

A Virtual Reality Training Application for Adults With Asperger's Syndrome

Diego Rojo

Centro Universitario de Tecnología y Arte Digital

Jesús Mayor

Centro Universitario de Tecnología y Arte Digital

Editor: Mike Potel

potel@wildcrest.com

José Jesús García Rueda

Centro Universitario de Tecnología y Arte Digital

Laura Raya

Centro Universitario de Tecnología y Arte Digital

Abstract—Asperger's syndrome is a disorder that involves a qualitative impairment in social interactions. While most treatments are aimed at children or adolescents, in this paper we present the development of a virtual reality training application in which adults with Asperger's syndrome can train in an autonomous and controlled way how to present in public.

■ **PRESENTATION AND COMMUNICATION** skills are increasingly important in professional and academic environments, e.g., making a public presentation of a student's final dissertation is a requirement in order to graduate in the European higher education area. Acquiring these skills is difficult and requires specialized training, as shown by the many courses aimed at helping people tackle their difficulties with presenting in public or even overcoming phobia of public speaking. For those with Asperger's syndrome (AS), being able to make public presentations becomes even a greater challenge and can

be a determining factor in their academic or career progress.

The use of virtual reality (VR) as a tool for rehabilitation of different types of phobias and disorders has been studied since the 1990s, proving its effectiveness in overcoming different phobias such as acrophobia or arachnophobia. However, the use of technology to improve the social skills of Asperger's syndrome has mainly focused on research and development of three-dimensional (3-D) virtual environments,^{1,2} but without using VR head-mounted displays (HMDs).

The report by Kothgassner *et al.*³ supports the hypothesis that the physiological reaction to speaking in front of an audience is similar, whether in a real-life environment or in a virtual reality environment using an HMD. This suggests

Digital Object Identifier 10.1109/MCG.2018.2884272

Date of current version 22 March 2019.

that the use of VR is a good approach to overcome a phobia of public speaking.

To this end, we have developed CicerOn VR: Virtual Speech Coach, a project of the Research Chair in Accessible Technologies created by Indra Sistemas S.A., Fundación Universia, and Centro Universitario de Tecnología y Arte Digital. The project's objective is to support people with Asperger's syndrome in the process of overcoming their difficulties when public speaking, by designing and developing an immersive virtual reality serious game. In this environment, users can train in a gradual and controlled way how to interact with others and present in public.

ASPERGER'S SYNDROME

Asperger's syndrome is an Autism Spectrum Disorder characterized by qualitative impairment in social interaction. Individuals have no clinically significant general delay in language, but show restricted, repetitive, and stereotyped patterns of behavior, interests, and activities.⁴ All this leads to a deterioration of the individual's social, academic, and work activity.

Most studies of treatments to improve symptoms of AS are aimed at children and adolescents because, according to scientific evidence, an intensive early intervention (between 0 and 6 years) can modify the poor prognosis generally associated with children with AS.^{5,6} It should be noted that most of these treatments do not make use of digital technologies, although in recent years interest in technology has increased due to promising results from relevant research projects.⁷

Despite this, Fuentes-Biggi *et al.*⁶ emphasize the need to continue supporting people with AS into their adult years, following a continuing education plan. The development of CicerOn is focused precisely on serving as a support tool during this adult stage, helping college students and adults with AS to improve their communication skills.

APPLICATION DESIGN AND DEVELOPMENT

CicerOn is a serious game that allows the user to talk to different avatars in different virtual environments. It is an experience that supports a virtual reality exposure therapy in a

user-friendly way due to its gamified design. The different levels of the game provide a hierarchy of increasing exposure to the phobia of public speaking, encouraging motivation to continue the training. An automatic speech recognition system evaluates the speaking of the users, showing feedback on their tasks. This allows for a gradual improvement of their speaking ability.

The application is aimed at users with Asperger's syndrome aged 15 and over. People can use it without supervision, at any time and without having to go to a specialized center. The strong feeling of presence when using VR with an HMD allows users to more easily transfer what they learn within the virtual world to the real world.

Hardware

We evaluated different alternative virtual reality platforms to determine the most appropriate one for our purposes. Since the aim of the project was to expose the user to a systematic desensitization of public speaking, the preferred head-mounted display should have a microphone and a pair of headphones. In the end, we decided to use mobile virtual reality platforms, as they allow us to reach the largest number of users thanks to their lower price and ease of use.

Although we have implemented a multiplatform application, user testing has been done with Samsung Gear VR from Oculus (see Figure 1). In addition, the users have a controller that allows them to move and interact with the virtual environment in an intuitive way, using the joystick for navigation and the buttons for interaction with the objects.

Application Design

CicerOn is a serious game whose goal is to make users speak aloud before different audiences.

During the design process of CicerOn, we carried out several meetings with both psychologists and people with AS. The first meetings were aimed at eliciting the main features to be included when designing a virtual reality application for the target users. Regarding interests and motivations, these include puzzle and riddles, which in the end became the main mechanisms implemented in CicerOn. Other particularities to be taken into account are the typical AS user's



Figure 1. Samsung Gear VR head-mounted display and a Bluetooth controller used for testing.

need for detailed instructions, lack of empathy, inability to understand irony and sarcasm, and a literal understanding of messages and propositions. These characteristics vary greatly from one individual to another, but they are widely present in this group.

This application is structured in six levels. Each level is associated with a higher level of exposure to the audience, implementing an incremental exposure hierarchy. From an initial level of exposure to phobia (public speaking on a stage but without spectators), the complexity of the phobia exposure increases level after level. The increasing rate of difficulty, in terms of exposure to the phobic stimulus, has been designed by psychologists specializing in patients with Asperger.

CicerOn's storyline is based on Egyptian mythology. Figuring out riddles and puzzles, users travel through different countries to find a lost object as if they were explorers. To complete each level, users must read aloud a final text. These texts contribute to the narrative, unveiling some key elements of the game's story.

The mechanics and structure of the game are designed to be repetitive, in order to foster the creation of a routine in players. Sudden changes of scenarios without preview explication are avoided. Routines are made explicit to the players by means of checklists and transition videos. The length of these videos and their appearance are always the same.

Users' main objective in each level is to find all the fragments of a ripped postcard (see Figure 2),

hidden all over the level. Each fragment contains a riddle that provides the user with a clue for finding the next one so that the user can finally complete the postcard. The last fragment will take the user near the final reading area, such as a stage or a lectern. Once all the fragments have been found, the complete postcard offers users a longer text to read. The user must read the text of the postcard aloud. The final text contains information on how the game story progresses and what to do in the next level. This text is divided into several blocks or paragraphs. After each block, users will receive the first feedback on their performance. We group



Figure 2. Fragments that compose ripped postcards containing individual clues which when reassembled lead the user to a final text to be read. The aesthetics of the postcards are closely related to the specific city where the action of the game is taking place.



Figure 3. Images of the comfortable base camp for the users with Asperger's syndrome to rest after completing each individual level and see the progress of the game.

the performance scores into three categories (improvable, quite good, and fantastic). Negative messages have been removed from our application and visual feedback using icons complements quantitative information. Once all the blocks have been read, an average score is displayed. To facilitate the reading flow, our application underlines the text while the user is reading.

The automatic speech recognition system included in our application evaluates the user's final reading aloud. It scores the reading according to the volume and correct pronunciation of the words. Several reports show feedback on the user's progress. If players pass a minimum score, the system will move on to the next level.

We designed an additional stage (a lounge) with a familiar and comfortable look to be used as a base camp. It is the user's safe place in the virtual world to go when they feel overwhelmed. With this additional environment, we want to prevent the user from removing the VR device under conditions of over-stress due to excessive exposure to the phobic stimulus. In addition, each time users complete a level, our application takes them to this base camp, where they can also check their progress (see Figure 3).

The different levels that make up the application are detailed below. Figure 4 shows some of the scenarios associated with these levels.

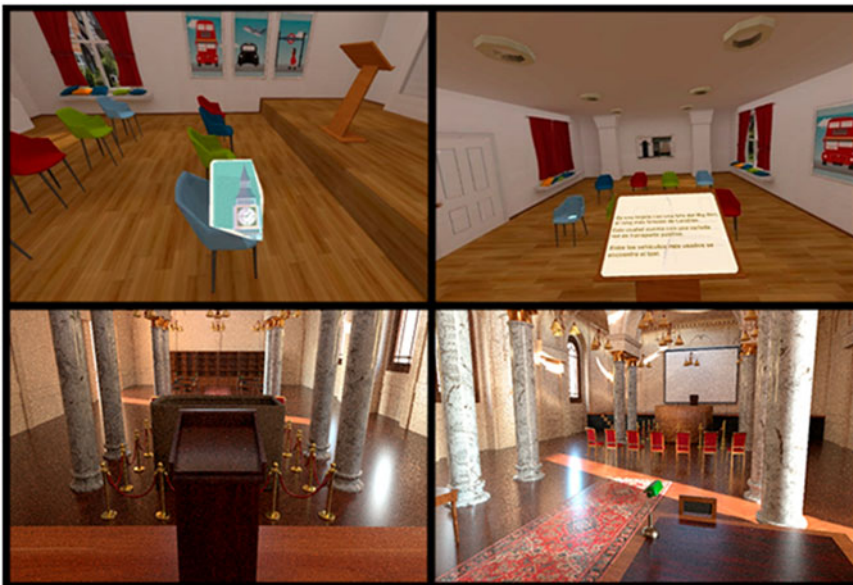


Figure 4. Images of scenarios used in the application: a school auditorium in London (upper left and upper right) and a lecture hall in the Louvre museum in Paris (lower left and lower right).

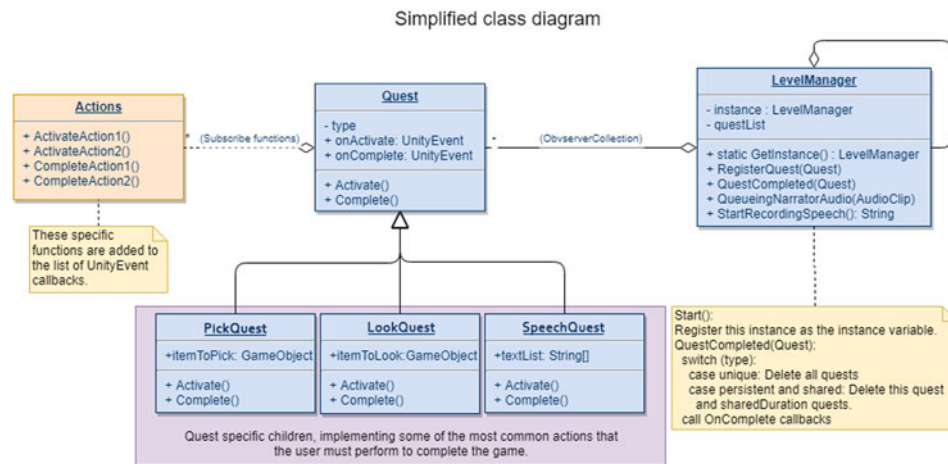


Figure 5. Object-oriented design approach facilitates the creation and reuse of multiple treatment scenarios.

- **Level 0.** Takes players to the base camp in Madrid, where the story is introduced, and the basic navigation and interaction mechanics are explained.
- **Level 1:** Located in a school auditorium in London. The initial final text to be read aloud by users explains the mythological background of the game and the goal to be achieved. The place is empty and there is no audience. A new character is introduced, who will continuously guide the user through the training.
- **Level 2.** A transition video carries the user to Paris. Once all the postcard's fragments have been found, users must read the final text in front of a nonhuman cartoon character.
- **Level 3.** Again, a transition video takes the user back to Madrid. In this case, users have to read the final text in front of three people who are not paying attention to them.
- **Level 4.** Users travel to Alexandria this time. As in the other levels, users have to discover the place where each fragment of the postcard is hidden. At each level, the number of pieces that make up the postcard is different (see Figure 2). In this case, the player makes the final reading before a group of 10 people, who are visibly, but not constantly, paying attention to the user.
- **Level 5:** The final level takes place in the city of Bubastis, where players find the lost object and the game comes to an end. Up to now, the

software has increased the difficulty of the training process by modifying the size of the audience and their attitude towards users. However, the highest level of exposure to the phobic stimuli for people with Asperger's syndrome reported by psychologists is reading before a single person, who is paying all their attention to the users. In this level, users become the entire focus of attention of the spectator.

Technical Design

CicerOn is developed with Unity 2017 (unity3d.com), which provides us with a solid game engine foundation. Maya 2017 (www.autodesk.com/products/maya) is used for modeling and animating virtual environments.

Due to the repetitive narrative-action loop of the game, we have generated a system of quests that allows level designers to easily create new levels. This system is implemented by a state machine that will control the execution flow of each scene within the application. The most significant design pattern under this structure is a Singleton programming design pattern, which provides a static referencing of the script using a class called *LevelManager*. To facilitate development, there are several subclasses of class *Quest* that use an Observer design pattern, so the state machine builds itself inside the *LevelManager*. In Figure 5, we provide a class diagram that explains the basic architecture of our application.

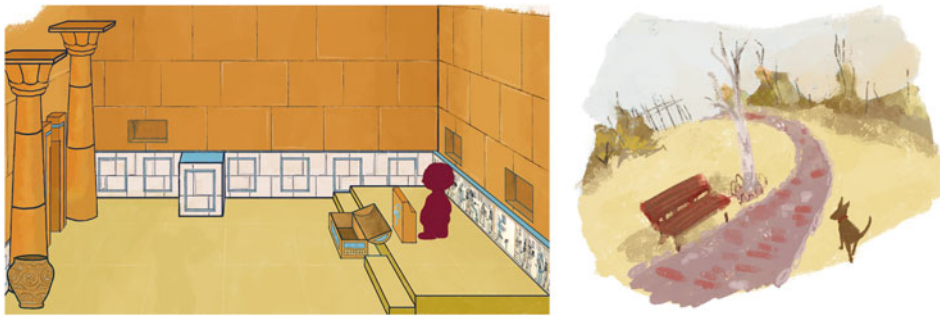


Figure 6. Sketches validated as acceptable to AS patients before construction of the 3-D models.

The developer or level designer, once they include the *Quest* script on the scene, can easily add functionality thanks to the callbacks *ActivateQuest* and *CompleteQuest*, which manage the actions that take place in any given *Quest*. This way, using the Unity editor, designers can easily add new behaviors. When several quests are available at a time, the *LevelManager* is responsible for the activation of quests according to the properties of the latter (e.g., persistence, uniqueness, and shared duration).

The *LevelManager* class also manages narrative resources and speech recognition, which can only be used sequentially. Thus, when a *Quest* asks for the playing of a piece of audio, it is queued. And if users perform an action that requires speech recognition, our easy-to-access Singleton will handle this request.

In the process of integrating an automatic speech recognition system, we encountered a major issue. At first, we tried to use an off-line recognizer called Pocket Sphinx (cmusphinx.github.io), to avoid the need for an Internet connection. Unfortunately, the quality of recognition was not good enough and caused confusion among users. The lowest results were obtained when used by people with high-frequency voices, such as women.

Due to these problems, we decided to use another Android-compatible recognizer called Wit.ai (wit.ai), which provides free speech recognition through its online services. This recognizer provided us with recognition accurate enough for the needs of the application.

Finally, good performance in VR applications is essential to prevent cybersickness in users. A requirement of at least 60 frames per seconds (fps) was determined to be necessary to obtain a

smooth motion within the virtual world. Optimizing the number of polygons and the texturing of 3-D models has also been a challenge for CicerOn. For example, we have reduced to 29 856 the number of polygons of level 1, with only 29 draw calls. We achieve a realistic and detailed appearance (see Figures 3 and 4) with a continuous performance above 100 fps due to the different optimization techniques and shaders used.

Perceptual Adapted Design

Adapting functional and visual design to students with AS is essential for better use of the experience. But most of the existing applications have a children's appearance and content, which leads to rejection by adult users with AS. Our application has a cartoon or caricatured appearance but not infantilized look. Avoiding saturated and bright colors, strident shapes or overwhelming environments have been some of the guidelines followed by artists. The sketches and models used in CicerOn have been validated both by people with AS and psychologists (see Figure 6).

When designing game interactions, it is necessary to consider the possible repetitive motor mannerisms of the users. Simple interactions have been designed to improve usability and reduce the learning curve. In order to increase the users' reading-aloud time, in the first design of the application, the interactions with the objects in the game were carried out through voice. According to the first tests conducted with real users, the gameplay was perceived as boring, lowering the users' motivation, what led to a redesign of the application. This new design provides a more dynamic and intuitive interaction through a gamepad.



Figure 7. VR headsets (Samsung Gear VR) enable users to move about freely and practice realistic movements in different scenarios.

Regarding level design, the increase of difficulty during the gameplay is defined by the number and nature of spectators, ranging from none to many. Some variability has to be taken into account here, as most people with AS perceive one-on-one reading as more stressful than reading before several people. Therefore, we have defined the one-to-one reading activity as the last challenge in the game.

EVALUATION

An iterative design evaluation process has been at the heart of CicerOn's development from the very beginning. Following a "with them and for them" philosophy, some members of the development team themselves have Asperger's syndrome. They have contributed to the testing of the experience, looking for performance improvements, locating bugs and proposing adaptations in functionality, script, and appearance to make it more adapted to people with Asperger's syndrome.

In addition, a qualitative evaluation has been conducted by psychologists, pedagogues, teachers, adults with Asperger's syndrome, and their families in each development cycle (see Figure 7). As a result, modifications in the narrative, in the number of levels or in their temporal arrangement and other extra functionalities have been made.

During the design and implementation process, we worked hand-in-hand with different associations specialized in Asperger's syndrome.

Among them are the PAUTA Association (Psychopedagogy of Autism and Associated Disorders) and the Asperger Madrid Association. They have also tested the application, giving us very positive and encouraging feedback regarding both the software and its potential to support phobia treatment therapies in people with Asperger's syndrome. The use of VR lowers the entry barriers for users to begin the therapy. The selected VR device (see Figure 1) is easy to handle and is not heavy or uncomfortable. The application's playful design motivated patients and caregivers to continue with the treatment. Patients claimed that, without noticing, they found themselves speaking from a stage.

After the development of CicerOn, Indra Sistemas S.A. and Fundación Universia are confident in the power of virtual reality applications to help overcome other phobias of people with Autism Spectrum Disorder. A video of our system in use is available at youtu.be/zFokOkoOtUM.

REFERENCES

1. J. A. Ehrlich and J. R. Miller, "A virtual environment for teaching social skills: AViSSS," *IEEE Comput. Graph. Appl.*, vol. 29, no. 4, pp. 10–16, Jul./Aug. 2009.
2. M. R. Kandalaft, N. Didehbani, D. C. Krawczyk, T. T. Allen, and S. B. Chapman, "Virtual reality social cognition training for young adults with high-functioning autism," *J. Autism Dev. Disorders*, vol. 43, no. 1, pp. 34–44, 2013.

3. O. D. Kothgassner *et al.*, "Salivary cortisol and cardiovascular reactivity to a public speaking task in a virtual and real-life environment," *Comput. Hum. Behav.*, vol. 62, pp. 124–135, 2016.
4. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, (4th ed., text rev.). Washington, DC, USA: American Psychological Assoc., 2000.
5. F. Mulas, G. Ros-Cervera, M. G. Millá, M. C. Etchepareborda, L. Abad, and M. Téllez de Meneses, "Modelos de intervención en niños con autismo," *Rev. Neurol.*, vol. 50, no. 3, pp. 77–84, 2010.
6. J. Fuentes-Biggi *et al.*, "Guía de buena práctica para el tratamiento de los trastornos del espectro autista," *Rev. Neurol.*, vol. 43, no. 7, pp. 425–38, 2006.
7. A. L. Wainer and B. R. Ingersoll, "The use of innovative computer technology for teaching social communication to individuals with autism spectrum disorders," *Res. Autism Spectrum Disorders*, vol. 5, no. 1, pp. 96–107, 2011.

Diego Rojo is a Full-Time Lecturer with Centro Universitario de Tecnología y Arte Digital (U-tad), Madrid, Spain. His research interests include data visualization, human–computer interaction, and machine learning. He received the Master's degree in computer graphics and simulation from U-tad. Contact him at diego.rojo@acm.org.

Jesús Mayor is currently working toward the Ph.D. degree in computer science at Universidad Rey Juan Carlos, Madrid, Spain, with an interest in virtual reality, user experience, computer graphics, and acoustics simulation. He is a Full-Time Lecturer with Centro Universitario de Tecnología y Arte Digital (U-tad), Madrid, Spain. Mayor received the Master's degree in computer graphics and simulation at U-tad. Contact him at j.mayor.m90@gmail.com (www.jesusmayor.com).

José Jesús García Rueda is currently working toward the Ph.D. degree in telematics at Universidad Politécnica de Madrid, Madrid, Spain, with an interest in virtual worlds, hypermedia, and educational technology. He is a Full-Time Lecturer and Researcher with Centro Universitario de Tecnología y Arte Digital, Madrid, Spain. Contact him at jose.rueda@u.tad.com.

Laura Raya is currently working toward the Ph.D. degree in computer science at Universidad Rey Juan Carlos, Madrid, Spain, with an interest in virtual reality, visualization, gamification, and education. She is a Full-Time Associate Professor with Centro Universitario de Tecnología y Arte Digital, Madrid, Spain. Contact her at laura.raya@u-tad.com (gmr.v.es/~lraya).

■ Contact department editor Mike Potel at potel@wildcrest.com.