Virtual Reality Music Intervention to Reduce Social Anxiety in Adolescents Diagnosed with Autism Spectrum Disorder

Ali Adjorlu
Multisensory Experience Lab
Aalborg University Copenhagen
adj@create.aau.dk

Nathaly Belen Betancourt Barriga
University of Trento
nb.betancourtbarriga@studenti.unitn.it

Stefania Serafin
Multisensory Experience Lab
Aalborg University Copenhagen
sts@create.aau.dk

ABSTRACT

This project investigates the potentials of Head-Mounted-Display (HMD) based Virtual Reality (VR) that incorporates musical elements as a tool to perform exposure therapy. This is designed to help adolescents diagnosed with Autism Spectrum Disorder (ASD) to deal with their social anxiety. An application was developed that combines the possibility of singing in VR while a virtual audience provides feedback. A pilot test was conducted on four adolescents diagnosed with ASD from a school for adolescents with special needs in Denmark. All four participants had shown signs of social anxiety according to their teachers. The initial results from this pilot study indicate that despite the participants' were capable of singing in front of the virtual audience without reporting a major level of social anxiety.

1. INTRODUCTION

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder, characterized by deficits in social communication and interaction across multiple contexts [1]. These deficits include difficulties with gestural and verbal communication, keeping eye contact and understanding facial expressions. Additionally, phobias such as social anxiety have been described to be common among individuals diagnosed with ASD since Leo Kanner first described a group of children with autism in 1943 [2, 3].

Social anxiety is described as the feeling that arises in particular situations evoked by the real or imagined concern of being evaluated by others [4]. With a prevalence rate of up to 18 %, social anxiety is one of the most common psychiatric disorders while being ranked among the top 10 chronic disorders that negatively affect the general quality of life [4].

Specifically, social anxiety in children diagnosed with ASD can promote further isolation from interaction with peers, avoidance of social situations [5] and school refusal behavior [6]. These social and interpersonal problems can reduce the chances for independent adulthood. Studies show that a vast majority of adults diagnosed with ASD are unemployed and a majority of them depend on parents or social agencies for support [7]. Furthermore, there is a correlation between anxiety, loneliness, and depression among children diagnosed with ASD [8], which has a negative effect on their general quality of life. The high prevalence of ASD (1 out of 50) further underlines the importance of developing interventions to help children and adolescents diagnosed with ASD to cope with their social anxiety [9].

Incorporating music into interventions for individuals diagnosed with ASD have illustrated benefits such as promoting engagement in social interaction and increasing self-esteem [10, 11]. In the paper "Autism and Music Therapy - is change possible, and why music?" Brown argues that music is intimately interwoven to our structure as human beings and our relationship to others [12]. Even newly born infants move in synchrony with the rhythms of human voice regardless of the language [13]. Since musical elements can establish social interactions even as early as infancy, it makes sense to use these elements to children diagnosed with ASD to overcome their social anxiety.

One of the most established non-pharmaceutical methods to treat anxieties is exposure therapy [14]. By gradually exposing the patient to a stimulus that seeks to provoke anxiety (e.g., musical performance in front of an audience) without the presence of the feared outcome (e.g., being negatively judged by the audience) has proven to help individuals to overcome their phobias.

Virtual Reality (VR) can be used to create a sense of presence in a virtual environment by replacing the real-world sensory information with digitally created audio and visuals. This enables the possibility to develop controllable simulations in which therapeutic interventions such as exposure therapy can take place. Additionally, VR-based exposure therapy does not include some of the practical and logistic issues associated with real life in vivo exposure therapy. As an example, feared stimuli such as performing in front of a live audience might not be easy to access in real life while being difficult to manipulate and control (e.g., managing the audience to avoid an adverse judgment of the user).

One of the first known published studies on VR exposure therapy investigated the effectiveness of virtual environments to help individuals with agoraphobia: fear of crowded

places (see [15] for a review of the topic). Since then, a variety of studies have examined the effectiveness of virtual environments to conduct exposure therapy on individuals diagnosed with phobias such as acrophobia [16], zoophobia [17], and social anxiety [18]. However, there is a shortage of studies investigating the effectiveness of VR exposure therapy to help children and adolescents diagnosed with ASD to deal with their social anxiety [19, 20].

Young individuals diagnosed with ASD have displayed a positive attitude towards head-mounted display (HMD) based virtual reality experiences, showing a high-level of enjoyment and immersion [21,22]. Furthermore, there have been studies illustrating the potentials of VR to teach a variety of everyday living skills to children and adolescents diagnosed with ASD [21] such as money skills [23] and social skills [24].

This study will investigate the potentials of combining the advantages of HMD-based VR exposure therapy with music and its ability to promote social interaction and connections.

2. METHODS

The VR intervention was designed and developed by the authors and evaluated at STUEN, a school for adolescents with mental disabilities in the Rødovre municipality.

2.1 The VR intervention

Traditional exposure therapy involves exposing the client to specific situations that provoke anxiety without the presence of the feared outcome. In this study, a VR application was developed to expose its users to a context within which she or he will have to sing a song in front of a virtual audience. The idea of singing in front of an audience was supported by the teachers at STUEN who informed us that some of their students with ASD had shown signs of social anxiety when they had to present a school topic in front of their peers or participate in music sessions. The application was developed using Autodesk Maya and Unity 3D, designed to run on the Oculus Rift HMD. A 3D virtual concert hall was designed containing a stage with a microphone (see Figure 1). Additionally, the virtual concert hall contained chairs for the virtual audience, a gate from which the audience would enter the hall as well as a screen (see Figure 2) on which the lyrics of the song to be performed would appear. In order to increase tension, a smaller screen was placed on the stage which was used to visualize the count down from 3 to 0 before the user had to start singing in front of the virtual audience (see Figure 1). The user was placed in front of the microphone on the stage facing the empty chairs, in order to build up the tension while providing time to familiarize herself with the virtual environment. A non-diegetic voice will then announce in Danish: "Hello. You are in the concert hall. Are you ready to make some noise? If you are ready, say come in out loud. Once you say come in, the audience will enter the hall.". The word noise is directly translated from the Danish word "larm" which means noise and is a commonly used word in Danish. Windows keyword recognizer is used

to detect the key phrase "come in" via the Oculus Rift microphone. If the user does not say "come in" after 20 seconds, the announcer will say out loud "please say come in once you are ready so we can start." Once the user has said "come in", the virtual audience will walk in through the gate of the concert hall and find their seats. Footstep sounds will further emphasize the presence of the virtual audience in the scene, having the purpose to induce the user with a bit of thrill and anxiety required in an exposure therapy intervention. A non-diegetic voice will once again announce in Danish "The audience is now seated and ready to start. Once you are ready, say start so we can begin the show". Once again, if the user does not say the keyword "start" after 20 seconds, the announcer will ask him to say "start" once more. Saying "start" will initiate a countdown from 3 to 0 which is visually presented to the user via the screen seen in figure 1 accompanied with an earcon signal for each step in the countdown. Once the countdown is over, the song will start playing, and the lyrics will be visible on the screen as seen in figure 2. The music and lyrics to be included in the VR music intervention can be changed before each session by the authors. This is done so that each user can perform their favorite song instead of a song that is chosen for them. A set of eight different virtual audience animations was developed including a variety of facial expressions and body movements (see Figure 3). These animations are triggered according to the users' singing amplitude. If the user is not singing along, the audience will sit still and look bored. If the user is singing, the virtual audience will clap their hands and look happy. If they start singing louder, the virtual audience will stand up and clap their hands. Once the song is over, the non-diegetic announcer will say in Danish "Thank you, that was awesome."

Figure 1. Screenshot from the virtual simulated room without any virtual audience. The user is placed in front of the 3D microphone. The countdown screen can be seen behind the chairs.

2.2 Evaluation

Four students from STUEN Rødovre diagnosed with ASD participated in this study. All participants have shown some characteristics of social anxiety. The Four participants age ranged from 18 to 20 years old and were all male. Each participant was asked if they would like to sing in front of a virtual audience in an HMD based VR application and all four accepted to participate in the study, each signing

Figure 2. The lyrics appeared on a screen behind the virtual audience.

Figure 3. The virtual audience reacting to the performance. Each member of the audience switched between eight different animations depending on the users' singing amplitude

a consent form describing the experiment. They were told that they could stop the VR session at any time. Additionally, each participant was asked which song they would like to sing, which was then added to the VR intervention by the authors before each session.

A simplified version of the Liebowitz Social Anxiety Scale (LSA) [25] was used in order to measure the extent of social anxiety in each of the participants. The LSA was simplified due to the communication difficulties observed in individuals diagnosed with ASD. The simplified version of the LSA contained only 5 items compared to the 24 items in the full version. The 5 items were chosen based on guidelines from the teachers who believed that the students would be able to understand and relate to the chosen questions. Each item in the LSA is designed to assess social anxiety in a variety of situations by asking two questions. The first question asks how anxious the participant usually is in a described situation from a scale ranging from none to mild, moderate and severe. The second question of the item asks how often the participant tries to avoid the described situation on a scale ranging from never to occasionally, often and usually. The five items included in this study are:

1. Acting, performing or giving a talk in front of an audience

2. Meeting strangers

3. Entering a room when others are already seated

4. Taking a exam

5. Looking at people you don't know very well in the eyes

All of the above situations were translated into Danish with simple and descriptive words to help the participant better understand the context. Additionally, each question was accompanied by an image to help the participant interpret the sentences. Furthermore, smiley face Likert scales (Smileyometer) were used when asking how anxious the participant usually is in the described situations as seen in figure 4. Smiley Likert scales have been reported to be effective with children. [26].

None   Mild   Moderate   Severe

Figure 4. Smiley Likert scale for the question on how anxious the participant usually is in a described situation from a scale ranging from none to mild, moderate and severe

Finally, a timeline was designed to illustrate the answers never, occasionally, often and usually as an attempt to help the participant understand the question.

Following the simplified version of the LSA, the students were introduced to the VR application and started their task of performing a song in front of an interactive virtual audience. Data were gathered during each VR session via screen recordings of the participants' performance in the virtual environment. One author remained in the room observing the participants behavior during each session.

After the VR session, the participants were asked four questions from the Witmer Singer presence questionnaire (PQ) and four questions from their Immersive Tendency Questionnaire (ITQ) [27]. This was done in order to explore whether the participants were sufficiently immersed in the virtual environment for it to be effective s an exposure therapy tool. Once again the questions were chosen based on the guidelines from the teachers and translated to Danish using simple descriptive words. The four chosen questions from the ITQ questionnaire were:

1. Do you easily become deeply involved in movies or TV dramas?

2. Do you ever become so involved in a movie that you are not aware of things happening around you?

3. How often do you play arcade or video games?

4. Do you ever become so involved in doing something that you lose all track of time?

The four questions chosen from the PQ questionnaire were:

1. How responsive was the environment to the actions that you initiated?

2. How natural did your interactions with the environment seem?

3. How involved were you in the virtual environment experience?

4. How much did the auditory aspects of the environment involve you?

The above questions on the simplified versions of the PQ and ITQ are to be answered via a 7-point Likert scale, once again communicated to the user via a smiley face Likert scale as seen in figure 5 . Exposure therapy interventions
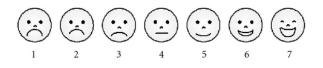

Figure 5. 7-point Smiley Likert scale for the simplified versions of the PQ and ITQ questions

expose their users gradually to a stimulus that seek to provoke the targeted anxiety without the presence of the feared outcome. Therefore, the participants were asked to rate how scary it was to sing in front of the virtual audience on a smiley Likert scale ranging from 0 to 4. In addition to the four participants, one of the teachers participated in a short unstructured interview, giving his input on the VR application as a tool to perform exposure therapy on his students. We also discussed the findings of our survey with the teacher to confirm the validity of the participants' answers.

3. RESULTS

All four participants tried out the VR intervention. However, P4 was not capable of understanding the questions on any of the surveys, despite them being simplified by the authors and the teacher. Therefore, only the observation data from him is presented in this paper.

3.1 Liebowitz Social Anxiety Scale

The results of the simplified version of the Liebowitz Social Anxiety scale can be seen in table 1 and 2. P1 reported no fears at all from all five scenarios described to him. Furthermore, he stated that he never tried to avoid these situations. According to the teacher, this participant tries to avoid having to present in front of his classmates. P2 reported moderate fear from having to perform in front of others while he reported that he occasionally tries to avoid this kind of situations. He also stated a moderate fear from having eye contact with strangers, but he never tries to avoid it. P2 answered 'none' to the question of how much he fears meeting strangers, however, he chose occasionally when asked how often he tries to avoid situations where he has to meet new people. Entering crowded rooms is reported to be mildly feared by P2, but he indicated that he never tries to avoid having to enter them. P3 answered 'none' to how much he feared all five situations.

He also reported 'never' on how often he tries to avoid the situations in the simplified version of the LSA, except for exams which he avoids occasionally.

Participant	P1	P2	P3
Performing	None	Moderate	None
Meeting Strangers	None	None	None
Crowded rooms	None	Mild	None
Taking a exam	None	Severe	None
Eye contact	None	Moderate	None

Table 1. Responses from the simplified version of the Immersion Tendency Questionnaire asking how anxious or fearful the participant feels in different situations ranging from none to mild, moderate and severe.

Participant	P1	P2	P3
Performing	Never	Occasionally	Never
Meeting Strangers	Never	Occasionally	Never
Crowded rooms	Never	Never	Never
Exams	Never	Usually	Occasionally
Eye contact	Never	Never	Never

Table 2. Responses from the simplified version of the Immersion Tendency Questionnaire asking how often the participant avoids the situation ranging from never to occasionally, often and usually

3.2 Observations during VR sessions

Each VR session lasted approximately ten minutes. P1, P3 and P4 chose to sing a Danish song called "Papirsklip" by a Danish artist called Kim Larsen. P2 chose the song "Born to be yours" by Imagine Dragons.

As previously mentioned, there were two vocal commands: "come in" which activated the animation in which the virtual audience would enter the concert hall and "start" which would initiate a countdown followed by the music starting. The users were asked to pronounce the voice commands by an announcer in the VR music intervention. P1, P2 and P3 had no problems pronouncing the commands. P4 struggled to understand the announcer and remained quiet after being asked to say "come in". The VR application was programmed to ask the user to say come in once again if the correct voice command was not detected. However, P4 remained quiet and looked shy while wearing the HMD. This resulted in the observer asking the user to say come in which was followed by P4 saying you may enter instead. Therefore, the observer had to say the voice command himself which activated the virtual audience entering the concert hall animation. P4 also struggled with the second voice command resulting in the observer once again having to activate it by saying out loud start. Once the Kim Larsen song started playing, P4 started singing along, struggling from time to time to pronounce the lyrics correctly, humming most of the song instead.

P1, P2 and P3 showed confidence when singing, looking to enjoy the experience while having no problems saying the voice commands. P2 hummed most of the song.

3.3 simplified version of the Immersive Tendency Questionnaire (ITQ)

The results from the simplified version of the ITQ can be seen in table 3. P1 and P2 reported a high level of immersion tendency. P3 reported that he is never so immersed in a movie that results in him not being aware of his surroundings. Furthermore, he declared that he never loses track of time when performing any activity. Questions regarding how often he plays video games and how easily he becomes deeply involved in a movie was answered with a 4 which is the midpoint on the Likert scale from 0 (never) to 7 (often).

Participant	P1	P2	P3
Easily Involved?	7	7	4
Aware of surrounding?	5	7	1
Video games	6	5	4
Loose track of time?	6	7	1

Table 3. Responses from the Immersive Tendency Questionnaire. Answer options ranged from 0 to 7.

3.4 Presence Questionnaire

The results from the simplified version of the PQ can be seen in table 4. All three participants reported that the VR environment was completely responsive to their actions. P1 and P3 both stated that the interaction with the VR environment seemed completely natural by choosing 7 on the smiley face Likert scale. P2 picked 6 which is just below the highest possible choice. P1 also chose the lowest

Participant	P1	P2	P3
Responsive environment?	7	7	7
Natural interacetion?	7	6	7
Involvement?	6	4	7
Audio?	5	5	7

Table 4. Responses from the presence questionnaire. Answer options ranged from 0 to 7.

smiley face Likert scale out of the three participants when asked about how involved he was with the virtual environment. P1 chose 6 while P3 once again chose 7. P3 also chose 7 on whether the audio in the VR environment increased his involvement with the application. P1 and P2 both chose 5.

3.5 Level of anxiety

The results of the participants' answers to whether it was scary to sing in front of a virtual audience can be seen in table 5. Only P2 reported some level of anxiety. He was also the only participant to reported fearing situations where he has to perform in front of other people.

Participant	P1	P2	P3
How scary was the experience?	0	1	0

Table 5. Responses from the anxiety questionnaire. Answer options ranged from 0 to 4.

3.6 Teacher's comments

At the end, one of the teachers working with the participants on a daily basis tried out the VR intervention followed by a short interview with the authors. The teacher mentioned that he believed that singing and music increases what he called the fun factor of exposure therapy. However, he said that the participants he provided us with would never sing in front of other people, referring to music sessions they have had at the school. During those sessions, participants in this study often stayed away. He believed that VR might have given the illusion of being alone to his students, resulting in them singing fearlessly. When presented with the results from the Liebowitz Social Anxiety Scale, the teacher stated that his students do not want to seem scared, resulting in them answering that they do not fear the described situations. He added that P2 is probably the one with the highest level of social anxiety together with P4.

4. DISCUSSION AND CONCLUSION

This explorative study aimed to investigate the potentials of singing in VR as a tool to help children diagnosed with ASD to cope with their social anxieties. More specifically, the project investigates whether singing in front of a reactive virtual audience is a sufficiently immersive experience, which is one of the main requirements for exposure therapeutic interventions.

A cartoonish virtual concert hall was designed to be appealing to the user while avoiding the uncanny valley. The uncanny valley is a concept that describes the observers' revulsion towards humanoid objects or 3D models that appear nearly human [28]. Additionally, the virtual audience was programmed to react positively to the users' singing via a variety of animations such as clapping their hands and standing up while looking happy to motivate the users to sing. In contrast, if the user did not sing along, the virtual audience would remind seated, looking sad and bored. Instead of pressing a button on the VR controller to start the experience, the authors implemented audio commands to further expose the user to a situation where they had to speak in front of an audiance. Finally, the participants got the option to chose the song they wanted to sing to increase their motivation.

Level of social anxiety in the participants was evaluated before each VR session using the simplified version of the LSA questionnaire, while their level of presence in the VR environment was measured using the simplified version of the ITQ and the PQ questionnaire, both developed by Witmer & Singer [27] . These questionnaires were translated to Danish together with the teachers and simplified using easy to understand words as well as illustrations visual-

ising each situation. However, P4 still had problems understanding the questions and was not capable of providing any answers. One of the primary deficits in individuals diagnosed with ASD is reduced communication capability. According to the teachers, P4 is on the low functioning side of the autism spectrum which results in him having a hard time communicating. Despite these communicative deficits, he still sang and hummed during the VR session, activating positive feedback from the virtual audience. This behavior was also observed in P1, P2 and P3 who sang or hummed during the VR session, activating the virtual audience' positive feedback. The singing behavior of the participants was surprising to the teacher, who did not expect them to behave the way they did. However, according to the results from the simplified version of the LSA questionnaire, P1 and P3 reported having no stage fear, making their performance on the virtual stage less surprising. The teacher's comment on this result is that they did not want to be perceived negatively by the authors by stating that they were scared of certain situations. In general, the results from the simplified version of the LSA does not correlate with the comments from the teacher who stated that his students suffer from social anxiety. Therefore, the validity of the simplified version of the LSA as well as the PQ, and ITS questionnaires is debatable. In future iterations, a qualified expert clinician or behavioral psychologists should be involved in the design of the experiment, helping to gather more valid data from this target group. Additionally, to gain further information on the social characteristics of the participants, methods such as the Social Responsiveness Scale could be used [29].

In the post VR session questionnaires, P2 rated 4 on his self reported involvement in the VR experience, which was the lowest of all three participants. This low level of involvement with the VR music intervention can explain why he was singing along despite his moderate fear of performing in front of an audience (according to his teacher). He also rated 6 on whether the interaction in the VR environment was natural, which is also the lowest rating from the three participants. He was the one who hummed to most of the song compare to the other participants who sang the words which can be explained by the fact that he was the only one who chose an English song to sing. P2 was not asked any question on whether he was able to speak English. In future iterations, only danish songs should be included in the VR application. The teacher suggested the songs included in this version of the application after he had asked the participants which songs they would like to sing in VR.

P3 choose the smiley face Likert item number 7, the highest possible on all four questions on the simplified version of the PQ survey. This is despite him reporting the lowest immersive tendency out of the three participants who answer the surveys. During the singing session, he was the one who seemed to know most of the lyrics of the Kim Larsen song, and also seemed to enjoy himself. Therefore, the music could explain his high rating in the simplified version of the PQ survey.

P1 reported being easily involved in movies and loses tracks of time when performing certain activities. He also reported a high level of presence in the VR music intervention, which was also observed during his singing session. He sang the lyrics, putting effort into trying to read the lyrics and pronouncing them correctly.

P4 was not capable of answering neither the simplified versions of the LSA, the ITQ or the PQ surveys, even though they were simplified and had images explaining their content. The teacher explained that P4 has the lowest communication skills of all the participants. He also had problems understanding the announcer in the VR experiment which asked him to repeat the voice commands "come in" and "start.". However, when the music started playing, he did not hold back and started to sing along, correctly pronouncing the lyrics of the Danish song he chose. Music might have reduced the social barriers he experienced since he looked relaxed during the VR singing session. Future iteration of this study could involve the teacher gathering data from his students about their experience with the VR music intervention. The students feel more comfortable with their teachers. This might result in them providing more relevant and valuable insight about their experience with the VR intervention. Using the simplified version of the LSA, PQ, ITQ was an attempt to not overwhelmed the users with a lot of questions. However, doing this we have reduced the validity of these methods.

Despite of this, we believe that the explorative study showed that a VR music intervention could be an immersive experience for adolescents diagnosed with ASD and social anxiety. As Brown argued, music seemed to remove barriers for P4 who started singing along the song moments after he was timid and unable to communicate with the observer [12]. In future iterations of the study, information on the participants' reaction to the feedback provided by the virtual audience must be collected to shed some light on its effectiveness to increase the users' immersion.

Exposure therapy is defined by gradually exposing the user to a stimulus that provokes anxiety without the presence of the feared outcome. Out of the three participants who answered questions after the study, only P2 reported experiencing any form of anxiety. Future iterations of the application should focus on creating a more tense experience to provoke more social anxiety in its users by adding new levels in the VR application. The first level will keep the cartoonish background and 3D virtual audience. The next level will replace the environment with a real classroom or concert hall recorded via a 360 camera while keeping the virtual audience. The third level will only consist of 360 video footage of real audience and environment recorded with 360 camera, making it as realistic as possible.

Acknowledgments

5. REFERENCES

[1] A. P. Association *et al.*, *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub, 2013.

[2] A. Gillott, F. Furniss, and A. Walter, "Anxiety in high-functioning children with autism," *Autism*, vol. 5, no. 3, pp. 277–286, 2001.

[3] D. Spain, J. Sin, L. Harwood, M. A. Mendez, and F. Happé, "Cognitive behaviour therapy for social anxiety in autism spectrum disorder: a systematic review," *Advances in Autism*, vol. 3, no. 1, pp. 34–46, 2017.

[4] M. B. Stein and D. J. Stein, "Social anxiety disorder," *The lancet*, vol. 371, no. 9618, pp. 1115–1125, 2008.

[5] B. S. Myles, G. P. Barnhill, T. Hagiwara, D. E. Griswold, and R. L. Simpson, "A synthesis of studies on the intellectual, academic, social/emotional and sensory characteristics of children and youth with asperger syndrome," *Education and Training in Mental Retardation and Developmental Disabilities*, pp. 304–311, 2001.

[6] E. K. Munkhaugen, E. Gjevik, A. H. Pripp, E. Sponheim, and T. H. Diseth, "School refusal behaviour: Are children and adolescents with autism spectrum disorder at a higher risk?" *Research in Autism Spectrum Disorders*, vol. 41, pp. 31–38, 2017.

[7] D. Hendricks, "Employment and adults with autism spectrum disorders: Challenges and strategies for success," *Journal of Vocational Rehabilitation*, vol. 32, no. 2, pp. 125–134, 2010.

[8] M. O. Mazurek, "Loneliness, friendship, and well-being in adults with autism spectrum disorders," *Autism*, vol. 18, no. 3, pp. 223–232, 2014.

[9] G. Xu, L. Strathearn, B. Liu, and W. Bao, "Prevalence of autism spectrum disorder among us children and adolescents, 2014-2016," *Jama*, vol. 319, no. 1, pp. 81–82, 2018.

[10] S. M. Shore, "The language of music: Working with children on the autism spectrum," *Journal of Education*, vol. 183, no. 2, pp. 113–124, 2002.

[11] D. Wimpory, P. Chadwick, and S. Nash, "Brief report: Musical interaction therapy for children with autism: An evaluative case study with two-year follow-up," *Journal of autism and developmental disorders*, vol. 25, no. 5, pp. 541–552, 1995.

[12] S. M. Brown, "Autism and music therapyis change possible, and why music?" *Journal of British music therapy*, vol. 8, no. 1, pp. 15–25, 1994.

[13] W. S. Condon and L. W. Sander, "Neonate movement is synchronized with adult speech: Interactional participation and language acquisition," *Science*, vol. 183, no. 4120, pp. 99–101, 1974.

[14] A. M. Reid, M. I. Bolshakova, A. G. Guzick, A. G. Fernandez, C. W. Striley, G. R. Geffken, and J. P. McNamara, "Common barriers to the dissemination of exposure therapy for youth with anxiety disorders," *Community mental health journal*, vol. 53, no. 4, pp. 432–437, 2017.

[15] A. Adjorlu, "Virtual reality therapy," in *Encyclopedia of Computer Graphics and Games (ecgg)*. Springer, 2018.

[16] P. M. Emmelkamp, M. Krijn, A. Hulsbosch, S. De Vries, M. J. Schuemie, and C. A. van der Mast, "Virtual reality treatment versus exposure in vivo: a comparative evaluation in acrophobia," *Behaviour research and therapy*, vol. 40, no. 5, pp. 509–516, 2002.

[17] C. Botella, M. Á. Pérez-Ara, J. Bretón-López, S. Quero, A. García-Palacios, and R. M. Baños, "In vivo versus augmented reality exposure in the treatment of small animal phobia: a randomized controlled trial," *PloS one*, vol. 11, no. 2, p. e0148237, 2016.

[18] E. Carl, A. T. Stein, A. Levihn-Coon, J. R. Pogue, B. Rothbaum, P. Emmelkamp, G. J. Asmundson, P. Carlbring, and M. B. Powers, "Virtual reality exposure therapy for anxiety and related disorders: A meta-analysis of randomized controlled trials," *Journal of anxiety disorders*, vol. 61, pp. 27–36, 2019.

[19] D. E. Parrish, H. K. Oxhandler, J. F. Duron, P. Swank, and P. Bordnick, "Feasibility of virtual reality environments for adolescent social anxiety disorder," *Research on Social Work Practice*, vol. 26, no. 7, pp. 825–835, 2016.

[20] R. Bradley and N. Newbutt, "Autism and virtual reality head-mounted displays: a state of the art systematic review," *Journal of Enabling Technologies*, vol. 12, no. 3, pp. 101–113, 2018.

[21] N. Newbutt, C. Sung, H.-J. Kuo, M. J. Leahy, C.-C. Lin, and B. Tong, "Brief report: A pilot study of the use of a virtual reality headset in autism populations," *Journal of autism and developmental disorders*, vol. 46, no. 9, pp. 3166–3176, 2016.

[22] D. Strickland, L. M. Marcus, G. B. Mesibov, and K. Hogan, "Brief report: Two case studies using virtual reality as a learning tool for autistic children," *Journal of Autism and Developmental Disorders*, vol. 26, no. 6, pp. 651–659, 1996.

[23] A. Adjorlu and S. Serafin, "Head-mounted display-based virtual reality as a tool to teach money skills to adolescents diagnosed with autism spectrum disorder," in *Interactivity, Game Creation, Design, Learning, and Innovation*. Springer, 2018, pp. 450–461.

[24] A. Adjorlu, A. Hussain, C. Mødekjær, and N. Warming Austad, "Head-mounted display-based virtual reality social story as a tool to teach social skills to children diagnosed with autism spectrum disorder," in

2017 IEEE Virtual Reality Workshop on K-12 Embodied Learning through Virtual & Augmented Reality (KELVAR). IEEE, 2018.

[25] R. G. Heimberg, K. Horner, H. Juster, S. Safren, E. Brown, F. Schneier, and M. Liebowitz, "Psychometric properties of the liebowitz social anxiety scale," *Psychological medicine*, vol. 29, no. 1, pp. 199–212, 1999.

[26] J. C. Read, S. MacFarlane, and C. Casey, "Endurability, engagement and expectations: Measuring childrens fun," in *Interaction design and children*, vol. 2. Shaker Publishing Eindhoven, 2002, pp. 1–23.

[27] B. G. Witmer and M. J. Singer, "Measuring presence in virtual environments: A presence questionnaire," *Presence*, vol. 7, no. 3, pp. 225–240, 1998.

[28] M. Mori, K. F. MacDorman, and N. Kageki, "The uncanny valley [from the field]," *IEEE Robotics & Automation Magazine*, vol. 19, no. 2, pp. 98–100, 2012.

[29] J. N. Constantino, S. A. Davis, R. D. Todd, M. K. Schindler, M. M. Gross, S. L. Brophy, L. M. Metzger, C. S. Shoushtari, R. Splinter, and W. Reich, "Validation of a brief quantitative measure of autistic traits: comparison of the social responsiveness scale with the autism diagnostic interview-revised," *Journal of autism and developmental disorders*, vol. 33, no. 4, pp. 427–433, 2003.