Collaborative Virtual Environment to Encourage Teamwork in Autistic Adults in Workplace Settings

Ashwaq Zaini Amat[1](✉), Michael Breen[1], Spencer Hunt[1], Devon Wilson[2], Yousaf Khaliq[2], Nathan Byrnes[2], Daniel J. Cox[2], Steven Czarnecki[2], Cameron L. Justice[2], Deven A. Kennedy[2], Tristan C. Lotivio[2], Hunter K. McGee[2], Derrick M. Reckers[2], Justin W. Wade[2], Medha Sarkar[2], and Nilanjan Sarkar[1]

[1] Vanderbilt University, Nashville, TN 37212, USA
ashwaq.zaini.amat.haji.anwar@vanderbilt.edu
[2] Middle Tennessee State University, Nashville, TN 37132, USA

Abstract. The employment settings for autistic individuals in the USA is grim. As more children are diagnosed with ASD, the number of adolescent and young adult with ASD will increase as well over the next decade. Based on reports, one of the main challenges in securing and retaining employment for individual with ASD is difficulty in communicating and working with others in workplace settings. Most vocational trainings focused on technical skills development and very few addresses teamwork skills development. In this study, we present the design of a collaborative virtual environment (CVE) that support autistic individual to develop their teamwork skills by working together with a partner in a shared virtual space. This paper described the CVE architecture, teamwork-based tasks design and quantitative measures to evaluate teamwork skills. A system validation was also carried out to validate the system design. The results showed that our CVE was able to support multiple users in the same shared environment, the tasks were tolerable by users, and all the quantitative measures are recorded accordingly.

Note: We are using both identity-first and people-first language to respect both views by interchangeably using the term 'autistic adults' and 'adults with autism' [22].

Keywords: Collaborative virtual environment · Neurodiverse employability · Autism spectrum disorders · Adolescent and adult with ASD · Teamwork · Collaborative skills

1 Introduction

Autism spectrum disorder (ASD) is a range of disorder that can significantly affect a person's ability to socially interact and communicate with others. Based on a new study by Centers for Disease Control and Prevention (CDC), there are approximately 2.21% of adolescents and adults in the USA with ASD [1]. As ASD is considered a

M. Antona and C. Stephanidis (Eds.): HCII 2021, LNCS 12768, pp. 339–348, 2021.
https://doi.org/10.1007/978-3-030-78092-0_22

lifelong condition, there have been growing interest and concern of the challenges faced by individuals with ASD as they are transitioning into adolescence and adulthood [2]. Employability is one of the main challenges as the unemployment rate for individuals with ASD are between 50–85% [2, 3]. For those with employment, majority of them are either underemployed or are unable to retain their position for long [4]. Studies reported that soft skills and communication were the biggest challenges and reasons for unemployment in this population [5, 6]. Despite that, majority of job trainings offered to individuals with ASD are more focused on developing technical skills and rarely on communication and soft skills development [7]. Based on these findings, there is a need for a system that could support soft skills development and encourage social interactions at workplace for autistic individuals.

Stauch and Plavnick combined vocational training with social skills training using video modelling for individuals with ASD [8]. Participants were shown videos of vocational tasks embedded with social skills as part of the intervention while performing the task in real life. Although the study reported improvements in both vocational skills and social skills among the participants, there was limited availability of quantitative measures and the cost to operate such study was quite high. The use of virtual reality (VR) based systems for vocational training has shown potential in recent years. VR training engages the users [9], provides a safe environment for training [10] and provides quantitative measures of the skills they are learning [11]. As a result, several important virtual systems have been explored in the context of vocational training [9–11]. But these systems are not catered for individuals with disabilities. A recent study by Bozgeyikli et al. presented a vocational VR-based training system for individuals with disabilities focused on building technical skills such as cleaning, shelf sorting and point of sale skills [12]. Participants were able to interactively practice on the technical skills within the virtual environment while a job coach supervise their performance. However, these VR-based training systems still lack the support on social skills development, specifically teamwork and collaboration skills.

The use of collaborative virtual environment (CVE) in the last decade has provided autistic individuals the platform to interact with each other or a coach to perform a task together. CVE offers the same advantages as the VR-based interaction with the added ability of sharing the virtual experience with other users, thus expanding on the social interaction experience. iSocial, a distributed learning environment was designed for school students to collaborate and work with each other in virtual settings following a defined syllabus [13]. In one of the tasks, children were asked to communicate with each other to plan a vacation. Zhang et al. designed a CVE system that allowed two children to play collaborative puzzle games with each other [14]. The system was able to promote collaboration and communication between users while quantitatively measure the skills. As can be seen from the examples mentioned, majority of CVE studies focused on social skills development in children with ASD, in support of early interventions creating a gap on the outcome for adult with ASD [15].

In this work, we present a collaborative virtual environment (CVE) where a person with ASD and a neurotypical (NT) partner will interact with each other to complete two designated tasks. The objectives of the CVE are to (i) allow users to work together (teamwork) in a shared virtual space and (ii) encourage communication and collaboration between users through shared tasks. The rest of the paper presents the system architecture, collaborative tasks design, and result of system testing.

2 Collaborative Virtual Environment (CVE) System Design

A collaborative virtual environment (CVE) is a virtual environment that allows multiple users to share the same virtual space where they can interact with other users and the environment itself. In this work, our CVE system was designed for two users to work together as a team to achieve the same goal in each task. In this study we designed tasks within the CVE that simulate workplace environment, even though the system is not limited to only these tasks. The first is a computer assembly task and the second task is a fulfillment center. These tasks were chosen as we wanted a workplace environment that could encourage teamwork and collaboration.

Users access the CVE from two different physical location using their respective computer and join the shared virtual environment by connecting to the same server. In each task, users are presented with a set of instruction that included task objectives, a list and manuals. The users control their movements and interact with the virtual environment by using common computer peripherals such as keyboard keypads, a mouse and or a gamepad. to control movements. Once connected, users can communicate with each other through video and audio streaming component embedded within the CVE. They use the instruction set as reference to navigate within the environment together by collaborating and communicating to accomplish the set goals.

2.1 Tasks Design

The main design criteria we considered for the tasks was they must encourage teamwork and collaboration. Both users need to communicate with each other and synchronize their actions in the virtual space to achieve the goals set in the tasks. Without communicating with each other to perform the task, the users will not be able to proceed with the task. Each task takes between five to ten minutes to complete. When users are unable to proceed with the task, researchers will provide assistance to move to the next step.

Computer Assembly Task. In the computer assembly task, users are working together in a computer workshop room. As shown in Fig. 1, five computer components together with a computer chassis are placed on a table in front of both users. The objective of the computer assembly task is to put together a computer using the provided components. One user plays the role of an assembler while the other user plays the role of a supervisor and has access to the installation manual. The supervisor should provide verbal guidance to the assembler. The assembler then follows the instruction conveyed by the supervisor to put the computer together. Both the assembler and supervisor will take turn to place the component into the chassis as some components are accessible to only either the assembler or the supervisor. To pick up the component, users select the component they want to move with the mouse and use keyboard keypad buttons to move and rotate the component into the right location in the chassis. Both users need to communicate well with each other to fully assemble the computer within a specified time. Throughout the task, the CVE system will continuously monitor interaction and collaborative actions.

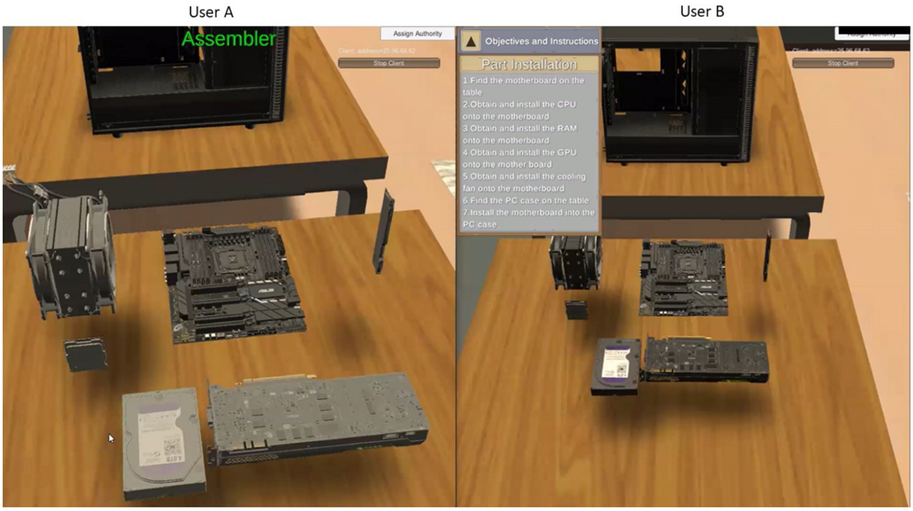
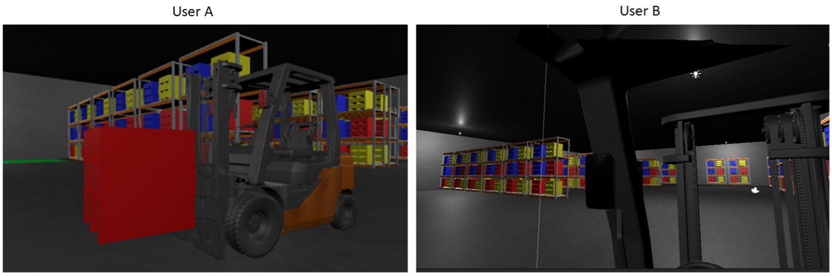


Fig. 1. Users' view of the computer assembly task.

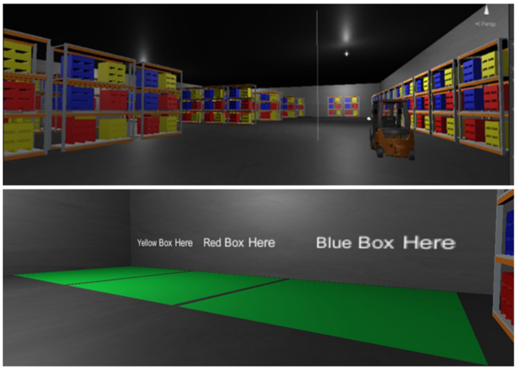
Fulfillment Center Task. The fulfilment center task is set in a warehouse. Users need to drive forklifts in the warehouse to retrieve pallets from the shelves to a collection area. The pallets are color-coded to match them to the different collection area. Figure 2 illustrates the different point of views of both users and overall view of the warehouse with the color-coded collection area. Both users are given a list of pallets they need to pick up from various location in the warehouse. One user is assigned as the trainee while the other user is the supervisor. The supervisor in this task will give directional verbal instruction (e.g., 'Move to the left', 'Move the fork down further') to guide the trainee when picking up and dropping off pallets as the trainee does not have the complete view of the pallet from inside the forklift. This is shown in Fig. 2(a). In this task, a gamepad controller is used to maneuver the forklift and move around in the warehouse. Figure 3 shows the controls used in this task. Additionally, at any time, only one forklift can access the shelves, as such the users have to communicate and coordinate their locations to avoid collision and efficiently transport the pallets. Both users will need to complete the task within a specified time. Similar to the computer assembly task, the CVE system will continuously monitor interaction and collaborative actions.

2.2 CVE Architecture

The virtual environment was designed and developed using Unity, a leading game development software that support interactive play [16]. Figure 4 presents the system interaction diagram and the five main components within the CVE architecture. The Network Communication Module is responsible to set up the shared environment when users start joining. Once connection is established, this module triggers the Player Controller to set up and assign players with their respective roles. At the same time, the module initializes the audio and video streaming component. When users start interacting and performing



(a)



(b)

Fig. 2. Fulfillment center task views. (a) Point of view of each player. User A can see User B driving the forklift carrying a red pallet. User B's view is from inside the forklift. (b) Overall warehouse view with pallets of different colors and the color-coded collection area.

Fig. 3. Gamepad configurations for forklift user.

a task, the Player Controller trigger events to the Network Communication Module to exchange player data between the users' so that both users' view are synchronized. At the same time, any user interactions with the environment are synchronized through the Environment Controller. Users' data are also stored locally in the Data Collection Module for post-task analysis.

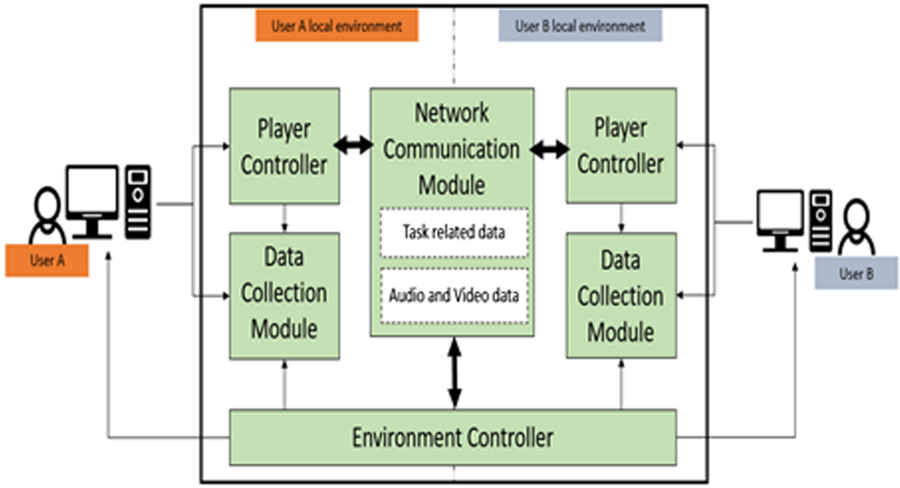


Fig. 4. System interaction diagram

Network Communication Module. The main component that enables a virtual environment to be shared among multiple users is the network component. For our implementation, we used two types of networking interfaces: Mirror [17] and WebRTC [18]. Both are compatible with Unity and purchased through Unity Asset store.

Using similar architecture to Unity Networking (UNet) application that has been deprecated [19], Mirror uses a server-client module. A server connects all the users together and manages data transmission in both server and client. A client is created for each user that is connected to the same environment through the server. In this CVE system, one of the users is connected as a host server which take the role of both a server and client and the other user is connected as client only. The assignment of the server client roles to the users are solely for network configuration and does not influence user experience at all. Mirror is used to exchange environment related data that are dynamically changing such as player actions and objects movements, and periodically exchanging data such as user performance and duration of the task across both user's environments. As our tasks involve moving virtual objects to specific locations and transporting large virtual objects around, we wanted the synchronization to be high but not overload the channel. The default update rate for Mirror was set to 60 messages per second or 60 Hz, which is suitable for high paced game [17]. To stream the audio and video data, we initially used Dissonance, an integration component in Mirror that allow video and audio streaming [20]. Due to incompatibility with other components in the development, we had to use an alternative network interface. We chose WebRTC to stream the user's audio and video data from a webcam and microphone, respectively, in real-time between users while they performed the virtual task. This plugin came with pre-configured network layer that makes establishing connection easy. An event handler within the plugin trigger updates to the receiving user when a new video frame was received, or an audio clip was received. For a smooth video transmission, a frame rate of 30 frame per second (fps) was chosen for WebRTC. In the computer assembly task, we are using both video and audio streaming service, while the fulfillment center uses only the voice streaming service.

Player Controller. The Player Controller configures the user movements within the environment and how they interact with the virtual objects. Within the controller we also defined two different roles for each user to encourage further collaboration and communication between them. The first role defined is of a supervisor, where they need to provide more verbal instruction to the other user and minimal involvement with interacting with the virtual objects. The other role is of a trainee where the user is expected to conduct most of the interaction with the virtual objects and communicate with the other user. The assignment of roles is done randomly and alternately applied between the two tasks. For example, if the first task was the computer assembly task and User A is assigned the supervisor role and User B is assigned the trainee role, then in the fulfillment center task, User A will be assigned the trainee role while User B will take the supervisor role.

This controller is also connected to the Network Communication Module to transmit and receive player's information which include player's location and actions such as moving or picking up virtual objects. Player representation in the shared environment is important to establish a sense of presence of each other and facilitate collaboration and team work better. It is also important to synchronize the players' information to ensure player representation is not delayed that can disrupt the teamwork and collaboration activity they are conducting.

Quantitative Measures. To evaluate the collaborative skills and teamwork, we are using tasks related data and social communication information gathered during the experiment. Task performance data include number of successful collaborative moves, duration of collaborative actions and scores are collected locally for each user in the Data Collection Module. Social communication data include speech and eye gaze data. Within the system we are transcribing the user's speech and including it in the Data Collection Module. Evaluation of speech and conversation are done post tasks where researchers view experiment recording and label the collaborative part of the speech. We then validate it against the transcription and add the label accordingly. The labelling can be used in future study to improve our system. For eye gaze data collection, we are using a TobiiEyeX [21] to track gaze points in the virtual space and specific virtual objects that the user looks at. Similar to speech data, gaze points are analyzed post tasks to obtain the gaze pattern in collaborative settings.

3 System Validation

We conducted system testing with 3 pairs of NT volunteers to validate that the system can support multiple users and allow users to perform the tasks together in the shared environment without any issues. Additionally, through the system testing we can validate that the data are correct and logged in the right format.

After the test, we asked the volunteers to fill out a system usability scale to rate their experience performing the tasks within the CVE. They commonly agreed that the system was tolerable and simple to understand. They could easily understand the instructions and objectives of both tasks. One user did comment that the use of different controllers across both tasks caused a bit of confusion when transitioning over to the other task. We took note of this and will add them into consideration when we revise the system for a user study.

As part of technical verification of the system, we observed two areas of the system: connectivity and recording of quantitative measures. To verify the system connectivity, we tested connecting two computers in the same building but in separate rooms 10 times a day and repeated the same process for five days. We achieved a success rate of 98% where only 1 trial failed to get connected. When testing with the volunteers, they accessed the shared environment six times where each pair connected twice, one time for each task. In all six times, there were no connectivity issues, and they were able to stay connected until they finished the tasks. Our Network Communication Module was robust, where players and all virtual objects in the computer assembly task and fulfillment center task were synchronized with unnoticeable latency. We have yet to test the network performance of this system in different buildings and observe the latency.

To verify the quantitative measures, the researchers reviewed and compared the recorded data against the video recording of the volunteers. As we did not have matching time stamp between the video recording and the log files, we relied on the duration of the task to align the data with the actual video. From the analysis of the 3 sessions with the volunteers, we were able to match the speech transcriptions to the utterances in the videos. Although we did not analytically measure the accuracy of the transcription, the system successfully discriminated the speech coming from both users (e.g.: User 1 Speech, User 2 Speech). As for the eye gaze data, we selected a short clip from the video recording and observe what was the volunteer doing in that clip. We then compared this to the gaze log file and find if there was a match between the object in the video to the object logged in the gaze log file. We found 100% match in eye gaze data against the video recording for the virtual object logging. We did not verify the raw gaze point data as that required more processing time. Both the speech and gaze data can be used to analyze collaborative level of the participants.

With the completion of our system testing, we can validate that our CVE system is tolerable by users, easy to understand, able to support multiple users connected to the same virtual environment and record meaningful quantitative data related to teamwork and collaborative skills. It is ready for user testing where we will analyze the user performance, teamwork skills and the system performance.

4 Conclusion

We designed a CVE that is aimed to support the development of teamwork and collaborative skills of autistic adults in vocational settings. This initial study addresses the design considerations for suitable tasks that would encourage individuals with ASD to communicate and work together with a partner to achieve a common goal. To quantitatively measure teamwork and collaborative performance, our system continuously collects user's actions in the task together with speech and gaze data.

There were a few limitations in our system validation. The main limitation is the system was tested with neurotypical users only. Next, there was not enough variation in the tasks to allow for more collaborative work between users. Also, the current design of the feedback is limited to only acknowledgement of successful attempts but does not prompt the users when they are stuck. Finally, availability of speech data that are labelled with collaborative tags are limited due to the limited number of study participants and shorter tasks length.

Our future work will address the limitations we identified from this round of testing. First, we would conduct a pilot study with autistic adults and NT partners. We will also introduce variations to the tasks. For example, we plan to have three different difficulty level for each task. The basic level will serve as a baseline to evaluate the users task performance while the second and third level introduces ambiguity to the tasks that would drive the users to work together to overcome the challenge. To further improve user experience using the system, a more robust feedback mechanism will be introduced to support users when they are not sure what to do. After the pilot study is finished, we plan to use the speech data to train a machine learning model that could classify the collaborative speech in real-time more accurately. With these improvements in plan, we are confident that our CVE system can benefit individuals with ASD by providing them with the right tool and support to secure and retain employment by preparing them to work better with others.

Acknowledgement. We are grateful for the support provided by NSF grants 1936970 and 2033413 for this research. The authors are solely responsible for the contents and opinions expressed in this manuscript.

References

1. Dietz, P.M., Rose, C.E., McArthur, D., Maenner, M.: National and state estimates of adults with autism spectrum disorder. J. Autism Dev. Disord. **50**(12), 4258–4266 (2020). https://doi.org/10.1007/s10803-020-04494-4
2. Levy, A., Perry, A.: Outcomes in adolescents and adults with autism: a review of the literature. Res. Autism Spectr. Disord. **5**(4), 1271–1282 (2011)
3. U.S. Bureau of labor statistics
4. Brooke, V., et al.: Employees with autism spectrum disorder achieving long-term employment success: a retrospective review of employment retention and intervention. Res. Pract. Pers. Severe Disabil. **43**(3), 181–193 (2018)
5. Hendricks, D.: Employment and adults with autism spectrum disorders: challenges and strategies for success. J. Vocat. Rehabil. **32**(2), 125–134 (2010)
6. Taylor, J., Seltzer, M.: Employment and post-secondary educational activities for young adults with autism spectrum disorders during the transition to adulthood. J. Autism Dev. Disord. **41**, 566–574 (2011)
7. Walsh, L., Lydon, S., Healy, O.: Employment and vocational skills among individuals with autism spectrum disorder: predictors, impact, and interventions. Rev. J. Autism Dev. Disord. **1**, 266–275 (2014)
8. Stauch, T.A., Plavnick, J.B.: Teaching vocational and social skills to adolescents with autism using video modeling. Educ. Treat. Child. **43**, 137–151 (2020)
9. Checa, D., Bustillo, A.: A review of immersive virtual reality serious games to enhance learning and training. Multimed. Tools Appl. **79**(9–10), 5501–5527 (2019). https://doi.org/10.1007/s11042-019-08348-9

10. Norris, M.W., Spicer, K., Byrd, T.: Virtual reality: the new pathway for effective safety training. Prof. Saf. **64**(06), 36–39 (2019)
11. Zhang, L., Warren, Z., Swanson, A., Weitlauf, A., Sarkar, N.: Understanding performance and verbal-communication of children with ASD in a collaborative virtual environment. J. Autism Dev. Disord. **48**(8), 2779–2789 (2018)
12. Bozgeyikli, L., et al.: VR4VR: vocational rehabilitation of individuals with disabilities in immersive virtual reality environments. In: Proceedings of the 8th ACM International Conference on Pervasive Technologies Related to Assistive Environments, pp. 1–4, July 2015
13. Stichter, J.P., Laffey, J., Galyen, K., Herzog, M.: iSocial: delivering the social competence intervention for adolescents (SCI-A) in a 3D virtual learning environment for youth with high functioning autism. J. Autism Dev. Disord. **44**(2), 417–430 (2013). https://doi.org/10.1007/s10803-013-1881-0
14. Zhang, L., Fu, Q., Swanson, A., Weitlauf, A., Warren, Z., Sarkar, N.: Design and evaluation of a collaborative virtual environment (CoMove) for autism spectrum disorder intervention. ACM Trans. Access. Comput. **11**(2), Article 11 (2018)
15. Glaser, N.J., Schmidt, M.: Usage considerations of 3D collaborative virtual learning environments to promote development and transfer of knowledge and skills for individuals with autism. Technol. Knowl. Learn. **25**(2), 315–322 (2020)
16. https://unity.com/
17. https://mirror-networking.com/
18. https://www.because-why-not.com/webrtc/#WebRTC_Video_Chat
19. https://docs.unity3d.com/Manual/UNet.html
20. https://dissonance.readthedocs.io/en/latest/Basics/Quick-Start-Mirror.html
21. https://www.tobii.com/
22. Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., Pellicano, E.: Which terms should be used to describe autism? Perspectives from the UK autism community. Autism **20**(4), 442–462 (2016)