

Understanding Video based Parent Training Intervention for Children with Autism

Deepti Aggarwal¹, Robyn Garnett², Bernd Ploderer¹, Frank Vetere¹, Patricia Eadie², Bronwyn Joy Davidson²

¹Dept. of Computing and Information Systems
The University of Melbourne, Australia
daggarwal@student.unimelb.edu.au
{ploderer, f.vetere}@unimelb.edu.au

²Dept. of Audiology and Speech Pathology
The University of Melbourne, Australia
garnettr@student.unimelb.edu.au
{peadie, bronwynd}@unimelb.edu.au

ABSTRACT

This paper investigates the challenges of delivering parent training intervention for autism over video. We conducted a qualitative field study of an intervention, which is based on a well-established training program for parents of children with autism, called Hanen More Than Words. The study was conducted with a Hanen Certified speech pathologist who delivered video based training to two mothers, each with a son having autism. We conducted observations of 14 sessions of the intervention spanning 3 months along with 3 semi-structured interviews with each participant. We identified different activities that participants performed across different sessions and analysed them based upon their implications on technology. We found that all the participants welcomed video based training but they also faced several difficulties, particularly in establishing rapport with other participants, inviting equal participation, and in observing and providing feedback on parent-child interactions. Finally, we reflect on our findings and motivate further investigations by defining three design sensitivities of Adaptation, Group Participation, and Physical Setup.

Author Keywords

Remote parent training; Autism Spectrum Disorder; Video communication;

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

Autism is a complex neurodevelopmental disorder that affects an individual's ability to communicate with the world around him or her (American Psychiatric Association, 2013). A person diagnosed with autism faces difficulty in social interaction and communication and has restricted and repetitive behaviours. Early signs of autism typically appear during a child's first 3 years of life. As a

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

OzCHI '15, December 07 - 10 2015, Melbourne, VIC, Australia
Copyright is held by the owner/author(s). Publication rights licensed to ACM. ACM 978-1-4503-3673-4/15/12...\$15.00
DOI: <http://dx.doi.org/10.1145/2838739.2838770>

result, early intervention through parent training has been recognized as a well-established approach to support child development (Ingersoll and Gergans, 2007; Oono et al., 2013). One aim of such training intervention is to educate and coach parents in strategies to address their child's social communication skills such as joint attention, taking turns, and making eye contact. Such parent training is often provided by speech pathologists to parents in a variety of ways e.g., in a structured multi-session program in face-to-face settings. Recent studies have shown that the number of children diagnosed with autism is growing rapidly, with prevalence rates reaching 1 in 119 in Australia (Barbaro and Dissanayake, 2010) and 1 in 68 in US (Baio, 2014). Consequently, the need for such parent training interventions has increased, making it difficult to provide face-to-face training particularly, to parents from remote geographical locations (Mackintosh et al., 2012). To fulfil the increasing demand, professionals are now looking at alternate ways of delivering these training interventions.

Rapid growth of the web-based services in households and easy access to high-resolution web-cameras has enabled video based parent training to become a viable alternative to traditional face-to-face delivery (Carter et al., 2011; Meaden & Daczewitz, 2015). During a video based parent training, parents receive the required coaching on improving their child's communication skills from speech pathologists over video. As video based training takes place in the home environment where a child with autism spends most of their time, it provides immediate advantages in terms of comfort and effectiveness to both the parents and children (Baharav & Reiser, 2010). However, since the professional is not present with the child and family during the video training, it remains unclear whether the video mediated interactions could provide an effective learning environment; which we explore in this paper.

This paper seeks to identify the benefits as well as challenges in delivering parent training over video. In this regard, we investigated a video based parent training intervention over a period of 3 months. The intervention is based on the *Hanen More Than Words Program*, which is designed to enhance communication skills in children with autism by providing education, training and social support to their parents. We observed 14 sessions of the intervention delivered by a speech pathologist to two mothers each having a son with autism. We investigated how the participants used the underlying technology

during the course of intervention. Various issues emerged as the interactions unfolded among the participants both on screen and away from the screen. In particular, we identified four key barriers to delivering video based parent training: 1) adapting software to suit the individual's needs and responsibilities; 2) encouraging equal participation from each parent; 3) setting up the technology without causing distraction to the child; and 4) observing and providing feedback on parent-child interactions.

This paper makes several contributions to both HCI and health communities interested in the intersection of video mediated communication and parent training. 1) We offer a first conceptual understanding of the practices of video based parent training for autism in action. 2) Our work illustrates the relationship between the spatial organization of participants and their interaction during the course of training (both on-screen and off-screen interactions). 3) By discussing the activities that participants performed during the intervention, we highlight the unique aspects that the context of parent training for autism brings to the video communication. In particular, our study has shown that video-based parent training is a complex scenario where the interactions among participants mainly happen away from the screen. 4) Finally, to address the raised issues and to guide future investigations, we extend our understanding and present three design sensitivities: Adaptation, Group Participation, and Physical Setup.

RELATED WORK

Video mediated communication has a long history in HCI. Several works have investigated video mediated communication for institutional and domestic settings (Isaacs & Tang, 1994; Veinott et al. 1999; Kaye et al. 2010; Olson & Olson, 2014). For instance, some early works from Heath & Luff (1991, 1992) provided foundational knowledge on video based communication to support collaborative work in office environments. Their work highlighted asymmetries in communication where the communicating parties have access to different sets of resources or they utilize the available resources differently- such asymmetry then influences the overall course of events. More recently, Ames et al. (2010) explored the use of video-mediated communication in people's homes to explore how family members coordinate communication among each other, while other works endeavored to explore how individuals mediate closeness and intimacy over-a-distance (Kirk et al., 2010; Neustaedter & Greenberg, 2012).

Video communication is also used to support distance training and classroom education between teacher and student where face-to-face interactions are not possible due to lack of resources and geographical distance (Dillenbourg & Baker, 1996; White et al. 2000; Wang & Hartley, 2003). For example, web platforms like Coursera (<https://www.coursera.org/>) and NPTEL (<http://nptel.ac.in>) provide access to pre-recorded video lectures by elite faculties on different courses. Similarly, there are online platforms that enable discussions among students from different locations on pre-recorded video

lectures (Cadiz et al., 2000). Success of these systems has prompted interest in investigating the use of video mediated training in other domains such as official training and webinars (Jancke et al., 2000). Taking inspiration from these works, this paper studies the viability of video mediated training for parents whose children are diagnosed with autism. Before discussing the study, let us first understand parent training for autism.

Parent Training for Autism

Autism manifests differently across people and may even be different in the same person at different times (American Psychiatric Association, 2013). As a result, every child diagnosed with autism behaves, reacts, and interacts with things differently. To support the development of children with autism in their early years, educating parents about autism and training them to understand their child's behaviour has been recognized as an effective approach (Ingersoll and Gergans, 2007; Oono et al., 2013). As a result, several parent training interventions have been developed *e.g.*, Hanen More than Words Program (Sussman et al., 2013) and Relational-Focused Intervention (Mahoney & Perales, 2003).

During a training program, parents learn about different strategies from a trained professional that could be useful for their child's development. Since each child with autism shows distinct behaviours, these strategies need to be customized to make them suitable for the child (Roberts & Kaiser, 2011). Parents therefore, engage and seek advice from the experienced professionals as well as from other parents facing similar problems on how to best customize the strategies for their child. Such parent trainings are often given in a face-to-face setting where multiple parents and professional meet in the same physical space and learn together. However, several factors such as geographical boundaries, service availability and cost often limit parents' participation in face-to-face trainings (Mackintosh et al., 2012).

To fulfil the growing needs for parent training across remote geographical locations, video communication has recently emerged as an alternate viable approach to deliver the training interventions. In this regard, some works have investigated the clinical outcomes of video based training for specific programs and mentioned the use of certain software to support their training (Boisvert & Hall, 2014; Carter et al., 2011; Meaden & Daczewitz, 2015). For instance, Baharav & Reiser (2010) reported the use of Skype and wireless Bluetooth camera to support supervision and coaching, while Suess et al. (2014) utilized Skype and Debut software for the same. However, what remains missing is a conceptual understanding of how professionals and parents use, adapt and adopt the underlying technology during the course of training; and it remains unclear how using different software and hardware facilitated the interaction and communication among the participants.

Video based parent training facilitates child development in an environment where the child spends the majority of their time and where they are likely to feel most comfortable. It is not surprising therefore, that such training often results in high levels of parent satisfaction

(Baharav & Reiser, 2010). Meaden & Daczewitz (2015) argued that it may become challenging to provide a similarly effective learning environment for the parents over video, as would be possible during face-to-face training where the professional is present to monitor and guide the parents. For instance, in a face-to-face training, parents can learn from the professional and from each other by exchanging their stories and performing group activities. The interactions in face-to-face settings are rich in nature, as the parents and professional can utilize gestures, physical interactions, non-verbal cues in addition to the verbal conversation. Hence, using video technology may not fully support the dynamic interactions that unfold among the participants during the course of training. It thus, requires careful analysis of the user needs to design appropriate technology for delivering parent training over video. Understanding how technology is enveloped in the context of video based parent training for autism is the core focus of this paper.

OUR STUDY

The aim of this research is to investigate the challenges of delivering parent training intervention for autism over video. To this end, we conducted a qualitative field study of a parent training intervention for autism. The intervention was delivered by the second author (who is a Hanen certified speech pathologist) to two mothers who each have a son with autism. We conducted observations of the sessions and semi-structured interviews with each participant. Ethics to conduct this study was approved by the HREC committee of The University of Melbourne, Australia.

The Intervention

We investigated a video based parent training intervention that was based on a well-established program called *Hanen More Than Words (HMTW)* (Sussman et al., 2013). The intervention aimed to support communication development in young children with autism by teaching and enhancing parental skills through intensive training. The intervention had two modules delivered across 3 months that are described below.

- 1) *Training Module*: The training module involves 8 sessions offered to a group of parents, with each session lasting for around 2.5 hours. These sessions are delivered on weekly basis to the parents and do not involve the child. In every session, parents learn specific strategies regarding the development of their child's communication skills. The training content is offered in the form of a PowerPoint slideshow and short video clips.
- 2) *Feedback Module*: The feedback module has 3 sessions per parent. The aim of this module is to understand how the parents are practising the strategies with their child. In this regard, each parent performs 1-3 activities (e.g., washing hands and reading books) with their child in front of the speech pathologist. These activities are planned by the parents around the strategies learned during the training sessions in accordance with their child's needs. The speech pathologist observes the parent-child interaction as it occurs in real time and provides

immediate feedback to enhance parent's learning about the use of strategies. Additionally, the speech pathologist videotapes the parent-child interaction for later discussion. Once the parent-child activities are finished, the parent and speech pathologist then watch the recorded video to further enhance the use of strategies. These individual video-feedback sessions are provided to each parent after the second, fourth and seventh group training sessions and typically last for around 1.5 hours.

Study Setup and Participants

Our study participants included a female speech pathologist, Clairra (second author), and two mothers Alice and Rachel. (We have used pseudonyms to represent our participants.) Figure 1 shows the demographic details of the participants and the study setup. The speech pathologist Clairra (47) has over 17 years of experience in offering face-to-face parent training programs for autism. This study was her first experience of video based training. Alice was 33 years old and her son Tom was 3 years and 9 months. Rachel was 38 years old and her son Sam was 4 years and 3 months old at the time of the study. Alice and Rachel were working part time and had experience in using video conferencing tools (Skype). However, this was the first time they took part in a training intervention for autism.

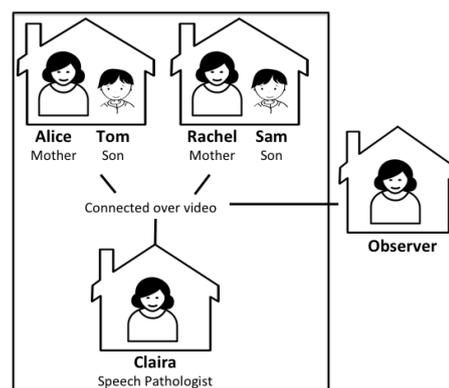


Figure 1: The intervention was offered by a speech pathologist to two mothers via VSee. The first author passively observed the sessions as another user on VSee.

The program sessions were organized through a video conferencing tool, VSee (<https://vsee.com/>). VSee connects remote parties and allows sharing of training material with real time control e.g., writing on slides, and stopping and replaying the video clips. Additionally, VSee provides separate windows for the video-stream of each participant and training material. Each window can be moved and re-sized locally by the participants to personalize the viewing experience. Clairra carefully chose VSee to deliver this intervention after conducting a separate comparative usability study of different video conferencing tools. (Details of the comparative study is beyond the scope of this paper.)

Data Collection and Analysis

Data was collected throughout the intervention period, i.e., over 3 months. The first author observed all 14 sessions of the intervention. She gathered data related to the physical environment, system usability, interaction

between users and technology, and the physical arrangement of users around technology. Field notes were taken during the sessions and extended after discussions with the speech pathologist. Semi-structured interviews were conducted before, during and after the intervention with each participant to reflect upon their experience with technology. All the interviews were audio recorded and transcribed later for the analysis.

Given the absence of prior work on video based parent training for autism in HCI, we conducted a qualitative analysis based on Grounded Theory (Strauss & Corbin, 1998). Data analysis was based on several rounds of coding of our field notes and interview transcripts. This was done iteratively, beginning from the start of data collection, in order to refine our interview questions and our focus for later observations. The first author coded the data on paper, and created memos to capture ideas and trends emerging from the data about challenges in the delivery and participation in video based training. While the first author performed coding, the emerging findings were discussed regularly with the second author to provide feedback on the intervention, and with other authors to reflect upon the observations and analysis. Using affinity diagrams, all authors worked together to structure key ideas into two themes that grouped around *Interactions during the Training Sessions* and *Interactions during the Feedback Sessions*.

FINDINGS

The parents found the video based training sessions comfortable and comparable to having a one-to-one training with the speech pathologist. Rachel mentioned, *“Usually I do not share my opinion when I am in group. I get anxious when I speak up in groups. But it’s easier to speak when you are at home just on your computer. It is not like that when 20 people are staring at you when you speak.”* Alice, on the other hand, appreciated the vis-à-vis arrangement of video conversation, which enabled her to have better discussions with Clair: *“Here [on video], you are always in front of each other. Clair can hear you all the time whereas in a room [face-to-face], sometimes it’s a bit hard to be heard.”*

Similarly, feedback sessions over video had several benefits. Firstly, the remote setting made the child more comfortable in performing activities during the feedback sessions. Parents did not mention any change in their child’s behavior when Clair was watching over video. The physical presence of the speech pathologist could have been intimidating for the children if they were co-located as mentioned by Baharav & Reiser (2010). Secondly, scheduling a session at a time convenient for everyone was easier than arranging a group meeting at a physical location. With video-based training, the parents could schedule or reschedule the session according to their child’s daily routine and mood. However, participants also encountered several issues during the session, which we discuss below.

Interactions during the Training Sessions

We first illustrate the training session findings, which involved the two mothers and the speech pathologist. We start by describing the use of VSee by each of the study

participants. The speech pathologist, Clair, facilitated the training session using VSee on a desktop computer with two screens. The first screen had an integrated web camera and the screen was used to make a VSee call to share PowerPoint slides and video clips. The second screen was used to collect feedback and note down discussion points. Alice and Rachel, on the other hand, used their laptops to connect to VSee.

Each participant had 6 windows via VSee on their Desktop: four corresponding to the four users and the other two for the training content (slides and video clips), as shown in Figure 2. Interestingly, all participants placed the windows of users on the upper portion of the screen, whereas the lower half of the screen was bigger and was dedicated to viewing shared training content. Every participant, however, used different sized windows at different times, according to their needs and comfort. Parents did not report any issues with having the observer in the sessions but they always made the observer’s window very small and placed it in a corner of the screen.



Figure 2: Each training session via VSee had 6 windows: 4 for the users and 2 for the training content.

Strategies of Arranging VSee Windows Differently

During the interviews, parents reported interesting patterns and behaviors in terms of arranging the windows on screen. Parents often adjusted the number and size of the windows in order to feel comfortable with themselves, the other parent and speech pathologist.

At first, seeing oneself on the screen during the training was awkward for the parents, but they gradually overcame it. For instance, Rachel preferred to minimize her own window completely so that she could not see herself. She said, *“It took a little bit to get used to seeing myself on the screen [laughed]. I just minimized my own window so as not to see my own face. Otherwise, I’d keep looking at myself, and then I’d feel like, ‘Oh my god! I look so terrible.’”* However, after a couple of sessions, she was fine with seeing herself on screen and therefore, she kept her window open along with the windows of other participants but with a reduced size. Self-consciousness was not an issue for Alice, however, she preferred to have her own window smaller than other windows in order to make more room for expanding other windows.

Parents also found sharing their personal experiences and expressing doubts in the presence of the other parent a

little awkward. As a consequence, parents devised their own strategies to feel more relaxed in participation during the sessions. For instance, Alice utilized the flexibility of VSee to make the setting personalized and to feel more comfortable in speaking, “*You can minimize or shrink your window or of others to feel that it’s a one-to-one session with Clairra*”.

Rachel also changed the window sizes and arranged the screen differently based on who was speaking at what moment. For example, she tried to have Clairra’s window wide for most of the time, but if Alice was talking, she increased her window size. If it was a group discussion, then she placed windows of Clairra and Alice together occupying most part of the screen. More interestingly, Rachel also arranged the window of video clips and made them fit to full screen even though the video clips were of low resolution and did not fit to full screen. However, she mentioned that maximizing video clips helped her to pay closer attention to them and to get their message clearly.

Challenges in Managing Activities Beyond Talking

To enhance parents’ understanding of the strategies being covered in the training sessions, Clairra made the sessions lively and interactive. Clairra asked the parents to perform activities related to strategies outlined in the training session. The aim was to make the strategies more meaningful and memorable for the parents by creating experiences that were similar to those experienced by the children. These activities were performed individually by the parents in front of the camera so that Clairra could easily see them and provide feedback, if necessary.

One such activity was to educate parents about the children’s sensory abilities. The activity involved the parents balancing a tissue on their head, applying slippery creams on hands and then taking a dictation from the speech pathologist in presence of loud music in the background. Once the dictation was completed, Clairra verified the written words of both the parents. Having all the senses activated concurrently made it difficult for the parents to concentrate on what the speech pathologist was saying, hence, their dictated words were full of errors. However, these errors were not caused solely by the augmented senses; the technology also contributed to the errors. For example, it was difficult for parents to understand how much cream they should apply, and how loud they should keep their system’s volume for the music played by Clairra. For instance, if the sound was

too loud for a parent, they did not know whether they could reduce it or if that was the required sound level for the activity. Being at a remote location, Clairra was not able to control these parameters, which she could have easily regulated in a face-to-face environment.

Difficulties in Encouraging Equal Participation

Clairra was also responsible for encouraging participation from parents so that they could learn from each other by sharing their personal experiences. Clairra asked questions to initiate discussions about the strategies. However, she found it difficult to encourage equal participation from both the parents. For instance, Rachel was highly motivated to discuss her personal experience with Sam and hence she was prompt in replying or raising different points. However, Alice was more introverted and did not initiate the conversation by herself. The different attitudes of the parents created some confusion in defining turns so that both parents were given equal opportunities to discuss their issues and receive feedback from Clairra and from each other. In this regard, Clairra found managing discussions over video bit difficult to organize: “*In face-to-face, the discussions are more natural where questions from one parent may invite some other parent to join in. However, here in the remote environment, as the group is small, parents might feel hesitant to initiate the discussion. I need to ask each parent explicitly about any questions. I will also need to check that everyone is getting a chance.*”

During the interview, Rachel mentioned her concern about being more conversational than Alice: “*It is difficult to understand when you should talk. I want to give Alice a chance to speak, and also want to listen to her ideas rather than me talking about me all the time.*” She further added that taking turns was bit unclear and confusing for her being in a remote environment: “*In a face-to-face setting, there are other mannerisms like raising hands to start talking, which is missing in remote sessions.*”

Interactions during the Feedback Sessions

Feedback sessions were challenging to manage because they involved interaction between the parent and child who were co-located remotely, and the speech pathologist who was present via video. We found asymmetries in the communication where parents were trying to communicate simultaneously with both child and speech pathologist, and the speech pathologist was only

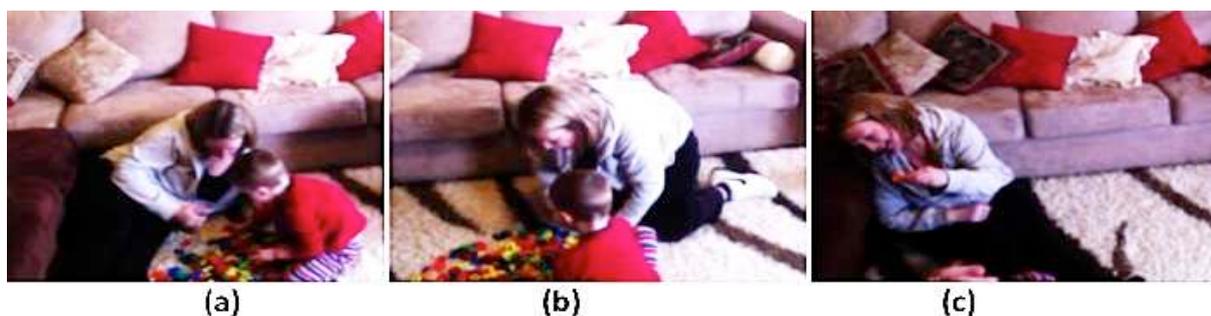


Figure 3: During a feedback session, Rachel and Sam played a word-making game. They started their interaction by sitting in front of the camera (figure a). However, their orientation with respect to camera changed as they got engaged in the activity (figure b). And finally, Sam was completely out of the camera focus (figure c).

communicating with the parent.

Difficulties in Observing Parent-Child interactions

The feedback activities were dynamic in terms of body postures, orientation and location in the home, whereas the camera remained fixed in one location. The static positioning of the camera made it difficult for Clairra to clearly observe and provide feedback on the activities. Figure 3 illustrates an example of Rachel and Sam performing an activity together. During one feedback session, Rachel and Sam performed an activity of making words with plastic letter blocks. At the beginning, both Rachel and Sam were facing the camera (Figure 3a) and Clairra could clearly see the interactions. However, as the activity unfolded with time, Rachel and Sam no longer faced the camera. Instead Sam sat with his back to camera (Figure 3b). Additionally, later, when Sam was very much engaged in the activity, he was completely away from the camera's field of view (Figure 3c). In another session, Sam even ran to another room to extend the play from one room to another. Similar difficulties were observed with Alice and Tom, when the parent or the child blocked the camera view unknowingly.

As a result, Clairra had to request that the parents adjust the computer in order to set a camera angle that allowed a clear view of the interaction. This, however, interrupted the interaction flow between the child and parent. On occasions, Clairra did not want to interrupt the child's play and she did not say anything. She waited for the correct moment (e.g., when the parent looked towards her or when there was a momentary pause in the activity) before requesting a repositioning of the camera. When Clairra could not find the correct moment, she explicitly asked the parents to come into the camera focus.

The camera angle also made it difficult for the speech pathologist to verify whether the parent was following the activity correctly and whether the child was making eye contact with the parent. (Developing the child's skills to maintain eye contacts is one of the aims of these sessions.) For instance, when Sam was not directly facing the camera (Figure 3b), it was unclear for Clairra to determine if he was making eye contact with Rachel or not. And the situation became worse when Sam was completely out of the camera's field of view (Figure 3c). In this regard, parents missed crucial feedback from Clairra that could have enhanced their learning.

Limitations in Audibility for Parents and Pathologist

Good quality audio was also an issue for both the parent and pathologist. During the sessions, Clairra provided immediate feedback to the parents, while observing the child-parent activities over VSee. Clairra gave short and direct vocal instructions to the parents. For instance, she said, "Wait more", "Come in the front", which guided parents on how to progress with a certain activity. Sometimes, the parents sought clarifications regarding the instructions, however these were not always clearly heard. The reason for the poor audio was in part due the change in the relative position of the parent and child with respect to the computer (microphone), while performing the activity. Often parents requested that Clairra repeat the instructions by saying, "Sorry", and "Can you please repeat?" To overcome the sound issue, the parent and pathologist often changed the volume level during the course of the session.

There were also moments when Clairra also could not listen to what the parent or child said to each other. In such scenarios, increasing the volume did not help Clairra because the loudness of the parent's voice was not the same as that of the child's voice. For instance, the children did not speak clearly or spoke with a very low voice, whereas the parent was clear and sufficiently loud. Increasing the system volume further increased the volume of the parent, but not that of the child.

Technological Distractions for Children

The computer sometimes became a distraction for the child. As the speech pathologist and parent were connected via VSee, parents placed the laptop nearby so that Clairra could observe the interaction clearly. Such placement of the computer produced a distraction for the children, and thus affected their concentration in the ongoing interaction with their parent. Figure 4 illustrates one such case. Tom was very fond of interactive technology (e.g., laptops and mobile phones) and wanted to play with them whenever possible. During one feedback session, Alice planned to engage Tom in play by singing his favorite song. Therefore, she sat on the dining table with Tom and placed her laptop in their vicinity so that Clairra could see their interaction (Figure 4). Alice started singing songs to Tom and attempted to gain Tom's attention by using several hand actions. However, Tom kept looking at the laptop and did not pay any attention to his mother. For later feedback sessions, Clairra supplied Alice with an external web-camera

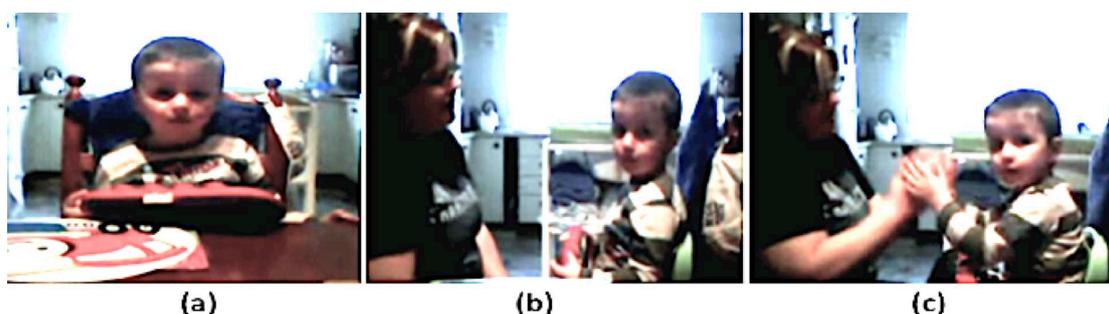


Figure 4: Tom getting distracted from the laptop placed on the dining table: While Alice was setting up for the feedback session; Tom kept looking towards the laptop (Figure a). Alice started singing a song and made hand gestures to engage Tom but he did not show any interest in the activity (Figure b & c).

(embedded with microphone) through which she observed interactions between Alice and Tom. Having an external webcam solved this issue of distraction and encouraged natural interactions during the session.

In the case of Rachel and Sam, Sam was also frequently distracted by his reflection on the laptop screen whenever the laptop was close by. To avoid such distractions, Rachel kept the laptop away from him, with its screen down.

DISCUSSION

This study highlighted two types of issues in supporting interactions over video. Firstly, on-screen issues are raised when participants converse with each other using a video conferencing tool, such as arranging VSee windows and negotiating turns. Secondly, off-screen issues happen away from the computer screen when parents performed activities by themselves or with their children. We now discuss these issues as different *design sensitivities*. Design sensitivities highlight the relevant topics that designers should consider while designing for the particular research context (Ciolfi, 2004; Dalsgaard, 2008). Below we articulate our understanding from the study as three design sensitivities: Adaptation, Group Participation and Physical Setup.

Adaptation

The first design sensitivity concerns the desire to adapt the technology to support different user needs during the course of sessions. In the study, we found that the speech pathologist and parents often adjusted the number, size and position of the windows to suit the given activity. For example, parents made the window containing the video stream of the speech pathologist bigger in order to simulate a face-to-face training session, whereas they kept the window containing their own stream smaller in order to overcome self-consciousness. The speech pathologist, on the other hand, had to manage the delivery of the training content in a smooth fashion. She had to keep a track of the content that was already shared and that needed to be shared with the parents. The VSee software used for offering the intervention did not cater to these predefined roles and needs of the participants. For instance, parents had to manually arrange the windows and constantly change them during the course of session. This sometimes led to distraction and confusion. Although manual arrangement provided the participants control over the software, the need for repetitively arranging the windows can be avoided by better design.

This could be addressed by providing multiple tailor-made templates. These templates would allow participants to have different arrangements of on-screen windows depending upon their needs. Participants should also be able to customize these templates over a period of time. For example, Figure 5 illustrates two templates, one each for the parents and speech pathologist. For parents, the important windows included the training content and the speech pathologist's video stream. These windows are larger than the others (Figure 5a). Similarly, more support can be provided to the speech pathologist by introducing two new windows, namely, 'links' and 'timeline' (Figure 5b). The 'links' window provides drag and drop access to

the training contents, while the 'timeline' window provides a temporal snapshot of the content previously shared and waiting to be shared next.

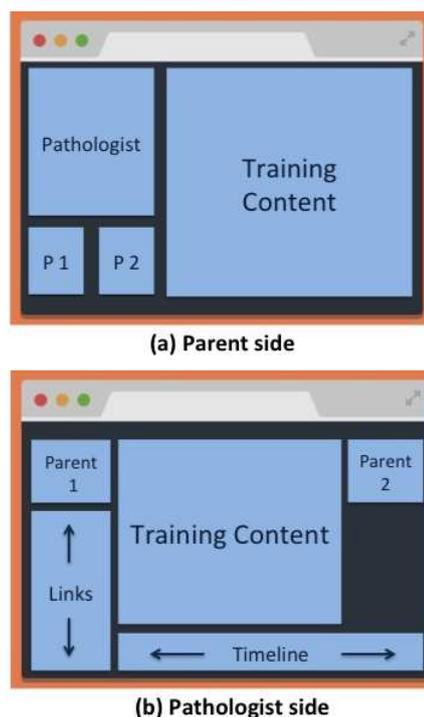


Figure 5: Mock-up templates demonstrating the possible arrangement of different windows opened during the training sessions at the parents' and speech pathologist's side.

Tangible user interfaces (Ishii and Ullmer, 1997) could also be explored to provide the participants with better control over multiple training resources. For instance, Brave et al. (1998) used physical blocks to support remote collaboration in workspace. A similar approach could also be utilized in this context where the speech pathologist and parents can arrange different resources physically.

Group Participation

The second design sensitivity concerns group participation. This relates to the inequality of participation by parents due to issues such as self-consciousness, and social awkwardness in public speaking. Since every case of autism is different, group interactions are important to generate shared knowledge and provide a better learning environment for every parent. In this regard, the training sessions are woven around the rich discussions among the parents related to the training content and their personal experience with their child. Such a discursive environment helps the parents become confident about sharing their stories, and often realizing that their struggle is not dissimilar to that experienced by others. Therefore, it is important that every parent feels sufficiently comfortable to participate in the discussion, and that they do not feel pressured or embarrassed.

The current arrangement of all participants facing each other (*vis-à-vis*) on the screen may make it difficult to generate a supportive environment. For example, on one hand, parents applauded the ability to see the speech

pathologist vis-à-vis as such an arrangement made them comfortable in initiating a conversation at any time. However, vis-à-vis arrangement also exposed them to other parents, which made them conscious about sharing their personal experiences. It may be possible that parents would prefer the vis-à-vis arrangement only with the speech pathologist, while they would rather choose a side-by-side with other parents. For instance, Seligmann et al. (2004) highlighted that providing a side-by-side view of the remote participants, helped in reducing the psychological distance between the participants. Therefore, future video conferencing tools for parent training should consider the spatial arrangement of participants on the screen.

Additionally, there are also concerns about how much parents should talk such that each parent gets a fair opportunity to share their personal experiences. In this regard, devices such as a physical table created by Rogers et al. (2009) can be used to make an individual aware of their participation in the form of bubbles. Furthermore, on-screen indicators and markers can also be used to display and define turns. Designers, therefore, should carefully design systems that provide self-reflection and awareness of everyone's participation in the discussion.

Physical Setup

The final design sensitivity describes considerations for the physical arrangement of the technology. This includes the setup of camera and microphone to obtain high quality video and audio, particularly, during the feedback sessions. During the feedback sessions, it was important that the speech pathologist could clearly see and hear the parent-child interactions. However, the static arrangement of the camera made it challenging for the speech pathologist to clearly follow the dynamic interactions between the child and parent. For example, the static camera could only capture those interactions when both the parent and child faced the camera directly. This is shown schematically in Figure 6a. However, if they oriented their body away from the camera, the speech pathologist would not be able to follow their interaction, as illustrated in Figure 6b and 6c. Additionally, as the parent and child kept moving away from the camera, the clarity of sound would also decrease. As a result, neither the parent nor the speech pathologist would be able to hear each other clearly.

A solution to maintain good sound and visual quality could be to ask the parents to face the camera all the time and to stay very close to it. However, this solution is unlikely to be successful because all activities cannot be performed in front of the camera. An alternative solution could be to equip the parent or child (or both) with wearable cameras like GoPro (Pirsiavash & Ramanan, 2012) and a headphone. However, doing so might make the interactions unnatural. More importantly, this setup may also distract the child's attention; just as we found that close proximity of technology distracted the children.

To this end, there remains a tension between achieving the desired quality of video and maintaining comfort of the participants. Designers, therefore, should think of ways to disguise the potential technology in the

surroundings in such a way that it is not a distraction but rather captures the interactions. For instance, having a camera sewed in the parent's clothing such that it identifies eye contacts between the child and parent, and in turn provides feedback to the speech pathologist, would enable more accurate parent coaching. Furthermore, Microsoft Kinect (Zhang, 2012) could be another potential resource to cover the interactions during the feedback sessions. Its depth camera could cover the dynamic interactions between parent and child along with a facility to receive volume based upon the interpersonal distance between the participants and camera.

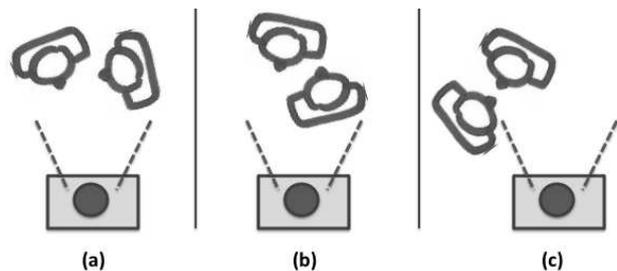


Figure 6: A static camera can only capture interactions where the user is facing the camera (figure a). However, it fails to capture other interactions (figure b & c).

CONCLUSION

In this paper, we presented a detailed qualitative study that outlines benefits as well as barriers in delivering a video based parent training intervention for autism. Participants appreciated the comfort and convenience of video as a communication medium, but also faced various challenges in managing the interactions with others. These challenges included technical concerns like adapting the software to suit individual's needs, social concerns like establishing rapport and encouraging participation from participants, as well as concerns arising from communicative asymmetries between the participants, i.e., limited visibility and audibility, as well as distractions. By highlighting the challenges and by presenting design sensitivities, we hope to provide guidance for other researchers, designers and facilitators of multi-party video based training programs.

Having the study conducted with only 2 parents may raise questions on the wider applicability of the presented findings. We would like to highlight that the findings discussed here are formulated from observing multiple sessions over an extended period of 3 months. Furthermore, this study was a pilot to test the feasibility of the method and the insights gained are used as a precursor to a larger study (happening next year) from which more broadly representative results could be gathered. We anticipate that the challenges in conducting video based training will further magnify as the number of participants increases. Thus, further work is required to understand the implication of video based training on the needs of families and children with autism.

ACKNOWLEDGEMENTS

Icons used in the Figure 1 are created by Lorena Galletosky and Snehal Patil and are available at <https://thenounproject.com> under CC license 3.0.

REFERENCES

- American Psychiatric Association. Diagnostic and statistical manual of mental disorders (5th ed.). Washington: American Psychiatric Association.
- Ames, M. G., Go, J., Kaye, J. J., & Spasojevic, M. Making love in the network closet: the benefits and work of family videochat. In Proc. CSCW 2010, ACM Press (2010).
- Baharav, E., & Reiser, C. Using telepractice in parent training in early autism. *Telemed J E Health* 2010, 16(6), 727-731.
- Baio, J. Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years - Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2010. *MMWR Surveillance Summaries* 2014, 63(2), 1-21.
- Barbaro, J. & Dissanayake, C. Prospective identification of autism spectrum disorders in infancy and toddlerhood using developmental surveillance: The social attention and communication study. In *Developmental and Behavioral Pediatrics* 2010, 376-385.
- Boisvert, M., and Hall, N. The use of telehealth in early autism training for parents: a scoping review. In *Smart Homecare Technology and Telehealth*, 2014: 2, 19-27.
- Brave, S., Ishii, H., & Dahley, A. Tangible Interfaces for Remote Collaboration and Communication. In Proc. CSCW 1998, ACM Press (1998), 169-178.
- Cadiz, J. J., Balachandran, A., Sanocki, E., Gupta, A., Grudin, J., & Jancke, G. Distance learning through distributed collaborative video viewing. In Proc. CSCW 2000, ACM Press (2000), 135-144.
- Carter, A. S., Messinger, D. S., Stone, W. L., Celimli, S., Nahmias, A. S., & Yoder, P. A randomized controlled trial of Hanen's 'More Than Words' in toddlers with early autism symptoms. In *Child Psychology & Psychiatry* 2011, 52(7), 741-752.
- Ciolfi, L. Situating 'Place' in Interaction Design: Enhancing the User Experience in Interactive Environments, 2004. Dept. of Computer Science and Information Systems, University of Limerick, Ireland.
- Dalsgaard, P. Designing for inquisitive use. In Proc. DIS 2008, ACM Press (2008), 21-30.
- Dillenbourg, P. & Baker, M. Negotiation Spaces in Human-Computer Collaborative Learning. In Proc. COOP 1996.
- Heath, C., & Luff, P. Disembodiment Conduct: Communication through video in a multi-media office environment. In Proc. CHI 1991, ACM Press (1991), 99-103.
- Heath, C, and Luff, P. Media Space and Communicative Asymmetries: Preliminary Observations of video-mediated interaction. In *Human Computer Interaction* 1992, 315-346.
- Ingersoll, A., & Gergans, S. The effect of a parent-implemented imitation intervention on spontaneous imitation skills in young children with autism. *Research in Developmental Disabilities*, 2007, 28, 163-175.
- Isaacs, E. A., & Tang, J. C. What video can and cannot do for collaboration: A case study. *Multimedia Systems* 1994, 2, 63-73.
- Ishii, H., & Ullmer, B. Tangible Bits: towards seamless interfaces between people, bits and atoms. In Proc. CHI 1997, ACM Press (1997), 234-241.
- Jancke, G., Grudin, J., & Gupta, A. Presenting to local and remote audiences: design and use of the TELEP system. In Proc. CHI 2000, ACM Press (2000), 384-391.
- Kaye, J. J., Go, J., Ames, M., and Spasojevic, M. The Joys and Frustrations of Family Videoconferencing. In Proc. CHI 2010, ACM Press (2010).
- Kirk, D. S., Sellen, A., & Cao, X. Home video communication: mediating 'closeness'. In Proc. CSCW 2010, ACM Press (2010).
- Mackintosh, V. H., Goin-Kochel, R. P., & Myers, B. J. "What Do You Like/Dislike about the Treatments You're Currently Using?": A Qualitative Study of Parents of Children with Autism Spectrum Disorders. *Focus on Autism and Other Developmental Disabilities* 2012, 27(1), 51-60.
- Mahoney, G., & Perales, F. Using relationship focused intervention to enhance the social-emotional functioning of your children with autism spectrum disorders. In *Topics in Early Childhood Special Education*, 2003, 23, 77-89.
- Meadan, H., & Daczewitz, M. E. Internet-Based Intervention Training for Parents of Young Children with Disabilities: A Promising Service-Delivery Model. *Early Child Development and Care*, 2015, 185(1), 155-169.
- Neustaedter, C., & Greenberg, S. Intimacy in long-distance relationships over video chat. In Proc. CHI 2012, ACM Press (2012).
- Olson, J. S. & Olson, G. M. Working together apart: Collaboration over the Internet. Morgan & Claypool, 2014.
- Oono, I. P., Honey, E. J., & McConachie, H. Parent-mediated early intervention for young children with autism spectrum disorders (ASD). *Cochrane Database of Systematic Reviews*, 2013.
- Pirsiavash, H., and Ramanan, D. Detecting activities of daily living in first-person camera views. In Proc. CVPR 2012, IEEE (2012), 2847-2854.
- Roberts, & Kaiser, A. The Effectiveness of Parent-Implemented Language Interventions: A Meta-Analysis. *American Journal of Speech-Language Pathology*, 2011, 20(3), 180-199.
- Rogers, Y., Lim, Y.-K., Hazlewood, W. R., & Marshall, P. Equal Opportunities: Do Shareable Interfaces Promote More Group Participation Than Single User Displays?. In *Human-Computer Interaction* 2009, 24 (1), 79-116.

- Seligmann, D. D., Roussel, N., Evans, H., & Hansen, H. Proximity as an Interface for Video Communication. In Proc. MultiMedia 2004, IEEE (2004), 11(3),12-16
- Silver, M., & Oakes, P. Evaluation of a new computer intervention to teach people with autism or Asperger syndrome to recognize and predict emotions in others. *Autism* 2001, 5(3), 299-316.
- Strauss, A. L., & Corbin, J. M. Basics of qualitative research: Techniques and procedures for developing grounded theory (2nd ed.). Thousand Oaks, 1998, CA: Sage.
- Suess, A., Romani, P., Wacker, D., Dyson, S., Kuhle, J., Lee, J., ... Waldron, D. Evaluating the Treatment Fidelity of Parents Who Conduct In-Home Functional Communication Training with Coaching via Telehealth. *Journal of Behavioral Education* 2014, 23(1), 34-59.
- Sussman, F., Honeyman, S., Lowry, L., & Drake, L. More Than Words - The Hanen Program for Parents of Children with Autism Spectrum Disorders (Third ed.). 2013. Toronto, Canada: The Hanen Centre.
- Veinott, E., Olson, J. S., Olson, G. M., & Fu, X. Video helps remote work: Speakers who need to negotiate common ground benefit from seeing each other. In Proc. CHI 1999, ACM Press (1999), 302-309.
- Wang, J., & Hartley, K. Video Technology as a Support for Teacher Education Reform. In *Technology and Teacher Education 2003*, 11(1), 105-138.
- White, S. A., Gupta, A., Grudin, J., Chesley, H., Kimberly, G., Sanocki, E. Evolving Use of A System for Education at a Distance. In Proc. HICSS-33, 2000, IEEE (2000).
- Zhang, Z. Microsoft kinect sensor and its effect. In Proc. Multimedia 2012, IEEE(2012), 19(2), 4-10.