5-2014

One space: Shared visual scenes for active free play

Maayan COHEN

Kody DILLMAN

Haley MACLEOD

Seth HUNTER

Anthony TANG
Singapore Management University, tonyt@smu.edu.sg

Citation

COHEN, Maayan; DILLMAN, Kody; MACLEOD, Haley; HUNTER, Seth; and TANG, Anthony. One space: Shared visual scenes for active free play. (2014). *CHI '14: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Toronto, Canada, 2014 April 26 - May 01*. 2177-2180.
Available at: https://ink.library.smu.edu.sg/sis_research/7994

OneSpace: Shared Visual Scenes for Active Free Play

Maayan Cohen[1,2] **Kody Dillman**[2] **Haley MacLeod**[3] **Seth Hunter**[4] **Anthony Tang**[2]
[1]Ontario College of Art and Design [2]University of Calgary [3]Indiana University [4]MIT Media Lab
mc11om@student.ocadu.ca {kody.dillman | tonyt}@ucalgary.ca hemacleo@indiana.edu hunters@mit.edu

ABSTRACT
Children engage in free play for emotional, physical and social development; researchers have explored supporting free play between physically remote playmates using videoconferencing tools. We show that the configuration of the video conferencing setup affects play. Specifically, we show that a shared visual scene configuration promotes fundamentally active forms of engaged, co-operative play.

INTRODUCTION
Many early psychologists emphasized the importance of physical and social play in early childhood development [7,10]—both for the development of practical motor skills, and cognitive and social reasoning skills. Yet play is frequently mediated through video game technology, where technology is used to bridge the distance between physically remote playmates (e.g. [3,4,11-13]). Using video-based technology to support free play has recently become of interest (e.g. [13,14]). Much of this work concerns sharing activities [3], or artefacts such as books (e.g. [11]) or a surface (e.g. [13])—much along the lines of Montessori's notion of "prepared materials" to help structure and focus interaction [7]. In contrast, our interest has been to engage children in more active forms of free play—where children can actively explore space, and devise rules for interaction with a remote participant.

Our approach is based on the idea that for children, *play* is not strictly limited to toys in a "play area"; instead that a child's body, and his/her interaction with his environment functions as the play space [7]. Thus, our interest is in allowing the child and his/her entire environment to be the subject of play, along with the play partner. As illustrated in Figure 1, where a mother and child play "paddycake", OneSpace (see [6] for technical details) is a video conferencing tool that merges the video feeds of two remote sites into a single shared visual scene (like [3,4,8]). Consequently, the visual scene, inhabited by the child *with* his/her partner, becomes a shared play area.

CHI 2014, Apr 26 - May 01 2014, Toronto, ON, Canada
ACM 978-1-4503-2473-1/14/04.
http://dx.doi.org/10.1145/2556288.2557117

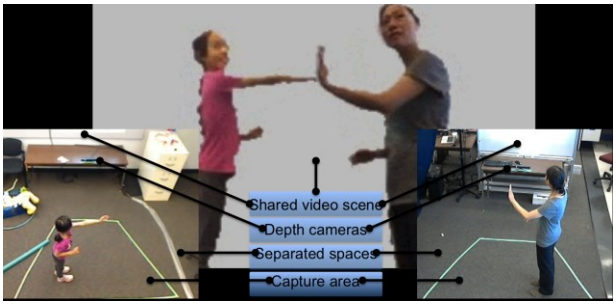

Figure 1. Illustration of OneSpace configuration.

We evaluated OneSpace as a tool to support play between 14 participants (two child-child pairs, and five child-parent pairs), comparing it to a conventional video conferencing configuration. Our findings demonstrate that the OneSpace configuration produces an entirely different "flavour" of interaction—one whose nature is fundamentally physical, and with high levels of engagement (as opposed to a "show-and-tell" and "pretend" interaction as in [12]).

VIDEO WITH CHILDREN: OBSERVATIONS
Several prior authors have explored the use of video conferencing tools with children. While an extensive review is beyond the scope of this note, we synthesize findings from prior work relevant to our purposes here.

Mental model & visibility. Children have a difficult time understanding "visibility" in conventional video conferencing. In real life, if I can see grandpa, he can see me; however, this is not necessarily the case in video conferencing: while I may see his face on screen, I may not be visible to the camera on my side [12,13]. *Managing attention*. Some prototypes have introduced multiple displays—perhaps one on a play space, and one on the person's face (e.g. [12,13])—and while this helps to focus activity to certain screens, it means that a remote partner is sometimes "looking at the wrong screen" (e.g. when I show you something, I want you to see it, rather than look at my face). *"Person space is play space"*. Yarosh et al. explore a "rug play area" condition, where the rug was captured and projected to a remote rug. The children co-opted both the projected rug and the image of the remote person's body into their play, suggesting that a person's body as an object (in addition to their facial reactions) contribute to play. *Field of play.* Size of the capture area is important—likely, if we capture and present more of a person's body and environment, we would expect to see increased physical activity. For example, art installations (e.g. [5]) typically

capture the entire body, whereas conventional videoconferencing focuses on capturing the face.

The OneSpace configuration was designed to address these challenges [6]. Drawing on early work [2,8], we merge two remote spaces, placing participants in a shared visual scene, where their entire bodies (and toys) are represented. The system is implemented using commodity depth-cameras, which allows us to address the "who's in front" problem of earlier green-screen approaches (e.g. [8]). It simplifies issues of visibility, and reduces the number of screens that are important to pay attention to. Most related to our work is WaaZaam, which allows people to personalize/customize the environment and gesturally transform themselves [4].

OBSERVATIONAL STUDY

How does the OneSpace video configuration change the nature of interaction between remote parties? We conducted an observational study where we compared free play between the OneSpace configuration, and a more conventional configuration (akin to Skype). Our goal was to understand the nature of free play in these mediated environments, in terms of the kinds of play (play types), the nature of engagement between the participants, and the level of physical activity participants would engage in. We were interested in whether OneSpace offered alternative ways for movement and play with the physical body.

Characterizing Play and Interaction. To support our efforts, we rely on previous work studying play types and engagement. While many schemes exist for describing play types, we adapted Bronson's scheme to comprise six types of play activity among children [1]: active; make believe; manipulative; creative; learning; object play. As a measure of engagement between participants, we rely on Parten's articulation, where each level describes the level of engagement (or involvement) for a child with regard to the play of another [9]: unoccupied; solitary; onlooker; parallel; associative; cooperative. For each scale, we also categorized "transitions" between activities. While most researchers aim to design systems that really engage children (i.e. at cooperative levels of engagement), our intention was just to study whether there would be differences between the configurations.

Participants. Table 1 shows ages and sex of the 14 people we recruited (two child-child, and five child-parent pairs). We selected only children between 6 and 10 years.

Design & Method. We employed a simple within-subjects design with two conditions: OneSpace configuration and conventional configuration. Participants were recruited as pairs (who knew each other), and after minimal warm-up (5-10 minutes of collocated play) were separated into two independent spaces. They were then asked to play with each of the two configurations (at least 10 minute sessions each, though some groups played longer), where presentation order was counterbalanced across groups. To illustrate each system, children stood within a marked trapezoid, and were told "When you stand here, your partner can see you." We provided participants with a range of toys (about two to three for each of Bronson's play types, e.g. stuffed animals, masks, books, drawing materials, large lego, etc.), and when they were separated, given roughly equivalent sets of toys to play with. We then interviewed the participants.

Group	P1 (sex-age)	P2
G1	M-9	M-10
G2	M-8	M-8
G3	M-6	M-Dad
G4	F-9	F-Mom
G5	F-8	F-Mom
G6	M-6	F-Mom
G7	M-6	M-Dad

Measures & Analysis. We captured video data (three angles) during sessions, recorded audio of interviews, and collected field notes. The video data was coded according to both Bronson's types of play framework, and Parton's engagement framework at a sampling interval of 5s, with the dominant play type/engagement level coded. One rater coded all the video, while a second rater coded two groups' data to ensure reliability (Cohen's κ=0.75 for 7 types of play; Cohen's κ=0.65 for 6 engagement levels).

FINDINGS

Play in the two configurations was different in terms of depth of engagement, and overall type of play. Even though the quality of the video in OneSpace was poorer (image quality and framerate), we saw many activities in OneSpace that were impossible in the conventional condition.

Engagement. The OneSpace setup seemed to encourage a more engaged co-operative and organized play, where both participants would take on an active role in structuring the play with one another. Since participants' actions would have visual consequences in the shared video scene, this would tend to stimulate additional discussion and action. Figure 2 provides an illustrative example of G3, where the child and the father have set up a scene where the blocks (set up by the father) are now being virtually destroyed by the child's dinosaur. The father stands out of the "visible" area, and responds to the child's movements by taking apart various parts of the tower. Here, both the father's and son's actions in their independent physical spaces is entirely

Figure 2. The boy is using a dinosaur to destroy a Lego tower his father built. The father stands outside the capture area and takes pieces off to complete this illusion.

guided by the merged video scene—their interactions are mutually relating to one another. This kind of interaction between body, physicality, and action was extremely common in the OneSpace condition (e.g. Figure 1).

In contrast, as illustrated in Table 2, the conventional configuration exhibited fairly high levels of parallel and associative play. Particularly striking was that while participants *began* playing with one another, they would usually disengage with one another: their play devolving into parallel or associative play—where their interactions were strictly mundane conversation, unrelated to the "play" taking place (i.e. individuals' play did not influence the other's play). This was particularly striking in our parent-child pairs, since parents were actively trying to engage their child. As a fairly instructive example, in G7, a child sat to draw a picture, while his father did the same. This continued for five minutes, and while the father initiated some small talk, neither looking up during the activity. Parten's phrasing of Associative Play really resonates with our observations here—in our child-parent participant groups, when the child "disengaged" in joint activity, the parent would try to strike up conversation with the child.

Play Type. The type of play also varied substantially between the two conditions, with more active play in OneSpace compared to make-believe being the main driver in the conventional configuration (Table 3). The nature and affordances of these video configurations have a strong impact on the style of play that participants choose. Figure 3 provides an illustrative account of an active and learning episode from the OneSpace condition. Here, the child is playing with a hulahoop. Her mother tries to join in, but has lost her hulahooping skills. The child then begins instructing her mother on how to properly hulahoop by standing next to and in front of her mother, demonstrating the correct motion for her mother to mimic.

Active (and physical) play was characteristic of play that we observed in OneSpace—participants (particularly the children) generally moved about the scene far more. For instance, groups played "tag" in OneSpace, or would run from one side of the shared scene to the other, and so forth. Figure 1 illustrates a pair playing "paddycake," while others would also play fight, miming punches and kicks, and feigning injury. This active play went beyond merely moving one's limbs around—indeed, our coding revealed that participants physically moved around far more (i.e. took steps) in the OneSpace condition. What is striking about these physical activities is that it seemed representative of the same physical activity one would expect of children in collocated scenarios. Of course, how they would "touch" one another in OneSpace and watch one another was different, but the character of the movement was natural.

Figure 3. The daughter teaches/demonstrates the hulahoop motion for her mother to mimic.

Perhaps because the conventional configuration enforces the separation of space, participants playing across the sites *needed* to engage in make believe to feel closer. As an illustrative example, G4 decided to play through a fashion show together, dressing up dolls in various costumes and creating a narrative between the characters. The physical separation between the spaces is extremely evident in these instances—the only real bridge between these spaces is the conversation to create a shared imaginary domain, and a video link that shows the facial reactions of one's partner.

Affordances of the Configurations. The OneSpace configuration seems to afford physical activity by virtue of the shared visual space that encompasses the entirety of both participants' bodies. The fact that both shared the same visual space means they are, in a sense, competing for the visual stage. Because the system handles this by respecting their "depth relationship" (i.e. whoever is closer to the camera is displayed), one could easily "upstage" or "take over" the scene by simply moving closer over to the other person, and closer to the camera. Recognizing this seemed to spark instances of playfulness, where participants would take turns trying to upstage one another. One common behavior was that participants would sometimes hide from view *in the scene*—this meant hiding from the camera (i.e. going out of view), but also sometimes hiding behind one another, and integrating this into their play. G5 used this technique to make a four-armed mosquito, where the child pretended to be the mother's arms by standing "in front of" her mother, but flapping her arms around. In this way, their bodies became objects of play in this shared space. The shared representation also allowed participants to use the space to gesture and point at body parts and objects in what would otherwise be only the remote space (e.g. Figure 3).

	Trans.	Solitary	Onlooker	Parallel	Associative	Cooperative
OneSpace	146	1	52	66	145	411
Conventional	91	4	55	132	361	194

Table 2. Engagement levels on 5s intervals across groups. Scheme adapted from [10].

	Trans.	Active	Make-believe	Manipulative	Creative	Learning	Object
OneSpace	150	477	110	16	7	34	27
Conventional	91	222	228	83	97	16	41

Table 3. Play types on 5s intervals across groups. Scheme adapted from [1].

In contrast, the conventional configuration seems to afford more solitude among participants. Even though the captured video space (i.e. the entire body) is the same, we saw substantially less "active" play. Instead, a lot of play centered on activities that each could engage in by themselves, and instances of "show and tell" where one would ask the other to look at something on their side.

Attention & Visibility. Children did seem to understand in the OneSpace condition what could be seen—they would regularly play with their partners in the video scene. Yet, we still observed instances in the OneSpace condition where children would ask, "Can you see me?" when they were in/out of the frame, though this seemed to occur markedly less than in the conventional configuration. There, we observed far more instances of the parents (for instance) coaching the child, "Why don't you back up so I can see you all the way? [G3]" Of interest is that for children, it was clear that the details presented by the video scene really mattered. In one instance, a child asked, "What are you looking at […] it doesn't look like you are [looking at me] [G4]," even though the parent was looking at the screen. Here, the child had noticed that the eye gaze of the parent was incorrect (due to camera positioning). In this case, the parent explained to the child how cameras worked, etc. Similarly, one pair discounted OneSpace entirely, citing its poor, blocky resolution and low framerate as reasons why they would simply not use it.

DISCUSSION & CONCLUSION

OneSpace's design builds on considerable prior work, not only in the research space [2,4,8,12], but also in the artistic space (e.g. [5]), where artists have long recognized the role of bodies—not only as subjects, but also as objects. Our application of these ideas to the domain of child freeplay is a natural extension of this, recognizing that children also consider their bodies and environments as objects of play.

Although the size of our subject pool is small, we believe the results to be generally representative. Like Yarosh et al. [12], we found that in conventional setups, children engaged in make-believe/pretend play; in our OneSpace configuration, which was related to the rug condition in [12], there was a lot more active/movement play. The differences in the quality of the play and the type of engagement between OneSpace and conventional video conferencing was immediately striking—not only to us, but to at least some of participants, one who noted: "Before [in OneSpace] we felt really close; now using Skype, it feels like we're really really (*sic*) far apart. [G4]"

It seems clear that the problem of managing attention and understanding visibility remains in OneSpace. While it does address attention across two disjoint displays, the subtleties and details of interaction (particularly eye gaze) are meaningful and powerful cues to children. This is perhaps unsurprising, since this age group is learning to read and respond appropriately to non-verbal social cues.

One concern with generalizability of these results is the novelty effect. Our lab-based study design did not allow us to address this effect; however, none of our children had ever had experience with video conferencing (i.e. not even a conventional setup). In future work, we will conduct long-term deployments of OneSpace into participant homes. We will also explore digital entities such as costumes or objects (to augment physical ones) in the shared play space.

This work contributes to the ongoing discussion in the research community about how to provide playful environments to connect children at a distance. Rather than focusing on a shared activity [3,11] or play surface [12,13], we focused here on giving children a visual representation that shows themselves with a playmate. Based on our study, we suggest two major lessons for future work in this space: first, for children, the body is actually a "play area" and that this should be an area of exploration, and second, creating a shared space environment for remote play can lead to a more active and physical play—thus forcing us to ask questions about what closeness and separation actually mean to people that inhabit these spaces.

REFERENCES

1. Bronson, M. B. (1995) The Right Stuff for Children Birth to 8: Selecting Play Materials to Support Development. National Association for the Education.
2. Cullinan, C. & Agamanolis, S. (2002). Reflexion: a responsive virtual mirror. In Ext Abs of UIST 2002.
3. Follmer, S., et al. (2010). Video play: playful interactions in video conferencing for long-distance families with young children. In Proc IDC 2010, 49-58.
4. Hunter, S., et al. (2014). WaaZam! Supporting creative play at a distance in customized video environments. In Proc CHI 2014.
5. Krueger, M. (1983). Artificial Reality. Addison-Wesley.
6. Ledo, D. et al. (2013). OneSpace: shared depth-corrected video interaction. In CHI EA 2013, 997-1002.
7. Montessori, M. (1995). The Absorbent Mind. Henry Holt and Company: New York.
8. Morikawa, O., and Maesako, T. (1998). HyperMirror: toward pleasant-to-use video mediated communication system. In Proc CSCW 1998, 149-158.
9. Parten, M. (1932). Social participation among preschool children. J. of Abnormal and Soc. Psyc. 28 (3): 136–147.
10. Piaget, J. (1926). The Language and Thought of the Child. Harcourt, Brace & Company, Inc., NY, USA.
11. Raffle, H., Revelle, G., et al. (2011). Hello, is grandma there? let's read! StoryVisit: family video chat and connected e-books. In Proc CHI 2011, 1195-1204.
12. Yarosh, S., Inkpen, K. M., and Brush, A. J. (2010). Video playdate: toward free play across distance. In Proc CHI 2010, 1251-1260.
13. Yarosh, S., Tang, A., et al. (2013). "Almost touching": parent-child remote communication using the sharetable system. In Proc CSCW 2013, 181-192.