagreement, notable accuracy or inaccuracy in the resulting placements, and the relationship of the placements to the participant sketches.

RESULTS

Sketch Similarity

We considered that similar models of the virtual space would be useful to coordinate work. We assigned a similarity score to *all* pairs of sketches (within groups and across groups) by summing the number of nominal characteristics that were different (i.e., a lower score indicated more similarity). Wilcoxon tests were run on within-group and between-group similarity scores, for both itinerary and image search. Results for itinerary showed no difference in similarity scores within groups (M=4.3) vs. between (M=5.15), Z= -1.49, p = .1353, while results for image search showed a significant increase in similarity within groups (M=1.03) compared to between groups (M=2.71) Z= -4.57, p <.0001.

Cognitive maps appear to converge only when they are used: similar cognitive maps were formed in image search, during which all groups employed a spatial left-to-right or numbering scheme at least once for each activity. By contrast, spatial reference schemes were used (even partially) in only 6 itinerary activities (out of the 24 completed by all groups). All groups used the non-spatial content type (map, text, image, schedule) to reference displays in itinerary—not possible in image search since all displays presented the same type of content.

Impact of Layout and Activity

We considered the placement of the content that appeared on display D (i.e. the schedule in the itinerary activity and leftmost image set in image search) relative to the remaining content in the post-activity sketches. We classified the content on this display as "differentiated" from the other content if it was depicted at a different orientation than *all* other content, and/or separated in space relative to *all* other content.

Z-tests were conducted to compare differentiation in each layout condition across both activities, and then within each activity. An expected main effect was found for activity (Z=2.754, p<.003), with less differentiation in the image search activity (which used all displays in the same way) than in the itinerary activity (which used display D for a separate scheduling subtask), regardless of room layout. No main effect was found for layout (Z=1.252, p=.105), but there was a significant interaction effect of layout and activity, such that the schedule (presented on display D) was differentiated more in the 2-table layout condition than in the 1-table condition for the itinerary activity (Z=1.919, p<.027), but not in the image search activity (Z=0, p=.5). This result suggests that room layout matters when the split in regions corresponds with a split in activity subtask, as was the case in the 2-table itinerary condition.

We also looked in the sketches for a regular placement of content in the virtual world to correspond with the organization of displays in the physical room. Placement didn't need to precisely match the display layout: content might be placed along a straight line, arc, or circle in the sketch, so long as is was positioned in a regular, contiguous pattern. There was a main effect of activity (Z=1.817,p=.035), with the image search activity (using all displays in the same way) encouraging more regular placement of all content (a total of 34/48 image search sketches) than itinerary. There was no main effect of room layout on the presence of such a pattern in the sketches (Z=0,p=.5), and no interaction effect of layout and activity.

After reviewing the sketches for image search, it appeared that the *type* of regular pattern might differ based on layout condition. We found a significant effect of layout (Z=1.873, p<.031) on type of pattern, with more curved and circular patterns in the sketches for the 2-table condition. In the 1-table condition participants had a straight-on view of the displays. In the 2-table condition participants were seated amongst displays A-C, and display D was to the left of and slightly behind their field of view. This suggests that participants mapped their egocentric relationship with the displays into content orientations in the virtual world. We did not capture discussion of curved vs. straight layouts during the image search activity, during top-down placement or in interviews—from this we observe that physical layout can also impact cognitive maps in ways that don't obviously impact collaboration.

Impact of Other Factors

Content Structure

Cognitive maps were influenced by content attributes in the itinerary task. 18/48 itinerary sketches grouped content by type (map, text, image) only, even though the content was organized by category (nature, culture, fitness) then by type in the virtual world (Figure 8). Grouping by content type was also evident in the placement exercise. It is unclear if all such groupings were purely semantic [1], or influenced by the (incorrect) perception that switching between content categories replaced the content on displays A-C instead of jumping to different virtual world regions:

"When turning the dial, I didn't notice that the background was changing"—P7.

Four sketches grouped content by type in a 2-D way with no virtual world, suggesting that a semantic grouping was important. Two in-room participants expressed that a spatial model was not necessary for the itinerary task:

"I… did not need to recall where the [displays] were located, since the four had very specific content that could be… identified as such" –P31.

Interface

Other factors also influenced the cognitive maps of our participants. The dial controller used to move between content categories in the itinerary task rotated clockwise and counter-clockwise; several participants cited this as an indication that the itinerary content was laid out in an arc or circle in the virtual world:

"When we rotated the dial, I assumed we were just rotating around. [The remote participant told] us they were all in a line, but that didn't seem intuitive to me, so I thought he was wrong" –P19.

This is an example of an interface element and layout cue *reinforcing* each other. Many participants also acknowledged the influence of the cove-like orientation of displays A-C:

"I was biased by how the [displays] were laid out in the

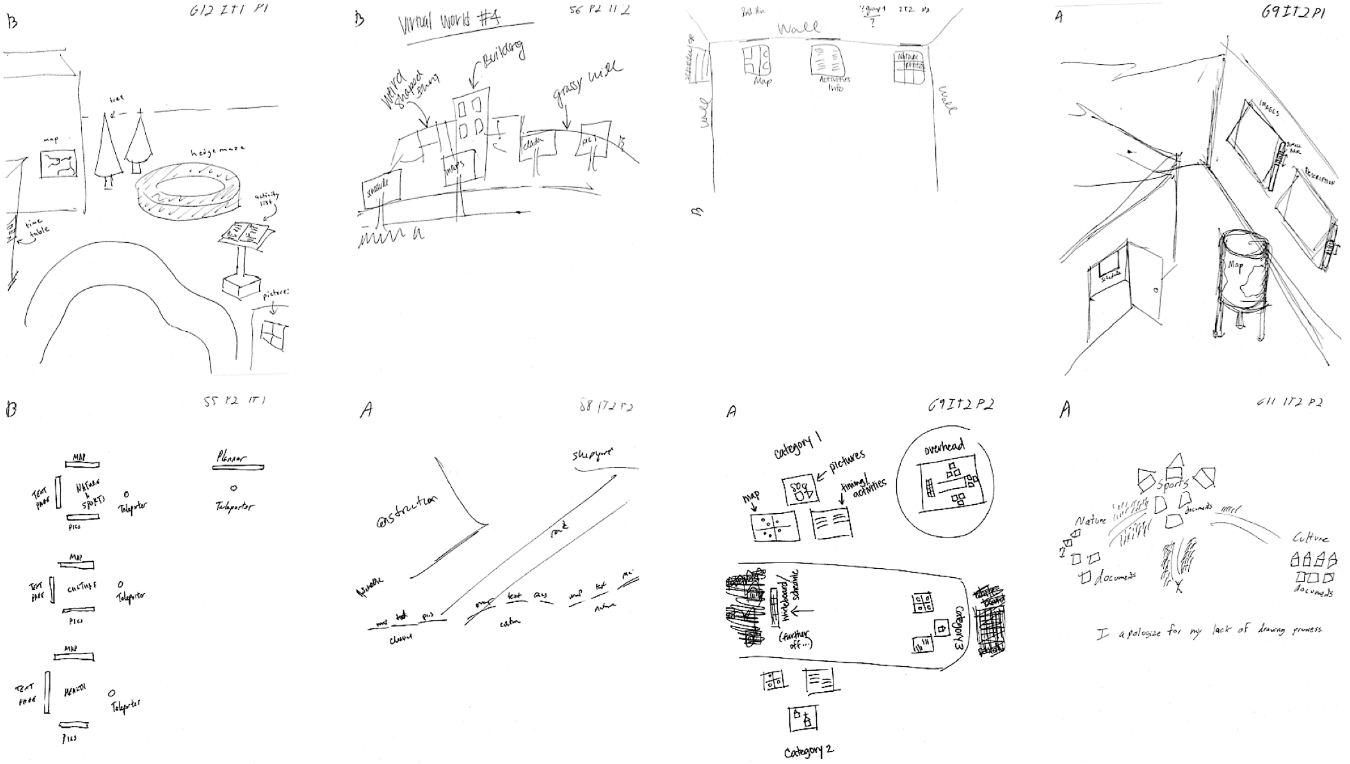

Figure 8 Post-condition sketches, itinerary activity. Top row: no indication that content categories were in different locations in the virtual world. Bottom row: more accurate depictions of how content was laid out.

room in my perception of how things were positioned in the world." – P26

In the image search activity, audio contact played a role in cognitive map formation. Several participants indicated in their sketches or during the interview that the content on displays without audio links was further away in-world, despite the fact that content was presented in l-r order and at the same scale across the in-room displays. This is an example of an interface element and layout cue *conflicting* with each other.

Virtual World

Occasionally, the remote collaborator's avatar would come into view on a display. This gave an indication of the remote participant's location, but only one group used the avatar to coordinate work, and only in image search, while one other group tried to use avatar movement to understand the virtual world layout. If the avatar transited between two displays this didn't necessarily indicate proximity in the virtual world due to the teleport links, and several groups noted this during the activities.

Finally, the virtual world background was sometimes used to infer relationships between the virtual world perspectives on each display. This varied from participant to participant:

> *"I would look at the background.. the only one [it] didn't work on is the one with the funky houses"—P19.*

> *"It wasn't until the last [condition] that I noticed that there was a background"—P25.*

Influence on Collaboration

Image Search

We saw no differences in collaboration style in the 2-table vs. 1-table layout image search conditions, corresponding with our cognitive map results. In-room collaborators could see all four screens while seated in either condition. Groups generally first established a scheme for staying in sync, often testing that locals and remotes could access all image sets, trying to determine whether there was a left-right correspondence, and labelling each image set.

As discussed, all groups used a left-right ordering scheme to refer to image sets. Other approaches were attempted on occasion, including references based on history ("the grid you were just at"), audibility ("the first one you can hear us at"), and image properties ("the image with the apple on the top-right"). Despite the simple document layout in image search and the prevalence of left-right ordering, this did not always structure participants' cognitive maps:

> *"We would use the first, second, third screen -- we were able to name them. We didn't know what his setup was, but since he never complained, it seemed okay"—P32.*

As mentioned, most groups struggled with the "audio holes" at displays B and D, adding reluctance to assume a l-r mapping. Conflicting cues also posed issues for the remote collaborator, particularly when the itinerary task had been done first: since the in-room participants could "jump" from category to category in itinerary and the remote collaborator could see where they were (and where they could be communicated with) based on the camera avatar, remote collaborators sometimes felt the same should be possible in image search ("can you move to set 4?"–P27), becoming confused by the stationary camera avatars in front of sets 1 and 3. The conflicting cues given by audio, camera avatars, and the display layout negatively impacted collaboration as well as cognitive map formation.

Itinerary

Synchronizing transitions between item categories was a source of confusion for some groups, due to the different interfaces used by remote and local collaborators. Remote collaborators usually saw the need to synchronize, since a video camera avatar indicated the item category their collaborators were currently at, while some in-room collaborators did not get the sense that there was movement at all, instead thinking that content was replaced as they turned the dial. Others just didn't seem to realize that coordination was required for transitions. For example, in one instance in-room collaborators in group 3 move from health to nature without saying anything to the remote collaborator, who notices and jumps to nature. When she gets there the in-room collaborators navigate back to health. She jumps again to health and says "hiii… " to re-engage. The local collaborators don't acknowledge, seemingly unaware that contact was lost. The changing virtual world background and the rotation of the dial were not strong enough cues for some in-room collaborators to realize that they were in fact shifting their virtual world location by using the dial.

Most groups used the item category names to manage transitions, although l-r ordering was also used. Again, different interfaces could be a source of confusion when coordinating transitions. In this excerpt, group 2 have finished selecting items in the culture category:

Remote: ok, so you want to go to the other ones?
Local 1: yeah, sure
Remote: ok… number 3 .. number 3
Local 1: number 3?
Remote: nature… nature
((local and remote collaborators transition simultaneously))
Local 2: ((garbled)) nature [((looking at item list))
[where's 3? …there is no 3
Remote [(looking at item list)
[ok, let's see what we have here, so…
Local 1: [there is no 3
Remote: 3? - no - no - I meant button number 3. Ok… sorry
I have a keyboard… which uh…"
Local 1: ok – which section are you on. Nature?
Remote: nature – nature… I can see you guys

The physical separation imposed by the 2-table condition impacted transitions between item selection and the calendar. Even though all groups selected six items before moving on to the calendar in both layout conditions, seven pairs of in-room collaborators moved back and forth between the calendar and the item content in the 1-table condition vs. one group in the 2-table condition. This clearer split between the two subtasks in the 2-table condition may have influenced (or have been influenced by) the Layout × Activity interaction on cognitive maps discussed earlier. Here, we see how physical layout can either solidify or blur the boundary between subtasks. Other cues can reinforce the boundary and facilitate transition: in the 2-table condition spatialized audio worked alongside layout to establish where the remote participant was during the transition; the only case where in-room collaborators returned to item selection in the 2-table condition was due to hearing the remote collaborator back in that region.

DISCUSSION

Our study explores how layout, activity, and other factors impact cognitive maps – but also sheds some light on how collaboration informs, and is informed by cognitive maps. In-room collaborators' cognitive maps converged when they used them to coordinate. In the 2-table itinerary condition participants were less likely to transition back and forth between calendar and item selection. This simple barrier coincided with a tendency to consider the calendar work area "elsewhere" in the virtual world. If cognitive maps are incomplete, additional time for grounding and coordination is required to ensure that work continues. This is made worse if cues conflict: in image search, collaborators struggled with the relationship between image sets that had a left-right correspondence but different audio connectivity. If cognitive maps are incorrect, coordination can suffer: in the itinerary task, some in-room collaborators would switch categories assuming that their remote collaborators would still be with them.

The results also suggest that when mixed reality is used for collaborative work in spaces like project rooms, the process of building cognitive maps does not strictly follow the physical – sensed – virtual sequence described by Dix et al. [6]. In-room collaborators did form a model of the virtual space based in part on its manifestation in the real world, but this was mediated in part by the room's layout and their egocentric perspective. For *artifact-centric* mixed reality, content semantics can influence cognitive maps, as collaborators use what is most effective to coordinate work [1]. The influence of task on mapping [10] was evident in our artifact-centric activities such that content used in the same task tended to be grouped together; this may also extend to other forms of mixed reality collaboration.

Implications for WYSIWIS In Collaborative Mixed Reality

The utility of real-virtual spatial correspondence can be considered at different scales within a project room. *Between work regions,* if the regions are allotted distinct tasks, and transitions between tasks are infrequent, then spatial correspondence is less important. In such cases it is less important to know "where" a collaborator is than "what" they are doing: permitting relaxed WYSIWIS approaches that concentrate only on connecting individual work regions without considering the room as a whole. Some basic spatial correspondence may still be useful, however: in our itinerary activity the spatialized audio was a useful cue that helped groups manage the transition between subtasks. This relaxed correspondence model is seen in the final inSpace lab setup (shown in Figure 2). If transitions are frequent, spatial correspondence may be more important, to facilitate communication and coordination.

Within a work region (and within a task), the structure and meaning of artifacts are paramount. Where artifacts can be categorized ("health, nature, culture"), spatial correspondence might be less important. If there is a 1-many mapping between an in-room work region and in-world (as with the item selection task in the itinerary activity), anchors for communication and coordination are needed. Document semantics is one possible anchor, as is correspondence in ordering. Clear cues should be available in both physical and virtual workspaces to facilitate coordination.

Within an active set of artifacts (i.e. content that is visible at the same time), if the individual documents can be easily categorized ("map, text, images"), spatial correspondence may be less important, although collaborators may expect that there is one. When there are no distinguishing markers (as in image search), order correspondence is more important. This may be spatial, but needn't be strict – e.g. a line of content in-world can map to a curve in-room.

We employed a form of relaxed WYSIWIS in our study: content placement followed the same left-right order as the displays, but their relative positions and orientations were not equivalent. In both activities, these differences didn't obviously impact collaboration – for example, in image search most groups recognized the left-right correspondence, using spatial language (e.g. "next grid to the right") when coordinating with each other and/or numbering the screens in left to right order. In general, as with other collaborative systems, designers of mixed reality collaborative systems need to prioritize visibility (of action and intent), and facilitate coordination. Spatial correspondence is one tool that must be considered among other aspects (e.g. task workflow, document content, interface cues) in achieving visibility and coordination support.

CONCLUSION

Our study demonstrates that the physical layout of a project room and the activity conducted within it impacts in-room collaborators' cognitive map of the connected virtual world, and that a range of other cues are used to understand real-virtual correspondence during tasks, including audio, input devices, document attributes and virtual world features. It was more important that collaborators could work together with content and coordinate next actions than to share an accurate model of the virtual space. However, conflicting or inadequate cues led to breakdowns in collaboration. Strict and relaxed WYSIWIS approaches are both feasible for mixed reality collaborative spaces, if designs consider the impact of cues influencing perception of the connected virtual space.

ACKNOWLEDGEMENTS

This work was supported by NSF grant IIS-0705569, NSERC grant 418612, and a research collaboration grant from Steelcase Inc.

REFERENCES

1. Aliakseyeu, D., Lucero, A., and Martens, J. Users' quest for an optimized representation of a multi-device space. Personal Ubiquitous Comput. 13, 8 (2009), 599-607.
2. Billinghurst, M., and Weghorst, S. The Use of Sketch Maps to Measure Cognitive Maps of Virtual Environments. In Proceedings of the Virtual Reality Annual International Symposium, (1995), pp 40-47.
3. Bly, S. A., Harrison, S. R., and Irwin, S. Media spaces: bringing people together in a video, audio, and computing environment. Commun. ACM 36, 1 (Jan. 1993), 28-46
4. Buxton, W. Living in Augmented Reality: Ubiquitous Media and Reactive Environments. In Finn, K., Sellen, A. and Wilber, S. (Eds). Video Mediated Communication: 363-384.
5. Colle, H. A., and Reid, G. B. The room effect: Metric spatial knowledge of local and separated regions. Presence:Teleoper.Virtual Environ.7(2)(1998),116–128.
6. Dix, A., Friday, A., Rodden, T., Koleva, B., Muller, H., Ranell, C. and Steed, A. Managing multiple spaces. In: Turner, Phil and Davenport, Elisabeth, (eds.) Spaces, Spatiality and Technology. Springer.
7. Golledge, R.G. Methods and Methodological Issues in Environmental Cognition Research. In *Environmental Knowing,* (Golledge, R.G. and Moore, G.T. Eds.), Dowden, Hutchinson and Ross, Inc., 1976, pp. 300- 313.
8. Harrison, S. and Dourish. P. Re-place-ing space: the roles of place and space in collaborative systems. CSCW 1996: 67-76.
9. Hindmarsh, J., Fraser, M., Heath, C., Benford, S., and Greenhalgh, C.. Fragmented interaction: establishing mutual orientation in virtual environments. CSCW 1998: 217-226.
10. Hommel, B., and Knuf, L. Action related determinants of spatial coding in perception and memory. In Spatial Cognition II, Integrating Abstract Theories, Empirical Studies, Formal Methods, and Practical Applications (2000), Springer-Verlag, pp. 387–398.
11. Huang, E., Mynatt, E., Trimble, J. Displays in the Wild: Understanding the Dynamics and Evolution of a Display Ecology. Pervasive 2006.
12. Huang, J. and Waldvogel, M. The swisshouse: an inhabitable interface for connecting nations. DIS 2004.
13. Hunter, S., Maes, P., Tang, T., Inkpen, K.M. and Hessey, S.M. 2014. WaaZam!: supporting creative play at a distance in customized video environments. CHI 2014: 1197-1206.
14. Luff, P., Heath, C., Kuzuoka, H., Yamazaki, K. and Yamashita, J. Handling documents and discriminating objects in hybrid spaces. CHI 2006: 561-570.
15. Mark, D., Freksa, C., Hirtle, S., Lloyd, R., and Tversky, B. Cognitive models of geographic space. International Journal of Geographical Information Science 13, 8 (1999), 747–774.
16. Newcombe, N. Methods for the Study of Spatial Cognition. In *The Development of Spatial Cognition,* (Cohen, R. Ed.), Hillsdale, New Jersey, Lawrence Erlbaum Associates, 1985, pp. 277-300.
17. Olson, G. M. and Olson, J. S. Distance matters. Hum.Comput. Interact. 15, 2 (Sep. 2000), 139-178.
18. Reilly, D., Mathiasen, N., Salimian, M., Edwards, W.K., Franz, J. and MacKay, B. (2014) SecSpace: Prototyping Usable Privacy and Security for Mixed Reality Collaborative Environments. EICS 2014.
19. Reilly, D., Rouzati, H., Wu, A., Brudvik, J., Hwang, J. Y., and Edwards, W. K. TwinSpace: an Infrastructure for Cross-reality Team Spaces. UIST 2010.
20. Reilly, D., Voida, S., McKeon, M., Le Dantec, C., Edwards, W. K., Mynatt, E. and Mazalek, A. Space Matters: Physical-Digital and Physical-Virtual Codesign in inSpace, IEEE Pervasive Computing 9(3):54-63, 2010
21. Stefik, M., Bobrow, D. G., Foster, G., Lanning, S. and Tatar, D.(1987) WYSIWIS revised: early experiences with multiuser interfaces. *ACM Trans. Inf. Syst.* 5:2, 147-167.
22. Streitz, N. A., Geißler, J., and Holmer, T. 1998. Roomware for Cooperative Buildings: Integrated Design of Architectural Spaces and Information Spaces. 1st intl. Workshop on Cooperative Buildings, London, 4-21.
23. Tang, A., Boyle, M., and Greenberg, S. (2004) Display and presence disparity in Mixed Presence Groupware. OzCHI 2004:73-82.
24. Teasley, S., Covi, L., Krishnan, M., and Olson, J. (2000) How does radical collocation help a team succeed? CSCW 2000: 339-346
25. Tversky, B. (1993) Cognitive maps, cognitive collages, and spatial mental models. In COSIT 1993:4-24.