VISTA: User-centered VR Training System for Effectively Deriving Characteristics of People with Autism Spectrum Disorder

Bogoan Kim
bogoankim@hanyang.ac.kr
Department of Intelligence Computing,
Hanyang University
Seoul, Republic of Korea

Dayoung Jeong
dayoungjeong@hanyang.ac.kr
Department of Artificial Intelligence,
Hanyang University
Seoul, Republic of Korea

Mingon Jeong
mingon21@hanyang.ac.kr
Department of Artificial Intelligence,
Hanyang University
Seoul, Republic of Korea

Taehyung Noh
yestaehyung@hanyang.ac.kr
Department of Artificial Intelligence,
Hanyang University
Seoul, Republic of Korea

Sung-In Kim
sunginkim@snu.ac.kr
College of Medicine,
Seoul National University
Seoul, Republic of Korea

Taewan Kim
taewan@kaist.ac.kr
Department of Industrial Design,
KAIST
Daejeon, Republic of Korea

So-youn Jang
soysoyee@snu.ac.kr
Department of Communication,
Seoul National University
Seoul, Republic of Korea

Hee Jeong Yoo
hjyoo@snu.ac.kr
Department of Psychiatry,
SNU Bundang Hospital
Seongnam, Republic of Korea

Jennifer G Kim
jennifer.kim@cc.gatech.edu
School of Interactive Computing,
Georgia Institute of Technology
Atlanta, Georgia, United States

Hwajung Hong
hwajung@kaist.ac.kr
Department of Industrial Design,
KAIST
Daejeon, Republic of Korea

Kyungsik Han*
kyungsikhan@hanyang.ac.kr
Department of Intelligence Computing,
Hanyang University
Seoul, Republic of Korea

Figure 1: The foreground (first row) and the background (second row) screenshots of *VISTA* from different angles. The foreground is the area in which the participants are mainly involved during training. *VISTA* provides a realistic background to help participants feel real and immersive during training.

*Corresponding author

VRST '22, November 29-December 1, 2022, Tsukuba, Japan

ABSTRACT

Pervasive symptoms of people with autism spectrum disorder (ASD), such as a lack of social and communication skills, are major challenges to be embraced in the workplace. Although much research has proposed VR training programs, their effectiveness is somewhat unclear, since they provide limited, one-sided interactions through

ACM ISBN 978-1-4503-9889-3/22/11…$15.00
https://doi.org/10.1145/3562939.3565608

fixed scenarios or do not sufficiently reflect the characteristics of people with ASD (e.g., preference for predictable interfaces, sensory issues). In this paper, we present *VISTA*, a VR-based interactive social skill training system for people with ASD. We ran a user study with 10 people with ASD and 10 neurotypical people to evaluate user experience in VR training and to examine the characteristics of people with ASD based on their physical responses generated by sensor data. The results showed that ASD participants were highly engaged with *VISTA* and improved self-efficacy after experiencing *VISTA*. The two groups showed significant differences in sensor signals as the task complexity increased, which demonstrates the importance of considering task complexity in eliciting the characteristics of people with ASD in VR training. Our findings not only extend findings (e.g., low ROI ratio, EDA increase) in previous studies but also provide new insights (e.g., high utterance rate, large variation of pupil diameter), broadening our quantitative understanding of people with ASD.

CCS CONCEPTS

• **Human-centered computing** → **Virtual Reality**; **Interactive systems and tools**; **User studies**.

KEYWORDS

Virtual reality, Autism Spectrum Disorder (ASD), Social skills training system, User study

ACM Reference Format:

Bogoan Kim, Dayoung Jeong, Mingon Jeong, Taehyung Noh, Sung-In Kim, Taewan Kim, So-youn Jang, Hee Jeong Yoo, Jennifer G Kim, Hwajung Hong, and Kyungsik Han. 2022. VISTA: User-centered VR Training System for Effectively Deriving Characteristics of People with Autism Spectrum Disorder. In *28th ACM Symposium on Virtual Reality Software and Technology (VRST '22), November 29-December 1, 2022, Tsukuba, Japan.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3562939.3565608

1 INTRODUCTION

One of the application areas of virtual reality (VR) technology is training and education [32, 54, 82]. This is because VR gives an opportunity to create an environment and conditions that correspond to those in the real world and to have users experience a specific situation in an immersive way [42, 66, 70, 73]. In VR, users can have great flexibility in interacting with virtual environments, agents, or objects, engage in repetitive learning, and receive feedback on their behaviors or decisions during training [38].

Such VR opportunities have been extended to those in need of social support [1, 14], including people with autism spectrum disorder (ASD) [14, 25, 52]. Autism Spectrum Disorder (ASD) is a complex developmental condition involving the challenges of social interaction, along with specific patterns of limited and repetitive behaviors [6]. Research on people with ASD has developed various types of VR content to support self-help skills (e.g., activities of daily living skills [3, 46, 65], driving [19], road crossing [61]) or social skills (e.g., job interview [68], facial expression recognition [10, 16], and social interaction [40, 60, 69]) that are necessary for one's daily life. Research has confirmed the improvement of target behaviors or educational effects through the VR experience. More recently, research has started to consider the information related to user body responses (e.g., physiological data, eye movement, body movement) to be implemented in the development of a training tool [10, 53, 60, 61]. This provides a way to quantitatively measure and understand the characteristics of people with ASD and to design additional solutions for their social independence.

Despite the advantages of VR content and technology in helping people with ASD improve their social skills, when it comes to VR content development, we identified two limitations in previous studies. The first is the one-sided interaction in previous VR solutions, in which users experience a pre-programed, fixed scenario (e.g., giving a presentation for a given time [23], answering an interviewer's questions in a job interview [68]) rather than offering them an explorable environment in which their situation could be understood, allowing them to take actions on their own, and receiving feedback on their responses to an avatar. Second, the design and development of the content have been conducted without careful consideration of people with ASD. Bozgeyiki and Katkoori [13] presented key design guidelines for VR system development for people with ASD. The authors emphasized that considering the characteristics of people with ASD (e.g., sensory issues [11], preference for expected interfaces [41, 47] and restricted and repetitive behaviors [48]) into account is critical in terms of the system's usability, effectiveness, and sustainability. With a VR system designed based on such careful guidelines, the characteristics of people with ASD can be better captured in sensor data through more genuine experience that can quantitatively represent their behaviors.

In this paper, we present *VISTA* (VR-based Interactive Social skills Training system for people with ASD) that was designed not only to reflect insights and feedback from ASD professionals and stakeholders but also to comply with design guidelines for people with ASD [13]. *VISTA* offers a training content that places people with ASD in an environment in which they need to understand given scenarios and interact with others. In particular, one key design of *VISTA* is the provision of training scenarios with incremental complexities, which has been considered an effective training method for people with ASD but has not been employed in previous studies. Lastly, *VISTA* was designed to collect various types of sensor data (i.e., eye movement, head movement, physiological signals, and voice) from people with ASD. We considered those signals that could be collected unobtrusively, given their high sensitivity to wearing extra devices.

We conducted a user study with 20 participants from two user groups (one group with 10 ASD participants and the other with 10 neurotypical participants). The objectives of the user study were to (1) measure the effectiveness of the design of *VISTA* on having people with ASD engage in training and helping them understand social situations and feel confidence in social interactions; (2) investigate the difference in sensory reactions between the two groups; and (3) discuss the characterization of people with ASD by comparing the results with those of previous studies and presenting new insights identified in our study.

Our study results highlight three main findings. First, the ASD participants showed positive perceptions of their training experience and increased self-efficacy after training through *VISTA*. Second, although there were no significant differences in the variation of sensor signals between the ASD participants and the neurotypical participants at the beginning, the differences became much more

Table 1: Examples of previous research on VR-based training systems for people with ASD. "ADL" indicates activities of daily living. AP/NP indicate ASD/neurotypical participants, respectively. 360-degree VR is an audio-visual simulation surrounding the participant, and its content can be viewed in all directions. "O": used/considered; "X": not used/not considered. For the last column (ASD consideration), "△": partially considered characteristics of people with ASD, missing detailed design guidelines.

Authors	Focus	Age, # of participants	Equipment	Sensor data	360-degree VR	Interaction in VE	ASD consideration
Robles et al. [60]	Social interaction	28.8 yrs (avg.), AP = 6, NP = 13	HMD, controllers, trackers	Eye-tracking, head & hand movement	O	O	△
Adiani et al. [2]	Job interview	22.5 yrs (avg.) AP = 9, NP = 8	Tobii EyeX/4C, a headset, E4	EDA, BVP, eye-tracking	X	O	O
Ke et al. [40]	Social skills	10-14 yrs, AP = 7	Mouse, headset	N/A	O	O	X
McCleery et al. [51]	Police interaction skills	12-38 yrs, AP = 60	Head-mounted iPhone	N/A	O	O	X
Herrero & Lorenzo [29]	Communication skills	8-15 yrs, AP = 7	HMD (Oculus Rift)	N/A	O	O	△
Rosenfield et al. [69]	Conversation skills	6-7 yrs, AP = 2	HMD (Oculus Rift), microphone	N/A	O	O	△
Simões et al. [65]	ADL (bus-taking)	No age info, AP = 10, NP = 10	HMD (Oculus Rift), BioNomadix	EDA	O	X	△
Adjorlu et al. [3]	ADL (shopping skills)	12-15 yrs, AP = 9	HMD, controllers	N/A	O	O	△
Bozgeyikli et al. [12]	Vocational training	25-29 yrs, AP = 9	HMD	N/A	O	X	△
Cheng et al. [16]	Social cognition	10-13 yrs, AP = 3	HMD	N/A	O	O	△
Saiano et al. [61]	Street crossing	19-44 yrs, AP = 7	Video projector, Microsoft Kinect	Eye-tracking	X	O	X
Smith et al. [68]	Job interview	18-31 yrs, AP = 26	Mouse	N/A	X	O	△
Kandalaft et al. [37]	Social cognition	18-26 yrs, AP = 8	Keyboard, mouse	N/A	X	O	X

prominent in most sensor signals as the task complexity in the scenarios increased. This highlights the effective design of *VISTA* for increasing the engagement and immersiveness of the ASD participants. Lastly, by interpreting our study findings in light of previous studies, we reaffirmed some of the key sensor signals as well as identified other signals that are somewhat neglected as indicators to understand people with ASD.

The main contributions of this paper are as follows:

- We developed *VISTA* which incorporated key design guidelines for VR content for people with ASD.
- We demonstrated the effectiveness of the design of *VISTA* in eliciting ASD characteristics and providing them with an immersive VR training experience.
- We extended our understandings of sensor signals generated by people with ASD during the VR experience by discussing our findings in the light of previous studies.

2 RELATED WORK

2.1 VR-based training for people with ASD

Many studies have used VR to train and educate the self-help skills (e.g., activities of daily living, driving, road crossing) of people with ASD [3, 65, 83]. For example, Saiano et al. [61] developed VR content to train the street crossing and path-following skills of adults with ASD. The participants started training by looking at the virtual environment projected on the large display front, in which their movements were reflected through the Kinect system. The results of the study showed a significant improvement in navigation performance (e.g., walking safely on a green light through a crosswalk). McCleery et al. [51] provided the training in interactions with police officers for adolescents and adults with ASD. The virtual police officer approached the participant and asked several questions for personal identification (e.g., name, address, identification card), and the participant was asked to answer the questions. The study results showed that the VR content helped the participants interact safely and effectively with police officers.

VR content has also been developed to help acquire and strengthen the social skills (e.g., job interviews, facial expression recognition, social interaction) [10, 29] of people with ASD. Ke et al. [40] used VR content for the training of social skills in children with ASD by offering four social interaction tasks (i.e., virtual schooling, social role-playing, artifact design, and social gaming). The authors found that the children with ASD showed improved social skills as VR training sessions were repeated. Adiani et al. [2] developed a VR-based job interview training system for adults with ASD. The participants experienced possible challenges in the career interview environment (e.g., interrupted by knocking into the interview room during the response).

Despite such opportunities for people with ASD, when it comes to design of content and system, previous studies did not seem to sufficiently consider the characteristics of people with ASD in terms of the design of content and systems, as summarized in Table 1. Bozgeyiki and Katkoori [13] presented design guidelines for

VR content for people with ASD, consisting of three main aspects (information presentation, task design, and VR system). Our VR training system, *VISTA*, employed those design guidelines. A salient gap in previous research is the somewhat limited user experience due to a pre-programmed, fixed scenario and one-sided interactions during training [23, 52]. Most existing systems were designed to support users only in situations in which they talked alone or only answered the agent's questions. Further, since the content is not structured for the user to understand his or her current situation and to respond, the application of such VR training experiences to the real world is somewhat unclear. *VISTA* allows people with ASD to experience in-the-moment feedback and provides task scenarios with incremental complexity. We will detail the development process of *VISTA* grounded in the design guidelines in Section 3.

2.2 Sensor signals to understand people with ASD

Much research has been conducted using various types of sensor data (e.g., body movement and orientation, eye movement, physiological signals, video and audio) to understand and characterize users and find ways to improve the VR user experience with quantitative evidence [31, 33, 49, 59].

Many previous VR studies for people with ASD employed general, casual contexts, such as game, emotion/facial expression recognition, and conversation, to collect sensor data [4, 9, 28, 45]. However, as shown in Table 1, the majority of VR studies in "training contexts" did not have the collection of sensor data [3, 12, 16, 29, 37, 40, 51, 68, 69]. Even for some studies that employed sensor data collection ([2, 60, 61, 65] in Table 1), their VR content is somewhat limited because they lack interactive components or because only few cases of the challenging moments frequently encounterd by people with ASD are presented. Thus, it is unclear whether their training experience through such training conditions is effective when it is applied to real-world situations.

Sensory issues must be carefully considered when collecting sensor data from people with ASD due to their high sensitivity to objects, environments, or atmosphere [11]. Obtrusive and excessive collection of sensor data for people with ASD will certainly degrade their VR experience or amplify their anxiety. Therefore, a comprehensive understanding of the characteristics of people with ASD and their application to content design is essential. Some studies have collected and analyzed brain signals [17, 21, 81], but the collection of such data can be highly difficult to implement and maintain. In this work, we attempted to leverage minimal sensor types that have been commonly used in previous studies (e.g., eye-tracking, head movement, voice, physiological signal) [10, 28, 45, 60, 61, 65].

3 VISTA DESIGN AND DEVELOPMENT

Being a barista assistant has been suggested as an appropriate job role for people with ASD because it allows continuous interactions with the barista or customers [5, 7, 77]. The role allows exposure to such interactive situations and provides people with ASD with the opportunity to gain situational awareness and interact with others. In a non-VR environment, barista vocational education has been widely conducted for people with ASD in the field of special education operated by welfare centers [34]. We had iterative discussions with medical professionals and ASD stakeholders (e.g., ASD experts, adults with ASD) to decide key components and the level of complexity that needed to be implemented in *VISTA* scenarios.

Table 2: Main categories of design considerations for people with ASD suggested in [13] and the strategies reflected in our training system.

Main Categories [13]	Strategies reflected in VISTA
Information presentation	Avoid sudden or loud sound
	Revisit the opening scenarios
	Provide adjustable sounds
	Provide clear foreground and background differences
Task design	Provide concrete and routinized tasks
	Provide gradual task complexities
	Provide in-the-moment feedback
	Provide short training scenarios
VR system	Provide the familiarization time with the equipment
	See the virtual environment before the training
	Avoid wearing/using or complex input devices
	Provide various interactions with human-like avatars

The design of *VISTA* was carefully decided based on three design guidelines of the VR system for people with ASD [13]: information presentation, task design, and VR system. We reflected on the first two for our VR training content development and the last for the experimental environment setting. Table 2 summarizes the design considerations of *VISTA*. We followed the information presentation guidelines by providing a clear foreground and background differences for more comfortable visual processing [43, 56, 62]. We then excluded abrupt visual and sound changes to consider the ASD participants' high affinity for predictable situations [39, 41] and audiovisual sensitivity [27, 64]. Further, the complexity of the scenario was increased incrementally to allow people with ASD to confront more challenging situations over time (not in a sudden). VR tracker- or joystick-based tasks were excluded, given the motor difficulties commonly seen in people with ASD [41, 78].

Figure 2 shows the key interaction steps in *VISTA*. We organized the content into three phases to increase effectiveness and to more closely examine ASD characteristics. In all phases, the participants began training with opening scenarios. In the opening scenarios, the participant was informed by the cafe manager (barista) that the customer's drink was ready, remembered the order information (e.g., order number, type of drink), and called the customer by the order number. From Phase 1 to 2, the complexity of the interaction increases as the number of drinks to be served increases by two. From Phase 2 to 3, complexity of interaction increases once again because an appropriate response is required to the customer's complaint about the served drink (The first scenario in Phase 3: cafe-side mistake, The second scenario in Phase 3: customer misunderstands that their drink is mis-served. When a cafe customer starts a complaint, the participant compares and checks the receipt on the table and then responds. The response in the first scenario in Phase 3 is judged by the researcher as (1) no response, (2) appropriate response, (3) inappropriate response, or (4) insufficient response. For the second scenario in Phase 3, (1)-(3) were used.

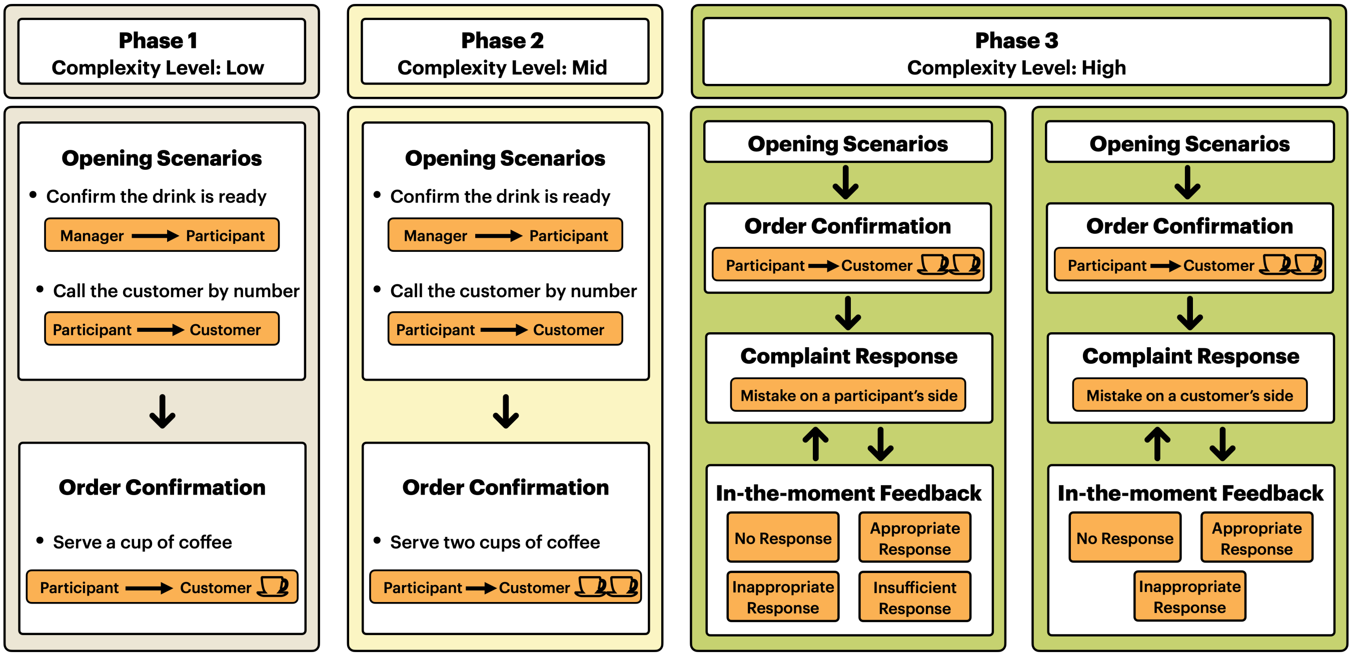


Figure 2: Training phases of *VISTA*. The participants conducted three consecutive training phases (Phase 3 includes two social interaction scenarios), and the participants could experience gradually increased complexity (from "Low" to "High") over phases. In "Complaint Response" (Phase 3), the participants repeatedly conducted training unless "in-the-moment feedback" on the response was "appropriate."

4 USER STUDY

4.1 Participants

We conducted a user study in an independent experimental space at the author's institution. We spent six months (from November 2021 to May 2022) recruiting 10 ASD participants (male = 8, female = 2) by distributing leaflets to organizations employing people with ASD and clinics, and posting recruitment notices on online ASD community websites. The inclusion criteria for participation were (1) 18 years of age or older, (2) diagnosed with ASD by a medical professional, and (3) capable of understanding the purpose of the study and independently participating in the study without any assistance of parents or caregivers. The age of the ASD participants ranged from 20 to 32 years (mean = 23.83, SD = 4.45). As the control group, we also recruited 10 adults without ASD from the authors' institutions through word-of-mouth or emails (22-31 years; male = 6, female = 4). We aimed to clearly identify the user experiences and sensor-related features displayed by people with ASD. We refer to the study participants in the remainder of the manuscript by specifying them as AP (ASD participants) and NP (neurotypical participants; individuals without ASD). A neurotypical person is an individual who does not have any pervasive developmental disorder.

4.2 Apparatus

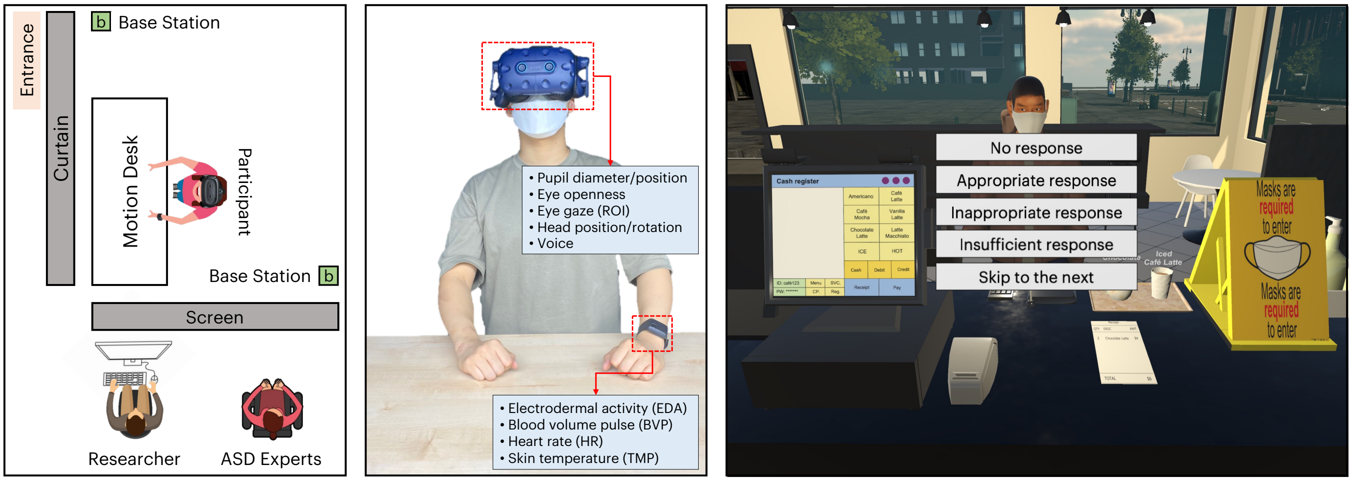
The HTC VIVE Pro Eye VR headset and Empatica E4 wristband [24] were used in the experiment and the virtual environment was run on a Windows 10 PC with an Intel Core i7 and GeForce RTX 2070 graphics card, RAM 16G. We developed the program using the Unity3D engine and the SteamVR plugin.

4.3 Study procedure

Two groups, the AP and the NP, joined the user study under the same study protocol. The researcher started by introducing the study before providing *VISTA*, and collected informed consent from each participant. We then asked the participants for a rehearsal, which covered Phases 1 and 2. The participants played the role of a barista assistant, and the two researchers played the role of a cafe manager and a customer. After a rehearsal, the researcher introduced VR and sensor collection equipment and helped the participants wear it. We gave the AP an additional 1-5 minutes to be familiar with wearing the equipment in consideration of their sensory issues. The researcher helped the participants calibrate view focus and adjust the level of sound and height of the motion desk before starting *VISTA*. The average training time was 8 minutes (min = 6.47, max = 12.13), and we collected sensor data (e.g., eye-tracking, head movement, voice, and physiological signals) through the E4 band and the HMD during the training. After training, the participants completed a short survey and an exit interview. The exit interviews consisted of open-ended questions investigating the overall satisfaction and the challenges regarding *VISTA* (i.e., How real the cafe environment was?, Did you feel dizzy or too stuffy?). The interview took an average of 11 minutes. Figure 4 shows the overall study procedure of the study. This study was approved by the Institutional Review Board (IRB) at the authors' institution (B-2202-736-302).

4.4 Measures

We collected responses to the survey, including the iGroup Presence Questionnaire (IPQ; 14 items with a 7-point Likert scale) [63] and



(a) Experiment environment **(b) Collected sensor data** **(c) Unity game view on a researcher's side**

Figure 3: Study environment settings and the types of sensor data collected during the experiment. Depending on the participant's response, the researcher selects the response type (only visible to the researcher), and the avatar (e.g., a customer) shows the subsequent reaction.

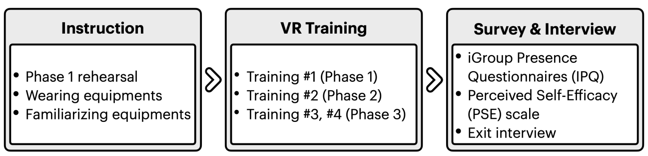


Figure 4: The procedure of the study. After rehearsing Phases 1 and 2 and wearing the equipment, the participants joined the training sessions. We then asked for the survey and exit interview.

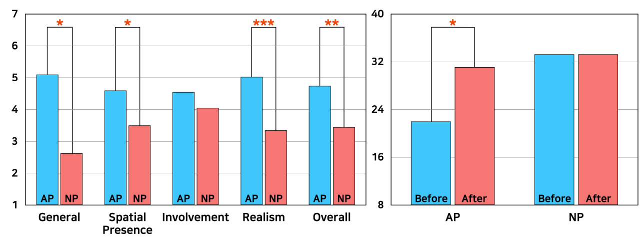
the Perceived Self-Efficacy (PSE) Scale (8 items with a 5-point Likert scale) [80]. The IPQ was used to measure participants' sense of presence in a given virtual environment with 13 questions for four categories: general (1 question), spatial presence (5), involvement (4), and experienced realism (4). The PSE was used to evaluate changes to self-beliefs on one's social skills through VR training. The items of the PSE were developed by referring to [57, 67, 72] based on Bandura's theory of perceived self-efficacy [8]. We asked the participants to complete the PSE before and after the training.

5 RESULTS

We collected eye-tracking, head movement, physiological signals, and voice from 10 AP and 10 NP. We collected physiological signals, including electrodermal activity (EDA), blood volume pulse (BVP), heart rate (HR), and temperature (TMP), through the E4 band [24], which has been widely used in prior research [2, 22, 35, 79].

We also employed the notion of entropy as a metric to investigate the state of randomness and irregularity of the time series data. Entropy has been employed in many studies, and its usefulness has been validated in various domains, including information science, economy, environment, etc. [55, 58, 74, 76]. High entropy means that the data are located sporadically, whereas low entropy means the concentration of similar data. In the entropy formula below, p_i indicates the probability that a particular value will occur.

$$Entropy = -\sum(p_i)log(p_i) \quad (1)$$



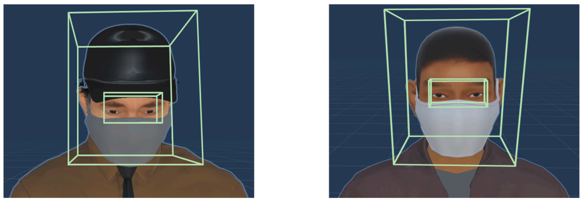
(a) iGroup Presence Questionnaire (IPQ) (b) Perceived Self-Efficacy Scale (PSE)

Figure 5: The survey results of (a) the iGroup Presence Questionnaire (IPQ) and (b) the Perceived Self-Efficacy (PSE) Scale ($^{***}p$ <0.001, $^{**}p$ <0.01, $^{*}p$ <0.05).

We provided participants with three phases of training, and the training complexity increased gradually. We investigated the statistical significance of two groups for each phase to examine the feasibility of training content complying with guidelines [13]. Considering the results of normal distribution from the Shapiro-Wilk test (p >0.05) and homoscedasticity from Levene's test (p >0.05), we performed the independent t-test. Figure 7 illustrates the differences between the two groups for each sensor signal over the three phases. Lastly, for the PSE scale analysis, we performed the paired t-test, since we asked for the PSE scale pre- and post-training. We employed Cohen's d as a measure of the effect size to investigate the strength of the differences between two groups, where 0.20, 0.50 and 0.80 indicate small, medium and large effect sizes, respectively [18].

5.1 Self-evaluations on VISTA

Figure 5-(a) illustrates the results of the IPQ responses. We found that the AP showed a higher level of satisfaction across all perspectives. There were statistically significant differences in general ($t(18)$ = 2.93, d = 1.40), spatial presence ($t(18)$ = 2.92, d = 1.52), realism ($t(18)$ = 4.50, d = 2.29), and overall scores ($t(18)$ = 3.42, d = 1.76). These results confirm that the AP felt greater VR presence



(a) Cafe manager (b) Customer

Figure 6: Box collider in Unity3D

and spatial immersion than the NP and positively perceived the virtual environment provided by *VISTA*.

Regarding the PSE (See Figure 5-(b)), the AP showed a significant difference ($t(18) = 2.64$, $d = 1.15$, $p < 0.05$) before and after the training (before: 22.00, after: 31.10), whereas the NP did not show a significant difference (before: 32.25, after: 32.25). Based on this result, we can conclude that *VISTA* provided the AP with opportunities to feel more confident in social interactions and task completion.

Furthermore, through interviews with participants, we confirmed that an immersive training environment was provided and there was no inconvenience (e.g., cybersickness) while performing the training. *"I didn't feel dizzy at all because I didn't move a lot during the study and usually looked at the front. ... Like the cafe I often go these days, there are seasonal beverage posters on the wall and COVID-19 guidelines on the desk, so I felt a sense of reality." (ASD participant).*

These results confirm the positive perception and experience of people with ASD in training through *VISTA*. Based on these findings, we expect that sensor data that reflect physical behaviors and states can be used as more reliable indicators for understanding people with ASD. In the following sub-sections, we present the results of a comparative analysis of sensor data between the two groups according to task complexity.

5.2 Eye-tracking

5.2.1 Pupil diameter and eye openness. Previous research has shown differences between the AP and the NP in pupil diameter in VR content with the same complexity [10, 45]. As shown in Fig. 7-(a), the averages of the two groups were similar in Phases 1 and 2 (Phase 1: 3.35 (AP); 3.34 (NP), Phase 2: 3.29 (AP); 3.24 (NP)), with no significant difference. However, in Phase 3, the entropy-based t-test results showed a significant difference (right eye: $t(18) = 2.22$, $d = 1.12$; left eye: $t(18) = 2.16$, $d = 1.09$; $p < 0.05$). We also measured the entropy of eye openness. The eye openness consists of a value between 0 and 1, where 0 indicates that the eyelid is completely closed and 1 indicates that the eye is wide open. As shown in Fig. 7-(b), the t-test results showed statistical significance in Phase 3 (right eye: $t(18) = 2.22$, $d = 0.75$; left eye: $t(18) = 2.16$, $d = 0.76$; $p < 0.05$). These results of pupil diameter and eye openness indicate that as the complexity of the required social interaction skills increased, the AP showed more active pupil dilatation/contraction and eyelids movement in both eyes compared to the NP as the complexity of the required social interaction skills increased. This also corresponds to some embarrassed behaviors exhibited by the AP during training (e.g., looking around and saying him/herself, "What should I do?").

5.2.2 Pupil position. The pupil position consists of x and y coordinates. The upper left corner of both lenses is $(x, y) = (0, 0)$, the lower right corner is $(x, y) = (1, 1)$, and the center is $(x, y) = (0.5, 0.5)$. Although there were no noticeable results in Phases 1 and 2 (Fig. 7-(c)), the t-test results support clear patterns with significant differences over the x-coordinates of the pupil position in Phase 3 (x of right eye: $t(18) = 2.36$, $d = 1.19$; x of left eye: $t(18) = 2.36$, $d = 1.19$; $p < 0.05$).

Furthermore, the AP generally showed higher entropy than the NP at all phases, and there was a significant difference in y coordinates in Phase 3 (y of right eye: $t(18) = 2.35$, $d = 1.18$; y of left eye: $t(18) = 2.31$, $d = 1.16$; $p < 0.05$).

5.2.3 Region of interest (ROI)-based gaze rate. One of the characteristics of people with ASD is their difficulty in making eye contact or looking at faces when communicating with others [15, 30]. To measure this characteristic, we used a region of interest (ROI), a visual sample within the dataset identified for a specific purpose. We set the faces of a customer and a manager as ROIs. We used the box collider of Unity3D (See Figure 6) and measured whether a user gazes at ROIs through Tobii G2oM [1], which is a machine learning algorithm that accurately predicts the objects a user focuses on. Fig. 7-(d) shows the percentage of time that the participants stared at the ROI during training. We investigated the percentages for the three timelines (i.e., the time of the entire scenario, the time when the manager speaks, and the time when the customer speaks).

As a result of the t-test, the two groups showed differences in Phases 2 and 3. In Phase 2, all timelines showed significant or marginal significant differences (entire: $t(18) = -2.84$, $d = 1.45$; $p < 0.05$, manager: $t(18) = -2.20$, $d = 1.09$, $p < 0.05$; customer: $t(18) = -1.93$, $d = 1.02$, $p = 0.06$). In Phase 3, there was a significant difference in the timeline of a manager and a customer (manager: $t(18) = -2.26$, $d = 1.10$, $p < 0.05$; customer: $t(18) = -2.02$, $d = 0.96$; $p < 0.05$), and a marginal difference in the timeline of the entire scenario (entire: $t(18) = -1.94$, $d = 0.97$, $p = 0.07$). Overall, the NP showed better eye contact than the AP throughout the entire training period.

5.3 Head movement

5.3.1 Head position. The head position includes values of x (left-right), y (up-down), and z (front-back). In all phases, the entropy of the AP for all directions of the head position was higher than that of the NP (Fig. 7-(e)). In Phases 1 and 2, there were no significant differences. In Phase 3, the t-test results showed significant differences in all axes (x: $t(18) = 2.31$, $d = 1.16$; y: $t(18) = 2.53$, $d = 1.26$; z: $t(18) = 2.23$, $d = 1.12$; $p < 0.05$). These results support empirically noticeable behaviors by the AP during training (e.g., asking the manager for help in customer complaints, moving to the coffee machine in the back, and pretending to make coffee).

5.3.2 Head rotation. Head rotation includes three types: roll, pitch, and yaw. Roll is the rotation to the x-axis and refers to the movement of the head (+/-: nodding forward/tilting one's head back). Pitch is the rotation to the z-axis and refers to the tilt of the head to the left or right (+/-: tilting head to the left/right). Yaw is the rotation to the y-axis, turning the face left or right and looking around (+/-: standing upright and looking to the right/left). As the complexity

[1] https://vr.tobii.com/sdk/solutions/tobii-g2om/

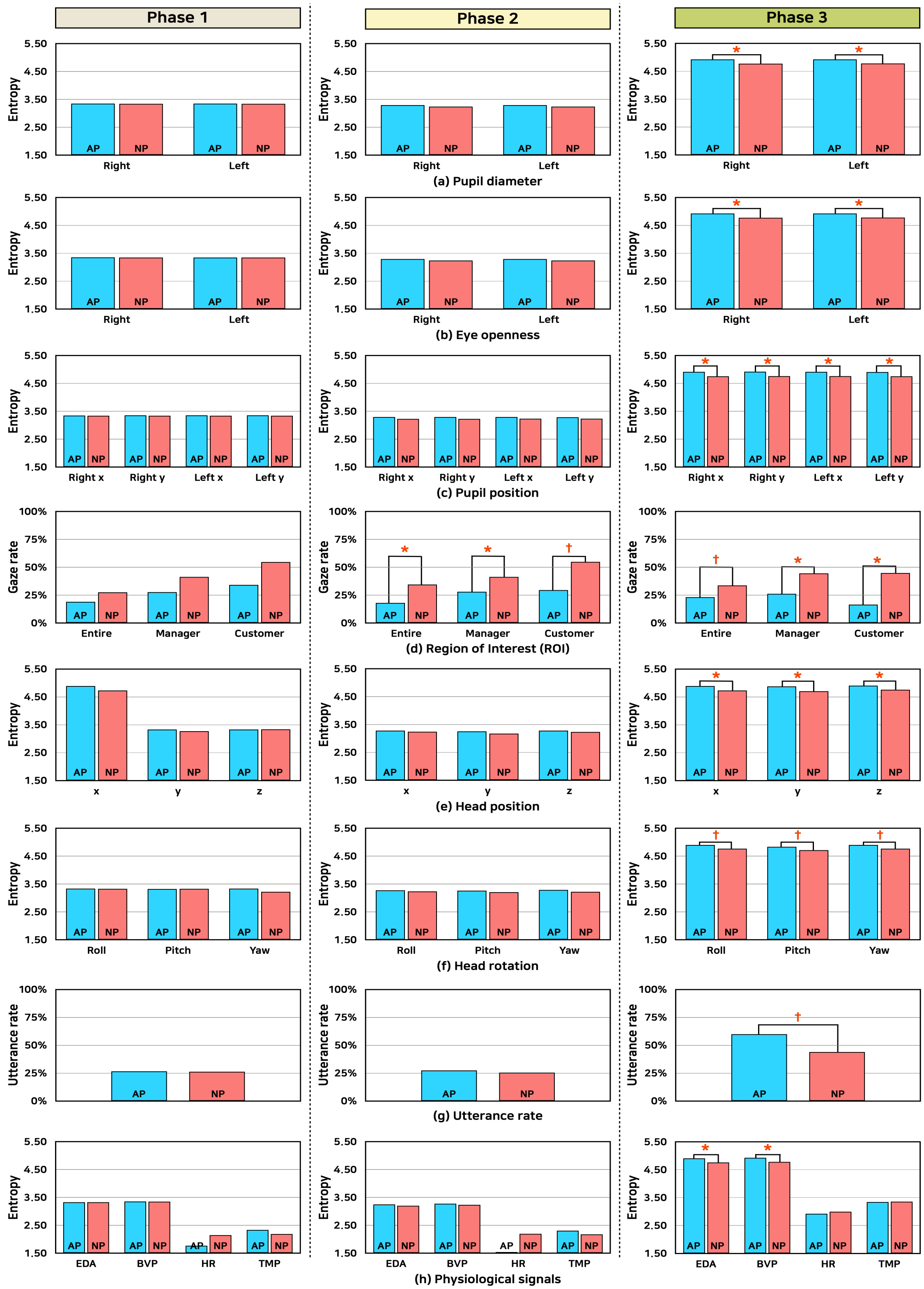


Figure 7: The mean differences in eye-tracking, head movement, physiological signals, and voice data between the AP and the NP across three phases ($^*p$ <0.05, †p <0.1). The complexity of training increases over phases, and the results showed the effectiveness of the *VISTA* design in eliciting the characteristics of people with ASD in VR-based training through a comparison with the NP.

increased, both groups showed a trend of entropy increases similar to the head position (Fig. 7-(f)). There were no significant differences between two groups in Phases 1 and 2. In Phase 3, the t-test showed significant or marginal significant differences between two groups for head rotation (roll: $t(18) = 2.11$, $d = 1.05$, $p = 0.05$; pitch: $t(18) = 1.78$, $d = 0.89$, $p = 0.09$; yaw: $t(18) = 2.03$, $d = 1.01$, $p = 0.06$).

These distinct differences between the two groups in head movements indicate that relatively more physical movements were displayed by the AP in performing the same given task. This may be due to their lack of social interaction skills, for example, exaggerated gestures or behaviors in customer response situations.

5.4 Utterance rate

We collected the participants' voice data using the microphone of the HMD. We extracted and analyzed the speech rate to measure utterances during the participant's responses to the avatars. We first extracted voice amplitudes from the recorded data. We then extracted only the voices of the participants using dBFS, the magnitude of the recorded sound in a digital environment. We empirically checked dBFS from -80 to -20 and set -45 dBFS to the threshold of being silence or not (threshold ≤ -45 means silence), which best extracts not only the spoken voice but also the murmur and filler words (e.g., um, uh, ah, okay). We calculated the utterance rate as the speaking time during the entire training period. We found that the two groups showed a marginally significant difference ($t(18) = 1.89$, $d = 0.19$, $p = 0.09$) in Phase 3. No differences were found in Phases 1 and 2 (Fig. 7-(g)).

Furthermore, we calculated the number of non-silence chunks, based on the same top dB. We confirmed that there was a significant difference between the two groups in Phase 3 ($t(18) = 4.37$, $d = 1.66$, $p < 0.05$). The NP showed more frequent moments of silence than the AP. The NP were silent in situations when they did not need to speak. By contrast, some AP tended to speak echolalia (e.g., repeating the customer's word), speak to themselves, or use filler words before serving a drink.

5.5 Physiological signals

We collected physiological signals to investigate emotional changes according to training stimulation. When stimulated, the electrical signal of the skin changes, which can be interpreted as a change in emotion. Research indicates that EDA and BVP generally increase when an individual becomes nervous [50, 65]. As shown in Fig. 7-(h), t-test results showed significant differences in EDA ($t(18) = 2.24$, $d = 0.53$, $p < 0.05$) and BVP ($t(18) = 2.21$, $d = 0.52$, $p < 0.05$) in Phase 3. There were no significant differences in HR and TMP.

6 DISCUSSION

The originality of *VISTA* is that it reflects the characteristics of people with ASD, follows design guidelines for them, and provides an interactive VR training environment. These aspects have been somewhat neglected or less considered in many previous studies. In our study, the AP showed positive responses regarding their experience in *VISTA*, and we found significant differences in sensor data between the two groups. All sensor data except the gaze rate and the utterance rate were analyzed based on the notion of entropy, which represents the degree of variations of sensor signals and has been used in many prior studies [55, 58, 74, 76] as an evaluation metric. We wanted to verify whether such variations were more prominent by the AP than the NP and well capture characteristics of AP. The next important step is to understand the meaning of the analysis results of the sensor data generated through *VISTA*. We interpret the results by referring to the findings of previous studies. Although our findings may not be directly interpreted relative to the findings in previous studies (because VR systems, VR content, and study conditions are quite different between our study and previous studies), it is possible to assess key sensor types that have been well characterized the AP across studies in different systems and experimental conditions and to obtain insights that have been less considered in previous studies.

6.1 Reaffirmation of ASD characteristics

In our study, three sensor signals – (1) ROI, (2) head rotation, and (3) EDA – showed similar results to those of previous studies.

People with ASD have a common difficulty in making eye contact with others due to their lack of social skills [15, 30]. Many studies have employed ROI and measured the ROI-based gaze rates of AP compared to those of NP [10, 60]. Similarly, in our study, the AP's gaze rate was lower than that of the NP at all phases, and the gaze rates of the two groups showed significant differences in Phases 2 and 3. Regarding head rotations, the AP showed higher results than the NP, similar to previous studies [60, 84]. Further, the EDA level was significantly higher in the AP than in the NP, similar to the results collected in VR-based bus-taking training [65].

The consideration of and similar results shown in those three sensor signals over many studies (including ours) indicate the importance and clear role of those sensor signals in reflecting the characteristics of people with ASD to some extent. Research that has the same goal and methodology as ours may need to consider those sensors. Apparently, with VR training systems that well follow design guidelines for people with ASD, it is also important to consider additional verification of those sensor signals with more subjects and in different training settings. We believe that such efforts are needed to make sensor data analyses useful in clinical assessment and application.

6.2 Additional insights from our study

Additional sensor signals showed significant differences between the two groups in our study. These sensors were either partially (due to limited study conditions) or not covered in previous studies.

First, studies on measuring pupil features have been conducted in a non-training context, such as the evaluation of emotional cognition [26, 71] and the diagnosis of light responses [20, 75] with a short period of time (e.g., 10-20 seconds). However, in our study, we investigated changes in pupil diameter over a few minutes. Two user groups showed significant differences in pupil diameter, pupil position, and eye openness in Phase 3. As the variation of pupil diameter is positively related to the level of cognitive load [36, 44], we could conclude that the AP generally felt higher cognitive load in Phase 3 than the NP. This also further suggests that the AP were more engaged in VR training, as shown in their survey responses.

Second, unlike head rotations, head positions have rarely been used because previous VR training studies on people with ASD did not consider the support of interactive and immersive components, limiting participants' head movement. Since *VISTA* supported participants in performing tasks in an interactive environment, head positions were identified as valid data points, and significant differences were found in Phase 3. For the utterance rate, we confirmed that the AP showed a higher rate than the NP in Phase 3. These findings seem all related and extractable from our study, because the participants were involved in a more flexible, interactive training environment through *VISTA*.

Since *VISTA* is an interactive VR training system that complies with key design guidelines, we discovered new insights into the aforementioned sensor signals. Further, those sensor signals that were highlighted in our study can be collected from the HMD; thus, the accessibility of those signals is high. Future studies are needed to verify the role of these sensor signals in quantifying and understanding the characteristics of people with ASD in different VR training scenarios.

6.3 Limitation and future work

The limitations of our study, which can be considered in future work, are as follows. First, due to the small sample size (10 ASD participants), it may be difficult to derive a generalizable interpretation of the sensor data collected in *VISTA*. While the number of participants in our study is comparable to that of previous studies [2, 40, 60], our study participants may not represent the ASD population; thus need more participants in various training scenarios to make our findings more convincing. Although our training content was decided with a series of discussions with ASD experts, another limitation is the lack of diversity in *VISTA*'s training content. Social skills training for people with ASD can be conducted in various workplace environments (e.g., library, grocery store). Therefore, it is necessary to additionally verify whether the VR training system that follow the design guidelines helps the participant feel immersive and engaged in training and positively affects their self-efficacy and whether the sensor signal results collected in various training situations are analogous to those of our study. We plan to recruit more participants, run the study, validate the findings in this study, and expand the experiment with more training scenarios.

Sensor signals generated during the VR training experience can be used to develop machine learning models for predicting the states of AP in the future. Specifically, if we identify embarrassing moments that people with ASD experience during VR training and develop a model to predict such moments based on the collection and learning of sensor signals, we will be able to more comprehensively understand and support the situation in which people with ASD may suffer. Furthermore, the results of sensor signals can be used as important information for medical professionals or caregivers to identify the physical or psychological changes in people with ASD.

7 CONCLUSION

This paper presents *VISTA*, a VR-based interactive social skills training system for people with ASD. To investigate the feasibility of *VISTA* that built on VR system guidelines for people with ASD, we conducted a user study with 10 people with ASD and 10 neurotypical people. Our study results revealed that *VISTA* provided positive training experiences, and the sensor signals collected through *VISTA* well described ASD characteristics. These results confirmed the effective design of *VISTA* in offering gradual task complexities, immersive VR experiences, and interactive components. Furthermore, we expanded our understanding of sensor signals generated by people with ASD during VR training. In summary, our study emphasizes the importance of identifying and following key design guidelines for the development of VR training systems. Our study's methodology and findings can be used to guide other studies in designing a tool to support social independence for people with ASD or people who are in need.

ACKNOWLEDGMENTS

This research was supported by the National Research Foundation (2021M3A9E4080780) and Institute for Information & Communication Technology Planning & Evaluation (IITP-2020-0-01373, IITP-2022-2018-0-01431, IITP-2022-2022-0-00240).

REFERENCES

[1] Dane Acena and Guo Freeman. 2021. "In My Safe Space": Social Support for LGBTQ Users in Social Virtual Reality. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–6.

[2] Deeksha Adiani, Aaron Itzkovitz, Dayi Bian, Harrison Katz, Michael Breen, Spencer Hunt, Amy Swanson, Timothy J Vogus, Joshua Wade, and Nilanjan Sarkar. 2022. Career interview readiness in virtual reality (CIRVR): a platform for simulated interview training for autistic individuals and their employers. *ACM Transactions on Accessible Computing (TACCESS)* 15, 1 (2022), 1–28.

[3] Ali Adjorlu, Emil Rosenlund Høeg, Luca Mangano, and Stefania Serafin. 2017. Daily living skills training in virtual reality to help children with autism spectrum disorder in a real shopping scenario. In *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*. IEEE, 294–302.

[4] Zeeshan Akhtar and Tanaya Guha. 2019. Computational analysis of gaze behavior in autism during interaction with virtual agents. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 1075–1079.

[5] The Autism Society Philippines (ASP). 2017. *Barista Training for Adults with Autism*. https://bit.ly/3zq0PsT

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

[7] Autism Spectrum Australia. 2022. *Coffee anyone? - Aspect's Central Coast receives funding for new education and employment program*. https://bit.ly/3cEA7F3

[8] Albert Bandura. 2012. On the functional properties of perceived self-efficacy revisited. *Journal of management* 38, 1 (2012), 9–44.

[9] Esubalew Bekele, Joshua Wade, Dayi Bian, Jing Fan, Amy Swanson, Zachary Warren, and Nilanjan Sarkar. 2016. Multimodal adaptive social interaction in virtual environment (MASI-VR) for children with Autism spectrum disorders (ASD). In *2016 IEEE virtual reality (VR)*. IEEE, 121–130.

[10] Esubalew Bekele, Zhi Zheng, Amy Swanson, Julie Crittendon, Zachary Warren, and Nilanjan Sarkar. 2013. Understanding how adolescents with autism respond to facial expressions in virtual reality environments. *IEEE transactions on visualization and computer graphics* 19, 4 (2013), 711–720.

[11] Olga Bogdashina. 2016. *Sensory perceptual issues in autism and asperger syndrome: different sensory experiences-different perceptual worlds*. Jessica Kingsley Publishers.

[12] Lal Bozgeyikli, Evren Bozgeyikli, Andrew Raij, Redwan Alqasemi, Srinivas Katkoori, and Rajiv Dubey. 2017. Vocational rehabilitation of individuals with autism spectrum disorder with virtual reality. *ACM Transactions on Accessible Computing (TACCESS)* 10, 2 (2017), 1–25.

[13] Lal Bozgeyikli, Andrew Raij, Srinivas Katkoori, and Redwan Alqasemi. 2017. A survey on virtual reality for individuals with autism spectrum disorder: design considerations. *IEEE Transactions on Learning Technologies* 11, 2 (2017), 133–151.

[14] Lal" Lila" Bozgeyikli, Evren Bozgeyikli, Srinivas Katkoori, Andrew Raij, and Redwan Alqasemi. 2018. Effects of virtual reality properties on user experience of individuals with autism. *ACM Transactions on Accessible Computing (TACCESS)* 11, 4 (2018), 1–27.

[15] Vincent J Carbone, Leigh O'Brien, Emily J Sweeney-Kerwin, and Kristin M Albert. 2013. Teaching eye contact to children with autism: A conceptual analysis and single case study. *Education and treatment of children* 36, 2 (2013), 139–159.

[16] Yufang Cheng, Cheng-Li Huang, and Chung-Sung Yang. 2015. Using a 3D immersive virtual environment system to enhance social understanding and social skills for children with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities* 30, 4 (2015), 222–236.

[17] S Chrisilla, Tharun R Ragav, S Vidhusha, and A Kavitha. 2021. Investigating Cognitive Global Coordination in normal and autistic children using virtual reality environments–An EEG Study. In *2021 10th International IEEE/EMBS Conference on Neural Engineering (NER)*. IEEE, 1016–1019.

[18] Jacob Cohen. 1988. The effect size. *Statistical power analysis for the behavioral sciences* (1988), 77–83.

[19] Daniel J Cox, Timothy Brown, Veerle Ross, Matthew Moncrief, Rose Schmitt, Gary Gaffney, and Ron Reeve. 2017. Can youth with autism spectrum disorder use virtual reality driving simulation training to evaluate and improve driving performance? An exploratory study. *Journal of Autism and Developmental Disorders* 47, 8 (2017), 2544–2555.

[20] Chathuri Daluwatte, Judith H Miles, Shawn E Christ, David Q Beversdorf, T Nicole Takahashi, and Gang Yao. 2013. Atypical pupillary light reflex and heart rate variability in children with autism spectrum disorder. *Journal of autism and developmental disorders* 43, 8 (2013), 1910–1925.

[21] Rosaria De Luca, Antonino Naro, Pia Valentina Colucci, Federica Pranio, Giuseppe Tardiolo, Luana Billeri, Maria Le Cause, Carmela De Domenico, Simona Portaro, Giuseppe Rao, et al. 2021. Improvement of brain functional connectivity in autism spectrum disorder: an exploratory study on the potential use of virtual reality. *Journal of Neural Transmission* 128, 3 (2021), 371–380.

[22] Cédric Di Loreto, Jean-Rémy Chardonnet, Julien Ryard, and Alain Rousseau. 2018. Woah: A virtual reality work-at-height simulator. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 281–288.

[23] Renee R Dijkhuis, Tim Ziermans, Sophie van Rijn, Wouter Staal, and Hanna Swaab. 2019. Emotional arousal during social stress in young adults with autism: Insights from heart rate, heart rate variability and self-report. *Journal of autism and developmental disorders* 49, 6 (2019), 2524–2535.

[24] Empatica. 2021. E4 Wristband. Real-Time Physiological Data Streaming and Visualization. (2021).

[25] Noah Glaser and Matthew Schmidt. 2022. Systematic literature review of virtual reality intervention design patterns for individuals with autism spectrum disorders. *International Journal of Human–Computer Interaction* 38, 8 (2022), 753–788.

[26] Katherine O Gotham, Greg J Siegle, Gloria T Han, Andrew J Tomarken, Rachel N Crist, David M Simon, and James W Bodfish. 2018. Pupil response to social-emotional material is associated with rumination and depressive symptoms in adults with autism spectrum disorder. *PloS one* 13, 8 (2018), e0200340.

[27] Temple Grandin. 2002. Teaching tips for children and adults with autism. *Fort Collins* (2002).

[28] Javier Hernandez, Ivan Riobo, Agata Rozga, Gregory D Abowd, and Rosalind W Picard. 2014. Using electrodermal activity to recognize ease of engagement in children during social interactions. In *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing*. 307–317.

[29] Jorge Fernández Herrero and Gonzalo Lorenzo. 2020. An immersive virtual reality educational intervention on people with autism spectrum disorders (ASD) for the development of communication skills and problem solving. *Education and Information Technologies* 25, 3 (2020), 1689–1722.

[30] Bogseon Hwang and Carolyn Hughes. 2000. The effects of social interactive training on early social communicative skills of children with autism. *Journal of autism and developmental disorders* 30, 4 (2000), 331–343.

[31] Seunggon Jeon, Seungwon Paik, Ungyeon Yang, Patrick C Shih, and Kyungsik Han. 2021. The More, the Better? Improving VR Firefighting Training System with Realistic Firefighter Tools as Controllers. *Sensors* 21, 21 (2021), 7193.

[32] Seung-Gon Jeon, Jaeho Han, Yonggeol Jo, and Kyungsik Han. 2019. Being more focused and engaged in firefighting training: Applying user-centered design to VR system development. In *25th ACM Symposium on Virtual Reality Software and Technology*. 1–11.

[33] Weina Jin, Jianyu Fan, Diane Gromala, and Philippe Pasquier. 2018. Automatic prediction of cybersickness for virtual reality games. In *2018 IEEE Games, Entertainment, Media Conference (GEM)*. IEEE, 1–9.

[34] Young-Hwan Jun, Yong-Hyeon Nam, Eona Kim, Myeng-Mo Yeon, and Min-Su Byeon. 2014. *Vocational Training of Persons with Disabilities in Korea*. https://bit.ly/3S3kmru

[35] Sungchul Jung, Richard Li, Ryan McKee, Mary C Whitton, and Robert W Lindeman. 2021. Floor-vibration vr: mitigating cybersickness using whole-body tactile stimuli in highly realistic vehicle driving experiences. *IEEE Transactions on Visualization & Computer Graphics* 27, 05 (2021), 2669–2680.

[36] Daniel Kahneman, W Scott Peavler, and Linda Onuska. 1968. Effects of verbalization and incentive on the pupil response to mental activity. *Canadian Journal of Psychology/Revue canadienne de psychologie* 22, 3 (1968), 186.

[37] Michelle R Kandalaft, Nyaz Didehbani, Daniel C Krawczyk, Tandra T Allen, and Sandra B Chapman. 2013. Virtual reality social cognition training for young adults with high-functioning autism. *Journal of autism and developmental disorders* 43, 1 (2013), 34–44.

[38] Sam Kavanagh, Andrew Luxton-Reilly, Burkhard Wuensche, and Beryl Plimmer. 2017. A systematic review of virtual reality in education. *Themes in Science and Technology Education* 10, 2 (2017), 85–119.

[39] Fengfeng Ke and Sungwoong Lee. 2016. Virtual reality based collaborative design by children with high-functioning autism: Design-based flexibility, identity, and norm construction. *Interactive Learning Environments* 24, 7 (2016), 1511–1533.

[40] Fengfeng Ke, Jewoong Moon, and Zlatko Sokolikj. 2022. Virtual reality–based social skills training for children with autism spectrum disorder. *Journal of Special Education Technology* 37, 1 (2022), 49–62.

[41] Bogoan Kim, Daehyoung Lee, Aehong Min, Seungwon Paik, Georgia Frey, Scott Bellini, Kyungsik Han, and Patrick C Shih. 2020. PuzzleWalk: A theory-driven iterative design inquiry of a mobile game for promoting physical activity in adults with autism spectrum disorder. *PLoS One* 15, 9 (2020), e0237966.

[42] Jack A Kincaid, Fengchen Gong, Tianjie Jia, Hong Z Tan, Casey Kohr, and Gary Bertoline. 2021. A Study on VR Training of Baseball Athletes. In *International Conference on Human-Computer Interaction*. Springer, 187–204.

[43] Evdokimos I Konstantinidis, Magda Hitoglou-Antoniadou, Andrej Luneski, Panagiotis D Bamidis, and Maria M Nikolaidou. 2009. Using affective avatars and rich multimedia content for education of children with autism. In *Proceedings of the 2nd international conference on pervasive technologies related to assistive environments*. 1–6.

[44] Krzysztof Krejtz, Andrew T Duchowski, Anna Niedzielska, Cezary Biele, and Izabela Krejtz. 2018. Eye tracking cognitive load using pupil diameter and microsaccades with fixed gaze. *PloS one* 13, 9 (2018), e0203629.

[45] Uttama Lahiri, Esubalew Bekele, Elizabeth Dohrmann, Zachary Warren, and Nilanjan Sarkar. 2012. Design of a virtual reality based adaptive response technology for children with autism. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 21, 1 (2012), 55–64.

[46] Liron Lamash, Evelyne Klinger, and Naomi Josman. 2017. Using a virtual supermarket to promote independent functioning among adolescents with Autism Spectrum Disorder. In *2017 International Conference on Virtual Rehabilitation (ICVR)*. IEEE, 1–7.

[47] Daehyoung Lee, Georgia C Frey, Aehong Min, Bogoan Kim, Donetta J Cothran, Scott Bellini, Kyungsik Han, and Patrick C Shih. 2020. Usability inquiry of a gamified behavior change app for increasing physical activity and reducing sedentary behavior in adults with and without autism spectrum disorder. *Health Informatics Journal* 26, 4 (2020), 2992–3008.

[48] Susan R Leekam, Margot R Prior, and Mirko Uljarevic. 2011. Restricted and repetitive behaviors in autism spectrum disorders: a review of research in the last decade. *Psychological bulletin* 137, 4 (2011), 562.

[49] Phil Lopes, Nana Tian, and Ronan Boulic. 2020. Eye thought you were sick! exploring eye behaviors for cybersickness detection in vr. In *Motion, Interaction and Games*. 1–10.

[50] Dong Wei Lu, Jenn Wu Wang, and Andrew Chih Wei Huang. 2010. Differentiation of Internet addiction risk level based on autonomic nervous responses: the Internet-addiction hypothesis of autonomic activity. *Cyberpsychology, Behavior, and Social Networking* 13, 4 (2010), 371–378.

[51] Joseph P McCleery, Ashley Zitter, Rita Solórzano, Sinan Turnacioglu, Judith S Miller, Vijay Ravindran, and Julia Parish-Morris. 2020. Safety and feasibility of an immersive virtual reality intervention program for teaching police interaction skills to adolescents and adults with autism. *Autism Research* 13, 8 (2020), 1418–1424.

[52] Patricia Mesa-Gresa, Hermenegildo Gil-Gómez, José-Antonio Lozano-Quilis, and José-Antonio Gil-Gómez. 2018. Effectiveness of virtual reality for children and adolescents with autism spectrum disorder: an evidence-based systematic review. *Sensors* 18, 8 (2018), 2486.

[53] Peter Mundy, Kwanguk Kim, Nancy McIntyre, Lindsay Lerro, and William Jarrold. 2016. Brief report: Joint attention and information processing in children with higher functioning autism spectrum disorders. *Journal of autism and developmental disorders* 46, 7 (2016), 2555–2560.

[54] Michael W Norris, Kristen Spicer, and Traci Byrd. 2019. Virtual reality: the new pathway for effective safety training. *Professional Safety* 64, 06 (2019), 36–39.

[55] Walid Osamy, Ahmed Salim, and Ahmed M Khedr. 2020. An information entropy based-clustering algorithm for heterogeneous wireless sensor networks. *Wireless networks* 26, 3 (2020), 1869–1886.

[56] Nikolay Pavlov. 2014. User interface for people with autism spectrum disorders. *Journal of Software Engineering and Applications* 2014 (2014).

[57] Monica Pedrazza, Elena Trifiletti, Sabrina Berlanda, and Gian Antonio Di Bernardo. 2013. Self-efficacy in social work: Development and initial validation of the self-efficacy scale for social workers. *Social Sciences* 2, 3 (2013), 191–207.

[58] Xue-Yang Pei, Ting-Hua Yi, Chun-Xu Qu, and Hong-Nan Li. 2019. Conditional information entropy based sensor placement method considering separated model error and measurement noise. *Journal of Sound and Vibration* 449 (2019), 389–404.

[59] Nimesha Ranasinghe, Pravar Jain, Nguyen Thi Ngoc Tram, Koon Chuan Raymond Koh, David Tolley, Shienny Karwita, Lin Lien-Ya, Yan Liangkun, Kala Shamaiah, Chow Eason Wai Tung, et al. 2018. Season traveller: Multisensory narration for

enhancing the virtual reality experience. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–13.
[60] Marta Robles, Negar Namdarian, Julia Otto, Evelyn Wassiljew, Nassir Navab, Christine M Falter-Wagner, and Daniel Roth. 2022. A Virtual Reality Based System for the Screening and Classification of Autism. *IEEE Trans. Vis. Comput. Graph.* 28, 5 (2022), 2168–2178.
[61] Mario Saiano, Laura Pellegrino, Maura Casadio, Susanna Summa, Eleonora Garbarino, Valentina Rossi, Daniela Dall'Agata, and Vittorio Sanguineti. 2015. Natural interfaces and virtual environments for the acquisition of street crossing and path following skills in adults with Autism Spectrum Disorders: a feasibility study. *Journal of neuroengineering and rehabilitation* 12, 1 (2015), 1–13.
[62] Harini Sampath, Ravi Agarwal, and Bipin Indurkhya. 2013. Assistive technology for children with autism-lessons for interaction design. In *Proceedings of the 11th Asia Pacific Conference on Computer Human Interaction*. 325–333.
[63] Thomas W Schubert. 2003. The sense of presence in virtual environments: A three-component scale measuring spatial presence, involvement, and realness. *Z. für Medienpsychologie* 15, 2 (2003), 69–71.
[64] Trisha Self, Rosalind R Scudder, Gamal Weheba, and Daiquirie Crumrine. 2007. A virtual approach to teaching safety skills to children with autism spectrum disorder. *Topics in Language disorders* 27, 3 (2007), 242–253.
[65] Marco Simões, Miguel Bernardes, Fernando Barros, Miguel Castelo-Branco, et al. 2018. Virtual travel training for autism spectrum disorder: proof-of-concept interventional study. *JMIR serious games* 6, 1 (2018), e8428.
[66] Christian Sinnott, James Liu, Courtney Matera, Savannah Halow, Ann Jones, Matthew Moroz, Jeffrey Mulligan, Michael Crognale, Eelke Folmer, and Paul MacNeilage. 2019. Underwater virtual reality system for neutral buoyancy training: Development and evaluation. In *25th ACM Symposium on Virtual Reality Software and Technology*. 1–9.
[67] Heather M Smith and Nancy E Betz. 2000. Development and validation of a scale of perceived social self-efficacy. *Journal of career assessment* 8, 3 (2000), 283–301.
[68] Matthew J Smith, Emily J Ginger, Katherine Wright, Michael A Wright, Julie Lounds Taylor, Laura Boteler Humm, Dale E Olsen, Morris D Bell, and Michael F Fleming. 2014. Virtual reality job interview training in adults with autism spectrum disorder. *Journal of autism and developmental disorders* 44, 10 (2014), 2450–2463.
[69] Natalia Stewart Rosenfield, Kathleen Lamkin, Jennifer Re, Kendra Day, LouAnne Boyd, and Erik Linstead. 2019. A virtual reality system for practicing conversation skills for children with autism. *Multimodal Technologies and Interaction* 3, 2 (2019), 28.
[70] Dorothy Strickland. 1997. Virtual reality for the treatment of autism. *Virtual reality in neuro-psycho-physiology* (1997), 81–86.
[71] Sai Sun, Paula J Webster, Yu Wang, Hongbo Yu, Rongjun Yu, and Shuo Wang. 2022. Reduced Pupil Oscillation During Facial Emotion Judgment in People with Autism Spectrum Disorder. *Journal of autism and developmental disorders* (2022), 1–11.
[72] Connie Sung, Annemarie Connor, June Chen, Chien-Chun Lin, Hung-Jen Kuo, and Jina Chun. 2019. Development, feasibility, and preliminary efficacy of an employment-related social skills intervention for young adults with high-functioning autism. *Autism* 23, 6 (2019), 1542–1553.
[73] Yonghang Tai, Lei Wei, Minhui Xiao, Hailing Zhou, Qiong Li, Junsheng Shi, and Saeid Nahavandi. 2018. A high-immersive medical training platform using direct intraoperative data. *IEEE access* 6 (2018), 69438–69452.
[74] Jimmy Tjen, Francesco Smarra, and Alessandro D'Innocenzo. 2020. An entropy-based sensor selection algorithm for structural damage detection. In *2020 IEEE 16th International Conference on Automation Science and Engineering (CASE)*. IEEE, 1566–1571.
[75] Chiara Tortelli, Marco Turi, David C Burr, and Paola Binda. 2021. Pupillary responses obey Emmert's law and co-vary with autistic traits. *Journal of Autism and Developmental Disorders* 51, 8 (2021), 2908–2919.
[76] Furui Wang, Zheng Chen, and Gangbing Song. 2020. Monitoring of multi-bolt connection looseness using entropy-based active sensing and genetic algorithm-based least square support vector machine. *Mechanical Systems and Signal Processing* 136 (2020), 106507.
[77] Casey Weldon. 2021. *Just Brew: Coffee shop offers jobs, training to people with developmental disabilities*. https://bit.ly/3PyPugX
[78] Yee-Pay Wuang, Chien-Ling Huang, and Hsien-Yu Tsai. 2020. Sensory integration and perceptual-motor profiles in school-aged children with autistic spectrum disorder. *Neuropsychiatric Disease and Treatment* 16 (2020), 1661.
[79] Tong Xue, Abdallah El Ali, Tianyi Zhang, Gangyi Ding, and Pablo Cesar. 2021. CEAP-360VR: A Continuous Physiological and Behavioral Emotion Annotation Dataset for 360 VR Videos. *IEEE Transactions on Multimedia* (2021).
[80] Hee-Jeong Yoo, Geonho Bahn, In-Hee Cho, Eun-Kyung Kim, Joo-Hyun Kim, Jung-Won Min, Won-Hye Lee, Jun-Seong Seo, Sang-Shin Jun, Guiyoung Bong, et al. 2014. A Randomized Controlled Trial of the K orean Version of the PEERS® Parent-Assisted Social Skills Training Program for Teens With ASD. *Autism Research* 7, 1 (2014), 145–161.
[81] Lian Zhang, Joshua Wade, Dayi Bian, Jing Fan, Amy Swanson, Amy Weitlauf, Zachary Warren, and Nilanjan Sarkar. 2017. Cognitive load measurement in a virtual reality-based driving system for autism intervention. *IEEE transactions on affective computing* 8, 2 (2017), 176–189.
[82] Dong Zhao and Jason Lucas. 2015. Virtual reality simulation for construction safety promotion. *International journal of injury control and safety promotion* 22, 1 (2015), 57–67.
[83] Wenbing Zhao, Xiongyi Liu, Tie Qiu, and Xiong Luo. 2020. Virtual avatar-based life coaching for children with autism spectrum disorder. *Computer* 53, 2 (2020), 26–34.
[84] Zhong Zhao, Zhipeng Zhu, Xiaobin Zhang, Haiming Tang, Jiayi Xing, Xinyao Hu, Jianping Lu, Qiongling Peng, and Xingda Qu. 2021. Atypical Head Movement during Face-to-Face Interaction in Children with Autism Spectrum Disorder. *Autism Research* 14, 6 (2021), 1197–1208.