Lessons Learned in Designing AI for Autistic Adults

Andrew Begel
andrew.begel@microsoft.com
Microsoft Research
Redmond, Washington

John Tang
johntang@microsoft.com
Microsoft Research
Redmond, Washington

Sean Andrist
sandrist@microsoft.com
Microsoft Research
Redmond, Washington

Michael Barnett
mbarnett@microsoft.com
Microsoft Research
Redmond, Washington

Tony Carbary
tony@microsoft.com
Microsoft Research
Redmond, Washington

Piali Choudhury
pialic@microsoft.com
Microsoft Research
Redmond, Washington

Edward Cutrell
cutrell@microsoft.com
Microsoft Research
Redmond, Washington

Alberto Fung*
alberto.fung@gmail.com
Department of Health and Human Performance, University of Houston
Houston, Texas

Sasa Junuzovic
sasajun@microsoft.com
Microsoft Research
Redmond, Washington

Daniel McDuff
damcduff@microsoft.com
Microsoft Research
Redmond, Washington

Kael Rowan
kaelr@microsoft.com
Microsoft Research
Redmond, Washington

Shibashankar Sahoo*
sahooshib1@gmail.com
Umeå Institute of Design
Umeå, Sweden

Jennifer Frances Waldern
jennifer.waldern@microsoft.com
Microsoft
Redmond, Washington

Jessica Wolk
jessica.wolk@microsoft.com
Microsoft Research
Redmond, Washington

Hui Zheng*
hzheng5@gmu.edu
Information Technology and Science, George Mason University
Fairfax, Virginia

Annuska Zolyomi†
annuska@uw.edu
Information School, University of Washington, Seattle
Seattle, Washington

ABSTRACT

Through an iterative design process using Wizard of Oz (WOz) prototypes, we designed a video calling application for people with Autism Spectrum Disorder. Our Video Calling for Autism prototype provided an Expressiveness Mirror that gave feedback to autistic people on how their facial expressions might be interpreted by their neurotypical conversation partners. This feedback was in the form of emojis representing six emotions and a bar indicating the amount of overall expressiveness demonstrated by the user. However, when we built a working prototype and conducted a user study with autistic participants, their negative feedback caused us to reconsider how our design process led to a prototype that they did not find useful. We reflect on the design challenges around developing AI technology for an autistic user population, how Wizard of Oz prototypes can be overly optimistic in representing AI-driven prototypes, how autistic research participants can respond differently to user experience prototypes of varying fidelity, and how designing for people with diverse abilities needs to include that population in the development process.

*Work was done while author was an intern at Microsoft Research.
†Work was done while author was a contractor at Microsoft Research

ASSETS '20, October 26–28, 2020, Virtual Event, Greece

ACM ISBN 978-1-4503-7103-2/20/10…$15.00
https://doi.org/10.1145/3373625.3418305

CCS CONCEPTS

• **Human-centered computing → Accessibility design and evaluation methods**.

KEYWORDS

Video calling, Autism Spectrum Disorder, design process, Artificial Intelligence, AI in design

ACM Reference Format:
Andrew Begel, John Tang, Sean Andrist, Michael Barnett, Tony Carbary, Piali Choudhury, Edward Cutrell, Alberto Fung, Sasa Junuzovic, Daniel McDuff, Kael Rowan, Shibashankar Sahoo, Jennifer Frances Waldern, Jessica Wolk, Hui Zheng, and Annuska Zolyomi. 2020. Lessons Learned in Designing AI for Autistic Adults. In *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '20), October 26–28, 2020, Virtual Event, Greece.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3373625.3418305

1 INTRODUCTION

Video calling applications offer real-time video, audio, chat, and desktop sharing channels that enable people to establish and maintain relationships over a distance, which has become important during this COVID-19 pandemic response. Our prior research [23] has shown that autistic adults experience significant stress during video calls, caused by sensory overstimulation, cognitive overload, and anxiety. To reduce some of the cognitive load and anxiety caused by each conversant reading and reacting to their conversation partner's emotional state, we developed an AI computer vision system to detect facial expressions and display them live, on-screen during the call. We developed two modes; the first, Expressiveness Mirror, reads the autistic person's facial affect and displays it to (just) them to give feedback about the way their affect may be perceived by others to help them feel more comfortable and better understood. The second, Expressiveness Prosthetic, reads the facial affect of the autistic person's conversation partner and displays it to the autistic person to help them more easily read and react to it during the conversation. In this paper, we focus solely on our experiences designing, developing, and evaluating the Expressiveness Mirror.

Figure 1 shows the Expressiveness Mirror giving the user on the left feedback on how a neurotypical person is likely to interpret his facial expressions. We engaged in an iterative design approach using Wizard of Oz prototyping techniques to show five autistic adults the prototype concept. However, when we built a working prototype and tested it with autistic users, their negative feedback caused us to reconsider its design. We reflect on the design issues around artificial intelligence (AI) and human-computer interaction (HCI), AI in Accessibility, Wizard of Oz prototypes in designing for people with diverse abilities, and general issues around designing for accessibility.

2 RELATED WORK

Several research projects have the goal of supporting autistic people in functioning more comfortably in neurotypical social and physical environments. For example, traditional therapies teach autistic children skills in social communication, speech and language, facial processing, and emotional recognition. More recent efforts utilize technology to facilitate socio-emotional interactions, including a robotic dinosaur that expresses emotion through movement and non-verbal sounds [11] and Jibo, an emotional storytelling robot whose use is facilitated by a caregiver [16]. In the areas of emotion and cognitive skills development, technology supports have been developed to teach autistic people to read emotions (happy, sad, anger, fear, surprise, disgust) [8, 18] and identify others' states of mind (agreement, concentration, disagreement, interest, in thought, confusion) [12, 15, 25]. However, effectively visualizing faces for autistic people to look at has been difficult, as studies have shown that reducing realism does not improve the recognizability of emotions [17]. A few projects help autistic people in real-time scenarios, for example, wearable technology that assists autistic people [1] in providing live feedback for vocal affect [4] and expression recognition [6, 21]. Virtual reality has been used to help autistic children learn to consider how to keep socially appropriate personal distance from others [3] and for autistic youth and adults to practice meeting new people and conduct job interviews [9]. Experiments with various modalities, such as video calls [14] or shared virtual spaces [24], have been shown as valuable spaces for autistic children to communicate, interact, and play with others, with technology-mediated interaction strategies such as drawing them away from distractions.

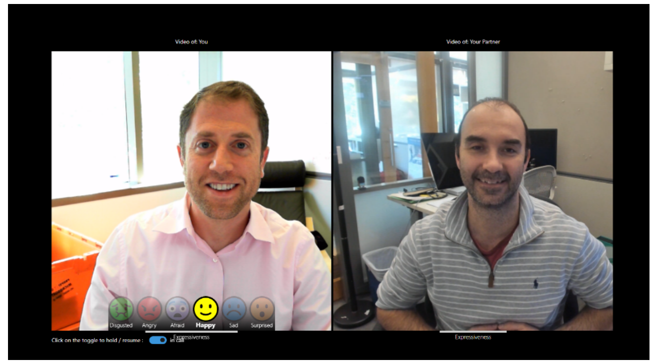

Figure 1: The Video Calling for Autism (VC4A) prototype shows the Expressiveness Mirror for the person on the left, whose expressions are detected by the AI computer vision agent and represented by six emotion emojis.

Tornblad et al. [19] highlight the need for communication tools and expression aids for people with ASD to engage in effective social exchanges and information sharing with neurotypical people. They also identify that autistic people have an affinity toward technology. Their work supports exploring technology that could scaffold communication between autistic and neurotypical people. Our work explores how to support communicating with autistic people in real time through video calling, and highlights some of the design challenges around using a Wizard of Oz design process to incorporate AI-driven expression recognition technology in a prototype supporting an autistic population.

3 THE DESIGN OF VIDEO CALLING FOR AUTISM

The Video Calling for Autism (VC4A) prototype was designed using a typical iterative design process using Wizard of Oz (WOz) prototypes with people in our user population (with Autism Spectrum Disorder). Building on our earlier formative study in which we interviewed 22 autistic adults about their experiences with video calling [23], we identified a number of design ideas that could support them. Through agile design sprints [10] and co-creation methodologies with the user experience (UX) researchers, developers, and the whole team, we generated several design sketches and prototypes that were evaluated by the priorities that emerged from the formative study, feasibility based on the technology we had available, and commercial product viability. This iterative conceptual design phase helped us to converge on supporting emotional expressiveness in video calling. Difficulty interpreting others' facial expressions (and vice versa, emoting their feelings effectively to others) are ideal challenges for an AI solution. To help autistic users feel more connected, we wanted to support emotional interpretation, either their partner's or their own. For the first version, we elected to pursue the Expressiveness Mirror concept, which would enable autistic users to manage or clarify their intentions, if they wish, to bridge the gap between neurodiverse and neurotypical differences in communication styles. We also felt that given the current inaccuracy in AI-inferred facial expressions, it was important to show them only to the person generating them, so they could contextually interpret any inaccuracies and disregard them without sharing with others.

We conducted a WOz study to evaluate the design before building it. We recruited five autistic adults (3 men and 2 women, between 18-40 years old, all employed full-time) to view five pre-recorded dyadic conversations (using neurotypical actors) that used a combination of auras, emojis, and text labels, shown in Figure 2, to render video conversants' facial expressions as interpreted by our designer (i.e., the wizard). All the participants indicated a preference for the emojis because they felt they were familiar and easy to interpret. P5 commented, "*I like these Emojis…easy to understand and less distracting*." P2 concurred, "*Emoji was most helpful.…achieved the right balance, gave enough information.*" P4 explained, "*There's more of a connection between emoji and the concept.*"

They also told us that our conceptualization to visualize facial expressions would be valuable in supporting their video calling process. P2 commented, "*1000% I want this. I want this in my life.*" And P5 commented, "*I would absolutely use it…I would like to have it built tomorrow.*" The WOz study indicated that the emoji were simple to use and interpret, which are criteria that Tornblad et. al [19] suggested for this user population.

Encouraged by these positive responses to the concept, we proceeded to complete the design and implementation of a working prototype. Based on the feedback from the WOz study, we focused on the emoji design, resulting in the UX shown in Figure 1. Five months later, we were able to conduct a user study of the VC4A working prototype. We planned a within-subjects, two-condition study (a video call with and without the Expressiveness Mirror feedback), in which we would employ a neurotypical, confederate conversation partner to try to elicit a broad range of emotions from the autistic study participant. To evaluate the user study process even before the working prototype was ready to deploy, we conducted a pilot test using a human wizard to drive the VC4A UX. We recruited a neurotypical participant, who happened to be a machine learning expert. Her feedback was that she was surprised at how responsive and accurate the expressiveness detection was, and that it was an interesting concept. Encouraged by these promising results, we completed building the working prototype and scheduled a user study with 21 autistic participants over a 2-week period.

Figure 2: Design ideas using an aura surrounding the face, emojis, or text to convey expressions within a video call UX.

To accommodate running a user study among a user population that is hard to fulfill, we developed the prototype on a web-based platform using WebRTC that would enable us to conduct the study remotely with the participant wherever they could connect to a web-based video call. This study design also enabled having the user connect from their own home, using their own device, and without the physical presence of researchers, which are usability testing guidelines recommended for this user population [19]. Our system supported audio, video, and text-based chat functionality. In the basic prototype we created, a user logs onto the system and "calls" another by clicking on their name at the top of the screen. Once a connection is established, the user can see the other person's video in equally-sized displays, unlike typical video calling systems which minimize the preview display of the user's own video. This enables visualization of either conversation partner's facial affect immediately adjacent to their image without requiring changes for size.

In the background, an additional connection was established to a backend server running an AI computer vision pipeline [13] that captures the perceived association between facial expressions and six emotional states. The states selected in this case were those identified by Paul Ekman [7]. In real-time, the pipeline would receive video frames from both conversants in a video call and return six values describing the likelihood (averaged over 5 seconds) the person in the video was expressing one of six non-neutral expressions: happiness, sadness, fear, surprise, anger, and disgust, and another value describing the 1-minute moving average of the sum of non-neutral expressions to illustrate the amount of expression the user had recently displayed with their face. The time period for the moving averages was determined through empirical iterations with the goal to minimize noise over short time periods due to fleeting facial muscle activations and match the change in value to real changes in affect occurring over longer periods of time.

We worked with a design team to explore how to render the facial expression detection in the UX, explicitly trying to minimize the cognitive overhead required to interpret it in real-time during a conversation, which is a concern for the autistic user population. Each design explored a different rendering technique to visualize the six expressions: aura, emoji, and text (see Figure 2). As the likelihood of each expression reported by our computer vision pipeline increases, the corresponding rendering increases in color saturation and size. Below a 20% threshold, the rendering either turns off (aura and text) or dims (emoji). Only the top two expressions above the threshold are displayed at the same time. This representation was intended to convey how a neurotypical person would be interpreting the autistic person's facial expressions. By enabling the autistic person to be aware of these interpretations, we hoped they would be able to clarify their communicational intent, either by augmenting or modifying their facial expressions or by verbally articulating what they were feeling, especially if it did not match how their facial expressions were being perceived. We hoped that this UX would help bridge the communication gap between neurotypical and neurodiverse conversation partners.

A conversant's overall expressiveness is displayed as a double-sided, horizontal, white meter displayed at the bottom of their video (see Figure 1). Beginning at a default of 50%, the conversant's meter grows longer (from the center) as their expressiveness increases over time. The meter shrinks as expressiveness decreases, dropping to a minimum, but non-zero, length representing a complete lack of detectable non-neutral facial affect. This representation was intended to give the autistic person feedback about their overall amount of expressiveness, relative to their conversation partner. Since autistic people are often perceived as showing flat, unexpressive affect, we hoped this display would help them be aware if they were being perceived as lacking expression, again enabling them to mitigate this perception during the conversation.

4 UNEXPECTED RESULTS

In the first week of the user study, we completed 7 of 12 scheduled sessions. The 7 completed sessions involved 4 men and 3 women between 18-40 years old, one employed full-time, and 3 part-time. Five of the participants were interviewed in the formative study [23], but none participated in the WOz study. Participants were first asked to answer some demographic questions and to speak about their prior experiences with video calling. Users then engaged in two 10-minute conversations with the confederate, one with the UX and one without (with the order counter-balanced between participants). Using an IRB-approved protocol, the confederate was asked to engage the participants in topics that could elicit happy, sad, surprised, and angry emotional reactions (without being insulting or cruel) and then bring them back to a baseline, neutral emotional state. Finally, users were asked about their level of distraction, confidence, emotional awareness, control, and comfort during the conversations, and for their feedback about their perceived performance, accuracy, and utility of the prototype. Reactions from the first 7 sessions were largely negative. So, we canceled the second week of study and reflected on the reasons for the disconnect.

Participants were critical about the accuracy and responsiveness of the expression detection AI agent. P12 said that "*It didn't seem to be accurate and a little distracting.*" and P12 responded to the question about accuracy, "*For the most part no. Happy was lighting up all the time.*" Furthermore, they were confused when some ostensibly conflicting emojis illuminated at the same time. P13 explained that he "*More paid attention to it when it was weird: happy and sad at the same time.*" This caused them to lose trust in the prototype.

Beyond the performance of the prototype, participants were critical of the concept. They observed that the six basic emotions rendered in the emojis of VC4A's UX were not the expressiveness feedback they wanted. P14 observed, "*Passion and happiness were confused because it was using basic emotions, but there are the more complicated ones. It doesn't detect stress.*" They were more interested in seeing anxiety, stress, frustration, confusion, or sarcasm, which are much more complex emotions. Just over half of the participants noticed the Expressiveness Meter, but its design did not communicate how to react to the information. P11 said of her conversation partner's meter ' '*hers was much longer, but, I did not know if I was supposed to make mine bigger or not–am I supposed to be more expressive?*" While we had gotten encouraging feedback on the prototype concept throughout the iterative design process with the WOz prototype, our user study with the working prototype elicited discouraging reactions.

5 DESIGN PROCESS REFLECTIONS

Reflecting on our design process raised several issues.

5.1 Applying AI to Accessibility for Autism

Although we know that autistic people often emote differently than neurotypical people (flat affect or overly expressive), we lacked enough facial expression data from autistic people to train an AI. Instead, our AI was trained by (presumed) neurotypical raters on (presumed) neurotypical facial expression data. We tried to finesse this by choosing a user scenario in which the autistic person was offered feedback on the way a neurotypical person would interpret their autistic facial expression, but our users instead believed that the system was trying to show them their true emotions. They fixated on why the system was inaccurate or doubted their ability to understand their own emotional state. Most of the users turned out to be comfortable interpreting most of the six basic emotions on their own but wanted our system to identify much more complex cognitive and emotional states that would help them mediate their conversation.

5.2 AI meets WOz

WOz prototypes can too easily gloss over AI implementation details that affect its value as a simulation tool. For example, our working prototype exhibited some noticeable delays in detecting and visualizing an expression, and sometimes triggered more than one emoji at a time (e.g., happy and sad). Our human wizard avoided both issues; thus, negative reactions to these issues were not evoked during the WOz testing. Conversely, the wizard could not keep up with the live video as well as the AI could. Nor could a single wizard replicate the fractional likelihood rating given by the AI because the model learns probabilistic estimates for face images during supervised training on data labeled by multiple raters. Finally, wizards would find it difficult to make the same kinds of mistakes an AI would, especially if trying to generate false positives when no identifiable expression is seen. Though a wizard that is too good may provide an upper bound on the value of a scenario, it cannot help to identify the minimum accuracy required to offer a useful AI-driven scenario. Our experiences with a WOz design process add to the challenges being identified of designing Human-AI interaction [22].

5.3 Conceptual Optimism

Recognizing and reacting appropriately to another's emotions can be difficult, but people do this all the time! It becomes problematic when they (or the AI) overtly proclaim what the other person is feeling. Despite this (and all the other problems with our prototype), our study participants were still optimistic about the concept and wanted us to develop an Expressiveness Prosthetic scenario to show them an interpretation of their conversation partner's facial expressions. This optimism was based on a verbal description of the Expressiveness Prosthetic and their experiences seeing their own emotions. As we learned throughout the VC4A process, the level of fidelity of a prototype affects the ability of a research participant to imagine what the user experience will actually be like in practice. Given our experiences with VC4A, we think an Expressiveness Prosthetic would be even more problematic, as it could show expressiveness inferences to others without the person being aware of what inferences were being made about their expressions. This optimism, which is a common reaction in the accessibility design space [20], contributed to leading us down a design garden path that ultimately resulted in a prototype that was not useful.

5.4 Autism in Design

Designing for the spectrum of diverse abilities among the autistic user population is an interesting challenge, and our experiences highlighted the importance of social context in interacting with autistic people. Autistic people are often misconstrued as appearing angry during conversations, but we do not think that is because their facial expressions connote anger. Reviewing the video recordings of the conversations showed that the anger emoji rarely registered during the conversation. However, the autistic participants tended to maintain a neutral facial expression more than their conversation partner, as evidenced by shorter Expressiveness Meters. Even though their facial expressions are not demonstrating anger, they may be socially interpreted as being angry since people tend to suppress visible expressions of anger. So, prolonged neutral expression may be considered as being angry. The AI agent that we used did not consider any social context of interpreting expressions.

Although we involved autistic people in the design process as user study participants, they did not serve on our design and development team. Thus, our well-intentioned design process did not abide by the "nothing about us without us" principle [5]. In the spirit of a Participatory Design process [2], we are setting up an advisory board of autistic people to provide expertise and guidance through all phases of the design process.

6 CONCLUSION

Our experiences designing VC4A ended up teaching us more about the design process for the autistic population, especially when involving AI technology. We tried to adapt an AI agent trained on neurotypical data to a neurodiverse population, in part because of the challenge of training new expression recognition models on much scarcer data from people who are neurodiverse. Developing AI models for populations that have more limited data remains a challenge for the future.

More generally, building design probes that leverage AI functionality can be difficult. We encountered ways that WOz techniques overestimate the actual AI capabilities in ways that affect the viability of the design concept. In the current stage of development of AI technologies, it is very important to develop a fully working prototype to accurately evaluate users' reactions to the concept. Now that we have built a platform for a working prototype, we can build on what we have learned to explore other AI components or different ways of surfacing them to the user. We want to share these experiences with the community to encourage more discussion about how to shape our research and design processes with AI and populations with diverse abilities.

ACKNOWLEDGMENTS

We thank our anonymous study participants for their time, opinions, and insights that guided us through this research journey.

REFERENCES

[1] E. M. Benssassi, J. Gomez, L. E. Boyd, G. R. Hayes, and J. Ye. 2018. Wearable Assistive Technologies for Autism: Opportunities and Challenges. *IEEE Pervasive Computing* 17, 2 (2018), 11–21.
[2] Laura Benton, Hilary Johnson, Emma Ashwin, Mark Brosnan, and Beate Grawemeyer. 2012. Developing IDEAS: Supporting Children with Autism within a Participatory Design Team. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 2599–2608. https://doi.org/10.1145/2207676.2208650
[3] LouAnne E. Boyd, Saumya Gupta, Sagar B. Vikmani, Carlos M. Gutierrez, Junxiang Yang, Erik Linstead, and Gillian R. Hayes. 2018. VrSocial: Toward Immersive Therapeutic VR Systems for Children with Autism. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada)

(CHI '18). Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173778

[4] LouAnne E. Boyd, Alejandro Rangel, Helen Tomimbang, Andrea Conejo-Toledo, Kanika Patel, Monica Tentori, and Gillian R. Hayes. 2016. SayWAT: Augmenting Face-to-Face Conversations for Adults with Autism. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 4872–4883. https://doi.org/10.1145/2858036.2858215

[5] James I. Charton. 2000. *Nothing About Us Without Us: Disability Oppression and Empowerment* (first ed.). University of California Press, Berkeley, CA.

[6] Jena Daniels, Jessey N. Schwartz, Catalin Voss, Nick Haber, Azar Fazel, Aaron Kline, Peter Washington, Carl Feinstein, Terry Winograd, and Dennis P. Wall. 2018. Exploratory study examining the at-home feasibility of a wearable tool for social-affective learning in children with autism. *npj Digital Medicine* 1, 1 (Aug. 2018), 32. https://doi.org/10.1038/s41746-018-0035-3

[7] P. Ekman and W. V. Friesen. 1975. *Unmasking the Face: A Guide to Recognizing Emotions from Facial Clues.* Prentice-Hall, Englewood Cliffs, NJ.

[8] Valerie Gay and Peter Leijdekkers. 2014. Design of emotion-aware mobile apps for autistic children. *Health and Technology* 4, 1 (May 2014), 21–26. https://doi.org/10.1007/s12553-013-0066-3

[9] Michelle R Kandalaft, Nyaz Didehbani, Daniel C Krawczyk, Tandra T Allen, and Sandra B Chapman. 2013. Virtual reality social cognition training for young adults with high-functioning autism. *Journal of autism and developmental disorders* 43, 1 (Jan. 2013), 34–44. https://doi.org/10.1007/s10803-012-1544-6 Publisher: Springer US.

[10] Wally Keijzer-Broers and Mark de Reuver. 2016. Applying Agile Design Sprint Methods in Action Design Research: Prototyping a Health and Wellbeing Platform. In *Proceedings of the International Conference on Design Science Research in Information System and Technology*. St. John's, NL, Canada, 68–80. https://doi.org/10.1007/978-3-319-39294-3_5

[11] E. S. Kim, L. D. Berkovits, E. P. Bernier, D. Leyzberg, F. Shic, R. Paul, and B. Scassellati. 2013. Social robots as embedded reinforcers of social behavior in children with autism. *Journal of autism and developmental disorders* 43, 5 (2013), 1038–1049. https://doi.org/10.1007/s10803-012-1645-2

[12] Miriam Madsen, Rana el Kaliouby, Matthew Goodwin, and Rosalind Picard. 2008. Technology for Just-in-Time in-Situ Learning of Facial Affect for Persons Diagnosed with an Autism Spectrum Disorder. In *Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility* (Halifax, Nova Scotia, Canada) *(Assets '08)*. Association for Computing Machinery, New York, NY, USA, 19–26. https://doi.org/10.1145/1414471.1414477

[13] Daniel McDuff, Kael Rowan, Piali Choudhury, Jessica Wolk, ThuVan Pham, and Mary Czerwinski. 2019. A Multimodal Emotion Sensing Platform for Building Emotion-Aware Applications. arXiv:1903.12133 [cs.HC]

[14] Sanika Mokashi, Svetlana Yarosh, and Gregory D. Abowd. 2013. Exploration of Videochat for Children with Autism. In *Proceedings of the 12th International Conference on Interaction Design and Children* (New York, New York, USA) *(IDC '13)*. Association for Computing Machinery, New York, NY, USA, 320–323. https://doi.org/10.1145/2485760.2485839

[15] Kathryn E. Ringland, Christine T. Wolf, LouAnne E. Boyd, Mark S. Baldwin, and Gillian R. Hayes. 2016. Would You Be Mine: Appropriating Minecraft as an Assistive Technology for Youth with Autism. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility* (Reno, Nevada, USA) *(ASSETS '16)*. Association for Computing Machinery, New York, NY, USA, 33–41. https://doi.org/10.1145/2982142.2982172

[16] Brian Scassellati, Laura Boccanfuso, Chien-Ming Huang, Marilena Mademtzi, Meiying Qin, Nicole Salomons, Pamela Ventola, and Frederick Shic. 2018. Improving social skills in children with ASD using a long-term, in-home social robot. *Science Robotics* 3, 21 (2018). https://doi.org/10.1126/scirobotics.aat7544 arXiv:https://robotics.sciencemag.org/content/3/21/eaat7544.full.pdf

[17] Marc Spicker, Diana Arellano, Ulrich Schaller, Reinhold Rauh, Volker Helzle, and Oliver Deussen. 2016. Emotion Recognition in Autism Spectrum Disorder: Does Stylization Help?. In *Proceedings of the ACM Symposium on Applied Perception* (Anaheim, California) *(SAP '16)*. Association for Computing Machinery, New York, NY, USA, 97–104. https://doi.org/10.1145/2931002.2931004

[18] James W Tanaka, Samantha Lincoln, and Logan Hegg. 2003. A framework for the study and treatment of face processing deficits in autism. In *The development of face processing*, G. Schwarzer and H. Leder (Eds.). Hogrefe and Huber Publishers, Boston, MA, 101–119.

[19] McKenna K. Tornblad, Keith S. Jones, and Fethi A. Inan. 2019. Individuals with Autism Spectrum Disorder: A Review of Human Factors Considerations and Usability Testing Guidelines. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 63, 1 (2019), 1575–1579. https://doi.org/10.1177/1071181319631423

[20] Shari Trewin, Diogo Marques, and Tiago Guerreiro. 2015. Usage of Subjective Scales in Accessibility Research. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers and Accessibility* (Lisbon, Portugal) *(ASSETS '15)*. Association for Computing Machinery, New York, NY, USA, 59–67. https://doi.org/10.1145/2700648.2809867

[21] Peter Washington, Catalin Voss, Aaron Kline, Nick Haber, Jena Daniels, Azar Fazel, Titas De, Carl Feinstein, Terry Winograd, and Dennis Wall. 2017. SuperpowerGlass: A Wearable Aid for the At-Home Therapy of Children with Autism. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 112 (Sept. 2017), 22 pages. https://doi.org/10.1145/3130977

[22] Qian Yang, Aaron Steinfeld, Carolyn Rosé, and John Zimmerman. 2020. Re-Examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376301

[23] Annuska Zolyomi, Andrew Begel, Jennifer Frances Waldern, John Tang, Michael Barnett, Edward Cutrell, Daniel McDuff, Sean Andrist, and Meredith Ringel Morris. 2019. Managing Stress: The Needs of Autistic Adults in Video Calling. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 134 (Nov. 2019), 29 pages. https://doi.org/10.1145/3359236

[24] Annuska Zolyomi, Ankitha Bharadwaj, and Jaime Snyder. 2017. Let's Play (While Far Away)! Using Technology to Mediate Remote Playdates for Children with Autism. In *Design, User Experience, and Usability: Understanding Users and Contexts*, Aaron Marcus and Wentao Wang (Eds.). Springer International Publishing, Cham, 415–432.

[25] Annuska Zolyomi and Marc Schmalz. 2017. Mining for Social Skills: Minecraft in Home and Therapy for Neurodiverse Youth. In *Proceedings of the 50th Hawaii International Conference on System Sciences.* Manoa, HI. https://doi.org/10.24251/HICSS.2017.411