MeEmo – Using an Avatar to Improve Social Skills in Children with ASD

Sapna Patel[1], Darin Hughes[1], Charles Hughes[1]

[1] University of Central Florida

spatel@cs.ucf.edu, darin@cs.ucf.edu, ceh@cs.ucf.edu

Abstract

For individuals affected by Autism Spectrum Disorder (ASD), the inability to make eye contact is a significant barrier to their engagement in social environments. This lack of eye contact limits their ability to read social and emotional cues exhibited through facial expressions resulting in a corresponding decrease in social engagement. The use of interactive virtual environments (VEs) as a therapeutic protocol is a growing field of study. In recent studies, individuals with ASD were placed in VEs and engaged with avatars controlled by a human in the background, resulting in improvements in eye contact and engagement for some subjects. This paper is the first in a series of experiments exploring the potential of virtual avatars controlled through software agency, rather than human control, as a therapeutic tool for ASD. This paper examines if a subject could learn to make eye contact with an avatar and consequently recognize and respond to emotional cues expressed by the avatar. Results indicate that children with ASD can learn to recognize the emotional cues of the virtual avatar, and that their reactions to the avatar's needs as well as their eye contact with the avatar improved over the course of the VE experiment. This study sets the stage for future exploration into therapeutic use of agent-based virtual avatars, including transference of emotional cues from avatars to humans in the real world.

Index Terms: Avatar Based Therapy, Virtual Reality Environments, Software Agents, Individuals with ASD

1. Introduction

Autism Spectrum Disorder, or ASD, is a debilitating condition that impacts both the person with ASD and their family. It has historically been a condition with limited treatment options. People with Autism find it difficult to interact with others in social settings and are often limited in their ability to recognize emotional cues, speak, and communicate [1]. Although not necessarily the only reason for a lack of understanding of nonverbal cues of emotions, difficulty in making eye contact contributes to the challenges individuals with Autism face when engaging with other humans, limiting their social interactions.

Recent research has shown the promise of a new approach, which we will call Human-Controlled Avatar-Based Therapy or HC_ABT, in the treatment of Autism [2], [3], [4]. This approach uses avatars, or digital characters that are manipulated electronically by a remote human, often called an interactor. A form of HC_ABT called TeachLivE has been shown to be very effective at eliciting natural responses from individuals with Autism [5].

TeachLivE and other HC_ABT approaches have the potential to provide an extremely effective way to treat Autism, but because of the requirement for a human interactor, these approaches are best used only if there is evidence that a child will likely benefit from the therapy. Thus, while evidence has shown the effectiveness of HC_ABT for a specific individual, Agent-Controlled Avatar-Based Therapy or AC_ABT. can provide inexpensive, always available games that can be scaled for a larger audience. As an example, WUBeeS was explicitly designed to aid young children with ASD in perspective taking and empathy by placing them in the role of caregivers to an agent-controlled virtual avatar [6].

The goal of our research project is to leverage the WUBeeS and TeachLivE findings by exploring the opportunity to create a diagnostic and evaluation tool that can be used to determine the likelihood that a subject may benefit from an HC_ABT, Such a tool is inexpensive and always accessible as a first-line treatment/assessment protocol. In this study, we present a longitudinal study exploring the responses of children with ASD to the emotional cues of an agent-controlled virtual avatar. We measure each individual's ability to respond to an avatar's needs, as well as their eye contact and engagement with the facial expressions of the avatar.

1.1. Background

Autism Spectrum Disorder is a growing problem in the United States as well as across the world. According to the CDC, the prevalence of ASD has increased from 1 in 150 children in 2007 [8] to 1 in 88 children in 2012 [8]. 80% of the students in the US identified with ASD are the recipients of some form of special education services [8]. The budgetary challenges with the rising rates of ASD impact education in the near-term but also become long-term societal challenges that continue to have economic impact well into the future. People with ASD are often unable to lead productive and independent lives during adulthood, sometimes requiring continuous monitoring and support [8].

An avatar is an electronic character that can be manipulated by software agency, a human puppeteer or a combination of both. Past research has shown that children with ASD often prefer simulated environments over typical role-playing scenarios [9], [10], [11]. Avatar based therapy (ABT) is a relatively new approach that is being used to treat children with ASD. ABT environments can be controlled for facial expression complexity and can be less intimidating for the child with ASD. Early tests using avatars in a HC_ABT have been successful [3]. Extending that research into AC_ABT is an area of exploration that has the potential to create new treatment protocols for ASD [2]. Recent work [6] has looked at the use of agent-based virtual avatars as a way to assess and improve empathy and perspective-taking skills in children on the ASD spectrum.

TeachLivE is a virtual environment initially developed to assist in the training of teachers [12] and has been studied as a tool for measuring social responses for children with autism [4]. It is primarily a virtual training platform where teachers can practice and master the interpersonal skills they need to be

effective in the classroom. By using avatars that are placed in familiar environments but controlled by an interactor on the back end, the participant can be given a controlled opportunity to interact with the avatars in a way that is non-threatening and safe and that doesn't place real students at risk. This approach has gained significant momentum and has proven to be extremely successful [9], [10], [12], [13], [14].

One experiment conducted within the TeachLivE environment was to have students with ASD interact with the avatars. Each one of the avatars presented in the TeachLivE environment has a very unique personality and set of behaviors (http://TeachLivE.org/about/kids/). In one study, a student with ASD was exposed to the environment multiple times, The student developed a relationship with one avatar in particular, and learned to converse and communicate with the avatar. The subject, a teenage high schooler, previously only communicated in 2-3 word snippets but started conversing with the avatar in full sentences and actually having a conversation. More interestingly, as the subject developed more communication skills there was an observed transference of the same behaviors from the virtual environment to the real world [14]. This finding set the stage for the study being reported in this paper.

2. The Goal and Objectives

Current human-centered avatar based therapy (HC_ABT) methods for people with ASD have been shown to be effective, but are not highly scalable because of the constant need for a human interactor [4]. A diagnostic test to determine if an individual may benefit from an HC_ABT is a desirable and cost-effective first step to take before the HC_ABT approach is widely applied to help children diagnosed with ASD.

2.1. The Research Question

This research project is the first in a series of steps meant to extend the WUBees research [6] by asking the question "Is it possible to assess the emotional perspective-taking skills of children with Autism Spectrum Disorder using an inexpensive, on-demand avatar based simulation/game?" The specific research question posed in this paper is "can children with ASD learn to respond appropriately to the emotional facial cues of a virtual avatar in a game environment?"

As discussed above, research has shown that some children with ASD respond positively to a HC_ABT, developing improved social interactions and better communication skills. The initial goal of this research project is focused on using an AC_HBT as an indicator of the potential benefits an individual may gain in an HC_ABT, but the end goal is to remove the need for human interactors in avatar-based therapy, decreasing its cost and increasing the range of practical applications. This paper focuses on the ability of children with ASD to learn to associate avatar facial cues with game actions. If ASD individuals are able to learn to interact with the avatar based on the avatar's facial cues, future studies can explore how individuals differ in their ability to discern facial cues, as well as transference of these facial cue recognition skills to the real world. This will eventually set the stage for using a simple and inexpensive screening test for identifying which individuals are a good target for ABT.

2.2. Hypothesis

The hypothesis addressed in this paper is that if children with Autism Spectrum Disorder interact with the MeEmo virtual character on a regular weekly basis then, over time, they will be able to better recognize and respond to changes in MeEmo's emotions displayed via facial expressions. The secondary hypothesis tested was that, with repeated exposure to MeEmo, children with ASD would have increased focus on MeEmo's facial cues to judge emotional state.

3. Test Design

The software described here features a character called MeEmo. MeEmo is able to display emotional states ranging from happy to sad based on a combination of mouth and eye expressions, as well as health meters. The MeEmo is modeled after the WUBeeS project [6] with some minor modifications to show enhanced facial expressions and to enhance accessibility by making the software web-enabled. The MeEmo's emotional state was modeled to be a function of its level of hunger, how tired it is, how bored it is and if it needs a bathroom break. These four aspects of a MeEmo's current state reflect four key human needs. The screen also shows the "health meters" (bars at the top of the screen representing the level of happiness through colors – starting at green and gradually getting closer to red as the happiness level decreases) for each one of these four areas of focus.

The subjects had opportunities to react to each one of the MeEmo's states by moving it to an area of the game where a specific need could be met; for example, for a hungry MeEmo the game player would take the MeEmo to the kitchen and feed it. The program captured each interaction point between the person interacting with the MeEmo and transmitted that to a centralized web server.

Figure 1: For trials 1-10, subjects were shown visual health meters at the top of the screen. These health meters use a red-to-green scale to show the MeEmo's happiness in a given area, so that subjects can use visual cues to understand MeEmo's needs. In this example, the MeEmo wants to play in the backyard, as indicated by the red "Backyard" health meter as well as the crying MeEmo.
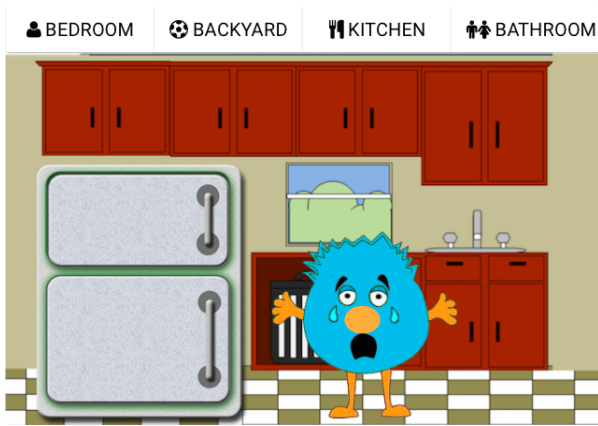
Figure 2: For trials 11-20, subjects were not shown visual health meters, and were only shown the emotional facial cues of the MeEmo as an indicator of MeEmo's happiness. Because of this, subjects must pay attention to facial cues (such as crying, in this case) in order to address the MeEmo's needs.

Subjects were taught how to play the game and what the goal of the game was. Subjects were tested two times a week, for a total of ten weeks. Each testing interaction was for five minutes. For the first five weeks of testing (condition 1), subjects were shown the health meters in the MeEmo game. In this stage, subjects could learn the dynamics of the game, including how to react to the MeEmo based on non-facial cues (the health meters). After five weeks of testing, the health meters from the game were removed (condition 2) and another five weeks of testing was conducted. This condition was meant to test if subjects could react to the needs of the MeEmo in the game based purely on facial expression.

Using an Eyetribe eyetracker, the study also measured what the subject was focused on in each MeEmo environment by exposing them to an image for 15 seconds. Starting with a baseline measurement, this eye-tracking study was repeated every five weeks. This data was used to see whether the subjects performed better with or without health meters – the latter process requiring the subject to recognize the MeEmo's emotional state purely through recognition of facial expressions.

Study subjects were recruited from middle and high schools in Seminole County, Florida. They were between 11 and 18 years of age. Test subjects were identified by placement in an ASD classroom based on their prequalification for an Exceptional Student Education plan. The teachers of these students were contacted prior to the study to outline the program and study. They then sent permission slips home with their students. Once the parents signed the permission slips, the students were then enrolled in the study. Control subjects were recruited in the same manner but from the regular class population. The overall study had 21 test subjects and 13 control subjects. Due to the limitations in recruiting subjects for the study, no attempt was made to control for gender or race.

4. Results

As described above, the experiments reported in this paper were meant to test if children with ASD are able to learn to respond to MeEmo's emotional facial cues over time. Subjects were tested over 20 trials; trials 1-10 included visual health meters as well as MeEmo facial cues, while trials 11-20 included only MeEmo facial cues. By analyzing trials 1-10, we can measure whether subjects were able to grasp the objective of the game (taking actions to keep MeEmo happy), revealing that both ASD and control groups were able to understand the gameplay. By analyzing progress over the course of trials 11-20, we were able to show that both ASD and control groups were able to learn to play the game based purely on MeEmo's emotional facial cues. Finally, comparisons between subject performance on trial 20 (based on MeEmo facial cues) and trial 10 (based on non-facial health meters) show that both ASD and control groups were able to respond to MeEmo's needs in both conditions.
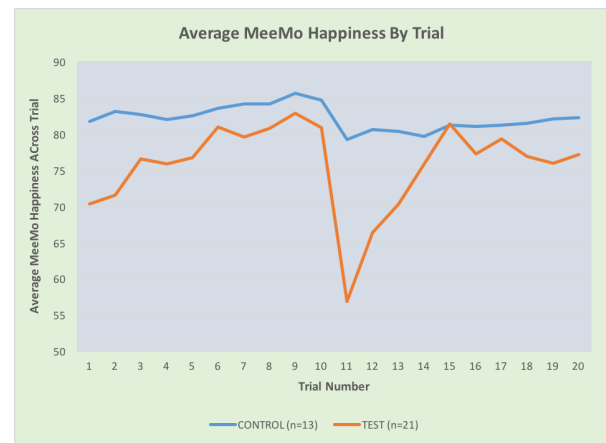

Figure 3: Average MeEmo happiness over the course of 20 trials for both ASD (test) and control groups. Both groups improved in performance over trials 1-10 (with visual health meters), as well as over trials 11-20 (with only facial cues). By the end of trial 20 (with facial cues), performance was similar to that of trial 10 (with visual health meters). The sharp decline at Trial 11 for the subject group with ASD is due to the removal of the health meters.

4.1. Learning the Objective of the Game: Trials 1-10

In trials 1-10, subjects were shown both visual health meters and MeEmo facial cues (as seen in **Figure 1**). The purpose of these trials was to evaluate each subject's ability to understand the gameplay and goal of the MeEmo's environment: responding to MeEmo's needs in order to maintain a high average happiness. The results indicate that both the ASD and control groups had increases in MeEmo happiness scores over the course of trials 1-10, indicating that they understood how to play the MeEmo game over the course of repeated exposure. Because subjects were shown both visual health meters as well as facial cues in these trials, these trials cannot show that subjects paid attention to facial cues, but rather that subjects understood the cause-effect relationship between their actions and keeping MeEmo happy.

For the ASD group, there was a significant increase in average MeEmo happiness scores from trial 1 (M=70.33, SD=9.85) to trial 10 (M=80.91, SD=5.91); t (df = 32) = -3.8098, p = 0.0003. Although the control group had an increase in average MeEmo happiness scores from trial 1 (M=81.78, SD=4.40) to trial 10 (M=84.74, SD=3.38), it was not statistically significant; t (df=32) = -2.196, p = 0.0176.

4.2. Playing The Game with Only Facial Cues: Trials 11-20

For trials 11-20, subjects were only shown emotional MeEmo facial cues (as seen in **Figure 2**). This means that in order to keep MeEmo happiness high, subjects need to focus on and respond to MeEmo facial cues. The results show that both the ASD and control groups had increases in MeEmo happiness scores over the course of these trials, indicating that the subjects were able to better understand and react to MeEmo's facial expressions over time.

For the ASD group, there was a significant increase in average MeEmo happiness scores from trial 11 (M=56.87, SD=6.41) to trial 20 (M=77.22, SD=4.12); t (df=32) = -11.01, p < 0.00001. The control group did not have a significant increase in average MeEmo happiness scores from trial 11 (M=79.22, SD=3.23) to trial 20 (M=82.29, SD=1.80); t (df=32) = -3.42, p = 0.99.

These results indicate that both groups were able to learn to respond to changes in MeEmo's emotions based on only facial cues.

4.3. Evaluating Performance With and Without Health Meters: Trial 10 vs Trial 20

The final comparison in this analysis was between trial 20 (the last trial based purely on facial cues) and trial 10 (the last trial based on visual health meters as well as facial cues). The purpose of this comparison is to determine if subjects were able to achieve similar levels of performance on the more difficult trial 20 (based on facial cues) as in the easier trial 10 (based on health meters) after repeated exposure.
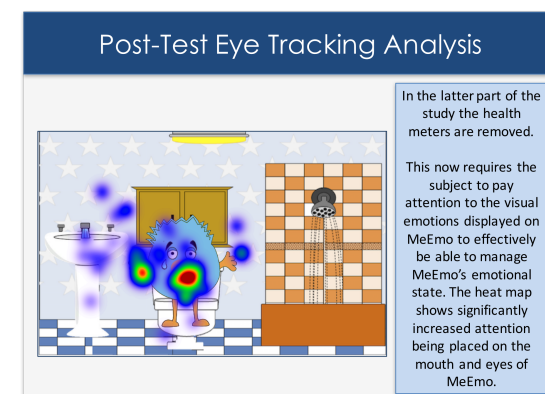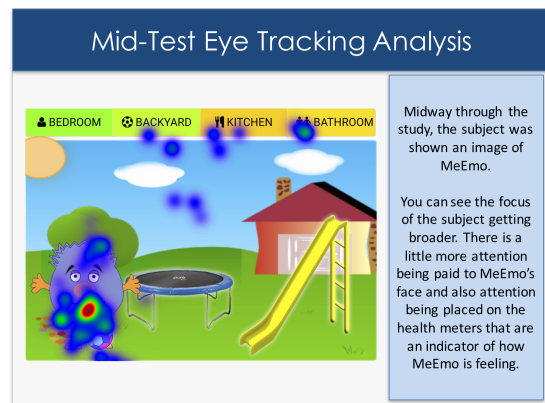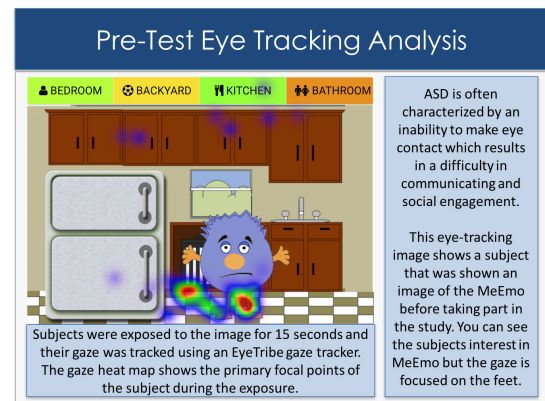
For the ASD group, there was no significant decrease in average MeEmo happiness scores from trial 10 (M=80.90, SD=5.91) to trial 20 (M=77.22, SD=4.12); t (df=32) = 2.11, p = 0.0221. Similarly, for the control group, there was no significant decrease in average MeEmo happiness scores from trial 10 (M=84.74, SD=3.38) to trial 20 (M=82.29, SD=1.80); t (df=32) = 2.2.64, p = 0.99. This indicates that both the ASD and the control groups could reach a similar level of performance in the MeEmo environment based purely on facial cues as they could with health meters.

4.4. Monitoring Eye-tracking

Subjects' eye movements and focus areas were monitored using an EyeTribe eye tracker. Participants were presented with static images from the MeEmo virtual environment for five seconds during which time their eye movements were measured.

This portion of the study is very interesting because it indicates that some subjects with ASD have the ability to recognize emotions as expressed through facial expressions rather than relying on the color-coded meters of happiness. Given that the core foundation of human interaction is based on the ability to recognize emotions as a precursor to showing empathy, this shows that some ASD subjects can make eye contact with an avatar and recognize emotions and react with empathy.

The following sequence of images shows how one subject's focus areas changed over the different trials.

Pre-Test Eye Tracking Analysis

Subjects were exposed to the image for 15 seconds and their gaze was tracked using an EyeTribe gaze tracker. The gaze heat map shows the primary focal points of the subject during the exposure.

ASD is often characterized by an inability to make eye contact which results in a difficulty in communicating and social engagement.

This eye-tracking image shows a subject that was shown an image of the MeEmo before taking part in the study. You can see the subjects interest in MeEmo but the gaze is focused on the feet.

Mid-Test Eye Tracking Analysis

Midway through the study, the subject was shown an image of MeEmo.

You can see the focus of the subject getting broader. There is a little more attention being paid to MeEmo's face and also attention being placed on the health meters that are an indicator of how MeEmo is feeling.

Post-Test Eye Tracking Analysis

In the latter part of the study the health meters are removed.

This now requires the subject to pay attention to the visual emotions displayed on MeEmo to effectively be able to manage MeEmo's emotional state. The heat map shows significantly increased attention being placed on the mouth and eyes of MeEmo.

5. Conclusions and Recommendations

In conclusion, the null hypothesis (Repeated interaction with MeEmo will not improve the child's ability to recognize emotions displayed visually on the avatar) was rejected. The experiment shows that, over a series of trials, subjects with ASD can learn to take better care of the virtual avatar's basic needs so that they can keep the MeEmo's overall happiness higher.

The results also show that, while test subjects with ASD score lower than the control subjects, there is continuous improvement over the duration of a trial and they are able to increase the overall happiness score of MeEmo over time.

The more interesting finding in the study comes from the eye-tracking observations, as well as observations over the course of trials 11-20 (where the subject was forced to rely on facial cues in order to react to MeEmo's needs). The data from MeEmo happiness scores shows that both ASD and control groups are able to react to MeEmo's emotional changes based purely on facial cues, without any other indicators of emotion. The early data from eye tracking measurements shows that some subjects have an increased focus on the MeEmo's face over the course of repeated testing. This is an important finding because it sets the stage for establishing the next phase of this research, i.e. to determine who is likely to benefit from HC_ABT in the future.

The next stage of research will focus on the applicability of the MeEmo game as a diagnostic test for HC_ABT. The results from the research reported here show that some of the subjects began to focus more on MeEmo's face as the sessions progressed. Next steps will include determining if this behavior can be used as an indicator to determine if a subject is a well-suited candidate for HC_ABT. Future research will also include the use of virtual auditory avatars as an interaction and engagement tool for subjects with ASD.

6. Acknowledgements

.

7. References

[1]   C. Kasari, S. F. Freeman, N. Bauminger and M. C. Alkin, "Parental perspectives on inclusion: Effects of autism and Down syndrome," *Journal of Autism and Developmental Disorder*s, Vol. 20, no. 4, pp. 297-305, 1999.

[2]   P. Mitchell, S. Parsons and A. Leonard, "Using virtual environments for teaching social understanding to 6 adolescents with autistic spectrum disorders," *Journal of Autism & Developmental Disorders*, Vol. 37, no. 3, pp. 589-600, 2007.

[3]   D. Moore, Y. Cheng, P. McGrath and N. J. Powell, "Collaborative virtual environment technology for people with autism," *Focus on Autism and Other Developmental Disabilities*, Vol. 20, no. 4, pp. 231-243, 2005.

[4]   E. Vasquez, C. Straub, C. Hughes, et. al., "A Comparison of Virtual Learning Environments on the Social Responses for Children with Autism," *The 2nd Annual TLE TeachLivE Conference, May 21-23, Orlando, FL, USA, Proceedings from Ludic Convergence*, University of Central Florida, 2014, pp. 40-44.

[5]   A. Nagendran, R. Pillat, A. Kavanaugh, G. Welch and C. E. Hughes, "A Unified Framework for Individualized Avatar-Based Interactions," *Presence: Teleoperators and Virtual Environments*, Vol. 23, no. 2, pp. 109-132, 2014.

[6]   D. Hughes, *The design and evaluation of a video game to help train perspective- taking and empathy in children with Autism Spectrum Disorder*, PhD. Dissertation, Modeling and Simulation Program, College of Sciences, University of Central Florida, Orlando, Florida, 2014.

[7]   D. L. Christensen, "Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years," *Autism and Developmental Disabilities Monitoring Networ*k, *11 Sites, United States 2012*," *MMWR Surveillance Summaries 2016*, Vol. 65, no. 3, pp. 1–23, 2012.

[8]   G. I. Orsmond, M. W. Krauss and M. M. Seltzer, "Peer relationships and social and recreational activities among adolescents and adults with autism," *Journal of Autism and Developmental Disorders*, Vol. 34, no. 3, pp. 245-256, 2004.

[9]   W. Bricken, "Training in virtual reality," *1st International Conference on Virtual Reality, November 5-7, Camps Bay, Cape Town, South Africa. Proceedings*, Meckler International, London, 1991, pp. 46-48.

[10]   Y. Cheng, & J. Ye, "Exploring the social competence of students with autism spectrum conditions in a collaborative virtual learning environment – The pilot study," *Computers & Educatio*n, Vol. 54, no. 4, pp. 1068–1077, 2009.

[11]   S. Parsons, A. Leonard and P. Mitchell, "Virtual environments for social skills training: Comments from two adolescents with autistic spectrum disorder," *Computers & Education*, Vol. 47, no. 2, pp. 186-206, 2004.

[12]   C. E. Hughes, A. Nagendran, L. Dieker, M. Hynes and G. Welch, "Applications of avatar-mediated interaction to teaching, training, job skills and wellness," In *Virtual Realities – Dagstuhl Seminar 2013*, Eds. G. Burnett, S. Coquillard, R VanLiere & G. Welch, LNCS, Vol 8844, Pp. 133-146, 2015.

[13]   J. J. Cromby, P. J. Standen and D. J. Brown, "The potentials of virtual environments in the education and training of people with learning disabilities," *Journal of Intellectual Disability Research*, Vol. 40, no. 6, pp. 489-501, 1996.

[14]   H. Neale, A. Leonard and S. Kerr, "Exploring the role of virtual environments in the special needs classroom," *4th international conference on disability, virtual reality and associated technologies (ICDVRAT 2002), September 18-20, Veszprém, Hungary. Proceedings*, Sharkey, P., Sik Lányi, C. and Standen, P. J., eds., University of Reading, 2002, pp. 259-266

[15]   W. Kathleen, "TeachLivE is for Lovers," *Autism Moves Out*, 2014.