Video Playdate: Toward Free Play across Distance

Svetlana Yarosh
Georgia Institute of Technology
School of Interactive Computing, GVU Center
85 5th Street NW, Atlanta GA 30332, USA
lana@cc.gatech.edu

Kori M. Inkpen & A. J. Bernheim Brush
Microsoft Research
One Microsoft Way
Redmond, WA 98052
{kori, ajbrush}@microsoft.com

ABSTRACT
We present an empirical investigation of video-mediated free play between 13 pairs of friends (ages 7 and 8). The pairs spent 10 minutes playing with each of four different prototypes we developed to support free play over videoconferencing. We coded each interaction for the types of play and the amount of social play observed. The children in our study were largely successful in playing together across videoconferencing, though challenges in managing visibility, attention, and intersubjectivity made it more difficult than face-to-face play. We also found that our prototypes supported some types of play to varying degrees. Our contribution lies in identifying these design tradeoffs and providing directions for future design of video-mediated communication systems for children.

Author Keywords
Children, free play, CMC, videoconferencing

ACM Classification Keywords
H5.3. Information interfaces and presentation: Group and Organization Interfaces: *Synchronous interaction.*

General Terms
Human Factors

INTRODUCTION
Free play is characterized as an unconstrained activity in which children initiate and direct their own interaction with each other and their environment [13]. Time spent in free play is key to a child's cognitive development [21] and to developing sociocultural and emotional competencies during the period between infancy and adolescence [19]. However, an extensive two-year survey of families across 16 different nations found that time spent in free play and experiential learning is decreasing alarmingly in favor of increased time spent watching television [18]. For example, a time study of American children showed that they spend an average of 12 hours per week watching TV [12]. Time spent passively watching television does not provide as many opportunities to develop social, verbal, and cognitive skills as play and most pediatricians recommend limiting TV time for young children [18].

The individual nature of television and other recent technology-mediated leisure pursuits have been implicated in the decline of the neighborhood and America's social capital [17]. However, a study conducted in the U.S. found the majority of respondents agree that the Internet and mobile telephone have actually brought them closer to their

Figure 1. Children attempting to play together via videoconferencing using a laptop (pilot study condition 1) and using a TV (pilot study condition 2).

CHI 2009, April 4–9, 2009, Boston, Massachusetts, USA.

friends and families [14]. Interestingly, increased use of these social technologies came at the expense of TV time and did not reduce in-person time with family and friends [14]. It is clear that at least for adults, social media provide an opportunity to replace passive entertainment with social interaction. However, young children usually have not mastered the communicational competencies to make full use of text- or audio-based technologies. Previous investigations note that video-mediated communication may be particularly appropriate for communication with children because it provides better resources for grounding conversation and supports playfulness in remote communication [1]. Our goal in this work is exploring how synchronous video technologies may better serve children as they engage in remote free play.

Synchronous video communication has been common in the workplace for many years, but has only recently become readily available for maintaining social relationships in the home as webcams become more prevalent and programs such as Skype and MSN Messenger support video chatting. Despite a great deal of media attention to the potential of this technology for families (e.g. [10]), it seems relatively few families have actually incorporated video into the ecology of CMC tools they use to stay in touch (e.g., 2 out of the 18 families studied in [1]; 2 out the 28 surveyed in [20]). The majority of homes with children have both the computer and the broadband Internet connection required for videoconferencing, but the current technology was designed to support remote work meetings rather than play. While face-to-face communication has many advantages for families, creating better opportunities for children to engage together in distributed play could allow extended family members (e.g., cousins) to play together, hospitalized children to stay in touch with their families, and geographically separated friends to maintain contact.

We provide an empirical investigation of video-mediated remote free play between pairs of friends, ages 7 and 8. We seek to answer three questions:

1. What are the current opportunities and challenges for remote free play using videoconferencing?
2. How can we better support social play mediated by synchronous video?
3. What are the tradeoffs between features of synchronous video technologies and the type of play they support?

We begin by situating our work in the context of previous research on children's play and Computer-Mediated Communication. We present an exploratory study of 3 pairs of children engaging in free play via currently available videoconferencing. We propose four potential approaches to address the challenges we observed. To understand the inherent tradeoffs for each of our solutions, we present an empirical study of 13 pairs of children participating in free play activities using each of our prototypes. Finally, we discuss the implications of our results for developers of synchronous video communication technologies.

RELATED WORK

We draw upon a rich history of previous investigations of play, video-mediated communication, and innovative designs from HCI fields.

Theory of Play

Social scientists have been exploring children's play for many decades, from the early investigations of Vygotsky and Piaget to the current work of the National Institute for Play. The National Institute for Play [16] identifies 7 patterns that constitute the elements of play: (1) attunement play is the interplay of affective feedback (e.g. returning a smile); (2) body play is exploring the possibilities of motion, contact, and place in space; (3) object play involves the manipulation of an object to explore its affordances and characteristics; (4) social play involves acting with and towards others with reciprocation, empathy, and regard; (5) pretend play focuses on taking on the role of another or assigning new roles to objects in the environment; (6) narrative play includes relating stories of real or imagined occurrences; and (7) transformative-integrative play is the act of generating and implementing new ideas. These elements are often combined during free play episodes.

In our study, we code for all of the above types of play, however we particularly focus on the social play that occurs. We draw upon the work of Parten and Howes, who observed that social play between children is characterized by 5 stages of mutual regard and reciprocity [13:156]. At the most basic level, children participate in parallel play—activities in proximity to one another, but without actually engaging in social behavior. At higher stages, children direct social behaviors to one another and respond to the behaviors of their play partners. At the highest level of social play, children engage in a complementary and reciprocal activity that requires both verbal and non-verbal coordination on their parts. During free play children may frequently switch between various types of social play, with more time at the higher levels generally demonstrating greater social skill development.

Designing to Support Children's Play

Many innovative prototypes have been designed to support social play between children. Most relevant to our study, several projects have looked at supporting social play between collocated children by creating immersive storytelling environments (e.g., [4]) or allowing them to participate as characters in a virtual world (e.g., [11]). However, interactions with such content are often limited to a set of constrained behaviors, while we are interested in exploring more child-driven free play. Cassell & Ryokai [6] emphasized this sort of interaction in creating StoryMat—a system that combined audio-recording, tangible toys, and projection on a special rug to allow children to create, modify, and share their own stories. We took inspiration from this work and share its focus on unstructured child-driven play, though we seek to expand this idea to interaction between remote participants.

A few investigations have looked at remote play with children. Bonanni et al., [5] explored play via networked tangibles by designing a set of wireless dolls where repositioning of the local doll would be mirrored by the remote doll. One of their main findings was that the dolls alone were not enough—co-play occurred only when the children were also provided with a synchronous audio connection. The Virtual Box [7] project explored asynchronous remote play by allowing a parent to place a virtual gift box on the floor plan of the child's home that the child could later try to find with the aid of a location sensitive PDA. Unlike both of these projects, we are not looking to create a new game or toy that can be played remotely, but rather we are investigating video as a lightweight way of sharing familiar free play experiences.

Video-Mediated Communication

Videoconferencing and video media spaces have had a long history of exploration in the work place [3], allowing us to draw on relevant findings from this domain. Several CMC theories point to the fact that synchronous video may be an appropriate medium for play tasks. Media Richness Theory emphasizes that ambiguous and uncertain tasks require more immediate feedback, more cues for communication, and more emotional awareness. Social Presence Theory suggests that video affords social awareness of the partner's state in a way that is more similar to in-person interaction than other media, and thus may be better at supporting tasks that are usually carried out in-person. However, there is significantly less theoretical grounding to suggest what *kind* of synchronous video may be good for supporting free play. Empirical work in the space highlights that videos of the collaborative task space and videos of the larger context of the remote room may often be more useful to remote participants than the face-to-face video view [9]. But, Gaver et al., [9] also emphasized that all three types of views (face-to-face video, task space video, and room context video) were useful, that switching between video views was challenging, and that it would be difficult to have a stationary arrangement of cameras meet the users' work needs. The role of available views and arrangement of cameras in supporting play is still an open question.

There have been a few investigations of play over synchronous video. Batcheller et al., [2] observed groups of college student playing the social game "Mafia" mediated by videoconferencing. They found that play over videoconferencing was fun for participants, but introduced new challenges in terms of managing attention, signaling to remote partners, and social distance. Another investigation examined parent-child pairs playing a board game together in a media space that included face-to-face video and a shared tabletop video task space [24]. They found that parents and children were able to socially negotiate rules and access to the physical artifacts in the remote space, but did not investigate social free play between peers. Mueller et al., [15] examined a class of prototypes called exertion interfaces which combine projection of full body video and computer vision techniques to allow remote partners to play sport-like games together. They discovered that exertion interfaces have a great potential to create and strengthen social bonds between adult strangers. However, all of these investigations asked participants to play games with pre-established rules, free play over videoconferencing was not investigated. The previous work shows that play is possible over videoconferencing, but also frames the open problem we seek to address: what features of synchronous video technology can support remote free play between children?

PILOT STUDY

To understand the challenges and opportunities for free play provided by current videoconferencing technologies, we conducted an exploratory study asking children to play together using a standard videoconferencing client (Windows Live Messenger). The children were asked to play in two different configurations (see Figure 1): (1) a laptop and (2) a large screen TV. The laptop provided the typical face-on view of the participant with the camera angled to also show the space on the table in front of the laptop. In the TV setup, the camera was focused on the floor space in front of the TV revealing an area of about 8 square feet. The signal was transmitted using standard videoconferencing software in VGA resolution, to represent currently available technology.

Three pairs of children were invited to play in separated rooms using each of these systems for 15 minutes. The participants were two boys (5 & 7, brothers), two girls (8 & 7, friends), and two boys (6 & 6, friends). We asked each pair to pick out some toys that they wanted to use from a set of action figures, dolls, etc. and invited them to pretend they were playing while at different houses. Their interaction was video recorded and the children were asked to answer a few quick questions about their experience after each condition. A limitation of this initial study was that several of the participants were related to the paper authors.

For our participants, we observed that free play was possible over videoconferencing, but limited to short periods of mutual play interweaved with longer periods of parallel play. Some of the examples of mutual play that we observed included: pretending to be TV characters, singing a song together, role playing using dolls, and narrating a scenario using action figures. In both conditions, children struggled to understand the unique communication asymmetries that videoconferencing presents. Children would assume that they were visible to their partner if they could see the screen and would assume that the sound volume was the same on both sides (so, they would speak louder if they had trouble hearing the other person). A big issue that all three pairs seemed to have was being able to see each other's toys clearly enough to play. Holding the toy up to the camera required a pause for meta-communication, but the children wanted to have a clear view of their partners' toys as they were being used.

Comparing the laptop and TV conditions, the children could understand each other better in the laptop condition and paid more attention to their partner. However, they also had to remain relatively immobile in front of the screen. In the TV condition, the two pairs of boys took the opportunity to move around the space more freely. All of the children seemed troubled by the amount of pixilation of the video that occurred in the TV condition (since it was a VGA image on a large screen). We also noticed that the TV condition introduced too much physical distance between the participants. Instead of using the floor space to play, as we anticipated, the children would often try to come right up to the screen (see Figure 1). When asked whether they would rather use the laptop or the TV, all four boys selected the TV, while the two girls preferred the laptop.

Many of the problems that we observed were related to finding the appropriate camera view for a particular situation, whether it was the need to see the partner closer in the TV condition, the ability to be seen while moving around the space in the laptop condition, or the need to better see each other's toys in both.

PROPOSED APPROACHES

We investigated four approaches to understand how different affordances to controlling view may influence free play. We asked pairs of children to play together for ten minutes using each of the following set ups (see Figure 2).

Condition 0: High Resolution Vanilla Videoconferencing

Many of the problems experienced by the children in our pilot study could be attributed to poor fidelity of the video. These problems may be resolved as network connections improve. To evaluate this possibility, we simulated high-resolution low-latency videoconferencing between two lab spaces. Figure 2 shows the setup of the space, including the camera (D) and display (E) used and the smaller side-screen (G) echoing the image currently being sent.

Condition 1: Smart Pan-Zoom-Tilt (PZT) Camera

In the pilot we observed that children wanted to be able to move freely about the space, have a clear view of their partner, and also be able to focus on the toys when appropriate. Though current pan/zoom/tilt cameras may be too complex for children to control while simultaneously attending to their play activity, computer vision may soon be sophisticated enough to accurately predict the ideal area of the room to display. To evaluate the potential success of this future possibility, we used a Wizard of Oz methodology with researchers controlling the PZT cameras (C, in Figure 2) to show what we anticipated to be the most salient view. The heuristic the researcher used was to keep

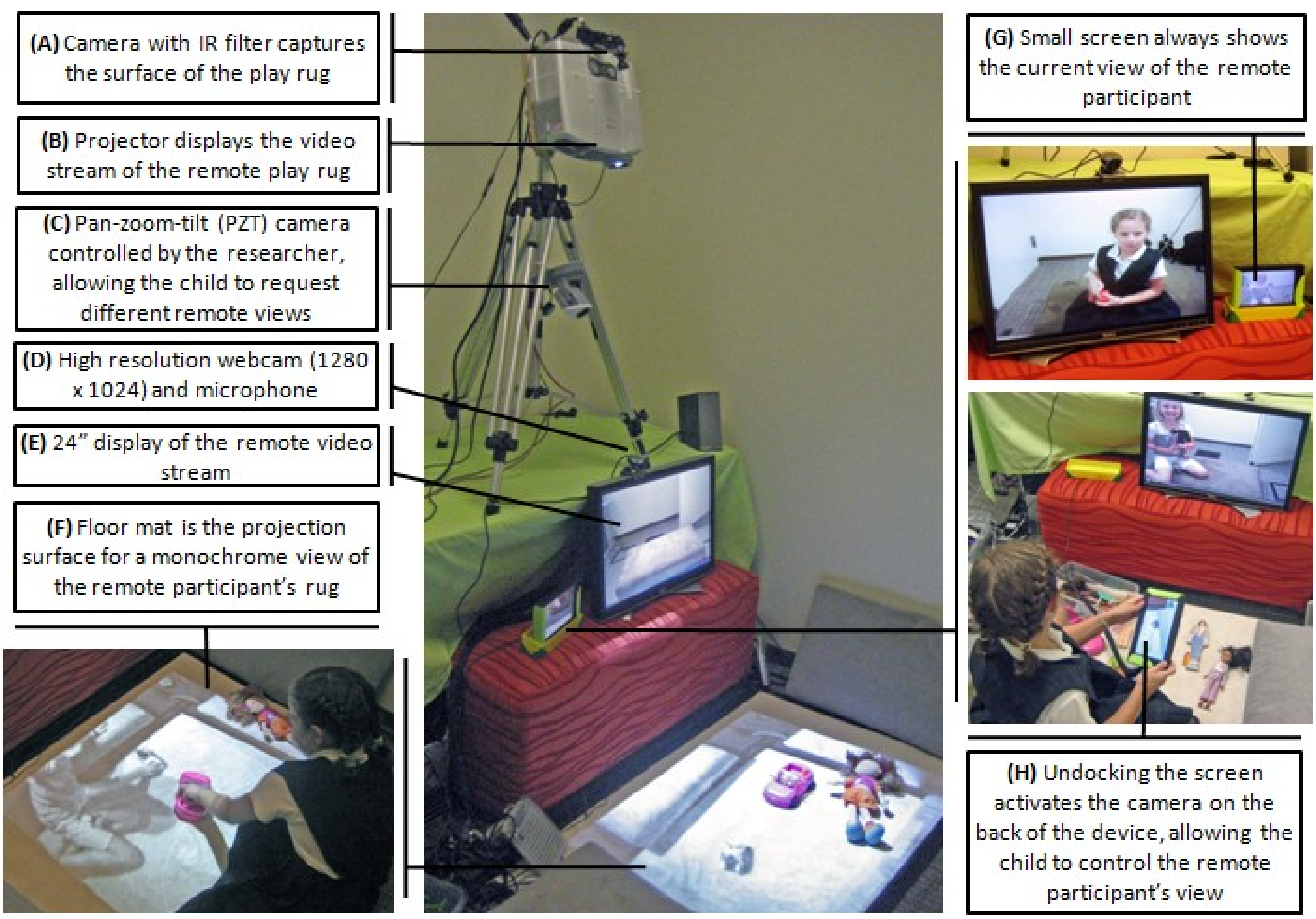

Figure 2. We installed the above setup in two adjacent lab spaces. Different components were activated in each of the four conditions: vanilla (D, E, & G), pan-zoom-tilt (C, E, & G), mobile (D, E, G, & H), and projector rug (A, B, D, E, F, & G).

#	A/S	Webcam Before?	Most Wanted?	Peer Play Rating (out of 5)				
				F2F	Mobile	Rug	PZT	Vanilla
1A	8/M	no	rug	4.22	2.74	4.05	4.06	4.13
1B	7/M	no	rug	4.30	3.75	4.24	4.06	4.10
2A	7/F	no	rug	-	-	-	-	-
2B	7/F	no	vanilla	-	-	-	-	-
3A	8/F	no	vanilla	3.94	2.65	2.81	2.21	3.11
3B	7/F	yes	rug	3.94	2.44	2.76	1.90	3.22
4A	8/M	no	mobile	0.94	3.95	2.95	3.15	1.64
4B	7/M	no	rug	0.62	3.43	2.05	2.52	1.33
5A	8/M	yes	rug	2.59	2.14	1.51	1.12	1.52
5B	8/M	no	mobile	2.56	2.44	1.59	1.24	1.79
6A	8/F	yes	mobile	5.00	4.54	4.59	4.97	4.88
6B	8/F	no	mobile	4.92	4.13	4.53	4.88	4.59
7A	7/M	no	PZT	3.71	2.37	1.12	1.92	1.68
7B	7/M	no	PZT	4.85	1.77	1.24	1.63	1.62
8A	8/F	yes	mobile	4.76	3.78	4.86	4.65	5.00
8B	8/F	yes	mobile	4.86	4.30	4.18	4.59	5.00
9A	8/M	no	rug	4.40	2.74	3.44	1.91	4.31
9B	7/M	no	PZT	4.40	3.87	3.41	2.36	4.81
10A	7/F	no	mobile	4.63	3.78	3.16	3.43	4.47
10B	7/F	yes	rug	4.68	3.19	3.41	3.70	3.53
11A	8/F	yes	mobile	4.87	4.06	2.60	2.29	3.76
11B	8/F	no	rug	4.73	3.39	2.76	1.93	3.41
12A	7/F	no	PZT	4.53	3.97	3.71	4.52	4.00
12B	7/F	no	mobile	4.59	4.08	3.85	4.19	4.07
13A	7/M	yes	PZT	3.91	4.23	4.74	4.31	4.76
13B	7/M	no	PZT	3.97	4.36	4.79	4.19	4.75
Avg. Across Conditions				4.00	3.42	3.27	3.16	3.52
Std. Dev. Across Conditions				1.18	0.81	1.16	1.27	1.28

Table 1. Description of the participants, including whether they have previously used a webcam, the condition they stated they would most want in their home, and their average Howes' Peer Scale score for face-to-face (F2F) play and in each of the four conditions (darker color indicates more social play).

the remote child in view as much as possible. The child was also supported in making explicit changes to their remote view by letting them verbally specify an area of interest to view (e.g., "zoom in on the toy car").

Condition 2: Mobile View Control Component

Mobile video technology will likely be a significant aspect of CMC once mobile networking improves. We gave the child the ability to control their partner's view with a simple mobile video device (G and H, in Figure 2). The mobile screen consists of a 7" monitor with a standard webcam attached to the back, facing away from the viewer. We wanted to maintain the metaphor that this feedback screen always shows the image as it appears to the partner. When docked, the device serves as a standard feedback window with the image being captured by the face-to-face camera (D, in Figure 2). Undocking the mobile device activates the camera on the back of it to allow the child to control the remote participant's view. The local child sees what they are transmitting on the feedback screen and can orient the device anywhere in the space. Docking the device returns the video to the standard face-to-face view.

Condition 3: Camera-Projector Shared Play Rug Space

A promising solution for supporting free play is augmenting the person-space view provided by the TV with a shared view of the task-space. We use a full-duplex camera-projectors system (most similar to the PlayTogether system [22]) to create a shared floor space for play. A camera suspended above the play rug (A, in Figure 2) captures a video stream of the rug surface and transmits it to the remote projector (B, in Figure 2). We project the video stream of the remote floor space directly on top of the local one and vice versa. Like the PlayTogether system, we solve the visual echo problem (i.e., re-projecting artifacts) by installing IR filters on the overhead cameras. This restricted us to monochrome video, but allowed us to use a standard rug rather than a specialized projection surface (unlike the polarization approach described in [24]).

METHOD

We invited 13 pairs of friends ages 7 and 8 to play together using each of the prototype systems to understand how specific characteristics of the video technology mediated free play.

Participants

Our 26 participants (see in Table 1) came from families in the area surrounding a large metropolitan city in the U.S. We explicitly recruited for same-gender non-sibling pairs of children who play well together (though, we later found out that 3A and 3B were brothers). We originally looked to recruit 6 pairs of girls and 6 pairs of boy for a total of 24 participants. However, an error in capturing the video prevented us from coding the peer interactions of Pair 2. This prompted us to recruit an additional pair of girls. We asked the children a few questions about their experience with media. All but one child (12A) reported that they watch TV at home and all but two children (1B & 12B) reported playing videogames at home. There was a mix of experiences regarding whether the child had previously used a webcam to speak with somebody (see Table 1).

Setting

The study took place in two adjacent identical laboratory rooms, with the setup as shown in Figure 2. Both rooms were equipped with a wide variety of age-appropriate toys, equally distributed between the rooms. These included action figures, fashion dolls, toy vehicles, construction blocks, rag dolls, dollhouse accessories, stuffed animals, doctor accessories, and various animal/dinosaur figurines. For safety, a researcher remained in the rooms with each of the children during the study. The parent could choose to observe the study from behind a half-silvered mirror.

Procedure

The children and any adults with them were given a tour of the labs where the study would be taking place. The children were then asked to play together in the same room for 10 minutes to get a baseline of their free play behavior. We encouraged the children to play together and explore the space, however we did not interfere or insist on social play taking place (e.g., Pair 4 often played apart). Following this, the children were taken into separate rooms and asked to play together for 10 minutes using each of the four technologies in counterbalanced order. Between conditions, the children were asked what they liked/disliked about the particular technology, provided with a break to visit each other's room, and given a short demo of the next condition. After completing all of the conditions, we asked the children to specify which condition they liked best/least, was easiest/hardest, found most fun, and would most want to have in their house (we report the most interesting of these in Figure 3). We asked each pair to set aside 2 hours for the study; most participants were done in 90 minutes.

Instruments

During the study, we recorded video of each child. To understand the type of play that occurred in each condition, we coded for the elements of play using the National Institute for Play patterns [16]. There is no validated coding schema for the patterns yet, so we simply coded for a Boolean assessment of whether the particular type of play occurred in each 10-minute play interval. However, we wanted a more detailed assessment of the levels of social play that occurred, so following the study, each video was coded for social play using the Howes' Peer Play Scale [13:157]. The Peer Play scale is a validated measure commonly used in the education domain. The researcher observes the child in 15 seconds intervals and makes a determination as to their level of social play for each interval of time. Level 1 represents parallel play—participants are involved in the same activity, but seemingly unaware of each other. Level 2 is parallel play with mutual regard—participant periodically attends to the others' play but does not interact directly. And so on, to Level 5 which represents mutual play where behavior and verbal communication are driven by the common task. Table 1 provides a summary of these results.

RESULTS

Though there was a great deal of individual variability, children were able to play together over videoconferencing, though not as easily as in-person (see Table 1). We review the findings from each condition and observations that held across conditions. Figure 3 displays counts of which conditions participants wanted, found hardest, found easiest, felt were the most fun. Figure 4 displays the number of participants observed in narrative, pretend, and movement play, separated by condition.

Vanilla Videoconferencing

Vanilla videoconferencing had an average social play rating second only to face-to-face play (see Table 1). There were no components of this system with which the children could actually interact, so many of them rated it as "the easiest" of the four conditions (see Figure 3). Participant 2A said that she liked how in this condition you "just pay attention to one place and just focus on playing." However, this condition was not without its difficulties. Framing was a problem for many of the children. As 8B stated "…you had to make sure [the toys] were at the right angle or it would be cut off." Even the children who were quite adept at holding toys up to the appropriate camera could only do so while attending carefully to the feedback screen. Luckily, most of the pairs came up with a successful strategy for playing together without constantly attending to what is being shown—creating a staging area. Several of the children would strategically place one of the cushions in the play area in such a way that they knew whatever they do on top of the pillow would be visible to the other person. This setup only had to be done once and several of the pairs verbally guided each other in setting up this staging area. This freed up cognitive resources for pretend and narrative play, but restricted physical play and movement for those that used this strategy (see Figure 4). While children recognized some of the benefits of "easy as pie" (3A) videoconferencing, it was not as highly rated on being fun or selected as the most wanted by most of the children.

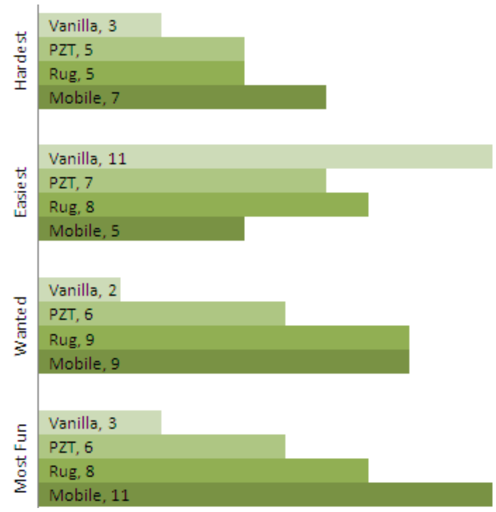

Figure 3. Number of participants (out of 26) listing a condition as hardest, easiest, most fun, or as the one the participant would most want in his or her home. Children could make 0, 1, 2, 3, or 4 selections (accounting for totals different than 26).

PZT

In the PZT condition, the researcher controlled a camera to try to keep the remote child in view and to follow explicit instructions from the local child about what they wanted to see on screen. We saw three types of interaction in this condition: passive, object zoom, and dodge-the-camera. Those who participated in this condition passively were content to let the camera choose the appropriate view and rarely or never attempted to direct their view. 8A appreciated the ease of this "no matter where you go, you can still see the things on the ground."

The second type of interaction involved those that wanted to direct the camera more actively. Children seemed to have trouble negotiating who should be doing the view control in which room. Often, they wanted to give directions for both

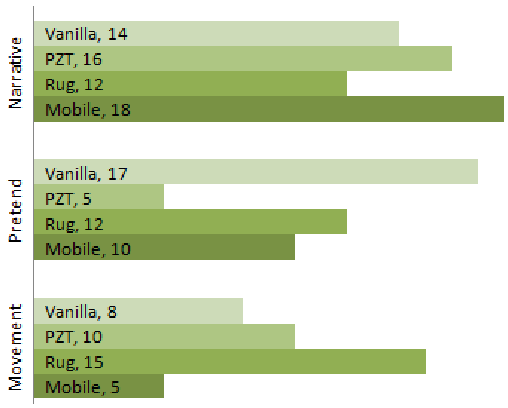

Figure 4. Number of participants (out of 26) observed engaging in narrative, pretend, & movement play by condition.

the cameras and ended up trying to coordinate this socially (e.g., "okay, ask yours to zoom in on the [toy]") or even plan complex sequences of views (8B: "so start out so we can't see them, and then we go here, and then ta-ta-da!"). The fact that they could not control what their partner may choose to view introduced tensions when wanting to keep an object or oneself hidden. Several pairs made statements similar to 2B: "don't look here, I want to do a surprise" while setting up a scene.

The tension between being followed by the camera and the desire to sometimes not be seen, lead some participants to invent a new game (called "dodge-the-camera" by pair 13). This consisted of moving around the room faster than the camera could follow you or finding hiding places in the room where the camera could not see you. Though the children who tried this game reported that it was great fun, it really amounted to a game between the camera operator and the child more than social play between the two children.

Projector Rug

The projector rug allowed children to view a monochrome projected video stream of overhead video from their partner's rug. Yarosh et al., [24] examined a similar technology for play between parents and children and found that the children were able to understand and manage the interweaving of physical and projected spaces with the parent's support. However, it seems that some of the children in this study were not able to manage that interweaving effectively. When a physical object was already occupying the space, monochrome projection of the remote activity over that space was too subtle to attract attention:

12B puts a stuffed turtle on the rug, on top of 12A's projected head. 12A doesn't notice or respond. 12B prompts: "there's a turtle on your head!" 12A is confused, looks up, down, but can't see any turtle until 12B directs her to move aside, but then the turtle isn't on her head anymore—it's just on the floor.

As 13A complained, you had to "look everywhere on the rug to see what [13B] was doing."

The always-on-top quality of the projection was confusing for many of the children. 13A was adamant that his partner's perception was in error "[13B] always said he was standing on me, not that I was standing on him!" While being able to occupy the same space allowed for some fun physical play (in fact, this condition had the most movement play), this feature also made it difficult to come to an agreement about the interaction between physical toys. As 6A&B played with cars, they could not agree on an interpretation of events (6B: "It's rolling over you!" 6A: "No, it's rolling under me!").

The main difficulty with the projector rug condition was pointed out by 12A: "it was hard to see the screen and the rug at the same time." This led to some missed opportunities for social play:

11A is dressing a doll, periodically looking up at the screen at 11B's activity. 11B had finished dressing her doll and placed it on the rug to show. She is attending to 11A by looking at her projection on the rug. 11A keeps looking at the screen to find 11B is looking down. Both partners think that the other is not attending to them and move on to parallel play for the rest of the session.

Some of the pairs became aware of this confusion regarding attention and several of the pairs made the explicit verbal agreement to "look at the screen, not at the floor" (2A). Unless the activity required occupying the same space (e.g., pretending to stomp on each other), the vertical display was consistently selected by the children over the rug surface. The main reason cited was the lower resolution and monochrome nature of the rug surface. As 11A described, "this is what TVs looked like a *long* time ago."

Despite these problems, children saw potential in this technology and often selected it as the one they would most want to have at home. 4A pointed out that the projection rug allowed him to do two things that he could not do otherwise: "we can fight, but [4B] can't hurt me," and "we can create our ideas *together* and make something."

Mobile

In the mobile condition, the child could pick up the small mobile screen and use the camera on the back of the device to control their partner's view. Many of the children considered this condition to be the hardest: you had to hold the device while composing your shots, and as 10B noted "it was hard to understand what [10A] was trying to do when she was moving it around." 6B noted an important shortcoming of this condition: "it did not look like you're looking at your friend … and you could not see their face or what they were paying attention to." The pairs that used it successfully often implemented turn-taking strategies ("first I show my doll, then you show your doll") and verbal attention-seeking signals ("ta-da!") to be able to play together. Because of these difficulties, 7 out of the 26 participants (3A&B, 4B, 7A&B, 11A&B) only used the mobile screen briefly or not at all.

Despite the challenges it presented, those who persisted with the mobile condition found it to be very compelling. 6A used the camera to give 6B a tour of a tent she set up for her doll and afterwards exclaimed: "you could *literally* be where the person was playing!" Most participants selected this condition as "the most fun" and it tied with the projector rug for being the "most wanted" condition (see Figure 3). Additionally, the mobile condition invited and supported narrative play (see Figure 4).

We did not anticipate that one of the most appealing points of the mobile condition would be the ability to turn the camera back on the TV. 10B loved it when 10A points the camera at her own TV "I can see what you're watching!" The biggest draw seemed to be the ability to take turns doing this to "see each other's faces on the big screen." Seven of the 13 pairs incorporated this activity into their play in the mobile condition. It seems that the children really wanted to see themselves and their friend together on the TV screen. 4A tried to share screen space by creating a makeshift tripod for his mobile device, pointing it at his TV, and then moving between the TV and mobile device, facing the camera. 4B was thus able to see both himself and his brother on screen at the same time.

Across Conditions

Despite the fact that children were generally successful at playing together over videoconferencing, two problems became apparent across all of the conditions: 1) the child's model of the other person's view was vague or incorrect at times, and 2) making two toys interact over videoconferencing was awkward.

Several participants provided evidence that maintaining mental models of one's own and the partner's view was a challenge. 1A pointed the mobile screen at his TV while asking "wanna see what you're looking at?" not realizing that he was actually showing what *he* was looking at himself. Pair 13 was by all other visible measures quite adept at holding up toys to the camera and yet they experienced confusion:

> *13A docks his small screen after having pointed it at his own TV. "Can you still see yourself?" he asks, clearly expecting a 'no' answer. 13B misinterprets and checks at his own small screen: "Yes." 13A looks at his own small screen to figure out what 13B is talking about. 13A insists: "No, I mean on your little screen!" 13B look again at his small screen, even more confused: "Yes, I can see myself!"*

Though the majority of the children were able to unravel this intersubjectivity well enough to keep playing, maintaining a mental model of what the other person sees was a cognitively demanding task that made play over videoconferencing more difficult.

Making two remote toys "interact" with each other over videoconferencing was most often accomplished by holding the toys towards the camera at the same time. As Pair 12 demonstrate while playing vet: "Hold your [puppy] up to the screen. I am going to comb you hair." However, this mechanism for interaction is also quite limiting:

> *9B holds his Cyborg action figure to the camera. 9A 'attacks' it with his Batman action figure by holding it to the camera. 9A moves Batman through a flip: "[9B], look at this!" But, 9B's view of the TV is blocked by his own outstretched arm: "I didn't see anything." 9A: "That's 'cause your Cyborg is in the way!"*

Though most of the children were able to play together despite these difficulties, Pair 7 was not able to overcome these barriers. When in the same room, the two boys played well together. When separated, 7B implemented the typical "staging area" for his action figures, but 7A consistently tried to line up his action figures immediately in front of the display (where they were out of view of the cameras). 7A would then manipulate his toys to interact with the on screen image of 7B's toys. To 7B, 7A's actions were unclear and seemed to be independent of his own. Though the boys continued providing each other with some mutual regard, they lapsed into parallel activity during remote play. It is clear that there is room for significant improvement in supporting children's free play over video.

DISCUSSION

Though most children in our study were successful at free play over videoconferencing, they also faced challenges to managing visibility, attention, and intersubjectivity during play that could potentially be addressed through design.

Managing Visibility (and Invisibility)

Seeing your partner's toys and being able to show your own is a prerequisite for many types of free play. Our pilot study found that low resolution and difficulty framing play within camera view were the most salient problems children experienced with playing over videoconferencing. In the follow-up study, our prototypes addressed the resolution concerns, but it was the children themselves who came up with the solution to the framing issue. Most of the pairs used the cushions around the play area to establish a stage for their toys that they knew was clearly visible to the other person. To simplify this process of setting up the stage and to provide support for children who do not independently arrive at this approach, video systems for supporting free play could include a movable tray that would be tracked by the camera and kept in-focus at all times. Unlike tracking the child, this approach resolves some of the tensions between wanting to be seen versus wanting to remain hidden. We saw that there were many situations where play involved hiding object or oneself from the partner. A shortcoming of the PZT condition was that it made it more difficult to coordinate hiding or reason about what your partner can or cannot see. Tracking a known tray rather than the person would simplify the tracking algorithm and provides a heuristic to help the child reason about what is visible and what is not. Objects can be moved to the staging area only when they are ready to be shown or the staging area can be blocked from view by shielding it from the camera with one's body.

Managing Attention

During free play, children frequently shift attention between individual and mutual activities. Witnessing your partner attend to your activity is an invitation for greater social engagement. For adults, being able to see the face-to-face view of the other person does not seem to be as important as having a good view of the shared task space [9]. However, for children, a face-to-face view of their partner was absolutely key to social play as it was the only reliable clue to the direction of their partner's attention. Seeing their partner attend to their activity led to greater social play, while perception of inattention led the children to play in parallel instead. We could encourage social play by using visual or audio cues to highlight to the child when their partner is regarding their work (or has recently done so).

Managing attention becomes significantly more complicated with multiple displays. Many of the children could only meaningfully attend to one display at a time. With the projector rug condition, we saw that multiple display areas made it difficult for some of the pairs to understand the direction of their partner's attention. The pairs often came to some understanding about which area would be their primary area of attention by taking into account the nature of their current task. There may be a benefit to making multiple areas *available* because each display type may provide different affordances (e.g., supporting movement play), but it is important that the child knows which area to attend to and which area signals that their partner is attending to them. One solution may be allowing only a single display to be active at a time, requiring the children to make an explicit switch. However, clever management of multiple cameras and video streams for each display may lead to a less restrictive solution. Elegant view management that both signals the direction of the partner's attention and lets the child appropriately direct their attention is an open challenge for designers.

Managing Intersubjectivity

Intersubjectivity is defined as the capacity for establishing and maintaining a common ground of engagement among participants involved in an activity together [23]. In the context of video-mediated play it involves understanding both what you and your partner see and determining how to act meaningfully towards each other. However, play is a cognitively demanding activity that leaves few attention resources available for maintaining a mental model of what the other person sees. Children who were most successful at framing their play made frequent use of the feedback screen, but many still seemed to get confused about who sees what. Rather than requiring such mastery of perspective-taking, it may be preferable to display the exact same view to all of the participants. While adults may think of videoconferencing as a window to the other room, children may be better served by thinking about videoconferencing as an opportunity to play together "inside the TV." One way this could be achieved is by splitting the screen in two equal sections and displaying each participant on their own side of the screen.

Adults may be self-conscious of seeing themselves on screen [8], but the children in our study had no such reservations—they loved seeing themselves on screen. Displaying both partners on screen at the same time may increase opportunities for pretend toy interaction. Currently, to make two toys "interact," both children would lift them up to their cameras at the same time. While this got the point across, it also required strong understanding of camera position and made it hard to actually see the gestures that the other child was making with their toy. With a shared view, it may be easier for the child to attend to the TV while presenting their toy to their partner's side of the screen. Though the children will not be able to feel their toys touch, they will be able to see and gesture with the toys in the same area of the screen. As computer vision researchers develop better algorithms for background subtraction and person-tracking, we may be able to combine the local and remote video streams in more sophisticated ways.

CONCLUSION

Supporting free play across distance may potentially increase the number of opportunities that children have to engage in social free play. However, current technologies for remote free play present some challenges: making sure that your partner can see what you are trying to show (and cannot see what you are trying to hide); understanding when your partner is attending to you and making sure they know when you are attending to them; and, determining how to act towards each other and each other's toys.

Though all four of our prototypes supported social play equally well, we found that different technologies for managing view lead to different types of play among pairs. A shared task space created through top-down projection supported movement and physical activities together, such as play fighting and tumbling. Providing the children with a mobile device for controlling the partner's view encouraged turn-taking and narrative play. However, when the children had to work less to interpret what their partner was showing them (i.e., vanilla and rug conditions), they could devote more cognitive resources to engaging in pretend play. Designers may be able to leverage the tradeoffs presented by each condition to achieve intended play outcomes.

Our investigation was necessarily limited by its scope, but revealed directions for future work. First, conducting this study in the lab provided us with the opportunity to control the setting and gather high-quality video for coding. However, it also introduced limitations, as some of the children may have been self-conscious about being so closely observed. Future field studies and ethnographic investigations of play across distance may help address this gap and lead to better understanding of the role of setting in investigating children's play. Second, we found that children had strong preferences for technologies that let them do activities together that they did not think they could do before: give doll-perspective tours of play spaces, fight without hurting each other, play dodge-the-camera, show funny close-ups of each other's faces, and see themselves on TV. We were limited in that we could only observe each child for a short time. It would be enlightening to examine how favorite conditions might be different after a week or a month of use. Examining how children's opinions of technology change may be able to provide significant insight to designers. There is still much work to be done before our technologies are viable in the real world, but we are enthusiastic about the potential of supporting free play across distance.

ACKNOWLEDGMENTS

We would like to thank all of the children who participated in our studies; Tom Bartindale, Andy Wilson, Michel Pahud, and Chris Harrison who aided with the hardware setup; and, all others who helped along the way.

REFERENCES

1. Ballagas, R., Kaye, J., Ames, M., Go, J., and Raffle, H. Family communication: phone conversations with children. *Proc. of IDC*, ACM (2009), 321-324.
2. Batcheller, A.L., Hilligoss, B., Nam, K., Rader, E., Rey-Babarro, M., and Zhou, X. Testing the technology: playing games with video conferencing. *Proc. of CHI*, ACM (2007), 849-852.
3. Bly, S.A., Harrison, S.R., and Irwin, S. Media spaces: bringing people together in a video, audio, and computing environment. *Commun. ACM 36*, 1 (1993), 28-46.
4. Bobick, A.F., Intille, S.S., Davis, J.W., et al. The KidsRoom: A Perceptually-Based Interactive and Immersive Story Environment. *Presence: Teleoperators & Virtual Environments 8*, 4 (1999), 369-393.
5. Bonanni, L., Vaucelle, C., Lieberman, J., and Zuckerman, O. PlayPals: tangible interfaces for remote communication and play. *Ext. Abst. of CHI*, (2006).
6. Cassell, J. and Ryokai, K. Making Space for Voice: Technologies to Support Children's Fantasy and Storytelling. *Personal and Ubiquitous Computing 5*, 3 (2001), 169-190.
7. Davis, H., Skov, M.B., Stougaard, M., and Vetere, F. Virtual box: supporting mediated family intimacy through virtual and physical play. *Proc. of OzCHI*, (2007), 151-159.
8. Filho, J.E.D.V., Inkpen, K.M., and Czerwinski, M. Image, appearance and vanity in the use of media spaces and video conference systems. *Proc. of GROUP*, ACM (2009), 253-262.
9. Gaver, W.W., Sellen, A., Heath, C., and Luff, P. One is not enough: multiple views in a media space. *Proc. INTERACT '93 and CHI '93*, ACM (1993), 335-341.
10. Harmon, A. Grandma's on the Computer Screen. *The New York Times*, 2008.
11. Hayes-Roth, B. and Gent, R.V. Story-marking with improvisational puppets. *Proc. of Conf. on Autonomous Agents*, ACM (1997), 1-7.
12. Hofferth, S.L. and Sandberg, J.F. How American Children Spend Their Time. *Marriage and Family 63*, 2 (2001), 295-308.
13. Johnson, J.E., Christie, J.F., and Yawkey, T.D. *Play and Early Childhood Development*. Scott, Foresman and Company, Glenview, Illinois, 1987.
14. Kennedy, T.L.M., Smith, A., Wells, A.T., and Wellman, B. *"Networked Families." Pew Internet & American Life Project Report.* Washington, D.C., 2008.
15. Mueller, F., Agamanolis, S., and Picard, R. Exertion interfaces: sports over a distance for social bonding and fun. *Proc. of CHI*, ACM (2003), 561-568.
16. National Institute for Play. *Play Science: The Patterns of Play*. Carmel Valley, CA, 2009.
17. Putnam, R.D. Bowling Alone: America's Declining Social Capital. *Journal of Democracy 6*, 1 (1995), 65-78.
18. Singer, D.G., Singer, J.L., D'Agostino, H., and DeLong, R. Children's Pastimes and Play in Sixteen Nations. *American Journal of Play 1*, 3 (2008).
19. Stafford, M. Communication Competencies and Sociocultural Priorities of Middle Childhood. In *Handbook of Family Communication*. Lawrence Erlbaum Associates, Mahwah, NJ, 2004, 311-332.
20. Tee, K., Brush, A. J., and Inkpen, K.M. Exploring communication and sharing between extended families. *International Journal of Human-Computer Studies 67*, 2 (2009), 128-138.
21. Vygotsky, L. Play and its role in the mental development of the child. *Voprosy psikhologii*, 6 (1966).
22. Wilson, A.D. and Robbins, D.C. PlayTogether: Playing Games across Multiple Interactive Tabletops. (2007), 53-56.
23. Winegar, L.T. and Valsiner, J. *Children's Development within Social Context*. Lawrence Erlbaum Associates, NJ, 1992.
24. Yarosh, S., Cuzzort, S., Müller, H., and Abowd, G.D. Developing a media space for remote synchronous parent-child interaction. *Proc. of IDC*, ACM (2009), 97-105.