Heterogeneous Device Arrangements Affect Both Partners' Experiences in Collaborative Media Spaces

Baris Unver[1,2](✉), Jasmine Jones[1], Alexander Thayer[2], and Svetlana Yarosh[1]

[1] University of Minnesota, Minneapolis, MN 55455, USA
{verxx003,jazzij,lana}@umn.edu
[2] HP Labs, Palo Alto, CA 94304, USA
huevos@alumni.washington.edu

Abstract. HCI has a history of developing rich media spaces to support collaboration between remote parties and testing such systems in investigations where each partner uses the same device setup (i.e., homogeneous device arrangements). In this work, we contribute an infrastructure that supports connection between a projector-camera media space and commodity mobile devices (i.e., tablets, smartphones). Deploying three device arrangements using this infrastructure, we conducted a mixed-methods investigation of device heterogeneity in media space collaboration. We found that the commodity devices provided a worse user experience, though this effect was moderated in some collaboration tasks. Collaborating with a partner who was using a commodity device also negatively affected the experience of the other user. We report specific collaboration concerns introduced by device heterogeneity. Based on these findings, we offer implications for the design of media spaces that use heterogeneous devices.

Keywords: Media spaces · Computer-mediated communication · Distributed collaboration · Projector-camera systems · Mobile devices · Asymmetry

1 Introduction

The proliferation of smart mobile devices [1] has led to a significant expansion in the availability of synchronous communication tools and diversity of device platforms. Each device used in collaboration may have significantly different capabilities and may provide different affordances for interaction and representations of the same information. For example, an online whiteboard drawing application may be simultaneously used by someone using a finger to draw on a 4-in. Android phone screen, someone using a stylus on a 13-in. iPad Pro, and someone using a mouse on a Windows laptop computer. Cross-platform compatibility and development to enable this kind of interaction are a major push in the industry development [2]. The impact of devices on the remote collaboration experience remains an open question and recent studies have shown that even minor differences in the presentation of digital content can lead to a

M. Kurosu (Ed.): HCII 2019, LNCS 11567, pp. 76–98, 2019.
https://doi.org/10.1007/978-3-030-22643-5_7

significant amount of miscommunication (e.g., [3]). The first contribution of this work is an empirical investigation examining the effect of device type on media space collaboration experiences of both partners.

In contrast to the cross-platform push in the industry, research investigations have almost exclusively focused on the interaction across homogeneous system setups. For example, a significant suite of studies has explored the role of the duplex pro-cam systems for communication and collaboration [4–6]. These systems typically provide a rich face-to-face media space and a shared tabletop or wall-wide projection surface for supporting collaboration on tasks, typically through the substantial use of custom hardware. In all the previous investigations, these novel systems were tested in a setup where each participant had identical access to the rich features of each system. As pro-cam communication systems become available on commodity hardware and show the potential for expansion into the real-world use (e.g., [7]), we need to consider how others may connect to and interact with such systems using a diverse suite of devices and how such an interaction may affect the collaboration experience. The second contribution of this work is the development of an infrastructure and system to support the collaboration between video-based media spaces running on heterogeneous device arrangements. We implemented three specific examples of how this infrastructure may be adapted to support collaboration between a projector-camera media space and commodity devices like tablets and smartphones.

In this paper, we describe the development of the ShareSuite infrastructure which addresses the research gap in understanding media spaces which use heterogeneous device arrangements. Through a mixed-methods investigation of three device arrangements and an analysis of qualitative feedback following the experiment, our work addresses three research questions:

- RQ1: How does the device (pro-cam vs. tablet vs. tablet & smartphone) used in a media space collaboration affect a one's experience?
- RQ2: How does the device used by one's communication partner in a media space collaboration affect one's experience?
- RQ3: What specific concerns and challenges are introduced with the use of heterogeneous device arrangements in remote collaboration?

We begin with a discussion of related work, describing previous investigations of novel media space collaboration systems and previous work on cross-device interaction. We describe the ShareSuite infrastructure. We report the methods and results of a 48-participant mixed-methods study where we compare collaboration experiences between multiple device arrangements. Our results demonstrate that using a commodity device negatively impacts the experience of the user and (to a lesser extent) their communication partner and introduces a new set of concerns and collaboration challenges. Finally, we offer future research and design directions based on our findings.

2 Related Work

We review relevant research and related fields that influenced both the system design and the research questions of this work.

2.1 Media Spaces for Remote Collaboration: Homogeneous Arrangements

Media spaces are collaboration technologies that combine video, audio, and other potential forms of synchronous media. This class of technologies was originally developed for an office context in the early 1990s, but has since become a common technological trope in home settings as well [1]. Media spaces may aim to increase general awareness of remote partners (e.g., [8]), enhance social contact and communication (e.g., [9]), or serve as a shared workspace where tasks may be synchronously accomplished (e.g., [10]). The latter is most relevant to this work. While a media space may consist of exclusively one type of media (e.g., just video [11]), media spaces for accomplishing synchronous tasks typically combine a "person-space" and a "task-space."

Previous work has pointed out that "media spaces are an inherently asymmetrical technology" with regard to several socio-technical dimensions, such as different levels of participation in, engagement with, and benefit from the collaboration [12]. These asymmetries have been investigated as side effects in previous collaboration studies. For example, in a study of a community awareness system, the users who chose not to share the video were neglected by others and felt like "2nd class citizens" [13]. In other investigations, asymmetries of participation arose from the specifics of a collaboration task (e.g., an expert instructing a worker [14]). Asymmetries of benefit have been discussed in media spaces aimed at connecting the parents and children in divorced [6] and/or work-separated families [15], where the child, the remote parent, and the co-present parent may all have different motivations, benefits, and costs. However, while these asymmetries have been acknowledged and briefly discussed, media spaces are typically deployed and evaluated in ways that artificially create homogeneity and symmetry. Our investigation addresses this gap by developing an infrastructure for heterogeneous media spaces and examining the role heterogeneous device arrangements on collaboration.

2.2 Device Ecosystems: Rise of Heterogeneous Arrangements

Research focus on homogeneous arrangements in media space development and testing is in stark contrast to other areas of HCI, which have sought to leverage the pervasiveness and diversity of available mobile devices by focusing on the cross-device interaction and diverse device ecosystems. Cross-device interaction allows multiple mobile devices to support different parts of a given task, leveraging the multi-device ecosystem available to most of the users [1]. Such cross-device interactions have been termed Distributed User Interfaces (DUIs) [16] and they may include interfaces distributed across multiple devices, multiple displays, etc. [17]. As an example of this work, Hamilton et al., investigated how to use multiple devices adjunct to each other for creating larger displays by developing a framework conductor which supports cross-device application interaction [18].

In comparison to these endeavors outside of a collaboration context, the role of mobile device ecosystems and the effects of heterogeneous device arrangements in media spaces has not been adequately investigated. Even as mobile devices become increasingly important to how people seek to share their experiences [19–24], video sharing applications are frequently poorly adapted for mobile platforms [25]. Our work contributes an infrastructure for and an empirical investigation of heterogeneous device arrangements (including DUIs) in dual-space media spaces.

2.3 Dual Space Media Spaces: Prototypes vs. Infrastructure

This investigation focuses on dual-space media spaces, which combine a person-space and a task-space. A person-space provides audio and video of a remote collaborator to increase workspace awareness and allow verbal conversations and expressions to be exchanged remotely [26]. Previous work has shown that such a "face-to-face" view is necessary and helpful in collaborative tasks like remote assistance [27]. A task-space provides a view of the artifacts in the workspace and ideally allows for some manipulation of these artifacts and some gestures over the shared task-space. The availability of such a shared space has been shown to provide necessary conversational grounding in remote tasks [28], resulting in better task performance [29].

HCI researchers have developed many dual space media spaces. One of the earliest such systems (VideoDraw [30]), allowed people to draw directly on a digital display and captured this activity with a top-down camera to transmit to a remote display. Many more recent systems use top-down projector-camera (pro-cam) setups to create a shared tabletop task-space [5, 6, 31, 32]. Other systems use horizontally-mounted (e.g., ClearBoard [33], C-Slate [34]) or vertically-mounted (e.g., ImmerseBoard [35]) displays that combine cameras, multi-touch, and styluses to support remote collaboration. The person-space in these systems is typically either shown on a separate vertical display (e.g., [6]), superimposed on the task space (e.g., [33]), or shown in a separate window on the same display as the task-space (e.g., [36]). Each approach to dual-space media spaces has different benefits for users and introduces a different set of technical challenges to resolve.

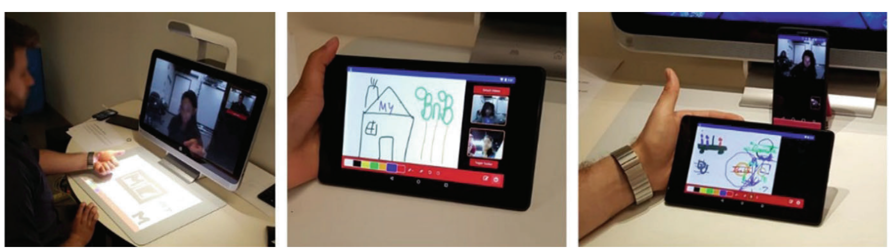

Fig. 1. Three device set used implemented to demonstrate the ShareSuite media space infrastructure: an HP Sprout duplex pro-cam system, a commodity tablet device, and a distributed user interface that combines a tablet and a smartphone.

These previous investigations share many common features. First, they all require custom hardware modifications and each system is developed only for the specific hardware that is utilized in the prototype deployment. In contrast, in this work, we offer an infrastructure that abstracts the idea of "task-space" and "person-space" and intelligently designates these to appropriate devices (or areas of devices) based on device availability. Second, the media spaces above have been tested exclusively in homogenous device arrangements. The present work leverages our infrastructure to offer an empirical investigation of collaboration in heterogeneous dual-space media space arrangements.

3 ShareSuite Infrastructure

The ShareSuite infrastructure supports collaboration across dual-space media spaces on a variety of devices. We designed this infrastructure to support the following motivating example of a heterogeneous device media space collaboration:

> *Simon wants some feedback from Zahra on a design sketch. His office is equipped with a projector-camera media space, and he logs into the ShareSuite, initiates a connection, and places some of his paper sketches on the table. Zahra is traveling and receives the contact request on her iPad Pro. As she answers the contact, her device splits the available screen real estate so that the Simon's person-space video appears next to the camera feed from his tabletop display. Zahra sketches on her iPad with a stylus and her contribution is projected directly onto Simon's table; conversely, Zahra can see Simon's sketches and gestures over the tabletop displayed as a video on her tablet. Wanting to have a bigger workspace while still seeing Simon's face, Zahra also logs into the ShareSuite on her Android smartphone. The ShareSuite readjust the workspaces so that the phone becomes the new person-space and the tablet becomes entirely task-space, allowing her to work with Simon more efficiently.*

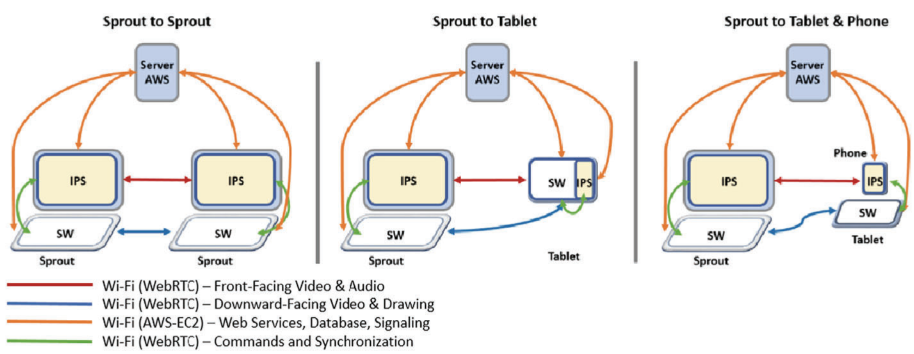

Fig. 2. Three collaboration arrangements supported by our system infrastructure shows Shared Workspace (SW) and Interpersonal Space (IPS) representation on each device; data transmission between devices and Amazon Web Services (AWS).

Below, we describe the device-side and server-side infrastructure of the ShareSuite, as well as three different device arrangements implemented over the ShareSuite infrastructure for the purpose of this investigation (see Figs. 1 and 2).

3.1 Device-Side Infrastructure and Functionalities

The ShareSuite infrastructure was designed to be adapted to multiple device types. Each device must have at least one screen, at least one camera, and the ability for manipulating the task-space (e.g., touch screen so that the user can draw on the screen). The device communicates the screen area available for the task-space and resolution of the screen to the server (see Sect. 3.2) allowing to send the appropriate video channel to the screen. To support a common task-space, we implemented a draw-over-video environment with standard digital tools (e.g., similar to Ou's et al. [37]). To support heterogeneous device arrangements, we implemented coordinate transformation to allow devices with different screen resolutions and dimensions. Each device receives a video stream of the task-space from the server along with the coordinate system describing the capture characteristics of the origin device. The device than converts the video to the local coordinate system (preserving the aspect ratio). In cases where a single screen must support both a task-space and a person-space, the device receives multiple video streams from the server. In the current implementation, the device splits screen area between task and person-space in a way that best preserves the aspect ratio while maximizing the task area, however the infrastructure supports dynamic adjustment in future instantiations.

3.2 Server-Side Infrastructure

ShareSuite's seamless connection between devices and its adaptive display reconfiguration is handled by a cloud-based server. The server infrastructure is hardware agnostic to ensure compatibility across commodity devices.

The ShareSuite clients and cloud server exchange audiovisual information using three simultaneous data channels, implemented as WebRTC-enabled web pages (see Fig. 2). The person-space channel carries the front-facing video, audio, and data for call initiation. The task-space channel can transmit an additional camera channel or a stream of digital drawing data from the shared workspace. The support channel provides synchronization, information, and command transfer between the audiovisual channels.

The ShareSuite provides a common authorization infrastructure across clients. Signing in on multiple devices (e.g., a phone and a tablet) intelligently divides the task- and person-spaces. Each instance connects to the AWS-EC2 web server to load the browser-based interface and video content. After loading the content, ICE/STUN server accessed over web socket to assign a unique Peer ID for the WebRTC connections to overcome cross-networking connections issues. All data and connections are requested, stored, processed, and delivered via a RESTful API from an Amazon Web Services Elastic Compute Cloud (AWS-EC2) server, running Apache, MySQL, and PHP.

3.3 Devices Implemented for Investigation

We implemented the ShareSuite to support three example device sets (see Figs. 1 and 2). In Sect. 4.2, we describe how these device sets are combined into heterogeneous arrangements order to investigate how collaboration may be affect.

Duplexed Projector-Camera Tabletop as Task Space, Vertical Display as Person Space. We implemented a full-duplex projector-camera system, over an off-the-shelf HP Sprout personal computer. We implemented the task-space on the Sprout's horizontal multi-touch mat with a top-down pro-cam display (HP DLP 1024 × 768 projection, 14.6 MP downward-facing camera, 20-point multitouch). Participants could draw directly on the multitouch mat using their finger. The person-space was placed on the Sprout's large vertical monitor (23″ 1920 × 1080 display, 1 MP front-facing camera), showing the face-to-face video of the collaborator. To enable access the Sprout's built-in hardware, we deployed our WebRTC ShareSuite solution in a Chromium CefSharp instance (rather than in a browser). We refer to this device as "Sprout" or "S" throughout this paper.

Commodity Tablet Split for Task and Person Spaces. The second device we implemented to connect with the ShareSuite infrastructure was a conventional tablet. We used a Nexus 7 tablet (7″ 1280 × 800 display, 1.2 MP camera) with a split screen to display the task- and person-spaces. We deployed the ShareSuite in a browser-instance as described above. We refer to this device as "Tablet" or "T".

Commodity Tablet Split for Task and Person Spaces. Finally, to demonstrate the DUI capabilities of the ShareSuite infrastructure, we implemented a dual-device set that combines a tablet and a smartphone. We deployed the ShareSuite in a browser-instance on both devices as described above. The tablet was a Nexus 7 (7″ 1280 × 800 display, 1.2 MP camera) and was fully dedicated to the task-space. The LG G2 Android smartphone (5.2″ 1920 × 1080 display, 2.1 MP camera) was fully dedicated to the person-space. We refer to this arrangement as "Tablet&Phone" or "TP".

4 Methods

To investigate our research questions regarding collaboration across heterogeneous device arrangements, we conducted a mixed-design lab experiment. Participants (N = 48) completed a collaborative task in pairs, with each pair completing one of two collaborative tasks using each of the three different device arrangements: Sprout-to-Sprout (S-S), Sprout-to-Tablet (S-T), and Sprout-to-Tablet&Phone (S-TP). In this section, we describe the study design, participant recruitment, data collection, and data analysis methods.

4.1 Participants

We recruited 48 participants internally in a large U.S. technology company through word-of-mouth and email solicitations and grouped them into 24 pairs for our study. Fifteen of the participants were female (33 male), and they ranged in age from 18 to 57 (M = 28.8, SD = 9.9). None of the pairs knew each other prior to the start of the study. Each participant received a $25 gift card.

4.2 Study Design

The study took place in a research lab with multiple separate partitioned spaces. Participants were randomly assigned into pairs upon recruitment and were given an overview of their roles on the tasks. We demonstrated all three device sets to each participant, ensuring that they were familiar with their functionality. At least one member of the research team was available in the room to answer questions and help resolve technical difficulties.

We chose to test two tasks in our experiment to observe the use of heterogeneous device arrangements in tasks that are typically carried out in a mutual way verses tasks that typically require an asymmetric collaboration. In the mutual collaboration task, the participant pairs were asked to work together to create a logo for a fictional company. A short, written scenario describing the company was provided to both the participants. Each of the three times they repeated the task, it had a similar goal, but the specific details of the fictional company differed. In the helper-worker task, participants were assigned to either a worker or a helper role. The helper was asked to instruct the worker to construct a specific Lego Duplo structure consisting of 13 bricks (based on a step-by-step manual provided to the helper). All tasks were piloted by confederates to insure consistent complexity and that each task took about 10 min to complete. Throughout the paper, we refer to these two tasks as "Logo" and "Lego" tasks respectively.

To address RQ1 (effect of the device on the user's experience), one partner was assigned to switch technologies (Sprout, Tablet, and Tablet&Phone) for each of the three tasks. The specific order of the three technologies was counterbalanced to avoid order effects. To address RQ2 (effect of the partner's device on the user's experience), the other partner was assigned to use the Sprout device throughout the three tasks. In other words, the first participant experienced switching devices with all other aspects held stable (task, partner, partner's device), while the second participant experienced their partner switching devices with all other aspects held stable (task, partner, their device). We refer to these two types of participant roles as "Switching" and "Non-Switching" throughout the paper.

Thus, this study constitutes a mixed-design: Lego/Logo and Switching/Non-Switching assignment was between-subjects, while the specific type of the technology arrangement was varied within-subjects. Upon completion of each round, the participants completed several short post-task questionnaires (see Sect. 4.3). After the participants had completed all three tasks, one researcher interviewed each participant about their overall experience.

4.3 Metrics and Analysis

We employed several complementary empirical strategies to understand our participants' experiences:

- **Video and Audio Recordings:** We collected video and audio recordings of the sessions as the basis for verbatim transcription and post-hoc coding. The lead author reviewed all the recordings and coded examples of miscommunications, struggles, and challenges.

- **Satisfaction and Preference Questionnaire:** To complement this qualitative understanding we deployed a questionnaire investigating users' holistic satisfaction, condition preference, and comparison to an unmediated interaction:
  - Comparison to Unmediated Interaction/Satisfaction with Collaboration: a single Likert-type item where a higher score indicates a higher level of similarity/satisfaction, answered for each condition and analyzed with a repeated-measure ANOVA.
  - Condition Preference: after completing all three conditions, we asked participants to specify a preferred device arrangement (or "None" if they had a similar experience with all three). These were analyzed using categorical comparison tests (Chi-Square, Fisher Exact) as appropriate.
- **Perceived Message Understanding (PMU):** We used the 6-question Perceived Message Understanding scale [38] to help address RQ1 and RQ2. The scale is scored by averaging the six Likert-type items within the scale—a higher score corresponds to a greater perception of understanding and being understood by the partner. We compared conditions using a repeated-measures ANOVA.
- **Open-Ended Debriefing Interview:** After the completion of the experiment, we interviewed each participant individually. The semi-structured interview contained five open-ended questions about the collaboration and lasted approximately 15 min. We asked participants to reflect and provide feedback about their experience using the ShareSuite, including their opinions of each device arrangement, how the devices compared, and any recommendations they might have for future collaborative systems. Interviews were audio recorded and transcribed. We followed a data-driven inductive process of analysis where we first open-coded the transcripts for individual units of meaning, clustered these codes using an affinity map, and iterated on the clusters to arrive at emergent themes. We are confident that we reached data saturation, as major themes began repeating and no new themes emerged after interview 19 for either task.

All sources are used for all three research questions. We focus primarily on qualitative data. Quantitative comparisons and descriptive statistics are used when appropriate as an additional source of evidence.

4.4 Method Limitations

We conducted a lab study with short, constrained tasks. There may be more emergent effects that could arise in a longer-term, less-controlled study. Nevertheless, similar previous studies showed that the task duration (10-min) is adequate to obtain reliable results [14, 29, 39]. We limited our study to two possible examples of heterogeneous device arrangements (Sprout-to-Tablet and Sprout-to-Tablet&Phone). Of course, the possible number of device arrangements grows exponentially with the number of supported devices and it is not feasible to test a complete crossing of all possible device arrangements (in this case, a full-crossing would have required us to test six different arrangements, which would have been impractical in a within-subjects study). Thus, we focused on three arrangements. The Sprout-to-Sprout arrangement provides an example of a homogenous pro-cam arrangement similar to those tested in previous work, serving as a point of comparison. We chose to explore connections to Tablet and Tablet&Phone as particular heterogeneous device arrangements because they represent the connection of a rich media space to two more limited (yet commonly available) devices (Table 1).

Table 1. Means and variance of the Satisfaction, Perceived Message Understanding (PMU), and Similarity to Unmediated measures for participants across the within-subjects (tech) and between-subjects (role, task) factors.

Role (each, N = 24)	Tech (each, N = 24)	Task (each, N = 12)	Satisfaction (out of 7)	PMU (out of 7)	Unmediated (out of 7)
Switching partner	S-S	Lego	5.67, *SD* = 1.23	4.83, *SD* = 1.03	5.83, *SD* = 1.33
		Logo	6.04. *SD* = 0.75	4.67, *SD* = 0.89	5.67, *SD* = 1.37
	S-TP	Lego	5.71, *SD* = 1.07	4.33, *SD* = 0.99	5.58, *SD* = 1.51
		Logo	5.13, *SD* = 1.33	4.25, *SD* = 1.29	5.08, *SD* = 1.38
	S-T	Lego	6.03, *SD* = 0.62	4.50, *SD* = 0.67	5.67, *SD* = 1.07
		Logo	4.92, *SD* = 1.21	4.08, *SD* = 0.90	5.17, *SD* = 1.70
Non-switching partner	S-S	Lego	5.97, *SD* = 0.88	4.67, *SD* = 0.49	6.08, SD = 0.67
		Logo	5.68, *SD* = 1.09	5.25, *SD* = 0.87	5.50, *SD* = 0.67
	S-TP	Lego	5.48, *SD* = 1.06	4.33, *SD* = 0.65	5.33, *SD* = 1.50
		Logo	5.61, *SD* = 1.00	5.33, *SD* = 1.16	5.42, *SD* = 1.38
	S-T	Lego	5.45, *SD* = 1.06	4.42, *SD* = 0.79	5.25, *SD* = 1.14
		Logo	5.53, *SD* = 1.64	5.00, *SD* = 1.21	5.25, *SD* = 1.22

5 Results

In the following sections, we present both qualitative and quantitative evidence regarding the effect of the device arrangement on satisfaction & preference, perceived similarity to unmediated collaboration, perceived message understanding, and asymmetries in participation.

Table 2. Stated preference between the three conditions and a Chi-Square Goodness of Fit comparison to an equal distribution. "None" specifies that the participant said that all conditions were equivalent for them (and was excluded from the Chi-Square).

Tech	Task	S-to-S	S-to-T&P	S-to-T	None	Chi-Square GoF
Switching partner	Logo task	12	0	0	0	$X^2 = 24.244$, $p < 0.001$***
	Lego task	9	1	2	0	$X^2 = 9.597$, $p = 0.008$**
Non-switching partner	Logo task	8	0	1	3	$X^2 = 12.796$, $p = 0.002$**
	Lego task	7	2	1	2	$X^2 = 6.264$, $p = 0.044$*

5.1 Satisfaction and Preference

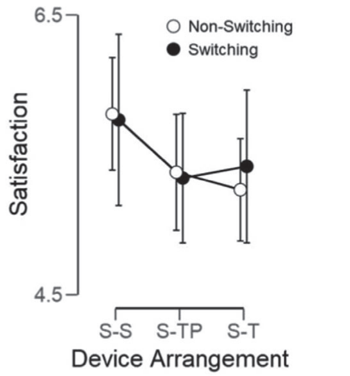

Fig. 3. Satisfaction ratings across device arrangements by participants in switching and non-switching roles. Error bars represent 95% confidence interval.

Given that we were interested in the holistic evaluation of the experience in each condition, we focused on qualitative feedback, self-reported satisfaction, and an explicit statement of preference between conditions to understand how the specific devices affected the experience of both participants.

The ShareSuite was overall well-received by participants, with the average satisfaction rating across devices at 5.49 (out of 7, *SD = 1.26*). Participants who switched devices mentioned that they appreciated the seamless transition between the three device arrangements afforded by the ShareSuite infrastructure: *"the app is really good because it worked fine also with the mobile devices."* Other participants who switched devices also note that the experiences across all three devices were comparable in that they and their partner *"were able to achieve the outcome desired in reasonable amount of time."*

While participants recognized that connection with mobile devices enabled previously-impossible contexts of collaboration, they also noted that it was not as rich as the connection afforded by the Sprout device. Participants indicated in the interviews that Sprout was the richer method of interaction due to screen size and ability to communicate gestures, e.g., *"Sprout was probably the best one because the screen was bigger"* and *"he pointed sometimes at the bricks that I've used, and this was very helpful."* When comparing across conditions, the Sprout-to-Sprout arrangement was the clear preference among all participants (see Table 2). It is not surprising the participants in the "switching" role would strongly prefer the Sprout condition, as it was the richest of three and provides the larger task- and person-spaces. It is perhaps more surprising that the non-switching partners (who always used the Sprout) had a similar preference towards the symmetric Sprout-to-Sprout condition, as nominally their experience with each condition featured the same partner, task, and device.

Descriptively, it may appear that S-S may have been particularly preferred by those doing the Logo task. While the difference in distributions is not statistically significant (Fisher Exact Test, $p = 0.157$), qualitative data points at a possible effect of the task factor. Namely, many participants found the low contrast lighting of the projector on the Sprout problematic in the Lego task, especially when the projection was onto a physical object rather than on the mat: *"I couldn't actually see the image of his hands out in the Legos."* This may have been exacerbated by the fact that this particular pro-cam solution used ambient light to address video echo that arises in full-duplex pro-cam systems [7], which washed out colors and reduced contrast. This was particularly problematic for depth perception in the Lego tasks:

> *I kept getting the feedback from my partner that the depth perception and other things were much better viewed [S-T] as compared to the first two and he was able to describe to me exactly what I needed to do.*

Overall, participants saw some positives and negatives in each device arrangement, though there was a marked preference towards the Sprout-to-Sprout setup. However, the satisfaction rating responses were less clearly divided, perhaps due to a ceiling effect (see Fig. 3). While descriptively, the S-S condition had greater satisfaction ratings and the device arrangement accounted for greatest variance in satisfaction ratings ($\eta 2 = 0.05$), the difference was not statistically significant ($p = 0.112$).

Table 3. Statistical significance and effect sizes for repeated-measure ANOVAs for PMU and "Similarity to Unmediated" measures. No values reported for any cases where p > 0.1

Factor	PMU	Similarity to Unmediated
Tech (S-S vs. S-TP vs. S-T)	$p = 0.074$ ($\eta^2 = 0.06$)	**$p = 0.036^{*1}$ ($\eta^2 = 0.08$)**
Role (Switch vs. Non-Switch)	n.s.	$p = 0.080$ ($\eta^2 = 0.07$)
Tech * Role * Task	**$p = 0.031^{*}$ ($\eta^2 = 0.08$)**	n.s.

Post-Hoc Paired t-Test w/ Bonferroni Correction	
S-S vs. S-T	**$p = 0.047^{*}$ ($d = 0.362$)**
S-S vs. S-TP	n.s.
S-T vs. S-TP	n.s.

[1] After applying Greenhouse-Geisser Sphericity Correction to correct for violation of Sphericity assumption

5.2 Similarity to Unmediated Interaction

While replicating unmediated face-to-face interaction is not always the "holy grail" of collaboration [40], it remains a reasonable comparison and baseline for synchronous collaboration technologies. We asked participants to compare their experience in each of the three device arrangements with that of collaborating in the same room. Qualitatively, participants described the ShareSuite as a *"a viable substitute"* for face-to-face collaboration. However, at least one participant reported that all three devices had low similarity to unmediated interaction due to the specifics of the dual-space media space setup:

> *Across all three, one thing I noticed is that me and my partner were both looking at the top of each other's heads a lot [as we bent our heads to draw on the surface] ... so to communicate with her I'd have to go through the canvas and not through the face.*

Aside from this participant, most others felt that different arrangements had different effects on comparison with unmediated interaction. Some participants made strong claims about the S-S arrangement being "just the way" in-person collaboration works:

> *When we are trying to collaborate across different offices, then this system is very useful, especially considering that we are not face to face but we are still able to see each other, and we are still able to communicate and collaborate just the way you would when you're sitting face to face with someone.*

One aspect of this comparison related to the specific orientation of the person and task spaces. Particularly, the fact that the Sprout-to-Sprout condition separated the two workspaces into the horizontal and vertical screens was viewed as particularly *"natural"* and *"helpful"* e.g.:

> *It felt like he is sitting in front of me, and also with his hands, he pointed sometimes at the bricks that I've used, and this was very helpful.*

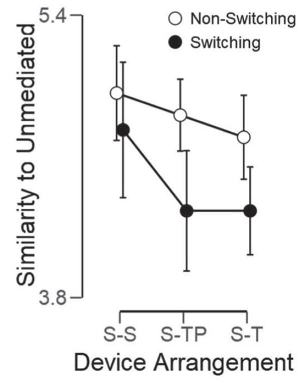

Fig. 4. Ratings of similarity to unmediated collaboration across device arrangements by participants in switching and non-switching roles. Error bars represent 95% confidence interval.

One switching participants expressed that the S-TP condition provided some similar affordances because of a similar division of spaces, stating that it helped to distinguish when she was focused on her drawing task and when she was talking to her partner: *"I wasn't drawing, I was looking at a person. It felt more natural."* In contrast, when using the Tablet, interaction could feel less natural. Non-switching participants using the Sprout explained that it was not clear whether a person was concentrating on the task or looking at them for more direction: *"It seemed that he wasn't looking at me, just looking on the surface of the tablet."*

Beyond the orientation of the two spaces, participants noted rating the Sprout-to-Sprout condition higher because it provided a natural size correspondence in the task-space, rather than scaling the task-space space to a smaller screen:

> *I loved the space availability that I had. I could draw on the Legos in 1:1 size analogy and my partner could easily follow my hints. I could use various ways of communication with my partner.*

The quantitative analysis of participant responses rating each condition based on its similarity to unmediated interaction corroborate these qualitative comparisons (see Fig. 4 and Table 3). There was a statistically-significant main effect of the device condition ($p = 0.036$) with a medium effect size ($\eta 2 = 0.08$). A pairwise comparison within device arrangement conditions revealed that the difference between S-S and S-T was statistically significant ($p = 0.047$), though descriptively the S-S condition was also rated as more similar to unmediated interaction than the S-T&P condition.

5.3 Perceived Message Understanding

Perceived Message Understanding (PMU) refers to each participant's perceived ability to understand their partner during the task and their estimation of how well their partner understood them, measured using the 6-item PMU scale from the Networked Mind Social Presence Measure [38]. Quantitatively, their responses had a statistically-significant ($p = 0.031$, $\eta 2 = 0.08$) three-way interaction between the device arrangement, task type, and participant role (see Table 3 and Fig. 5). The non-switching partners' PMU measures remain more-or-less consistent across conditions. However, the PMU of the partner who is switching devices seems to be affected by the task, with the S-S condition supporting PMU better than S-T on the Logo task. Given the complexity of this three-way interaction, we turn to the factors participants mentioned in the interviews to understand how PMU was affected by device type and participant role.

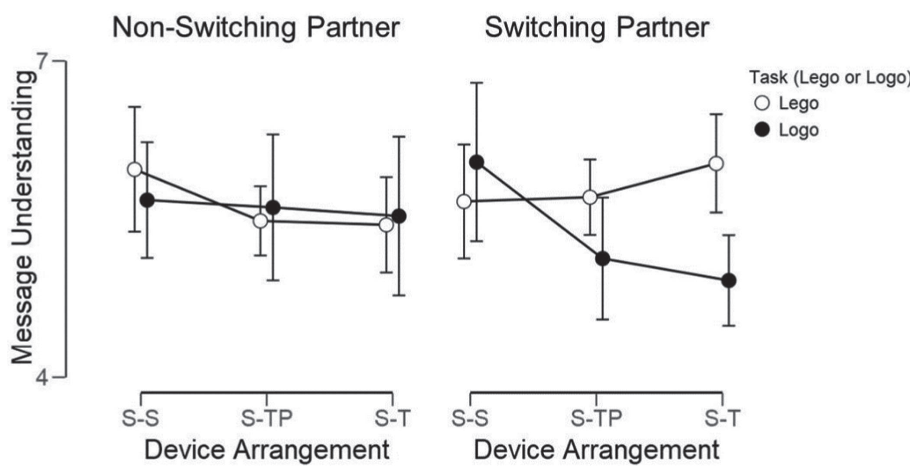

Fig. 5. Perceived message understanding ratings across participant roles (switching and non-switching), tasks (Lego and Logo), and device arrangements (S-S, S-TP, and S-T). Error bars represent 95% confidence interval.

Establishing Common Ground. Collaborators using the same device (S-S condition) reported little difficulty understanding their partner. As one non-switching participant remarked: *"You wouldn't have to worry about if they are seeing the same thing, because you know it's the same setup."*

However, in the heterogenous device arrangements, participants had to work to establish a shared reference and accurate understanding of what the other person was seeing. The differences across each device arrangement was especially salient to the Sprout user as their partner switched from device to device: *"the visual representation between our screens was not consistent and I had no idea what my collaborator could see."* The smaller-display commodity devices also conveyed less information about their users' actions, which sometimes left the non-switching user guessing about what their partner was doing and left the switching user having to find new ways of establishing common ground. One participant explained that when he switched to the Tablet and Tablet&Phone during a Logo task, his partner on the Sprout could no longer see the natural gestures he was making: *"My partner couldn't see my hands … she wasn't exactly able to anticipate what I was going to do and so there was sort of a gap in communication."*

Participants reported that it was difficult to establish common ground during the Lego task, which was particularly sensitive to the orientation and description of the pieces. One participant said she had a hard time interpreting the instructions that her partner was giving her about how to build their assigned structure, because she could not interpret the perspective he was communicating from*: "I didn't really understand what he was seeing on the other side"* and *"I don't know if she's seeing what I'm seeing."*

These qualitative reports conflict with the quantitative PMU score on one major aspect: the non-switching partners reported being affected by the heterogeneous device setup in their qualitative feedback but not in their PMU rating. As a follow-up analysis, we considered that the factor called out in the qualitative reports was not just about the difference between the specific devices but about the trouble introduced specifically when the arrangement is heterogenous (two partners using different devices) versus homogeneous (both partners using the same device). To do an exploratory test of this assertion, we pooled PMU averages across each participant's ratings of the S-T and S-T&P conditions and compared to the PMU rating of the S-S condition via a Repeated Measures ANOVA. Through this analysis, we found that PMU scores in the homogeneous condition were statistically significantly higher than in the heterogeneous conditions ($p = 0.041$, $\eta 2 = 0.09$) across participant roles. While our study design is not conducive to conclusively addressing questions of heterogeneity (see Sect. 6.2), it seems that heterogeneity of devices may better reflect the participant's qualitative experience and may quantitatively account for more variance in the PMU scores than device type alone role.

Verbal Interaction as a Work-Around. Across-condition PMU scores seem to suggest that device arrangement only mattered in the Logo task, even though participants mentioned several common ground issues in the Lego task condition. We observed that participants may have been able to overcome message understanding issues in the Lego task through verbal work-arounds, which may have smoothed PMU differences across Lego device arrangement conditions.

For example, we saw in our video observations that several of the participant pairs in the Lego task developed verbal practices (e.g., *"take the red 2 × 4 brick"*) that leveraged the simple physical characteristics of the Lego bricks to communicate in a detailed and specific manner:

> *It felt like when I had the mat to work with I started drawing on it, but it ended up not being very effective [in other conditions] … I think [I liked] being more specific with the dimensions verbally.*

However, though this collaboration style was effective in completing the Lego task, it could have limited applicability to other types of tasks where the there are fewer verbal shortcuts. This may explain the significant interaction with task and why the Lego task PMU appeared more resistant to condition affects.

Difficulty of Switching Between Displays. When discussing the merits of each device, participants pointed out several PMU issues that were particular to the Sprout and the Tablet & Phone devices. A key challenge was maintaining awareness of their partner's actions and attention across multiple displays, as they switched between talking, gesturing, and drawing.

One participant commented that she found herself ignoring all visual cues from her partner as she focused only on her drawing: *"What I noticed is you're just focusing down here, so you're not really focusing on the other person, even if it's up there."* Another participant reported that he missed when his partner was attempting to switch contexts between talking or drawing:

> *I didn't see when he was looking at the front camera or if he was drawing on this thing [the mat] ... when I was watching at this part of the Sprout, I couldn't see that he was doing something on the screen because it was out of my [peripheral vision] and until I figured out that he was doing something on this screen, it was already like 4 or 5 seconds.*

Participants felt that missing these attention cues and other non-verbal gestures would impact their ability to understand be understood:

> *It is sometimes important to see the other person, focus on the other person and see how he's reacting or how he's feeling, that might come out of the gestures he's making.*

In light of this, some participants reported that the side-by-side person- and task-space on the tablet provided the best awareness of their partners: *"it would have been much nicer if I could see the video ... on the mat."*

It seems that each device type introduced a different set of potential limitations to Perceived Message Understanding, leading complex interaction that a single PMU score may have failed to capture with appropriate nuance.

5.4 Asymmetries of Participation

Working on different devices introduced significant challenges for participants in both non-switching and switching roles. Participant interviews revealed equality of participation as another important aspect of collaboration across devices.

We observed that having less drawing space in the Tablet and Tablet & Phone conditions than on the Sprout had a negative effect on the participation level of the user in the switching role. When using different devices to collaborate, switching participants' activity shifted to make the non-switching Sprout user the primary driver of the task, or *"the main person drawing,"* and the switching user became *"more of a facilitator."* One reason for switching into this facilitator role was inability to precisely draw on the small screen:

> *Drawing on the tablet this time was very difficult because of small screen size. Several times, I had to tell my partner what to draw and she would draw it for me.*

While these issues were particularly pernicious in the Logo task, we saw aspects of participation affected by device used the Lego task as well. We noticed in our video observations that switching participants in the Lego task exhibited a similar trend of decreasing hands-on participation as they moved to the smaller devices. Participants used sketches and gestures more to direct their partner when they were using the Sprout, but relied more on verbal instructions and feedback when using a different device. This may have been one aspect that may have negatively influenced the non-switching participants experience with the collaboration even when they themselves continued using the richer Sprout media space.

All in all, participants seemed to strongly prefer an arrangement that would let both people see and experience the same thing, as that would best support and equality of participation and experience:

> *Sprout-to-Sprout, worked better just because they're the same technology in front of us. So, it was equal communication.*

While media space asymmetries of participation have been considered in previous investigations [12], our work suggests device heterogeneity may be an important factor amplifying such asymmetries.

6 Discussion

We enumerate our contributions and discuss opportunities for future research and design.

6.1 Contributions

Previous work has focused almost exclusively on homogeneous device arrangements in such systems. Our work addresses this gap by contributing a technical infrastructure for heterogeneous device media spaces and an empirical investigation that uses this infrastructure to answer three research questions regarding device heterogeneity in media space collaboration.

Empirical Contribution. We were motivated by three research questions. First, we asked how the device (Sprout vs. tablet vs. phone and tablet DUI) used in a media space collaboration affects the participants' experience. To answer this, we asked half of our participants to try three different devices to collaborate with their partner on three similar tasks (i.e., "switching" role). We found that switching participants were strongly affected by the device type, generally preferring the Sprout, expressing highest levels of satisfaction with the Sprout, finding the Sprout to be most similar to unmediated collaboration, identifying the Sprout as the condition with the highest level of perceived message understanding for drawing tasks, and participating more equally when using the Sprout. This is not particularly surprising, but is nonetheless important to consider and quantify as more collaboration moves to commodity mobile devices [1], as it may make the commodity device user into a second-class citizen when collaborating with a partner who uses a richer device.

Second, we asked the converse question: how does the device used by one's communication partner in a media space collaboration affect one's experience?

To answer this, we asked half of our participants to do the same task, with the same partner, and the same device three times (i.e., "non-switching" role), while their partner switched between a Sprout, table, and tablet & phone. We found that the partner's device affected the non-switching participants' experience in similar though more muted ways. For example, non-switching participants expressed a strong preference for the condition where their partner used the Sprout as well and exhibited similar metrics on measures of satisfaction to their partner in each of the three conditions. The non-switching partner was also affected by the asymmetry of participation in the commodity mobile device conditions, receiving less active help and participation from their collaboration partner. The effect of device choice on the collaborative partner is important to consider since many may think of one's device choice as an individual matter, without reflecting on how it may affect the collaboration as a whole.

Finally, we sought to uncover specific concerns and challenges introduced with the use of heterogeneous device arrangements in remote collaboration. We uncovered four specific concerns that are amplified by heterogenous device arrangements:

- As scaling and perspective across devices differ, collaboration feels less natural and more mediated
- Establishing common ground may require additional effort when working on different devices
- Adding multiple screens or displays (i.e., DUIs) may make it difficult to attend to the right signals
- Equality of participation may be affected by partners having devices with different affordances

These aspects are critical to consider when designing collaboration systems and infrastructures that may connect people via diverse devices or systems of devices.

System Contribution. Our second contribution is in the design and implementation of the ShareSuite infrastructure. The ShareSuite allows for device-agnostic implementation of media spaces that combine synchronous real-time video-based task- and person-spaces. The implementation consists of a backend signaling server and device-side WebRTC components that can be easily implemented on many commodity devices. To demonstrate this infrastructure's ability to support collaboration across multiple devices, we implemented three example device sets: a fully duplexed pro-com HP Sprout device, a commodity tablet, and a distributed user interface (DUI) that uses a commodity tablet and a commodity smartphone in concert.

6.2 Implications for Research

Almost all the previous studies of synchronous collaboration have focused on the deployment and evaluation of the user experience in homogeneous device arrangements (e.g., [5, 6, 14, 29, 41]). As remote collaboration systems and ideas move beyond research prototypes and onto commodity devices (e.g., as the HP Sprout has enabled a duplex pro-cam system to be instantiated on an off-the-shelf device), it is critical to consider the role of device heterogeneity. There is no way to ensure that all communication partners will be using devices with equal screen sizes, equal precision pointers, and similar hardware capabilities. This paper is a first step in this direction and we suggest two critical future research directions.

First, our findings on establishing common ground suggest that a salient feature of heterogeneous device systems may be not just in the specifics of the device, but in the factor of heterogeneity itself—collaboration may be affected by using different devices regardless of the specific features of the device. This effect is confounded in our study since the only homogeneous condition is also the richest media space. A future investigation may set up a 2×2 design where two interfaces are each investigated in heterogeneous and homogeneous arrangements. Based on our effect sizes, a comparison between a pro-cam and a tablet-based interface may yield most substantial results.

Second, we investigated device heterogeneity in a lab setting to be able to quantify the effects of the factors and gather recorded observational data. However, this controlled setup inherently advantages the richest condition since the main benefits of commodity mobile devices (e.g., portability, quick setup) are irrelevant in the controlled setting. Field studies of device heterogeneity in collaboration may reveal a different set of trade-offs and long-term effects. For example, the inequality of participation (see Sect. 5.4) observed in our study may be ameliorated by being able to contribute in previously unsupported contexts (i.e., contributing at 75% capacity but doubling the number of sessions where one is able to contribute would still be a net positive). We contribute the robust ShareSuite infrastructure for future field investigations.

6.3 Implications for Design

As media space systems move onto commodity devices and out of the lab, it is increasingly important to consider implications for design to help manage some of the challenges that arise from heterogeneous device use in collaboration.

Reveal the Constraints of the Partner's Device. The partner's device affects the collaboration experience. Aspects like screen size, available pointing precision, and camera perspective may all influence how the remote partner can collaborate. Once both participants are aware of each other's constraints, they may adapt their approach to avoid some of the challenges. Currently, most media space systems leave the participants guessing about the constraints of their partner's device on the unspoken assumption that everybody is connecting via an equivalent device setup. At the very least, systems should reveal the remote device display size(s), pointing precision, and other key capabilities to all collaboration parties.

Provide a Task-Space Feedback Window. Understanding exactly what one's partner is seeing was an important challenge that was amplified by the use of heterogeneous devices. Standard video chat typically provides a "feedback" window showing how the front-facing camera feed appears to the partner. Though the practice is not entirely uncontroversial [42], it nonetheless has been shown to be an important tool in managing one's actions towards the partner (e.g., [43]). It may be beneficial to provide a similar feature for the task-space in dual-space media spaces, particularly when it is possible to represent both the remote size and any degradation, latency, etc. due the transmission to help the partner on the richer interface be more aware of the others' experience.

Manage Attention Switching Between Multiple Displays. Our study, like previous work in other collaborative contexts [43], has revealed some challenges with managing attention between multiple displays, whether in pro-cam media spaces or in multi-device distributed user interface systems. One solution may be abandoning the idea of multiple displays and placing both the person-space and the task-space on the same display. However, this solution would reduce the screen real estate available for the task-space and undermine the natural division in orientation that people appreciated and rated as more similar to unmediated collaboration. An alternative solution may be to increase awareness of which display is currently attended by the partner. As webcam-based eye-tracking systems become more robust, one can provide a visual indication of the partner's gaze to help inform joint attention across multiple displays.

7 Conclusion

The assumption that each partner is using an identical device to connect in a media space collaboration may impact HCI's ability to transition research prototypes to commodity systems and into wide adoption. We contribute the ShareSuite infrastructure which supports implementation of device-agnostic dual-space media spaces. We demonstrate this infrastructure by instantiating it on three diverse devices sets. We conducted a mixed-methods investigation of device heterogeneity in media space collaboration, examining the effect of device on its user and the user's communication partner. We identify four core challenges introduced by heterogeneous device arrangements. First, scale and display arrangement may vary across device sets, which may make the interaction feel more mediated and less natural. Second, the two partners may now have different visibility and perspective of the task space, requiring additional effort to establish common ground. Third, diverse number and arrangement of displays and spaces may make it more difficult to achieve joint attention. Finally, asymmetries in the capabilities of each device (e.g., precision in pointing) may lead to asymmetries of participation, affecting the collaboration. We provide implications for future research and design to further investigate these issues.

Acknowledgements. We would like to thank to Mirjana Spasojevic, Mia Minhyang Suh, Mithra Vankipuram, Sabirat Rubya, Aysegul Kuzulu for their support and all the participants for their contributions. This work is supported by NSF grants #1526085, #1651575.

References

1. Anderson, M.: Technology device ownership: 2015 (2015). http://www.pewinternet.org/2015/10/29/technology-device-ownership-2015/
2. Xanthopoulos, S., Xinogalos, S.: A comparative analysis of cross-platform development approaches for mobile applications. In: Proceedings of the 6th Balkan Conference in Informatics, pp. 213–220. ACM, New York (2013)
3. Miller, H., Thebault-Spieker, J., Chang, S., Johnson, I., Terveen, L., Hecht, B.: "Blissfully happy" or "ready to fight": varying interpretations of emoji. In: ICWSM 2016 (2016)

4. Yarosh, S., Cuzzort, S., Müller, H., Abowd, G.D.: Developing a media space for remote synchronous parent-child interaction. In: Proceedings of the 8th International Conference on Interaction Design and Children, pp. 97–105. ACM, New York (2009)
5. Junuzovic, S., Inkpen, K., Blank, T., Gupta, A.: IllumiShare: sharing any surface. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1919–1928. ACM, New York (2012)
6. Yarosh, S., Tang, A., Mokashi, S., Abowd, G.D.: "Almost Touching": parent-child remote communication using the sharetable system. In: Proceedings of the 2013 Conference on Computer Supported Cooperative Work, pp. 181–192. ACM, New York (2013)
7. Unver, B., McRoberts, S.A., Rubya, S., Ma, H., Zhang, Z., Yarosh, S.: ShareTable application for HP sprout. In: Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems, pp. 3784–3787. ACM, New York (2016)
8. Greenberg, S., Rounding, M.: The notification collage: posting information to public and personal displays. Presented at the proceedings of the SIGCHI conference on human factors in computing systems, 1 March 2001
9. Saslis-Lagoudakis, G., Cheverst, K., Dix, A., Fitton, D., Rouncefield, M.: Hermes@Home: supporting awareness and intimacy between distant family members. In: Proceedings of the 18th Australia Conference on Computer-Human Interaction: Design: Activities, Artefacts and Environments, pp. 23–30. ACM, New York (2006)
10. Ishii, H.: TeamWorkStation: towards a seamless shared workspace. In: Proceedings of the 1990 ACM Conference on Computer-Supported Cooperative Work, pp. 13–26. ACM, New York (1990)
11. Judge, T.K., Neustaedter, C., Kurtz, A.F.: The family window: the design and evaluation of a domestic media space. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 2361–2370. ACM, New York (2010)
12. Voida, A., Voida, S., Greenberg, S., He, H.A.: Asymmetry in media spaces. In: Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work, pp. 313–322. ACM, New York (2008)
13. Romero, N., McEwan, G., Greenberg, S.: A field study of community bar: (mis)-matches between theory and practice. In: Proceedings of the 2007 International ACM Conference on Supporting Group Work, pp. 89–98. ACM, New York (2007)
14. Kirk, D., Stanton Fraser, D.: Comparing remote gesture technologies for supporting collaborative physical tasks. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1191–1200. ACM, New York (2006)
15. Yarosh, S., Abowd, G.D.: Mediated parent-child contact in work-separated families. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1185–1194. ACM, New York (2011)
16. Gallud, J.A., Tesoriero, R., Vanderdonckt, J., Lozano, M., Penichet, V., Botella, F.: Distributed user interfaces. In: CHI 2011 Extended Abstracts on Human Factors in Computing Systems, pp. 2429–2432. ACM, New York (2011)
17. Vanderdonckt, J.: Distributed user interfaces: how to distribute user interface elements across users, platforms, and environments. In: Proceedings of XIth Congreso Internacional de Interacción Persona-Ordenador Interacción 2010, Valencia (2010)
18. Hamilton, P., Wigdor, D.J.: Conductor: enabling and understanding cross-device interaction. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 2773–2782. ACM, New York (2014)
19. O'Hara, K., Black, A., Lipson, M.: Everyday practices with mobile video telephony. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 871–880. ACM, New York (2006)

20. Inkpen, K., Taylor, B., Junuzovic, S., Tang, J., Venolia, G.: Experiences2Go: sharing kids' activities outside the home with remote family members. In: Proceedings of the 2013 Conference on Computer Supported Cooperative Work, pp. 1329–1340. ACM, New York (2013)
21. Kim, S., Junuzovic, S., Inkpen, K.: The nomad and the couch potato: enriching mobile shared experiences with contextual information. In: Proceedings of the 18th International Conference on Supporting Group Work, pp. 167–177. ACM, New York (2014)
22. Oduor, E., et al.: The frustrations and benefits of mobile device usage in the home when co-present with family members. In: Proceedings of the 2016 ACM Conference on Designing Interactive Systems, pp. 1315–1327. ACM, New York (2016)
23. Unver, B., D'Angelo, S., Miller, M., Tang, J.C., Venolia, G., Inkpen, K.: Hands-free remote collaboration over video: exploring viewer and streamer reactions. In: Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces, pp. 85–95. ACM, New York (2018)
24. Venolia, G., Tang, J.C., Inkpen, K., Unver, B.: Wish you were here: being together through composite video and digital keepsakes. In: Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services, pp. 17:1–17:11. ACM, New York (2018)
25. Jones, B., Witcraft, A., Bateman, S., Neustaedter, C., Tang, A.: Mechanics of camera work in mobile video collaboration. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, pp. 957–966. ACM, New York (2015)
26. Gutwin, C., Greenberg, S.: The importance of awareness for team cognition in distributed collaboration (2001)
27. Watts, L., Monk, A.F.: Remote assistance: a view of the work and a view of the face? In: Conference Companion on Human Factors in Computing Systems, pp. 101–102. ACM, New York (1996)
28. Fussell, S.R., Kraut, R.E., Siegel, J.: Coordination of communication: effects of shared visual context on collaborative work. In: Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work, pp. 21–30. ACM, New York (2000)
29. Kraut, R.E., Gergle, D., Fussell, S.R.: The use of visual information in shared visual spaces: informing the development of virtual co-presence. In: Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work, pp. 31–40. ACM, New York (2002)
30. Tang, J.C., Minneman, S.L.: VideoDraw: a video interface for collaborative drawing. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 313–320. ACM, New York (1990)
31. Wellner, P.: Interacting with paper on the DigitalDesk. Commun. ACM **36**, 87–96 (1993). https://doi.org/10.1145/159544.159630
32. Tang, A., Neustaedter, C., Greenberg, S.: VideoArms: embodiments for mixed presence groupware. In: Bryan-Kinns, N., Blanford, A., Curzon, P., Nigay, L. (eds.) People and Computers XX — Engage, pp. 85–102. Springer, London (2007). https://doi.org/10.1007/978-1-84628-664-3_8
33. Ishii, H., Kobayashi, M.: ClearBoard: a seamless medium for shared drawing and conversation with eye contact. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 525–532. ACM, New York (1992)
34. Izadi, S., Agarwal, A., Criminisi, A., Winn, J., Blake, A.: C-slate: a multi-touch and object recognition system for remote collaboration using horizontal surfaces. In: Proceedings of the Second Annual IEEE International Workshop on Horizontal Interactive Human-Computer Systems (Tabletop 2007). IEEE (2007)

35. Higuchi, K., Chen, Y., Chou, P.A., Zhang, Z., Liu, Z.: ImmerseBoard: immersive telepresence experience using a digital whiteboard. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, pp. 2383–2392. ACM, New York (2015)
36. Weibel, N., Signer, B., Norrie, M.C., Hofstetter, H., Jetter, H.-C., Reiterer, H.: PaperSketch: a paper-digital collaborative remote sketching tool. In: Proceedings of the 16th International Conference on Intelligent User Interfaces, pp. 155–164. ACM, New York (2011)
37. Ou, J., Chen, X., Fussell, S.R., Yang, J.: DOVE: drawing over video environment. In: Proceedings of the Eleventh ACM International Conference on Multimedia, pp. 100–101. ACM, New York (2003)
38. Harms, P.C., Biocca, P.F.: Internal consistency and reliability of the networked minds measure of social presence. Presented at the international workshop on presence (2004)
39. Fussell, S.R., Setlock, L.D., Kraut, R.E.: Effects of head-mounted and scene-oriented video systems on remote collaboration on physical tasks. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 513–520. ACM, New York (2003)
40. Hollan, J., Stornetta, S.: Beyond being there. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 119–125. ACM, New York (1992)
41. Fussell, S.R., Setlock, L.D., Yang, J., Ou, J., Mauer, E., Kramer, A.D.I.: Gestures over video streams to support remote collaboration on physical tasks. Hum. Comput Interact. **19**, 273–309 (2004). https://doi.org/10.1207/s15327051hci1903_3
42. Miller, M.K., Mandryk, R.L., Birk, M.V., Depping, A.E., Patel, T.: Through the looking glass: the effects of feedback on self-awareness and conversational behaviour during video chat. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, pp. 5271–5283. ACM, New York (2017)
43. Yarosh, S., Inkpen, K.M., Brush, A.J.B.: Video playdate: toward free play across distance. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1251–1260. ACM, New York (2010)