Graphics Interface Conference 2015, 3–5 June, Halifax, Nova Scotia, Canada

Effects of Arm Embodiment on Implicit Coordination, Co-Presence, and Awareness in Mixed-Focus Distributed Tabletop Tasks

Andre Doucette, Carl Gutwin, and Regan Mandryk
University of Saskatchewan, Canada

Abstract

Mixed-focus collaboration occurs when people work on individual tasks in a shared space – and although their tasks may not be directly linked, they still need to maintain awareness and manage access to shared resources. This kind of collaboration is common on tables, where people often use the same space to carry out work that is only loosely coupled. At physical tables, people easily manage to coordinate access to the table surface and the artifacts on it, because people have years of experience interacting around other physical bodies. At distributed digital tabletops, however, where there is no physical body for the remote person, many of the natural cues used to manage mixed-focus collaboration are missing. To compensate, distributed groupware often uses digital embodiments. On digital touch tables, however, we know little about how these embodiments affect coordination and awareness. We carried out an empirical study of how four factors in an arm embodiment (transparency, input technique, visual fidelity, and tactile feedback) affected implicit coordination, awareness, and co-presence. We found that although some embodiments affected subjective feelings of co-presence or awareness, there were no changes in table behavior – people acted as if the other person did not exist. These findings show the possibilities and limitations of digital arm embodiments, and suggest that the natural advantages of tables for collaboration may not extend to distributed tables.

Keywords: Embodiments; digital tabletops; awareness.

1 Introduction

Mixed-focus collaboration – where people move back and forth between individual tasks and shared activities – is a primary way that group work occurs in the real world [15]. In mixed-focus settings, even when people are carrying out parallel or individual tasks, they still need to maintain awareness of others' work, in order to manage access to the workspace, to coordinate the use of shared artifacts in the space, and to keep track of the other person's progress [7]. This is particularly true on tables, which are one of the most common settings for mixed-focus collaboration. Tables provide a natural environment for group work, but many of their benefits are based on people's expertise in working with and gathering information from the arms, hands, and bodies on and around the table. This information is critical to the success of tabletop work, even when carrying out individual tasks – because people still share both the physical space of the table and many of the tools and artifacts on the surface.

Part of our expertise in these kinds of physical interactions arises from the many social rules that govern and guide touch and close-proximity interactions, learned through years of experience. Rules of personal space, for example, reduce behaviors such as stealing items from another person's work area, or interfering with other people by occluding their workspace or physically bumping into them [17]. These rules are also useful for guiding a group's close-proximity behavior, providing means for automatic coordination.

* [andre.doucette,carl.gutwin,regan.mandryk]@usask.ca

When people work on distributed tables, however, the other person's physical body is absent, and so we lose the main source of information for awareness, coordination, and social protocols. Without the information produced by the other person's body in the shared space, it becomes more difficult to stay aware of what others are doing, and more difficult to coordinate actions and access to shared items – leading to duplicated tasks and more conflicts (e.g., grabbing the same item).

In an attempt to replace the missing co-present body, designers of distributed tables represent remote participants through digital embodiments, such as cursors or virtual arms. These embodiments convey some level of information about the remote collaborator's actions in the shared space. However, digital embodiments are poor replacements for physical co-present bodies, because the social protocols that govern interaction often do not work with virtual representations. Digital embodiments are much less noticeable than real bodies [30], and rules about touch avoidance often do not hold with digital arm embodiments, even in co-located settings [5].

If distributed tabletop systems are to re-enable people's expertise in physical bodily interaction – something that is important even for the parallel individual work of mixed-focus collaboration – we need to understand how the design of embodiments affects awareness, coordination, and co-presence. There are several factors in an embodiment that could change its effects on these qualities – e.g., whether the embodiment uses touch or mouse input, the visual fidelity of the embodiment compared to real arms, whether the embodiment provides tactile feedback, and the degree to which the embodiment occludes the workspace [5,6].

Tabletop arm embodiments have been studied in co-located scenarios [30, 5], but little is known about them in distributed systems. To provide this information, we studied four design factors (visual fidelity, occlusion, input technique, and tactile feedback) in a controlled study. Pairs of people carried out a mixed-focus task across two networked tables. Although the individual tasks were not strongly coupled, they used the same shared workspace and task artifacts – a similar activity, for example, to building different parts of the same puzzle. We chose a loosely-coupled task intentionally, because coordination in these tasks is subtler and less overt – in contrast, tasks with explicit requirements for tightly-coupled coordination will often be carried out through verbal communication and explicit task structures or roles.

In our tasks, participants were represented on the other table with different arm embodiments. Visual representations included: a picture of the participant's arm, a translucent picture arm, and a video arm that showed live video of arm movements. Participants controlled these embodiments using either direct touch or a mouse, and for some embodiments we included a tactile feedback device that buzzed whenever the arms crossed. We also included one co-located condition where participants worked at the same table, and used their physical arms and touch input to manipulate artifacts.

We gathered several measures to investigate how the design of the embodiment affected participants' coordination in accessing the table (e.g., the number of times that people reached over one another), people's level of awareness (e.g., self-reports of noticing the other embodiment), and people's level of co-presence (i.e., the degree to which it felt that the other person was in the same room).

Our results show that although participants felt that some of the embodiments provided them with increased co-presence and awareness, there were no changes in the way that people acted on the tabletop. People acted essentially as if there was no other person in the space – even with live video of the other person's arm, and even with vibration feedback on arm crossing, the awareness and social protocols that are so easily evident in co-present work were completely missing in the distributed setting.

These results provide important information for designers of distributed tabletop systems. Distributed tables have been seen as a way to recreate some of the natural and facile collaborative behaviors that are seen when people work face to face – but our findings show that for the common scenario of mixed-focus collaboration, people behave as if there is no-one else at the table. This means that designers will be less able to depend on social protocols as a way of managing group processes and access to shared resources. Although it is also possible that people will be able to work more quickly when they do not need to worry about the other person in the space, these issues complicate the intended use of distributed tables as natural sites for collaboration.

2 Related Work

2.1 Mixed Focus Collaboration

CSCW researchers have noted that most group work in the real world is not solely composed of tightly-coupled shared activity (e.g., [7,15,28]). Instead, people move back and forth from individual tasks to periods of shared focus; this mixed-focus style is common, for example, when people divide labour, carry out assigned jobs individually, and then gather to merge their results.

Mixed-focus collaboration requires that people maintain a certain level of awareness of others' activities, even when tasks are only loosely coupled. When mixed-focus work occurs in the same location, people must coordinate access to space, shared tools, or artifacts that may be needed by multiple people. For example, reaching into the same space causes interference [41], and trying to take the same artifact causes a resource conflict. This meta-level activity can be considered as a type of "articulation work" [31] that is required regardless of the collaborative nature of the individual tasks. In co-located situations, however, this articulation work happens naturally and easily, due to people's long experience working near other people.

Researchers have also considered ways of supporting mixed-focus work for distributed collaborators. For example, change visualizations and enhancements to embodiments (see below) can help people to keep track of what has happened in the space, even if they have not been paying close attention. On tabletops, different directions have also been explored – e.g., providing people with separate views in order to provide better support for loosely-coupled (or even uncoupled) work on the same table.

2.2 Distributed Embodiments

When people are physically distributed, their physical bodies do not occupy the same space. Distributed shared digital spaces can connect remote users, providing a shared visual space that helps groups coordinate their actions by making the state of the task and others' actions visible [12,27]. It is common for distributed systems to represent the other person through a digital embodiment, a visual representation of remote people [1]. Research in distributed embodiments focuses on the transmission and interpretation of gestures as a means of communication (e.g., [7,10,11,19,25,26]). We are more interested in interactions in the shared space, where digital embodiments not only represent people's communicative gestures, but also their coordinative interaction.

Researchers have investigated different kinds of digital embodiments. Telepointers, the simplest embodiments, represent other people's locations with shapes and colours [13,15], and can be augmented with additional user information [34]. Though researchers identified that video loses much of the information of 3D interactions because it is projected onto a flat 2D display [10], many systems have provided richer embodiment visualizations with video [21,22,26,37], typically overlaying the remote user's video stream over the local workspace. A more recent technique uses digital video and masking to remove the background, leaving just the digital arms [15,35,36,39]. We know little, however of how people actually use and interpret these video arm embodiments when co-interacting in a shared spaces. One study of a distributed tabletop task that used Video Arm Shadows [38] showed that when co-located, people avoided occluding the other person with their arms, but when distributed, people regularly occluded one another without any verbal comment.

2.3 Digital Personal Space and Mediated Touch

Research in avatar-based systems suggests that people extend their own personal space [17] to surround their avatars (e.g., [23,31, 33]), and avoid invading the personal space of other's avatars. Researchers have shown that other embodiments do not necessarily convey the same social rules as physical bodies. For example, in a collocated system, people touch and cross digital arm embodiments, regardless of their visual design [5], something avoided when interacting with their physical arms. By augmenting the digital arm embodiments with touching feedback, researchers have shown that augmentations can cause people to treat digital arm embodiments more like physical arms by avoiding touching others [6]; however, little is known of how people interpret arm embodiments in distributed systems. Researchers have shown that digital arms may provide a mechanism for communicating intimacy through metaphorical touch [40], though other researchers have shown that distribution may change physical social protocols (e.g., people sit "in each other's lap" without issue [36]).

2.4 Social Presence and SoE in Distributed Systems

Distributed systems are more impoverished than collocated systems due to the lack of physically co-present bodies. First, the distance changes people's interactions and their feelings of sharing the same space [28]. Second, people miss simple physical cues that help inform others' actions because they are represented through an embodiment instead of their physical bodies [13]. Researchers have attempted to increase feelings of social presence (co-presence) – the sense of being with another in a mediated system ([2,9]).

A separate issue is whether the digital embodiments *are* the person. Sense of Embodiment (SoE) is when "…some properties of [an embodiment] are processed in the same way as the properties of one's body" ([4], p.3). It encompasses sensations of "being inside, having, and controlling a body" ([24], p.374), and has three components: sense of self-location (I'm *inside* the embodiment), sense of agency (I'm *controlling* the embodiment), and a sense of body ownership (the embodiment *is part* of my body) [24].

3 The Study

To understand how the design of distributed arm embodiments affects coordination, co-presence, and awareness in mixed-focus distributed tabletops, we carried out a controlled experiment. Based on previous work on co-located arm embodiments [5,6,30], we investigated four embodiment design factors.

1. *Occlusion*: the degree to which an arm embodiment blocks the view of objects underneath it.
2. *Input*: the input technique (e.g., direct touch or mouse) used to control the embodiment.
3. *Visual fidelity*: the degree to which the embodiment conveys the appearance and behavior of the real arm.

4. *Tactile feedback*: whether touching an arm embodiment provides a tactile sensation (e.g., through vibration).

3.1 Task and Study System

The mixed-focus task used in the study was the poem-building task used by Doucette et al. [5,6] for earlier arm embodiment research. This task involves people carrying out individual activities, but in the same workspace and using the same artifacts. Therefore, the requirements for coordination and awareness are based on the meta-activity of managing access to space and objects, rather than on the task itself. The task is similar in this regard to many mixed-focus tasks that can be carried out using a divide and conquer strategy (however, we focus here on the parallel-work phase of the activity, rather than on the shared-focus phase).

In the task, dyads sat side-by-side at a tabletop and created haiku poems about an assigned topic from a set of shared words on the tabletop. There were two "haiku papers" on which the poems were built – one in front of each person. Topic word locations were switched, such that the words on each side of the table were more appropriate for the haiku on the other side of the table [5]; this meant that people had to reach to the other side of the table, and were required to manage access to the shared space.

We developed a distributed table system for the study that linked two tables in different rooms across a network. The tables used 60" Sony HDTVs with PQ Labs multi-touch overlays, and the system allowed direct touch and mouse input. To ensure that all words were reachable while seated, people sat on the short side of the table, and the system used only the half of the display closest to their location (Figure 6). Skype was used for a voice connection. Experimental instructions were described over the Skype connection, so participants were aware they could speak freely with the remote participant.

3.2 Participants and Conditions

We tested 17 pairs, removing two outlier groups because these groups did not complete the task as instructed. Of the 30 remaining participants, 18 were men, median age was 24 years, and 16 reported English as their first language. Participants were paired with a stranger, which was intentional since previous work has shown that explicit management of a shared space is more pronounced with strangers 5. Gender pairings were: 5 male-male, 8 female-male, and 2 female-female.

We designed and evaluated five digital arm embodiments that instantiated our four design factors. We compared these distributed arm embodiments to each other, and also to a co-located touch-input condition. The embodiments were:

- *Transparent*: Showed an outline of the participant's actual physical arm, filled with purple or green and set at 70% opacity. The mouse controlled the tip of the embodiment's finger.
- *PictureMouse*: Showed the same outline as *Transparent*, but with the actual visual image of the participant's arm, at full opacity. This embodiment was also controlled with the mouse.
- *PictureArm*: Used the same arm as PictureMouse, but was controlled using direct touch: the tip of a person's physical arm (tracked using a Kinect) controls the tip of the embodiment finger. The "base" of the embodiment was fixed to the right side of their haiku paper.
- *VideoArm*: Showed live video of the participant's arm (which is more realistic than a picture, as people can articulate fingers, wrist, and elbow). We implemented a version of VideoArms [35] using KinectArms [11]. The embodiment's base moved with the participant's physical body, adding realism.
- *VideoArmVibe*: Showed the same visual representation as the VideoArm, but added tactile feedback when people touched embodiments. The effect was implemented using a vibrating box placed in each person's front pants pocket, following [6].
- *Co-located*: At the end of the study, groups completed one additional haiku while co-located and using touch input, providing a baseline measure of physical reaching behaviour for each group.

3.2.1 Embodiment latency

The KinectArms toolkit introduces latency in the display of the video image, due to the Kinect hardware, the video-manipulation software, and network transmission. We calculated latency through video analysis of a reciprocal movement task, and local latency was recorded as the time between a finger-down event to the moment when the embodiment arrived at the down location. VideoArms had the largest local latency (500ms), well above the threshold of noticeability [8]. The video processing for VideoArms also adds network lag of around a second. The PictureArm embodiment added a fifth of a second of local latency, and no additional network lag as compared to mouse-based techniques.

Table 1. Approximate system latency times

Latency	**Transparent**	**Picture Mouse**	**Picture Arm**	**Video Arm**
End-to-end	1050ms	1000ms	950ms	1300ms
Local	<100ms	<100ms	200ms	500ms

3.3 Design, Measures and Statistics

The study was a within-subjects design; all groups saw the entire set of embodiments (Latin-square counterbalanced). Groups were told they would complete a co-located haiku at the end of the study.

We use both quantitative and qualitative analyses to answer our research questions. We investigated people's explicit coordination in the table's shared space by recording the number of times their arms crossed. When reaching physically over a tabletop, people avoid crossing over and under other people's physical arms. People's aversion to interrupting other people provides an avenue for fast and automatic coordination such as taking turns or backing off when another person reaches into the space [5,6]. People's ability to avoid crossing embodiments also demonstrates an increase in awareness of the other person's actions [6].

In addition to this data collected automatically through log files produced by the system, we also investigated how embodiment design affects the sense of co-presence and subjective awareness of action through questionnaires [5,8].

3.3.1 Planned comparisons

We investigate the four design factors by comparing one pair of embodiments for each factor:

1. *Occlusion*: Transparent (partial occlusion of objects under the arm) vs. PictureMouse (complete occlusion)
2. *Input*: PictureMouse (mouse input) vs. PictureArm (touch input);
3. *Visual fidelity*: PictureArm (static) vs. VideoArm (live video);
4. *Tactile feedback*: VideoArm (no vibration) to VideoArmVibe (vibration when embodiments touch).

3.3.2 Quantitative analyses

The system recorded the number of times people crossed embodiments, which we use as a proxy of people's explicit coordination. The number of crossings is analysed through an RM-ANOVA (α=.05), using the Greenhouse-Geisser method to compensate for sphericity violations.

The mouse-input arm embodiments cannot bend (e.g., at the elbow), and thus a crossing event was triggered when the straight lines running through each embodiment crossed (see Figure 1 left, crossing lines denoted in red). The VideoArm embodiments allow people to move their shoulders (where the crossing line begins), as well as bend their elbows, wrist, or fingers. Thus, the crossing line may no longer even be within the arm embodiment (see Figure 1 right). A crossing event with VideoArms is triggered when any part of the two arm embodiments overlap. In principle, this is an over-count compared to the mouse-based arm embodiments, as there are "touches" of VideoArms that would not be a "crossing" with mouse-based arm embodiments (e.g., Figure 1 right). In summary, with touch-input embodiments, we count the number of times the visual embodiments intersect as a crossing, whereas with mouse-based embodiments, we count the number of times the lines running through the embodiments intersect as a crossing.

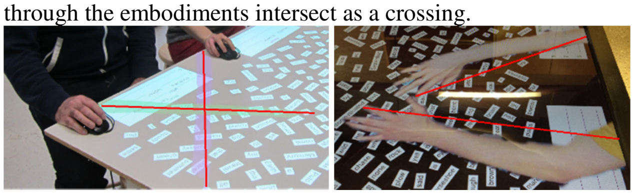

Figure 1: Crossings with arm embodiments. A crossing with Transparent embodiments (left), and a crossing of VideoArms without crossing lines (right).

We also collected subjective responses to questionnaires through 7-point Likert-style questions (from Strongly Disagree to Strongly Agree). The responses were analyzed using non-parametric analyses. We used Friedman tests to establish main effects, and use Wilcoxon Signed Ranks for our planned pairwise comparisons. There were two surveys:

- A *between-conditions questionnaire* collected people's feelings of co-presence; we asked questions about sharing the same space and questions related to the sense of embodiment (control of the arm and the sense of being *in* the arm).
- A *post-experiment questionnaire* collected people's feelings of awkwardness, their subjective awareness of positions and actions, and the subjective similarity to interacting with a collaborator at the same table.

3.3.3 Qualitative analyses

We video recorded each session and finished each session with a semi-structured interview. The videos were used as exploratory and explanatory analyses of a group's behaviours. Post-experiment, semi-structured interviews were used to follow up on observations from the sessions. Interview questions asked people to directly compare embodiments (e.g., picture to video and touch-based to mouse-based interaction), and to describe their sense of embodiment. Groups were asked "*Did it seem like the other person's embodiment was them?*" and "*Did your own embodiment seem like it was you?*"

3.3.4 Co-located condition

We include the co-located condition to provide a benchmark for the reader to compare the distributed conditions against. The results from the co-located condition are not used in any statistical analyses, as this condition is not included in any planned comparisons to answer our research questions (this condition was extensively explored in [5] and [6]). We also use the co-located video for video analyses.

4 Results

We report on our analyses of the effects of the four factors (occlusion, input, visual fidelity, and tactile feedback), grouped by coordination, co-presence, and awareness.

4.1 Coordination

Mixed-focus collaboration requires that people be able to coordinate access to the shared artifacts on the table, and at physical tables this is accomplished partly through being cautious about crossing the other person's arm. Therefore, we studied the effects of arm embodiment design on coordination by looking at people's willingness to cross embodiments (originally studied in [5]), coupled with observation of coordinated actions and people's subjective responses to questionnaires.

4.1.1 Crossings analysis

There was a main effect of embodiment on the number of crossing events ($F_{(2.25,31.45)}$=6.680, p=0.003, η^2=0.323, adjusted for sphericity using Greenhouse-Geisser). The pairwise comparisons in Figure 2 show there was an effect of *Input* (p=0.014): people cross less with touch input than with mouse input. All other comparisons showed no significant difference (all p>0.05).

Figure 2 shows a split between touch input and mouse input. People seem to cross more with the mouse than when interacting with direct touch. We observed little evidence of people coordinating more to avoid crossing with touch than mouse input, so we investigated whether the difference between touch and mouse input can been explained by different reaching behaviours, as crosses typically only occur during the reaching gestures.

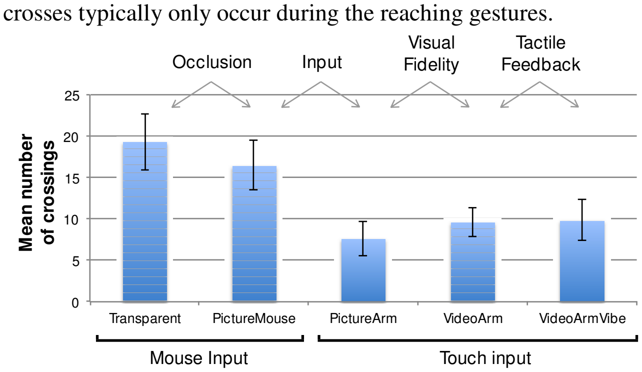

Figure 2: Mean number of crossings (± s.e.), with conditions grouped by input type (below) and question (above)

4.1.2 Follow-up reaching analysis

To explain the difference between touch and mouse input, we performed follow-up analyses on reaching behaviour.

One reason people cross less with touch input may be that there are fewer opportunities to cross. For example, if people reach fewer times, there will be fewer opportunities to cross. We performed a follow-up RM-ANOVA on the number of reaches and found there was no main effect of embodiment on the number of times people reached past their haiku papers (p>0.05). As shown in Figure 3 there was no overall effect of input on the simple number of times people reached for words. As the frequency of reaches does not explain the difference in crossings, we performed a second analysis on the reach durations.

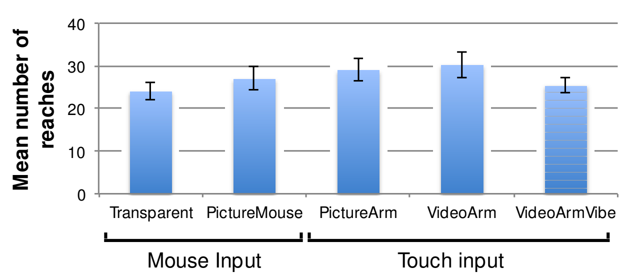

Figure 3: Mean number of reaches past haiku papers

There was a main effect of embodiment on the proportion of time spent reaching past the haiku papers ($F_{(4,56)}$=68.85, p≈0.000, η^2=0.831). The pairwise comparisons in Figure 4 show that there was a significant difference for the *Input* factor (p≈0.000), but no significant difference for the other factors (all p>0.05). As shown in Figure 4, people spent a larger proportion of time with their cursor (i.e., their embodiment's fingertip) past the haiku papers with mouse input than with touch input (discussed further next).

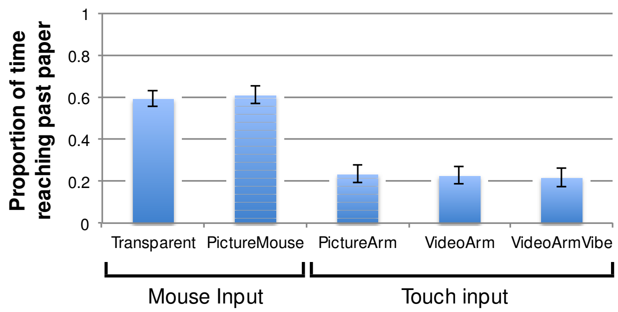

Figure 4: Proportion of time spent past haiku papers

4.1.3 Physical resting position (video analysis)

When using touch input, people generally did not leave their arms extended on the table, except when reaching for words. With mouse input, people often scanned the surface with their arm embodiment while looking for words [5]; this behaviour was observed only once with touch input. In addition, when they were done building their haiku, people sometimes flicked their mouse out, leaving their embodiment stretched out while the other person finished.

With physical arms, people had a natural arm resting position on the bezel near their haiku paper. We suspect this resting position, as well as scanning and flicked out behaviours when using mice, explain why people spent more time reached out with mouse input embodiments than with touch input embodiments (Figure 4), and contributed to the difference in the number of crossings (Figure 2).

4.1.4 Observations of coordination (video analysis)

In general, we observed very little evidence of people explicitly coordinating their reaching gestures: people just reached for the object they wanted. This mirrors previously reported results [5]. In a few cases in the vibration condition, participants appeared to consider the other person's location, but often this coordination seemed to be as a reaction to the vibrations, not to prevent the cross or vibration. People would respond to the vibration by pulling their arms back and monitoring what the other person was doing, but did little to predict when the initial vibration may occur (in contrast to previous research on vibration in co-located reaching [6]).

In summary, people cross more often with mouse-based embodiments than when physically reaching (touch input), likely because there are fewer opportunities to cross with physical input. In all conditions, people reached for words with the same frequency, but there is a substantial difference in the proportion of total time people spent reached out. With mouse-based input, people often scanned the surface of the table with their embodiment while searching for words and flicked their mouse out after finishing their haiku, leaving their embodiment stretched out over the tabletop. These behaviours contributed to the differences in the number of crosses between mouse and touch input.

4.2 Co-presence

We study the effect of arm embodiment design on co-presence by people's subjective questionnaire responses, coupled with observations of their body movements.

4.2.1 Sense of being in the same space (questionnaire)

Figure 5 shows agreement ratings to the statement "*I had a sense that I was in the same space as my partner*" from the between-conditions questionnaire. A Friedman test showed a main effect of embodiment on participants' sense of sharing the space (χ^2_4=29.26, *p*≈0.000). Wilcoxon Signed Ranks tests showed an effect of *Visual fidelity* (Z=-2.63, *p*=0.009) and *Tactile Feedback* (Z=-2.96, p=0.003): people had a greater sense of sharing the same space with video and with vibrations. There was a marginal effect of *Occlusion* (Z=-1.86, p=0.063) (elevated sense of sharing the space with occluding embodiments). There was no effect of *Input* (*p*>0.05).

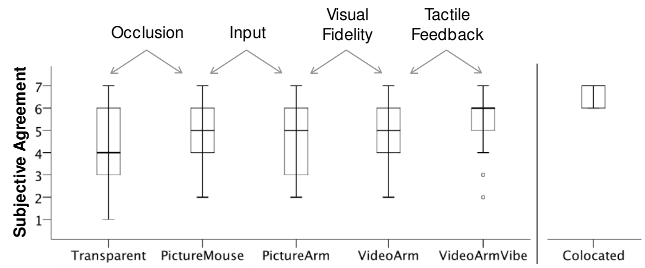

Figure 5: Subjective sense of being in the same space (dots are outliers, box bounds are upper and lower quartile with median as cross bar, and whiskers are min and max non-outliers)

4.2.2 "No other person there" (video analysis)

Overall, there was only a single vocalization that was intended for the other person over the 15 sessions (people sometimes spoke to the co-located researcher). During an occlusion incident, the person being occluded was trying to see under the other's embodiment, and vocalized an "umm" to get the other's attention (audible to the other person through Skype). The person occluding had no reaction, and continued their interaction as if nothing was wrong.

4.2.3 Use of horizontal space (video analysis)

When collocated, each person used about half the horizontal bezel space to avoid encroaching on the other person's personal space (Figure 6, bottom). When distributed, people on the right stretched out on the bezel, suggesting people had little feeling that they were in the other person's personal space (Figure 6, top) – note that people on the left stretched less because they used their mouse with their right hand. This behaviour is similar to previous work showing people had little issue sitting "in each others' laps" [36].

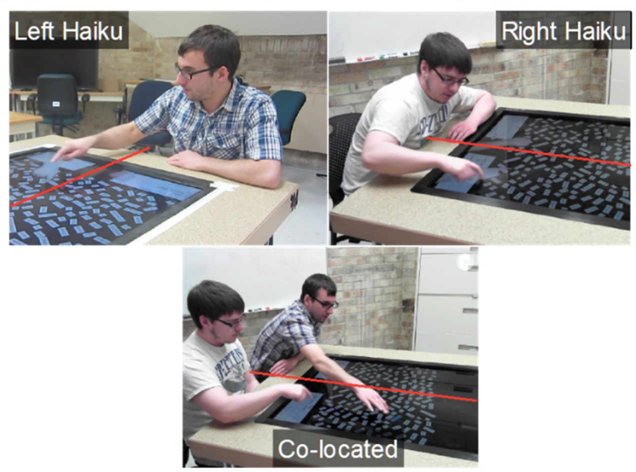

Figure 6: People's horizontal size when distributed (top) and when collocated (bottom). Lines split the table in half, showing how people stretch to the other side when distributed

4.2.4 Similarity to interacting at the same table

Figure 7 shows agreement ratings to the statement "*This embodiment was similar to interacting at the same table*" from the post-experiment questionnaire. A Friedman test showed a main effect of embodiment (χ^2_4=69.94, *p*≈0.000). Wilcoxon Signed Ranks tests showed an effect of *Input* (Z=-2.38, *p*=0.017), *Visual fidelity* (Z=-2.96, *p*=0.003), and *Tactile feedback* (Z=-4.05,

p≈0.000): people felt that the distributed embodiments were more similar to interacting at the same table with physical input, video, and vibrations.

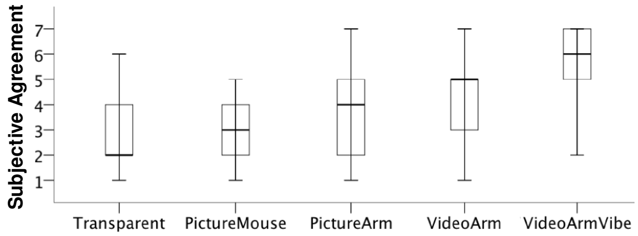

Figure 7: Responses to "similar to interacting at the same table"

4.2.5 Realness of VideoArms (interviews)

Nine people reported that VideoArms were the most realistic and the most like the participants' real bodies. For example, one person stated, "The video, it seemed more real, I thought about using my second arm as well." Another participant said: "It's better with live video than pictures. It is much more normal, more comfortable."

Subjectively, people reported that VideoArms were the most real and they treated them the most like physical arms. This suggests that people have a higher sense of embodiment 4,24 with video than with lower fidelity embodiments; however, people ignored their partner, and acted as if the other person was not even there. People freely crossed the other's embodiment, and occluded areas where the other person was interacting, suggesting the other person was not embodied in their remote arm embodiment.

4.2.6 Vibration reminded me of other person (interviews)

Doucette et al. showed that, in a co-located system, tactile feedback typically caused groups to begin coordinating in order to avoid crossing embodiments 6. In the distributed system, it appears that people do not actively try to avoid crossing, and only coordinate when *reminded* of the other person, through the tactile vibrations.

As one participant said, "Before the vibrating thing, I didn't even notice you're there; I just do my own work." Similarly, one group stated, "The vibrating one, you kind of noticed where their arm was" and "Yeah, other than the vibrating one, I didn't even pay attention to where her arm was."

4.2.7 Summary of Co-presence Results

People reported a greater sense of sharing the space with video and with vibrations, but this space may not be the physical space where the remote person "is". People completely ignored their partner, physically occupying the space where the other person would be. People felt the distributed embodiments were more similar to interacting at the same table with physical input and video. They reported higher feelings of "realness" of the VideoArms, but there are no substantial differences in behaviour by adding video.

The vibrations were interpreted very differently than in previous work 6. People ignored the other person, and made little effort to coordinate reaching. People reacted to the vibrations, but made no effort to track the other person to avoid a cross – instead, the vibrations just reminded them that the other person was there.

Overall, people reported higher feelings of co-presence with video and with vibrations. This co-presence did not extend to the local physical space where the body represented by the arm embodiment would be; there was little evidence that people thought they were co-interacting with another person.

4.3 Group Awareness

Mixed-focus collaboration has strong requirements for group awareness, even when people are carrying out individual tasks. We asked participants to rate their awareness; Figure 8 shows agreement ratings to the statement "*I was aware of my partner's actions on the table*" from the post-experiment questionnaire. A Friedman test showed a main effect of embodiment on participants' feelings of awareness of action (χ^2_4=39.75, p≈0.000). Wilcoxon tests showed that people felt more aware with *Visual fidelity* (Z=-2.31, p=0.021) and *Tactile feedback* (Z=-3.17, p=0.002).

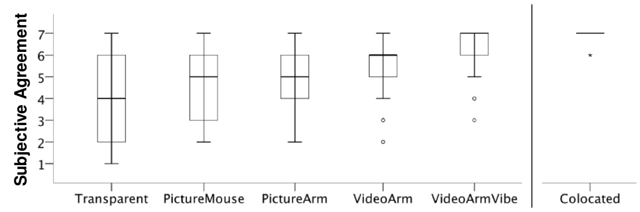

Figure 8: Subjective awareness of partner's action

Figure 9 shows agreement ratings to the statement "*It was awkward to cross my partner's embodiment*" from the post-experiment questionnaire. A Friedman test showed a main effect of embodiment on participants' feelings of crossing awkwardness (χ^2_4=49.25, p≈0.000). Wilcoxon tests showed that people felt more awkward with *Visual fidelity* (Z=-2.72, p=0.007) and *Tactile feedback* (Z=-4.35, p≈0.000).

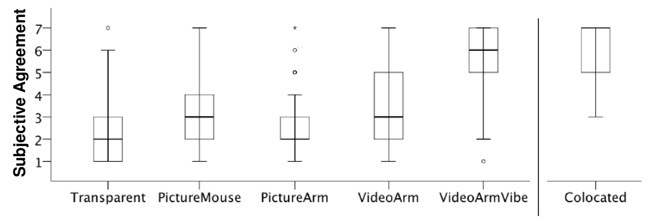

Figure 9: Subjective feelings of awkwardness to cross

5 Discussion

Our main result is that although video embodiments were subjectively preferred over simpler visual embodiments and increased people's sense of co-presence, there was very little effect on people's implicit coordinative behaviours – that is, on the subtle ways in which people manage access to a shared table space. This result is shown in several different ways: we found that participants generally ignored the remote person, freely occluding their personal workspace and crossing their embodiment. In addition, we found that there were substantial differences in how people use touch- and mouse-based embodiments, spending less time with their arm reached out into public space with touch-based input.

5.1.1 Interpretation of results

Although people reported feeling that visual fidelity and tactile feedback increased their sense of a present collaborator, people did not coordinate their actions to avoid crossing a remote mouse-based arm embodiment: instead, people reached as needed for their individual task, regardless of the location of the other person's embodiment. This follows results previously shown for co-located mouse-based arm embodiments 5. However, distributed tactile feedback did not replicate results previous shown for co-located arm embodiments (in which tactile feedback reduced crossing) 6. In general, we observed little effort to coordinate reaching, regardless of visual embodiment or input type. This also contradicts studies by Tuddenham and colleagues [38], who showed that people did use a remote embodiment for coordination – however, this was in a tightly-coupled design task.

Although additional research is needed to explore these issues further, our results suggest that one of the main reasons for using a table as a setting for collaboration – that is, that people already know how to work at tables – may not hold true for distributed table systems and mixed-focus collaboration. The results were more extreme than for any previous studies – none of the different embodiment designs made any difference to behavior in our loosely-coupled task, and it was clear that people acted as though the other person did not even exist. This suggests that it is may be more difficult to support natural awareness mechanisms and well-practiced social protocols at distributed tables than in other remote-groupware settings, because tabletop collaboration may be more dependent on these subtle cues. Although there may be some benefits of enhanced embodiments such as VideoArms, at least on subjective awareness and realism, but there is very little ability to affect behavior in a loosely-coupled task.

5.1.2 Physical reaching versus mouse reaching

There were large differences between mouse and touch input on the time spent reaching. It is physically tiring to keep an arm extended over a table, so most people keep their arms in a resting position near their seated location. This means there are fewer opportunities to cross: people spend less time reached out, and thus cross less.

Although this seems to be an obvious characteristic of the input type, the different behaviours may have a substantial effect on distributed tabletop interactions. People are attuned to changes in the environment, so an increased number of mouse-based actions in the public tabletop space (that may not be meaningful) means that people may start to ignore reaching actions. Large and less frequent physical reaches are more noticeable than the often quick and jerky gestures of mouse-based input. In addition, physical reaching is a more purposeful act than mouse-based reaching; people typically do not spend any more time with their physical arm reached out than they have to. Video embodiments are also subjectively reported to provide more awareness of action. Together, these results suggest that touch-based distributed tabletops may provide better awareness of the other person's actions than their mouse-based counterparts.

5.1.3 Predicting the other person's actions

It is difficult to predict what other people are about to do in distributed environments. In our system, reaching gestures are only captured and transmitted once they are over the table's surface, removing the subtle preparatory gestures that precede a reach. For example, people move their gaze towards where they are going to reach, they rotate their torso to orient themselves towards the target, and lean in to begin the reaching gesture. Visual cues, such as display trajectories 7, may alleviate some of these issues.

In addition, VideoArms require a lot of processing power, and introduced lag into both the local and remote embodiments due to the Kinect and the LAN connection. Visual lag has been shown to affect coordinative behaviours 12, which may help explain people's behaviours with VideoArms. The video quality of our VideoArms is also not perfect, due to the technical limitations of the Kinect. There are visual noise artifacts around the embodiments and blur during movement, giving them a ghosted appearance. These issues may also reduce people's sense of embodiment – the importance of temporal correspondence has been shown in other settings, such as the "rubber hand illusion", where a false hand can be interpreted as the person's own hand. Here is the temporal correspondence of the tactile and visual feedback that is key in creating the embodiment. Reducing this temporal correspondence may contribute to a reduced sense of embodiment.

Overall, these effects can lead people to ignore the other person's interactions in the shared space. This may increase the perceived distance between remote groups [28], and potentially increase people's separation of in-group and out-group [3], breaking the collaborative experience.

5.1.4 Distributed tactile feedback

Why did tactile feedback work so well in a co-located setting [6], but have little effect in a distributed setting? We suspect there are at least two reasons. First, the latency of our VideoArms may have made it harder to predict when other people were reaching (see previous section). The difficulty of predicting when a crossing might happen may have caused people to simply give up trying to avoid the tactile feedback.

Second, even though people reported that crossing with tactile feedback was more awkward, in practice it seemed that vibration was not awkward when the other person was not co-present. When people are co-located, the vibrations are a shared experience, with both people reacting and generally trying to avoid it. In a distributed setting, the vibrations become individual experiences: they become easier to ignore, and it becomes easier to forget about the other person's tactile experience. Thus, there seems to be something about seeing another's actions directly causing the tactile sensation that is lost when people are distributed.

5.1.5 What it means to be embodied with arm embodiments

True embodiment (a sense of being *inside* the embodiment, a sense of *controlling* the embodiment, and a sense the embodiment *is part* of physical body [24]) may require more than just a visual representation. The visual representation alone is not strong enough to cause people to treat arm embodiments as they treat physical arms, and augmented embodiments lose some of their power to promote awareness and coordination when deployed in distributed environments. Vibrations increase the feeling of sharing the same space, though they do not change people's willingness to touch the other embodiment. In the end, our results suggest that people do not extend their personal space to surround their arm embodiments. They do not avoid reaching near or through others' arms.

One reason we wanted to study distributed embodiments was to test the "realest thing in the room" hypothesis – that is, the embodiments might become more real because there was no overshadowing presence of the physical body. When co-located, people's physical bodies are the realest representations of others [5]; however, when augmentations are added in a co-located setting, the arm embodiments become more "real". People's actions have consequences in the physical world, and so people extend their personal space to encompass the digital arm embodiment.

Distribution removes the physical co-present body, so the realest thing in the room should be the other person's arm embodiment. However, people did not behave this way – people were generally oblivious to the other person's embodiment and even their personal workspace (see occlusion example in video analysis). Instead of becoming more real without a co-present body, the embodiments became less real. Even the tactile feedback did not cause the distributed VideoArms to encapsulate people's personal space, or change people's behaviour.

The best remote embodiment may end up being a physical device, such as a proxy robot or the robot arms used in remote surgical systems; however, we still know little of how people would treat the personal space of these physical representations.

5.1.6 Future work

The task studied in this work includes only symmetric interaction, as people are performing the same task and interact at the same time (symmetric and synchronous interactions). In everyday tasks, people can interact at different times (asynchronously) and can work on different tasks (e.g., gatherer and assembler). We believe arm embodiments may be useful as asynchronous visual traces [15], even sped up or aggregated.

At a higher level, this exploratory work leaves us with many questions. How would people's interactions be different in a cooperative (instead of parallel) task? Groups created a playful environment with the augmentations, poking at each other jokingly, opening up questions about the meaning of digital touch. Will digital touch one day have similar social norms as physical arms?

Will augmentations be required to induce behaviour change, or will a new medium or embodiment induce a higher SoE?

6 CONCLUSION

Digital embodiments have been considered an important component of distributed tabletop systems, because groups require support to replace the missing co-present body. We investigated how the design of distributed digital arm embodiments affects coordination, and subject awareness and co-presence in a distributed mixed-focus task. Our results showed that although video embodiments are preferred over simpler visual designs, none of the embodiments made any difference to behavior, and that people often completely ignore the remote person, even when virtual touch caused tactile feedback.

These results provide important information for designers of distributed tabletop systems. Distributed tables have been seen as a way to recreate some of the natural and facile collaborative behaviors that are seen when people work face to face – but our findings show that for the common scenario of mixed-focus collaboration, people behave as if there is nobody else at the table. This means that designers will be less able to depend on social protocols as a way of managing group processes and access to shared resources. Our findings question whether it is possible to recreate, over distance, the natural collaborative behaviour that is seen in face-to-face tabletop work, and they demonstrate the difficulty that designers will face in attempting to use distributed tables as natural sites for collaboration.

REFERENCES

1. Benford, S., Bowers, J., Lennart, E.F., Greenhalgh, C., Snowdon, D. User embodiment in collaborative virtual environments. *Proc. CHI'95*, 242 – 249.
2. Biocca, F., Harms, C., & Burgoon, J. K. Toward a more robust theory and measure of social presence: Review and suggested criteria. *Presence: Teleoperators and virtual environments 2003*, *12*(5), 456-480.
3. Bos, N., Shami, N.S., Olson, J.S., Cheshin, A., and Nan, N. In-group/out-group effects in distributed teams: an experimental simulation. *Proc. CSCW 2004*. 429-436.
4. de Vignemont, F. Embodiment, ownership and disownership. *Consciousness & Cognition*, 2010, 82-93.
5. Doucette, A., Gutwin, C., Mandryk, R.L., Nacenta, M., & Sharma, S. Sometimes when we touch: how arm embodiments change reaching and collaboration on digital tables. *Proc. CSCW '13*, 193-202.
6. Doucette, A., Mandryk, R.L., Gutwin, C., Nacenta, M., Pavlovych, A. The effects of tactile feedback and movement alteration on interaction and awareness with digital embodiments. *Proc. CHI '13*, 1891-1900.
7. Dourish, P., Bellotti, V., Awareness and Coordination in Shared Workspaces, *CSCW 1992*, 107-114.
8. Fraser, M., McCarthy, M., Shaukat, M., and Smith, P. Seconds Matter: Improving Distributed Coordination by Tracking and Visualizing Display Trajectories. *Proc. CHI 2007*, 1303-1312.
9. Garau, M., Slater, M., Pertaub, D., Razzaque, S. The Responses of People to Virtual Humans in an Immersive Virtual Environment. *Presence 2005*. 14,1. 104-116.
10. Gaver, W. The affordances of media spaces for collaboration. *Proc. CSCW'92*, 17-24.
11. Genest, A.M., Gutwin, C., Tang, A., Kalyn, M., and Ivkovic, Z. KinectArms: a toolkit for capturing and displaying arm embodiments in distributed tabletop groupware. *Proc. CSCW '13*, 157-166.
12. Gergle, A., Kraut, R.E., and Fussell, S.R. The impact of delayed visual feedback on collaborative performance. *Proc. CHI 2006*, 1303-1312.
13. Greenberg, S., Gutwin, C., and Roseman, M., Semantic Telepointers for Groupware. *Proc. OzCHI '96*, 54-61.
14. Gutwin, C. (2001). The effects of network delays on group work in real-time groupware. *ECSCW '01*, 299-318.
15. Gutwin, C., and Greenberg, S., Design for Individuals, Design for Groups: Tradeoffs between Individual Power and Workspace Awareness, *Proc. CSCW 1998*, 207-216.
16. Gutwin, C. and Penner, R. Improving interpretation of remote gestures with telepointer traces. *CSCW'02*, 49-57.
17. Hall, E.T. *The Hidden Dimension*. 1966.
18. Hayne, S., Pendergast, M., and Greenberg, S., Implementing gesturing with cursors in group support systems. *J.MIS*, 10, 3, 1993, 43–61.
19. Heath, C., and Luff, P. Disembodied conduct: communication through video in a multi-media office environment. *Proc. CHI '91*, 99-103.
20. Huisman, G., Darriba Frederiks, A., Van Dijk, B., Hevlen, D., & Krose, B. The TaSST: Tactile sleeve for social touch. *World Haptics Conference 2013*, 211-216.
21. Ishii, H., Kobayashi, M., ClearBoard: a seamless medium for shared drawing and conversation with eye contact. *Proc. CHI 1992*, 525-532.
22. Izadi, S., Agarwal, A., Criminisi, A., Winn, J., Blake, A., Fitzgibbon, A. C-Slate: a multi-touch and object recognition system for remote collaboration using horizontal surfaces. *Proc. Tabletop 2007*, 3–10.
23. Jeffrey, P., Mark, G. Navigating the virtual landscape: co-ordinating the shared use of space. In *Social Navigation of Information Space*, 112-131.
24. Kilteni, K., Groten, R., and Slater, M. The Sense of Embodiment in Virtual Reality. *Presence* (2012), Vol. 21, No. 4, 373–387.
25. Kirk, D., Crabtree, A., and Rodden, T., Ways of the hands. *Proc. ECSCW'05*, 1-21.
26. Kirk, D. & Fraser, D.S. Comparing remote gesture technologies for supporting collaborative physical tasks. *Proc. CHI '06*, 1191-1200.
27. Kraut, R.E., Gergle, D., and Fussell, S.R. The use of visual information in shared visual spaces: informing the development of virtual co-presence. *CSCW '02* 31-40.
28. Lissermann, R., Huber, J., Schmitz, M., Steimle, J., and Mühlhäuser, M. Permulin: mixed-focus collaboration on multi-view tabletops. *Proc. CHI 2014*, 3191-3200.
29. Olson, G.M., and Olson, J.S. Distance matters. *Human-Computer Interaction 2000.* 15, 2, 139-178.
30. Pinelle, D., Nacenta, M., Gutwin, C., Stach, T. The effects of co-present embodiments on awareness and collaboration in tabletop groupware. *GI 2008*. 1-8.
31. Schmidt, K., and Bannon, L., Taking CSCW Seriously: Supporting Articulation Work. *JCSCW* 1, 1992, 7-40.
32. Slater, M., Steed, A. Meeting People Virtually: Experiments in Shared Virtual Environments. *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments. 2002*.
33. Smith, M.A., Farnham, S.D., Drucker, S.M. The Social life of Small Graphical Chat Spaces. *CHI '00*. 462-469.
34. Stach, T., Gutwin, C., Pinelle, D., Irani, P. Improving recognition and characterization in groupware with rich embodiments. *CHI'07*, 11-20.
35. Tang, A., Neustaedter, C., Greenberg, S. VideoArms: Embodiments for Mixed Presence Groupware. *People and Computers 2006.* 2-18.
36. Tang, A., Pahud, M., Inkpen, K., Benko, H., Tang, J. C., and Buxton, B., Three's Company: Understanding Communication Channels in Three-way Distributed Collaboration. *Proc. CSCW'10,* 271-280.
37. Tang, J., and Minneman, S. VideoWhiteboard: Video Shadows to Support Remote Collaboration. *Proc. CHI'91*, 315-322.
38. Tuddenham, P., Robinson, P. Territorial coordination and workspace awareness in remote tabletop collaboration. CHI '09, 2139-2148.
39. Yamashita, N., Kaji, K., Kuzuoka, H., & Hirata, K. Improving visibility of remote gestures in distributed tabletop collaboration. *Proc. CSCW '11*. 95-104.
40. Yarosh, S., Tang, A., Mokashi, S., and Abowd, G. D. "Almost touching": parent-child remote communication using the sharetable system. *Proc. CSCW '13*, 181-192.
41. Zanella, A., and Greenberg, S. Reducing interference in single display groupware through transparency. *Proc. ECSCW 2001*, 339-358.