V-DAT (Virtual Reality Data Analysis Tool): Supporting Self-Awareness for Autistic People from Multimodal VR Sensor Data

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

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

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

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

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

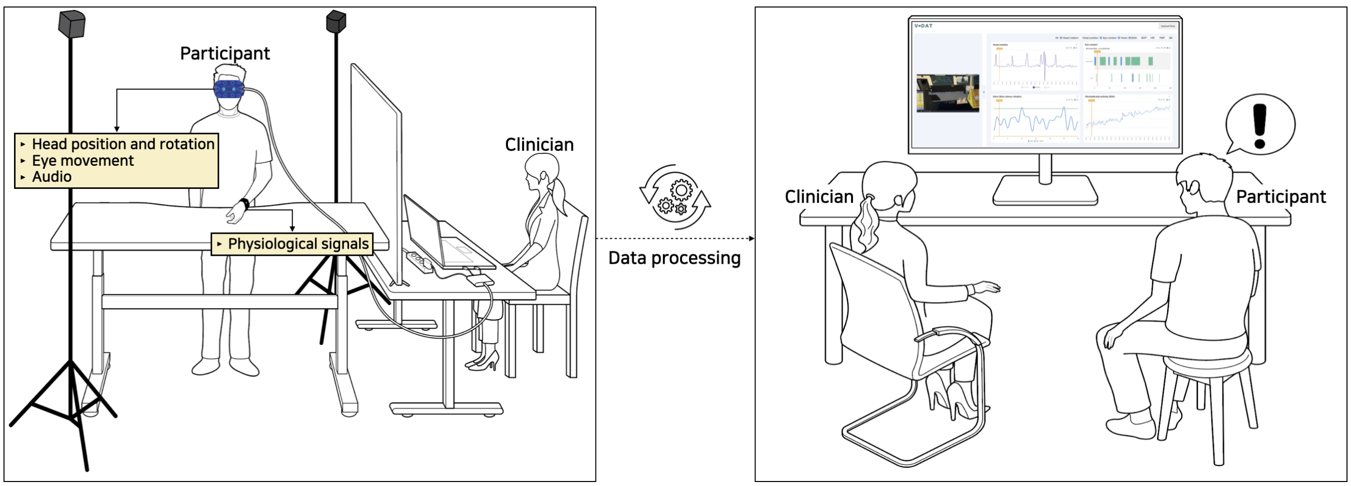

Figure 1: A use case example of V-DAT (Virtual Reality Data Analysis Tool). Autistic people can experience VR training content, and autism experts can provide sensor data-based training analysis results to autistic people in a holistic manner. The visualizations of the four sensor modalities collected from the VR content are provided, synchronized with the training video.

ABSTRACT

Virtual reality (VR) has become a valuable tool for social and educational purposes for autistic people, as it provides flexible environmental support to create a variety of experiences. A growing body of recent research has examined the behaviors of autistic people using sensor-based data to better understand autistic people and investigate the effectiveness of VR. Comprehensive analysis of the various signals that can be easily collected in the VR environment can promote understanding of autistic people. While this quantitative evidence has the potential to help both autistic people and others (e.g., autism experts) to understand behaviors of autistic people, existing studies have focused on single signal analysis and have not determined the acceptability of signal analysis results from the autistic person's point of view. To facilitate the use of multiple sensor signals in VR for autistic people and experts, we introduce *V-DAT* (Virtual Reality Data Analysis Tool), designed to support a VR sensor data handling pipeline. *V-DAT* takes into account four sensor modalities—head position and rotation, eye movement, audio, and physiological signals—that are actively used in current VR research for autistic people. We explain the characteristics and processing methods of the data for each modality as well as the analysis with comprehensive visualizations of *V-DAT*. We also conduct a case study to investigate the feasibility of *V-DAT* as a way of broadening understanding of autistic people from the perspectives of both autistic people and autism experts. Finally, we discuss issues with the process of *V-DAT* development and complementary measures for the applicability and scalability of a sensor data management system for autistic people.

*Corresponding author

UIST '23, October 29-November 1, 2023, San Francisco, CA, USA

ACM ISBN 979-8-4007-0132-0/23/10…$15.00
https://doi.org/10.1145/3586183.3606797

CCS CONCEPTS

• **Human-centered computing** → **Accessibility systems and tools**; *Visualization systems and tools.*

KEYWORDS

autism, virtual reality, accessibility, data handling pipeline

ACM Reference Format:

Bogoan Kim, Dayoung Jeong, Jennifer G Kim, Hwajung Hong, and Kyungsik Han. 2023. V-DAT (Virtual Reality Data Analysis Tool): Supporting Self-Awareness for Autistic People from Multimodal VR Sensor Data. In *The 36th Annual ACM Symposium on User Interface Software and Technology (UIST '23), October 29-November 1, 2023, San Francisco, CA, USA.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3586183.3606797

1 INTRODUCTION

Virtual Reality (VR) provides a flexible environment that allows users to experience immersion without spatial or temporal constraints. By virtue of this, most recent VR studies have focused on its use for specific purposes, including vocational education (e.g., industrial field practice, medical surgery practice, and firefighter training) and personal competency improvement (e.g., interview or presentation skills) [6, 30, 63, 66]. VR users can reflect on their behaviors and decisions and receive feedback on them through repetitive learning of training/education programs [32].

Such opportunities have been extended to autistic people [1] [12, 25]. Autism is a complex developmental condition involving social interaction challenges with specific patterns of repetitive behaviors [4]. Research on autistic people has developed various types of VR content to support self-help skills (e.g., activities of daily living skills [2, 43, 60] and driving [19]) or social skills (e.g., facial expression recognition [8, 16], job interview [63], and social interaction [33, 55, 64]) needed to support one's social independence.

Recently, studies on autism have started to consider body responses (e.g., physiological signals, body/head movement, and eye movement) of autistic people during VR experiences and employ the data to analyze and model various aspects of their behaviors or conditions [21, 28, 55]. For example, studies have considered changes in the autistic user's eye movement (e.g., focus point and gaze movement) based on the data collected from an eye-tracking device, those in users' head or body movement collected from a head-mounted display (HMD) or a tracker, and those in users' physiological signals (e.g., electrodermal activity (EDA) and blood volume pulse (BVP)) collected from a wearable device while the user interacts with VR content [7, 8, 42, 55, 60]. Furthermore, with the collected data, studies have tried to understand the relationships between the collected sensor data and several user aspects, such as stress, engagement, and anxiety, through data-driven analysis [9, 35, 60, 72]. These findings can help understand design choices in the virtual environment and improve VR content and systems better suitable for autistic people.

Despite the potential benefits, there are several challenges in using VR technology as a tool that gives autistic people a better understanding of themselves and the opportunity to advocate for themselves. To accurately interpret behavior patterns, it is crucial to understand the types, collections, and analysis methodologies of data that represent every observable response to internal and external stimuli. Many recent studies on VR for autistic people have proposed methods to utilize and analyze various types of sensor data [1, 7, 8, 42, 50, 55, 56, 67, 72, 74]. However, previous studies have only collected and analyzed one or two sensor modalities. Even studies that collected two sensor modalities analyzed the data from each modality individually, without systematically integrating the fragmented sensor signals. Although focusing on a single data modality yields valuable research insights for understanding autistic people, VR technology offers convenient access to multiple data modalities via a standard HMD and an additional device that poses no inconvenience to autistic people. By comprehensively analyzing multiple types of data in the same VR session, we can gain a more holistic understanding of the unique characteristics of each modality and explore ways to better leverage sensor data to understand autistic people. It is thus important to collect sensor signals that are accessible from the device as autistic people interact within the VR training content, examine the relationship between each signal, and present comprehensive analysis results to the training participants, autism experts, and other neurotypical people. This approach would give autistic people the opportunity to reflect on their behaviors/conditions and advocate for themselves with more evidence. Moreover, this would allow experts to gain a deeper understanding of the behavioral characteristics that emerged during the training and to support the improvement of challenging behaviors of autistic people from diverse perspectives. Despite the research and clinical significance, there has been little research into building a pipeline that addresses multiple data modalities for autistic people in the context of VR.

In this paper, we present *V-DAT* (Virtual Reality Data Analysis Tool), which is integrated with a pipeline for handling and analyzing data collected from autistic people (Figure 1). *V-DAT* deals with four representative data types – head position and rotation, eye movement, audio, and physiological signals – that have been widely used in prior VR studies for autistic people [1, 7, 8, 35, 41, 42, 74]. *V-DAT* is designed to not only collect and store the data but also comprehensively visualize and analyze the data. We provide a detailed explanation of the sensor data used in VR research and the structure and characteristics of *V-DAT*. Moreover, through a case study of 20 autistic people, we empirically demonstrate the feasible use of *V-DAT* and discuss the applicability of *V-DAT* in clinical settings with autism experts.

Our work has the following contributions:

- We articulate the characteristics, management, feature extraction, and analysis methods for the data from the four sensor modalities most widely used in VR research for autistic people.
- We present *V-DAT*, which aims to provide a comprehensive analysis of sensor data collected during the VR experience of autistic people for autism experts' effective clinical judgment.
- We demonstrate the effectiveness of the design of *V-DAT* in eliciting quantitative evidence of user experience in VR that is helpful to autistic people and experts and discuss key elements for using the VR sensor data handling pipeline.

[1]We use identity-first language due to a reported preference of autistic people [34] and recent movement in academia [10].

In summary, *V-DAT* aims to provide key information on the use of multimodal sensor data in a VR environment. It is expected that the data handling pipeline by *V-DAT* will help autism experts who wish to study data provided by sensors and VR as a tool to examine sensor patterns, prepare data analysis methods, and expand their understanding of autistic people. If the basic framework of research is established by *V-DAT*, autism experts and other neurotypical people will have an opportunity to broaden their understanding of autistic people and expand their studies by considering additional sensor data modalities or developing new visualizations.

2 RELATED WORK

2.1 Sensor Signals Used in VR for Autistic People

Much research has been conducted using various types of sensor data to understand autism characteristics and challenging behaviors better through VR systems. In this section, we summarize four types of sensor data commonly used in many VR studies for autistic people: head position and rotation, eye movement, audio, and physiological signals (see Table 1 for details). We also highlight the need to establish a sensor data handling pipeline in VR research.

2.1.1 Head position and rotation. Head position and rotation represent the user's physical head in the virtual space through the 3-DoF (degree of freedom) position and rotation. This data can be obtained from an HMD. Head movement has been employed as an important feature that shows autistic people's unique characteristics when experiencing VR content that requires social communication and interaction skills. For example, Simões et al. [61] collected head position data to compare the interpersonal distance of autistic adolescents to that of their typically developing peers. The results demonstrated that autistic adolescents had difficulties perceiving loss of non-verbal communication (e.g., body gestures) that caused the changed regulation of interpersonal distance. Similarly, Robles et al. [55] identified that head movement was a significant biomarker and developed a machine learning model to screen autism using head movement as one of the main features.

2.1.2 Eye movement. Eye-tracking data are generated based on the movement of the pupil of the eyes. Some devices used for this purpose include Tobii [2] and Pupil labs [3]. Recently, some HMD, such as HTC Vive Pro Eye [4] and Meta Quest Pro [5] have begun to support eye-tracking. Eye-tracking data from autistic people in VR offer ample opportunities, allowing researchers to more accurately notify the diverse socially challenging moments for autistic people through analysis of the region of interest (ROI). Eye-tracking data help researchers better understand their characteristics and find ways to improve autistic people's social skills by examining the interaction between the autistic person and the objects and content in the VR environment [1, 7, 8, 42, 50, 55, 56, 67, 72, 74]. Using eye-tracking data, Bekele et al. [8] confirmed a difference in how autistic people process and recognize emotional faces. Lahiri et al. [42] developed a VR system that supports social skill improvement for autistic people.

2.1.3 Audio. Audio data provide a meaningful perspective in reflecting the user's experience [15, 48, 69, 70], as each user's auditory stimuli may differ even when using the same VR content. With the increasing prevalence of AI technology, audio data can be presented or embedded in 2D space and used as input vectors for developing machine/deep learning models. Audio data can be collected through a microphone built into the HMD and is primarily useful from an interaction perspective. For example, the data can be used as an index to evaluate the quality of education and simulation by using voice data from an autistic user who is communicating with a virtual agent [35, 63, 73, 74]. This is useful in the case of, for example, validation and prediction of a VR simulation designed to mitigate public speaking anxiety and improve social interaction skills.

2.1.4 Physiological signals. In ubiquitous computing, much research has been conducted to collect physiological signals, such as EDA and heart rate (HR), from autistic people using sensor equipment [20, 22, 58, 68]. Some of the most commonly used devices for this purpose include the Empatica E4 wristband [6], Fitbit [7], Bitalino [8], and smartphones. Schoen et al. [58] confirmed that EDA is valid indicator of arousal and sensory reactivity in autistic children. Daluwatte et al. [20] demonstrated a significant negative correlation between pupillary light reflex constriction amplitude and average HR in autistic children. Ward et al. [68] used EDA to induce social engagement in a specifically designed theatrical workshop. Fenning et al. [22] demonstrated that the greater the magnitude of EDA variability, the higher the severity of autistic children.

Research on autism in VR has started to adopt sensor equipment to identify the relationship between physiological signals and VR users' perceptions (e.g., stress, valence, and anxiety) [1, 9, 60]. For example, Adiani et al. [1] analyzed the valence and stress level during VR job interview training content based on EDA and BVP. Simões et al. [60] investigated the anxiety moments of autistic adults during the bus-taking routine using EDA. Bian et al. [9] inferred the driving engagement of autistic people based on physiological signals, including BVP, HR, EDA, and breathing rate (BR).

2.1.5 Other physiological signals. There have been studies using other physiological signals to understand the characteristics of autistic people. For example, Zhang et al. [72] collected BR, BVP, EDA, skin temperature (TMP), electrocardiograph (ECG), electromyography (EMG), and electroencephalography (EEG) to measure cognitive load during driving. The participants wore a wireless EEG collection headset and tracking devices on their fingers and wrists while experiencing a VR driving simulation. Bekele et al. [7] collected physiological signals with similar additional sensor devices on the fingers and head while autistic people experienced VR social interaction content. Such inclusive sensors can provide researchers with participants' additional body response information. However, given the sensory issues of autistic people [4, 37, 45], participants may feel excessive sensory discomfort wearing all possible sensor equipment during the VR experience [11]. This will undoubtedly

[2]https://www.tobii.com/
[3]https://pupil-labs.com/
[4]https://www.vive.com/us/product/vive-pro-eye/overview/
[5]https://www.meta.com/quest/quest-pro/
[6]https://www.empatica.com/research/e4/
[7]https://www.fitbit.com/
[8]https://bitalino.com/

Table 1: Summary of prior VR studies that employed sensors to understand autistic people (H: Head position and rotation, E: Eye movement, A: Audio, and P: Physiological signals).

Authors	Year	Summary	H	E	A	P
Adiani et al. [1]	2022	Job interview training system for autistic people using physiological signals and eye gaze-based stress detection module		○		○
Robles et al. [55]	2022	System for screening and classification of autism based on head movement, hand rotation, and eye gaze	○	○		
Simões et al. [61]	2020	Investigating interpersonal distance of autistic adolescents based on head position data	○			
Zhang et al. [73]	2020	Collaborative virtual environment for assessing of social communication in autistic adolescents			○	
Bian et al. [9]	2019	VR driving simulator for autistic people by inferring one's engagement based on physiological signals				○
Simões et al. [60]	2018	System for training bus-taking routines of autistic adults based on physiological signals				○
Zhao et al. [74]	2018	System for improving communication skills of autistic children with a social interaction platform		○	○	
Fan et al. [21]	2017	Group-level classification models to recognize affective states and mental workload of autistic people				○
Zhang et al. [72]	2017	System for measuring cognitive load of autistic adults during driving based on eye-tracking data and physiological signals		○		○
Bekele et al. [7]	2016	Adaptive social interaction VR environment for social training of autistic people		○		○
Mundy et al. [50]	2016	Analysis of information processing during joint attention in autistic children based on eye-tracking data		○		
Kuriakose & Lahiri [41]	2015	Analysis of social communication skills in autistic people based on physiological signals during menu ordering tasks				○
Saiano et al. [56]	2015	System for training street crossing and path following of autistic adults based on eye-tracking data		○		
Hernandez et al. [28]	2014	SVM model to identify engagement in autistic children during social interaction by using a wearable sensing device				○
Wade et al. [67]	2014	System for assessing driving in autistic adolescents based on eye-tracking data and physiological signals		○		○
Bekele et al. [8]	2013	Analysis of eye-tracking data for emotional expression patterns of autistic people and general adolescents		○		○
Lahiri et al. [42]	2012	System for improving social skills by measuring the task engagement level of autistic people based on eye-tracking data		○		

reduce the reliability of the collected sensor data. Thus, carefully considering autism characteristics, we focused on the most representative sensors and reaffirmed that the sensor coverage of *V-DAT* is sufficient through expert interviews.

Based on the review of prior VR research, we selected these four types of representative sensor data—head position and rotation, eye movement, audio, and physiological signals (EDA, BVP, HR, TMP, and inter-beat interval (IBI)). This paper aims to construct a VR sensor-based, multimodal data handling pipeline and design a support system, *V-DAT*, that integrates and manages such data. Our goal is to examine how *V-DAT* provides autistic people an opportunity to reflect on their behaviors and helps autism experts understand autistic people through quantitative evidence.

2.2 Analysis of Multimodal Sensor Data for Autistic People

In a traditional (non-VR) environment, studies that collect multimodal sensor data from autistic people and monitor specific factors are mainly conducted in the physical and mental healthcare domains [1, 7, 35]. However, compared to other research domains (e.g., ubiquitous computing) that utilize various types of sensor data in real-world settings and have established data management pipelines from a sheer amount of research outcomes [24, 46, 49, 57], relatively little research has been conducted to investigate and present such a pipeline in VR. In particular, these pipelines are more necessary for autistic people in that integrating and providing analysis results for multimodal sensors is effective in understanding the challenges they face and advocating for themselves during VR content experiences.

Furthermore, these sensors differ in type, purpose, output signal, underpinning theoretical principles, and technical infrastructure. Thus, the separation and independence of the data could lead to inconsistency in (1) the selection of data on measuring user characteristics and behaviors, (2) the processing of data on the sampling frequency, synchronization, standardization, or feature engineering, and (3) the methods of analysis that consider characteristics of the collected data across studies. These are important for autism experts (e.g., clinicians, caregivers, and researchers) who want to leverage such data-driven methodologies along with quantitative evidence to expand the characteristics of autistic people, and neurotypical people who want to better serve the needs of autistic people. By establishing a sensor data handling pipeline, these issues of methodological reproducibility and data transparency can be mitigated, and the direction of future studies that employ sensor data can be better determined.

In this paper, we introduce *V-DAT*, which is intended to help autism experts collect, manage, and analyze data from VR via a web-based interface. *V-DAT* provides comprehensive visualizations of sensor modalities along with a training video, allowing autistic people to reflect on themselves with quantitative evidence. With *V-DAT*, autism experts can quickly examine the validity of training tasks and autistic participants' challenges, and discover new opportunities to better understand and support autistic people based on analysis results.

2.3 Computer-based Tools for Autism

Previous studies have demonstrated the effectiveness of a range of tools used by autistic people and clinicians to support training, diagnosis, and analysis. These tools include a clinician-parent communication tool [40], a data analysis tool [38], a caregiver-supporting tool [31], and job interview simulators [3, 14]. For instance, Kim et al. [38] employed BEDA, an analysis tool that addresses challenges of analyzing physiological and behavioral data by integrating analysis tasks, thereby improving data analysis and encouraging the combination of physiological measures with behavioral analysis. Jo et al. [31] proposed GeniAuti, a mobile-based data collection tool that helps caregivers document challenging behaviors and relevant contextual information, facilitating collaboration with clinical experts to develop tailored intervention strategies for autistic children. Breen et al. [14] implemented a data visualization dashboard based on a VR interview simulator with the aim of improving employment opportunities for autistic people. This tool provides analysis of visual attention data during job interviews, helping to overcome discomfort with eye contact.

However, it is worth noting that there are few computer systems that can effectively handle multimodal data, provide comprehensive visualized analysis, enable autistic people to reflect on their behaviors, and help autism experts make evidence-based decisions about diagnosis. Moreover, due to the lack of appropriate computer tools, most hospitals and autism care centers still rely on manual procedures to measure and assess the behavior of autistic people [35]. *V-DAT* can fill this research gap by analyzing sensor data that reflect the VR experiences of autistic people and facilitating effective clinical judgment.

3 V-DAT: DATA HANDLING PIPELINE

V-DAT is a system that is integrated with a pipeline for handling data collected from autistic people. To support this, we constructed a pipeline with the following two modules: (1) *data collection* considering different collection methods for each sensor modality and (2) *data processing* considering processing and synchronization methods suitable for each sensor modality of different data types. Algorithm 1 shows the VR sensor data handling pipeline supported by *V-DAT*.

Algorithm 1 V-DAT: Data Handling Pipeline

```
procedure DataCollection(m)                ▷ m stands for modality
    if m == head movement then
        Use OpenVR script
    else if m == eye movement then
        Use Eye Tracking SDK from Vive
    else if m == audio then
        Use AudioModule implemented in UnityEngine
    else if m == physiological signals then
        Use E4 Streaming server and communicate TCP/IP
    end if
end procedure
procedure DataProcessing(cm)  ▷ cm stands for collected m
    if cm == eye then
        Use Tobii XR SDK
        if g == ROI_n then
            Send conflict signal n
        end if
    else if cm == audio then
        Set empirical anomaly points of dBFS
        if v < X or v > x then
            Filter inappropriate voice
        end if
    else if cm == physiological signals then
        Use MS Azure API to detect anomaly points
    end if
end procedure
```

3.1 Data collection

According to the VR content design guideline for autistic people [11], given their sensory issues [4, 36, 37], the wearing of equipment must be simplified and complex interactions within the virtual environment must be excluded for their effective VR experience. Therefore, based on iterative discussions with autism experts and the guideline, we used an HMD (HTC Vive Pro Eye) to collect head position and rotation, eye movement, audio, and an Empatica E4 wristband to collect physiological signals. We applied OpenVR to Unity3D, an API and runtime that made it easy for users to access HMD. Table 2 shows the main features of each of the four sensor modalities.

We investigated and selected the data collection methodology suitable for each modality since each one requires different programming packages and functions for data collection. *V-DAT* collects all data types at 90 Hz intervals, following the default setting of the HTC Vive Pro Eye.

3.1.1 Head position and rotation. We measured the head position and rotation using only the HMD through scripts provided by OpenVR, following many previous studies [2, 13, 51]. The head position includes x (left-right), y (up-down), and z (front-back) values. There are three types of head rotation: 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).

3.1.2 Eye movement. Researchers often use an eye-tracker to collect various types of information as follows: eye gaze origin (the point from which the gaze ray of the eye is generated based on the center of the cornea), normalized gaze direction (expressed as a combined normalized vector for the direction of the eyes), pupil diameter, eye openness, and pupil position (the position of the pupil within the sensor area in which eye movement is tracked) [7, 8, 42]. The following features can be collected from Eye and Facial Tracking SDK provided by Vive [9].

3.1.3 Audio. Most VR HMDs have a built-in microphone, and the sound input to this microphone can be recorded through the microphone class of Unity. When this class is assigned to the audio source component, audio data can be collected.

3.1.4 Physiological signals. Physiological signals, including EDA, BVP, HR, TMP, and IBI, are commonly used in VR research for autistic people. We used the E4 wristband to collect physiological signals, but other devices can also be adapted to *V-DAT* if their data is accessible from those devices. To collect physiological signals while a user experiences VR content, we linked Unity3D and an E4 wristband. The E4 streaming server provided by Empatica enables

[9] https://bit.ly/3l094IH

Table 2: The main focus and required technical skills of each collected sensor modality to understand autistic people better.

Sensor	Focus	Technical skills
Head movement	Awkward or excessive interactions with virtual avatars and the environment	Extracting head movement by collecting head position and rotation with OpenVR
Eye movement	The degree of eye contact and duration when interacting with virtual avatars	Extracting ROIs by Tobii G2OM and collecting the gaze direction with Eye Tracking SDK
Voice	Voice volume control when communicating with virtual avatars	Separating non-silence chunks from wav files and extracting voice volume
Physiological signals	Training moments that exhibit abnormal physiological reactions (e.g., anxiety)	Collecting with E4 streaming server and extracting anomaly points using MS Azure anomaly detector

data collected by the E4 wristband to be transmitted to other engines (e.g., Unity and UNREAL) through TCP/IP communication. For data collection, we set Unity3D as a client. EDA represents the gradual changes in an individual's arousal level or the immediate changes in arousal level often elicited by momentary stimulus [28, 53]. TMP was measured using the infrared thermopile sensor in the E4 wristband, which converts infrared energy radiated from the skin into electrical signals. BVP is an initial value used for calculating IBI and HR and is collected by a photoplethysmograph (PPG) sensor. IBI is obtained by detecting a peak of BVP and then calculating the interval between adjacent peaks. HR can be obtained by dividing the IBI value by 60.

3.2 Data processing

3.2.1 Head position and rotation. Some autistic people may repeat certain words or actions (e.g., head/body rocking), such as stimming, which is an action that produces certain sensory stimuli on its own, and exhibit behavioral characteristics that may seem relatively exaggerated [4]. In our study, this aspect is related to head movement data. If we rely on the anomaly detection algorithm (see Section 3.2.4 for more details on the algorithm) for these data, we may end up with having many anomaly points that are not really abnormal but the natural head movement of autistic people. For this reason, we decided not to use any normalization or anomaly detection algorithm, but to use the raw signals for the analysis. This is expected to allow autism experts or autistic people to interpret the result together with other analysis results provided by *V-DAT*.

(a) (b)

Figure 2: (a) Gaze depth and (b) Region of Interest (ROI). Gaze depth was measured based on the field of view of a VR training participant.

3.2.2 Eye movement. A gaze direction, one of the collected eye movement raw data, refers to gaze depth information (Figure 2-(a)) as viewed by a user. Obtaining depth information is essential for accurately portraying the user's visual information because objects can overlap in one direction of 3D space [17]. ROI is defined based on gaze depth information. To measure eye contact, we used ROI and set the faces of virtual avatars in the training content as ROI. Tobii XR SDK [10] provides information about objects by accurately tracking ROI through these algorithms. We then used the Unity 3D box collider (Figure 2-(b)) to measure whether the user is gazing at the ROI. *V-DAT* stores specific signal data when the tracked object matches the object set by the researcher (Equation 1).

$$CS(g) = \begin{cases} 0 & \text{if, } g \neq ROI_n \\ n & \text{if, } g = ROI_n \end{cases} \tag{1}$$

where $CS(g)$ is a conflict signal function, and g is the gaze point of a user. ROI_n is the n number of ROIs defined by the researcher.

3.2.3 Audio. We extracted the total length of the sound, decibels relative to full scale (dBFS), and the number of non-silence chunks from the recorded audio. The dB concept is used by comparing the amplitudes of the two audio signals, and the dBFS is derived from dB. In a digital environment, the clipping (a form of waveform distortion) point and the amplitude of the acceptable signal varies depending on the microphone. Clipping occurs when the amplitude of the adequate signal is exceeded. The amplitude of the clipped signal is set to 0 dBFS. Thus, the amplitude of the dBFS was always a negative number [39, 59]. The formulas for dB and dBFS are as follows.

$$dB = 20log_{10}\frac{A1}{A2} \tag{2}$$

$$dBFS = 20log_{10}\frac{A}{Max} \tag{3}$$

where $A1$ and $A2$ are the amplitudes of the two audio signals to be compared. A is the amplitude of the signal to be measured, and Max is the maximum amplitude in a specific digital environment.

V-DAT uses dBFS to check the user's voice volume. We used the Python library, pydub [11], to analyze dBFS-based audio data. By using this, we could divide the utterance and silence parts and analyze the volume of the voice. We empirically checked dBFS from -80 to -20 and set -45 dBFS as the threshold for determining whether it is silent or not, which best extracts the spoken words as well as the murmur and filler words (e.g., um, uh, ah, and okay).

We checked the number of utterance chunks by dividing the silence and utterance parts. We also analyzed the cases where the user's voice was either "too small" or "too loud" for each chunk. To do this, the four researchers rehearsed the training scenario and established upper and lower dBFS limits of voice volume that were considered appropriate for social interaction.

[10] https://developer.tobii.com/xr/solutions/tobii-g2om
[11] https://pypi.org/project/pydub/

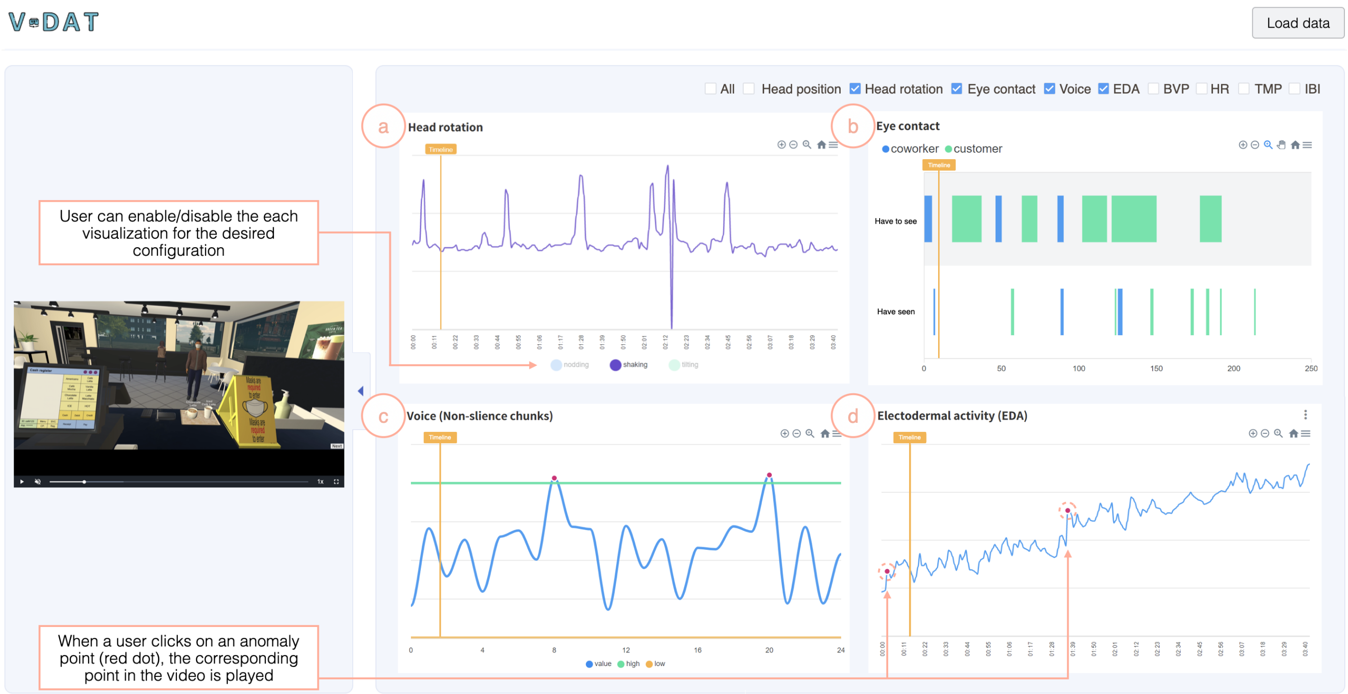

Figure 3: The V-DAT web interface for data visualizations, which consists of two sections (left: training video, right: visualizations). It provides the visualizations of four sensor modalities: (a) head position and rotation, (b) eye contact, (c) voice, and (d) physiological signals. With V-DAT, autistic people and experts can review training results by combining synchronized video and visualizations. The timeline flag (yellow vertical line) on the graph is synchronized with the video. The overall introduction of V-DAT is available at: https://youtu.be/qH6f5-jtjgk.

$$s < X, for\ all\ s \geq sup(\theta) \tag{4}$$

$$x < l, for\ all\ l \geq sup(\theta) \tag{5}$$

where s is small voice volume and l is loud voice volume. X is -35 dBFS, which refers to the least upper bound (supremum) for identifying small voice. x is -15 dBFS, which refers to the greatest lower bound (infimum) for identifying loud voice. θ is a set of silence and $sup(\theta)$ is -45 dBFS, which refers to the upper bound for identifying silence.

3.2.4 Physiological signals. Message from E4 streaming server consists of data type, timestamp, and data value. *V-DAT* extracts and uses only data value from the message.

V-DAT employs the MS Azure API service [12], widely used in anomaly detection through multivariate regression using body responses [54, 75]. This service captures the time point where abnormal physiological signals occur. Five types of sensors, including EDA, BVP, HR, TMP, and IBI, are passed to the anomaly detection API. The API then returns abnormal time points detected in any of the signals. *V-DAT* displays those points on the visualization, which allows autism experts to use those results as supplementary materials for their analysis.

[12]https://bit.ly/3CAle19

3.2.5 Synchronization of four sensor modalities. Synchronization is necessary when collecting different sensor modalities for utilization. Data extracted from the HMD (i.e., head position and rotation, ROI, and voice volume) have the same timestamp information, but data collected from the E4 wristband have a separate timestamp. To map the data from different devices to the same timeline, *V-DAT* uses Flask [27], a web framework. In addition, the E4 streaming server receives only one message per physiological signal and does not send data for some signals when requested simultaneously. Thus, *V-DAT* requests a message with a time interval between signals to ensure that all data is received intact before VR content starts.

4 V-DAT: WEB INTERFACE

V-DAT was designed to support multimodal sensor data collection and analysis in a VR environment. The representative functions of *V-DAT* include the collection and management of data by communicating with sensing devices, synchronization between data modalities, and visualization interfaces suitable for each data characteristic. Through *V-DAT*, we provide the visualizations of sensor modalities. The design rationales of *V-DAT* 's visualizations were based on the following two key characteristics of autistic people [26, 36]. First, autistic people often struggle to concretize their experiences, making it quite challenging to achieve self-awareness of their strengths and weaknesses only through a simple video replay. Second, autistic people tend to be more receptive to evaluation and feedback based on objective and quantitative data.

Based on these design principles, we finalized the visualization methods considering the diagnostic processes used by autism experts. The head position and rotation section displays raw signal changes over time, and the physiological signals section further shows anomaly points in raw signals. The sections for eye contact and voice visualize the characteristics of autistic people identified in previous studies (e.g., difficulty in making eye contact during communicating and difficulty in adjusting voice volume in public settings) [4, 29, 35].

4.1 Head position and rotation

We provide the visualizations of the head position and rotation so that autistic people and autism experts can easily accommodate the information provided by *V-DAT* without a deep understanding of raw sensor data (Figure 3-(a)). In *V-DAT*, the values for the head position x, y, and z are seen as movements in the direction of "left-right," "up-down," and "front-back," respectively. Similarly, the head rotation roll, pitch, and yaw values are seen as "head nodding," "head tilting," and "head shaking," respectively. In addition, depending on the type of data to be explored in detail, each visualization is implemented to enable/disable the desired configuration (e.g., visualization of head nodding only or of head nodding and shaking together).

4.2 Eye contact

Autistic people avoid active eye contact. When it happens, they look confused and anxious [4]. Based on this, *V-DAT* intuitively visualizes how properly autistic people made eye contact when interacting with avatars in a VR environment. The timeline boxes in Figure 3-(b) show the proportion of eye contact during the phase in which autistic people are asked to interact with virtual avatars. By clicking on each timeline box, autism experts can review all the video frames where autistic people made eye contact with virtual avatars.

4.3 Voice

In *V-DAT*, voice below -35 dBFS is considered "too small," while voice above -15 dBFS is considered "too loud." This range is found to be appropriate to determine whether the autistic people are speaking with an appropriate voice volume (Figure 3-(c), this range can be changed depending on an expert's decision). The green/yellow lines represent the upper/lower dBFS limits of voice volume, respectively, and a red dot indicates a non-silence chunk identified as "too loud" (above the upper limit). Through this visualization, autism experts can review how often an autistic participant speaks excessively loud or softly in social interaction situations during training and suggest effective interventions to improve the related challenging behaviors.

4.4 Physiological signals

V-DAT provides trends in physiological signals and marks anomaly points with red dots using the MS Azure API service [13]. In addition, *V-DAT* allows for simultaneous review of the training video and physiological signal trends. As a result, the changes in the five physiological signals (EDA, BVP, HR, TMP, and IBI) can be examined through the corresponding visualization and with the video synchronized by time (Figure 3-(d)).

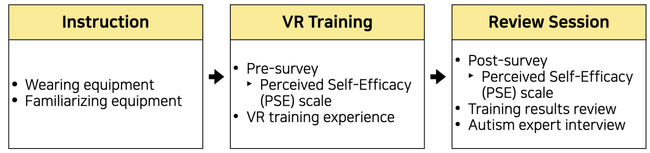

Figure 4: The procedure of the case study. After familiarizing the equipment, the participants attended the training sessions. We asked the participants to complete the survey and attend the review sessions. Finally, we conducted expert interviews to gather feedback on V-DAT from an expert's perspective.

5 CASE STUDY WITH AUTISTIC PEOPLE

To examine how autism experts can review and understand autistic people's training experience through *V-DAT*, we conducted a case study with 20 autistic people. We then interviewed five autism experts to acquire their insights and advice on how *V-DAT* could be used for clinical sessions in hospitals or centers for autistic people in the future. Figure 4 shows the overall study procedure of the study. Our study was approved by the Institutional Review Board (IRB) at the author's institution (B-2102-664-301, B-2202-736-302).

5.1 Participants

We spent eight months (from October 2021 to June 2022) recruiting 20 autistic adults (male = 16, female = 4) by distributing leaflets to public healthcare organizations, having sessions with autistic people, and posting recruitment notices on online autism community websites. Although the gender ratio of participants recruited was somewhat imbalanced, it was in line with that of a previous study, which estimated that there were approximately about four times as many autistic males as autistic females [47]. The inclusion criteria for participation were (1) 18 years of age or older, (2) diagnosed with autism by medical experts, and (3) capable of understanding the purpose of the study and independently participating in the study without assistance from parents or caregivers. We set the lower age limit to 18 based on feedback from autism experts, considering that the existing VR content is mainly designed for promoting social independence, which is more needed for young adults. The age of the autistic participants ranged from 20 to 39 years (mean = 26.46, SD = 4.35).

All 20 autistic participants experienced the same VR training content [35]. The content is a VR-based interactive social skills training system for autistic people that was carefully designed to reflect insights and feedback from autism stakeholders (i.e., autism experts, parents of autistic children, and autistic adults) and to follow design guidelines for autistic people [11]. The training content places autistic people in an environment where, as barista assistants, they need to understand given scenarios and interact with others. They also joined the review session under the same experimental protocols. After the case study, we gathered opinions from five additional autism experts who were not involved in the initial design process of *V-DAT* to receive objective evaluation and discuss their

[13] https://azure.microsoft.com/en-us/products/cognitive-services/anomaly-detector

Table 3: Demographics of autism experts who participated in interviews ("exp." = "years of experience").

Code	Gender	Age	Job Description
E-1	Female	42	Child & adolescent therapist (15+ exp.)
E-2	Male	39	Child & adolescent psychiatrist (10+ exp.)
E-3	Female	51	Child & adolescent therapist (25+ exp.)
E-4	Male	33	Child & adolescent psychiatrist (5+ exp.)
E-5	Male	37	Professor specialized in autism (10+ exp.)

overall feedback on *V-DAT* and its feasible use in clinical settings in the future. Those experts have 5+ to 25+ years of experience (Table 3).

We denote participant quotes using "P-X" that refers to "Participant-Participant Number X." For five autism experts, we denote their quotes using "E-X" which refers to "Expert-Expert Number X."

5.2 Study Procedure

Using *V-DAT*, participants experienced VR training content with fully considered autism characteristics and participated in individual face-to-face interviews with two researchers (one author and one autism clinician). Participants initially responded to the pre-survey, including the Perceived Self-Efficacy (PSE) scale (eight items with a 5-point Likert scale) [71]. The PSE was used to evaluate changes to self-beliefs in one's social skills through VR training. The items of the PSE were developed by referring to [52, 62, 65] based on Bandura's theory of perceived self-efficacy [5]. The participants were then instructed about the training content. After being given as much time as they needed to familiarize themselves with the equipment, they started the VR training. Before the experience, we obtained their consent for the study and informed them that any recorded content could be deleted if they wanted.

The interview was led by the clinician of our research team. The participants were asked to answer an open-ended question about the overall user experience of how effective the comprehensive display of collected data by *V-DAT* is in recognizing behaviors that reflect their training and improve their challenging behaviors. Finally, the participants responded to the PSE scale as a post-survey. To analyze the interviews, three authors of this paper independently coded and compared the coded results to resolve the conflict between the coders through repeated discussions. The Cohen's Kappa measurement [44] showed that each category scored higher than 0.78 (max = 0.86, average = 0.81). Total session time, including experience and training reviews, averaged 63 minutes, and expert interviews averaged 48 minutes. The participants received $50 upon completion of the user study.

5.2.1 VR training review sessions. We conducted an in-depth one-on-one interview to collect participants' feedback on *V-DAT*. We provided four sensor data results (i.e., head movement, eye contact, voice volume, and physiological signals) through *V-DAT* to understand in detail the emotions and behaviors of autistic participants in VR training. First, the clinician reviewed the training moments identified as anomaly points with the participants. Next, through the visualizations of eye contact and voice, he confirmed whether the participants responded appropriately to the social situation. He then continued to review additional check points considering each participant's characteristics (e.g., interest, severity, and problematic behavior). He finally asked the participants about how the data-driven review session provided by *V-DAT* helped them examine their training experiences.

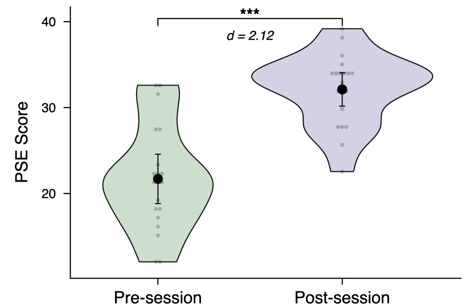

Figure 5: The survey results of the Perceived Self-Efficacy (PSE) Scale (*p <0.001).**

5.2.2 Autism experts interviews. We reviewed the entire process of *V-DAT* with five autism experts (two therapists, two psychiatrists, and one university professor) and collected their feedback on the interview responses from the autistic participants and the effectiveness of *V-DAT*. Specifically, we asked (1) Does *V-DAT* add clinical value to autism experts as well as autistic people?, (2) How reliable and easy are the results of the data-driven analysis provided by *V-DAT*?, (3) Is there any additional sensor-based information that you think is necessary to effectively understand autistic people?, and (4) Do you think *V-DAT* could be used in clinical settings?

5.3 Lessons learned from VR training reviews

The review session with the anomaly detection results and visualization of *V-DAT* elicited a self-explanation of which part the participants did well, which training scenarios they found challenging, and why they felt so. Participants reported that *V-DAT* allowed them to review their training content and reflect on their behavior. This positive perception was supported by the changes in the PSE scores measured before and after the training. As shown in Figure 5, the PSE scores significantly increased after the *V-DAT* experience (before = 21.40, after = 32.20; $t(38) = 2.73$, $d = 2.12$, $p < 0.05$). We also noticed that the distribution of the scores was positioned above the average in the post-session, whereas the proportion of the scores below the average was high in the pre-session. This indicates a significant improvement after the experiment. We used Cohen's d to assess the magnitude of differences between pre- and post-session, with values of 0.20, 0.50, and 0.80 indicating small, medium, and large effect sizes, respectively [18].

5.3.1 High reliability supported by comprehensive data visualizations. Participants reported that *V-DAT* was informative, as they were able not only to evaluate their behaviors against criteria that are commonly used/accepted in real life, but also to examine the differences between the visualization results and their own thoughts. Furthermore, they found the training analysis results of *V-DAT* to be reliable, as the clinician thoroughly presented them with the help of synchronized visualizations and a video. The high reliability of the analysis results provided by *V-DAT* enabled the participants to gain confidence in what they did well and to be aware of what needed more practice.

"[...] Some of the visualizations differed from what I thought. I thought I did a good job in some scenarios, but it actually needed more practice. Usually, I don't listen to my evaluations very carefully, but I was totally persuaded by showing various evidence and explaining it. Thanks to this review session, I learned more about what I should pay more attention to before getting a job (P-7)."

"I've often been scolded by my parents for speaking too loudly in public or not making eye contact when talking to someone. So, I was worried during this training that my voice might have been too loud in some parts or that I might not have made eye contact. However, I was able to gain confidence after seeing the objective evidence that I actually did quite well (P-15)."

5.3.2 Privacy concerns about possible data use. The participants were mostly positive about the results of the sensor data analysis and review session provided by *V-DAT*. Interestingly, some of them wanted to be informed specifically about the use and disclosure of their training videos or sensor data.

"Where else will the collected sensor data be used besides today's review session? Experiencing VR and reviewing like this is personally helpful for me, so it's okay to review it with me, but I don't want it to be used elsewhere (P-2)."

"[...] I'll keep using this kind of training method unless the results are shared with my family. This kind of training seems to be mostly for adults. Since I'm an adult, I don't want to share anything like what I did or didn't do well during training (P-11)."

5.4 Lessons learned from expert interviews

Through the expert interviews, we reaffirmed the need for VR training and the effectiveness of the visualizations of *V-DAT* to better understand autistic people from a more diverse perspective. Experts reported that wearing as little sensor equipment as possible would support participants in more immersive training. We also identified a possibility of using *V-DAT* for effective intervention and treatment.

5.4.1 Data-driven clinical decision supporting tool. All experts gave positive feedback on the VR training and *V-DAT* interfaces for autistic people. They noted that the comprehensive visualization results of *V-DAT* could help to confirm the difficulties and challenging behaviors of autistic people, accurately identify the factors or environments that cause them, and suggest more effective interventions based on data-driven evidence. In addition, some of the visualization results could be used to reduce human errors when studying challenging behaviors in autistic people compared to conventional protocols (e.g., reviewing the recorded video of social skills training role-plays).

"Many people find it difficult to make eye contact and have a lot of social anxiety. However, it's quite limited to knowing specifically what causes them to be nervous and what physical or psychological changes occur when anxiety is amplified. [...] Eye contact, voice, and excessive and awkward gestures are the tasks I usually observe or sometimes check with recorded videos. This system enables repetitive VR training in a fully controlled situation, visualizes, and provides training results at a glance. A system like this would be helpful and save me a lot of time (E-3, Therapist)."

However, some experts have pointed out that the practical use of *V-DAT* may still be challenging for experts who lack VR and computing knowledge/skills. In particular, they mentioned that even though *V-DAT*'s data collection and visualization delivery is automated without user input, it is difficult to deal with potential problems that may arise during data communication and loading.

"VR equipment is being considered for use. However, I think it is another problem actually to use such a system. Other web-based systems used in our institution have also had quite a lot of troubles because some data has been collected with loss or because of poor loading. In clinical settings, we typically have very limited time... this does not just mean that we don't know how to address these issues, but also that we don't have time to resolve them directly (E-2, Psychiatrist)."

5.4.2 Bridging the gap between outpatient clinics and autism care centers. Through interviews, we confirmed a gap between the doctors (mainly psychiatrists) who run outpatient clinics and the therapists in welfare centers who are more frequently confronted with autistic people. The doctors reported that, due to the short time available for outpatients, they had no choice but to focus mainly on regular prescriptions or prominent problems such as severe emotional anxiety and aggressive behavior. The therapists, on the other hand, often observe the challenging behaviors of autistic people and try various validated interventions for behavioral therapy (e.g., applied behavior analysis (ABA) [23] and video modeling). However, the therapists find it difficult to identify the exact cause of challenging behaviors because their method lacks objective and quantitative indicators. The therapists also mentioned that even some information could help with further diagnosis and prescription; without evidence, it is difficult for them to apply such information to the clinical stage.

V-DAT could fill this gap in three ways. First, it could allow doctors and therapists to quantitatively investigate the underlying causes of challenging behaviors in autistic people. Second, the results of the analysis, confirmed by therapists with quantitative evidence from *V-DAT*, can be discussed to see whether the results can be used to support the doctor's clinical stage, potentially leading to effective diagnosis or treatment. Finally, *V-DAT* can make training and treatment results more acceptable and reliable to autistic people by providing comprehensive visualizations synchronized with a training video.

"This behavioral therapy training is usually conducted in the center like a 1:1 role play with a therapist and an autistic person. The problem is that an autistic person easily loses interest, and it's hard for therapists to do exactly the same training every time. [...] In this way, it is also difficult to identify the underlying reason for challenging behavior. Also, what the therapist knows is

sometimes not well communicated to patients, [...] with the result that many autistic people repeatedly receive somewhat obvious suggestions based only on publicly known behaviors (E-5, Professor)."

"In fact, the service that V-DAT can provide is data-driven, not replacing what someone has been doing, but filling in the necessary areas that are not working well. [...] Therapists are mostly located in big cities, and it's practically impossible for them to cover many autistic people. It is also difficult for autistic people to get behavioral therapy as adults. [...] If the therapists examine the results of V-DAT, and if the examination becomes available in outpatient clinics, doctors can also benefit from a better understanding and potentially effective diagnosis (E-4, Psychiatrist)."

6 DISCUSSION

6.1 Summary of research

This paper presented *V-DAT* that supports the sensor data handling pipeline for autism research using VR. First, we identified four sensor data modalities and detailed the characteristics and collection and analysis methods of each one. Then, through a case study, we showed how the data of each modality from autistic people could be analyzed and presented through *V-DAT* and investigated the perception of autism experts and the potential of the application of *V-DAT* in clinical settings in the future. By establishing this pipeline, *V-DAT* is expected to allow autism researchers who are interested in using sensor data in a VR context to analyze data through visualizations and to examine the relationship between VR sensor signals and their target variables of interest. In the following subsections, we discuss some key points and next steps for *V-DAT*, which are closely related to the "methodological reproducibility" and "data transparency" highlighted in our work. These lessons may also be useful for studies considering the development of a data management system for people with other neurodevelopmental disabilities.

6.2 Multimodal data management

In Section 5.4.1, autism experts raised concerns about potential difficulties in dealing with problems that may arise during data processes due to their limited VR and computing knowledge/skills. We experienced a slight data loading problem when reviewing VR training through *V-DAT*, because processing large amounts of data takes a long time. Minimizing data loading time should be a priority when developing clinical systems for autistic people, as autism experts have noted that they cannot spend too much time on issues of computer systems. Given the characteristics of sensor data, some data redundancy may occur in the collection and processing procedures. In addition, this problem may be exacerbated because the majority of most HMDs store data with high Hz to support high-resolution content. A possible solution is to reduce the data size by eliminating unnecessary, redundant data by lowering the Hz to a level that does not adversely affect data interpretation (we believe that the investigation of an appropriate Hz level is also necessary as future work). In this way, we can mitigate the problems (e.g., delay in review sessions) that experts encounter when loading analysis results and improve the accuracy of the data provided by *V-DAT*.

6.3 Sustainable use of a VR-based tool

We investigated the applicability of the clinical aspects to *V-DAT* through a case study based on training content carefully tailored to autism characteristics. We also discussed with autism experts the potential clinical value of *V-DAT* in the future. However, the case study used only a single training content to improve social skills. There are many other scenarios (e.g., collaborating with co-workers and meeting strangers and acquaintances) that are related to social skills that autism experts have highlighted for autistic people. Therefore, it is important to develop various training content according to the VR content guidelines for autistic people. Future studies are needed to verify the effectiveness of *V-DAT* based on this content. Using different training content through *V-DAT* can help autistic people and their families explore challenging behaviors and support social independence. This can also facilitate appropriate support for clinicians such as psychiatrists and autism therapists in a variety of settings. For example, autistic people can practice understanding the context of a workplace meeting or casual conversation and participating in timely responses and conversations. Therefore, it is necessary to develop more VR content that adequately considers autism characteristics in order to support the sustainable use of *V-DAT*.

As highlighted in Section 5.3.1, we also confirmed that participants trusted and actively accepted the analysis results of the *V-DAT* experience. Autistic people are generally reluctant to engage in nominal activities (e.g., role-playing for social skills training) and tend to be conservative in accepting evaluations of themselves [36]. Given these characteristics, it is important to consider ways to provide objective and comprehensive evidence to encourage the sustained participation of autistic people so that they can be confident of receiving feedback to improve their behaviors and social skills. Furthermore, the clinician reviewed the training video and visualizations in the review session (Section 5.2.1) and then continued to review the visualizations for additional check points. We found that this flexible review process was important when considering the wide range of autistic people with different interests, severity, and problem behaviors. In this sense, future studies may need to collect logs of the use of *V-DAT* by autism experts so that their interaction patterns can be defined and identified.

Finally, we confirmed that privacy issues for autistic people should be considered first before using data from *V-DAT* in clinical settings (Section 5.3.2). Some autistic participants were quite concerned about how much information recorded in *V-DAT* would be available to others and how it would be used. Such issues should be considered when developing systems and using data from autistic people in clinical settings.

7 CONCLUSION

Research in VR for autistic people has moved beyond simply providing realistic experiences and has begun to analyze sensor data to identify multiple interactions in a virtual environment. This enables researchers to provide effective and valuable VR training and education to the autism population. In this paper, we presented *V-DAT*, integrated with a VR sensor data handling pipeline to facilitate both autistic people and autism experts to effectively use multiple sensor signals in VR. We identified four types of sensor modalities that have been investigated in many prior VR studies for autistic people and explained the characteristics and processing methods of each data modality. Autism experts can use comprehensive visualization analysis results of *V-DAT* to investigate behavior patterns of autistic people. We also conducted a case study of *V-DAT*, which highlighted that the data can be used to provide autistic people with potentially meaningful interventions or diagnostic support through comprehensive visualization results. Finally, we discussed the extensibility and sustainable use of *V-DAT* in clinical settings. We hope that our findings and insights will benefit autism researchers, practitioners, and designers who wish to use a variety of sensor data in their study of VR training experiences.

ACKNOWLEDGMENTS

This research was supported by the National Research Foundation (2021M3A9E4080780) and Institute for Information & Communication Technology Planning & Evaluation (IITP-2020-0-01373, IITP-2023-2018-0-01431).

REFERENCES

[1] 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.

[2] 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.

[3] Shameem Ahmed, Wesley Deneke, Victor Mai, Alexander Veneruso, Matthew Stepita, Anais Dawson, Bradley Hoefel, Garrett Claeys, Nicholas Lam, and Moushumi Sharmin. 2020. InterViewR: a mixed-reality based interview training simulation platform for individuals with autism. In *2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)*. IEEE, 439–448.

[4] American Psychiatric Association et al. 2013. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. American Psychiatric Pub.

[5] Albert Bandura. 2012. On the Functional Properties of Perceived Self-efficacy Revisited. *Journal of management* 38, 1 (2012), 9–44.

[6] Mayra D Barrera Machuca, Paul Asente, Jingwan Lu, Byungmoon Kim, and Wolfgang Stuerzlinger. 2017. Multiplanes: Assisted freehand VR drawing. In *Adjunct Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*. 1–3.

[7] 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.

[8] 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.

[9] Dayi Bian, Joshua Wade, Amy Swanson, Amy Weitlauf, Zachary Warren, and Nilanjan Sarkar. 2019. Design of a Physiology-based Adaptive Virtual Reality Driving Platform for Individuals with ASD. *ACM Transactions on Accessible Computing (TACCESS)* 12, 1 (2019), 1–24.

[10] Kristen Bottema-Beutel, Steven K Kapp, Jessica Nina Lester, Noah J Sasson, and Brittany N Hand. 2021. Avoiding ableist language: Suggestions for autism researchers. *Autism in Adulthood* 3, 1 (2021), 18–29.

[11] 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.

[12] Lal" Lila" Bozgeyikli, Evren Bozgeyikli, Srinivas Katkoori, Andrew Raij, and Redwan Alqasemi. 2018. Effects of Virtual Reality Properties on User Eexperience of Individuals with Autism. *ACM Transactions on Accessible Computing (TACCESS)* 11, 4 (2018), 1–27.

[13] Ryan Bradley and Nigel Newbutt. 2018. Autism and Virtual Reality Head-mounted Displays: A State of the Art Systematic Review. *Journal of Enabling Technologies* (2018).

[14] Michael Breen, James McClarty, Caleb Langley, Jamshid Farzidayeri, Kyle Trevethan, Brandon Swenson, Medha Sarkar, Joshua Wade, and Nilanjan Sarkar. 2021. 2D and 3D Visualization of Eye Gaze Patterns in a VR-Based Job Interview Simulator: Application in Educating Employers on the Gaze Patterns of Autistic Candidates. In *International Conference on Human-Computer Interaction*. Springer, 533–544.

[15] Kate Carey, Emily Saltz, Jacob Rosenbloom, Mark Micheli, Judeth Oden Choi, and Jessica Hammer. 2017. Toward Measuring Empathy in Virtual Reality. In *Extended Abstracts Publication of the Annual Symposium on Computer-Human Interaction in Play*. 551–559.

[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 Disorder. *Focus on Autism and Other Developmental Disabilities* 30, 4 (2015), 222–236.

[17] Viviane Clay, Peter König, and Sabine Koenig. 2019. Eye Tracking in Virtual Reality. *Journal of Eye Movement Research* 12, 1 (2019).

[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] Jing Fan, Joshua W Wade, Alexandra P Key, Zachary E Warren, and Nilanjan Sarkar. 2017. EEG-based Affect and Workload Recognition in a Virtual Driving Environment for ASD Intervention. *IEEE Transactions on Biomedical Engineering* 65, 1 (2017), 43–51.

[22] Rachel M Fenning, Jason K Baker, Brian R Baucom, Stephen A Erath, Mariann A Howland, and Jacquelyn Moffitt. 2017. Electrodermal Variability and Symptom Severity in Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders* 47, 4 (2017), 1062–1072.

[23] Richard M Foxx. 2008. Applied Behavior Analysis Treatment of Autism: The State of the Art. *Child and adolescent psychiatric clinics of North America* 17, 4 (2008), 821–834.

[24] Dimitris Giakoumis, Anastasios Drosou, Pietro Cipresso, Dimitrios Tzovaras, George Hassapis, Andrea Gaggioli, and Giuseppe Riva. 2012. Using Activity-related Behavioural Features Towards More Effective Automatic Stress Detection. *PloS one* 7, 9 (2012), e43571.

[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] Temple Grandin. 2008. *Thinking in pictures, expanded edition: My life with autism*. Vintage.

[27] Miguel Grinberg. 2018. *Flask web development: developing web applications with python*. " O'Reilly Media, Inc.".

[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] 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.

[30] 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.

[31] Eunkyung Jo, Seora Park, Hyeonseok Bang, Youngeun Hong, Yeni Kim, Jungwon Choi, Bung Nyun Kim, Daniel A Epstein, and Hwajung Hong. 2022. GeniAuti: Toward Data-Driven Interventions to Challenging Behaviors of Autistic Children through Caregivers' Tracking. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW1 (2022), 1–27.

[32] 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.

[33] 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.

[34] Lorcan Kenny, Caroline Hattersley, Bonnie Molins, Carole Buckley, Carol Povey, and Elizabeth Pellicano. 2016. Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism* 20, 4 (2016), 442–462.

[35] 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*. 1–11.

[36] Bogoan Kim, Sung-In Kim, Sangwon Park, Hee Jeong Yoo, Hwajung Hong, and Kyungsik Han. 2023. RoutineAid: Externalizing Key Design Elements to Support Daily Routines of Individuals with Autism. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–18.

[37] 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.

[38] Jennifer G Kim, Melinda Snodgrass, Mary Pietrowicz, and Karrie Karahalios. 2015. Visual analysis of relationships between behavioral and physiological sensor data. In *2015 International Conference on Healthcare Informatics*. IEEE, 170–179.

[39] Pascal Klein, Michael Hirth, Sebastian Gröber, Jochen Kuhn, and Andreas Müller. 2014. Classical Experiments Revisited: Smartphones and Tablet PCs as Experimental Tools in Acoustics and Optics. *Physics Education* 49, 4 (2014), 412.

[40] Ha-Kyung Kong and Karrie Karahalios. 2020. Addressing cognitive and emotional barriers in parent-clinician communication through behavioral visualization webtools. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–12.

[41] Selvia Kuriakose and Uttama Lahiri. 2015. Understanding the Psychophysiological Implications of Interaction with a Virtual Reality-based System in Adolescents with Autism: A Feasibility Study. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 23, 4 (2015), 665–675.

[42] 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.

[43] 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.

[44] J Richard Landis and Gary G Koch. 1977. The Measurement of Observer Agreement for Categorical Data. *biometrics* (1977), 159–174.

[45] 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.

[46] Lee Boon Leng, Lee Boon Giin, and Wan-Young Chung. 2015. Wearable Driver Drowsiness Detection System based on Biomedical and Motion Sensors. In *2015 IEEE SENSORS*. IEEE, 1–4.

[47] Rachel Loomes, Laura Hull, and William Polmear Locke Mandy. 2017. What is the male-to-female ratio in autism spectrum disorder? A systematic review and meta-analysis. *Journal of the American Academy of Child & Adolescent Psychiatry* 56, 6 (2017), 466–474.

[48] Paul McIlvenny. 2019. Inhabiting Spatial Video and Audio Data: Towards a Scenographic Turn in the Analysis of Social Interaction. *Social Interaction. Video-Based Studies of Human Sociality* 2, 1 (2019).

[49] Amir Muaremi, Bert Arnrich, and Gerhard Tröster. 2013. Towards Measuring Stress with Smartphones and Wearable Devices during Workday and Sleep. *BioNanoScience* 3, 2 (2013), 172–183.

[50] 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 Disorder. *Journal of autism and developmental disorders* 46, 7 (2016), 2555–2560.

[51] Nigel Newbutt, Ryan Bradley, and Iian Conley. 2020. Using Virtual Reality Head-mounted Displays in Schools with Autistic Children: Views, Experiences, and Future Directions. *Cyberpsychology, behavior, and social networking* 23, 1 (2020), 23–33.

[52] 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.

[53] Hugo F Posada-Quintero and Ki H Chon. 2020. Innovations in Electrodermal Activity Data Collection and Signal Processing: A Systematic Review. *Sensors* 20, 2 (2020), 479.

[54] Hansheng Ren, Bixiong Xu, Yujing Wang, Chao Yi, Congrui Huang, Xiaoyu Kou, Tony Xing, Mao Yang, Jie Tong, and Qi Zhang. 2019. Time-series Anomaly Detection Service at Microsoft. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*. 3009–3017.

[55] 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.

[56] 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 Disorder: A Feasibility Study. *Journal of neuroengineering and rehabilitation* 12, 1 (2015), 1–13.

[57] Akane Sano and Rosalind W Picard. 2013. Stress Recognition Using Wearable Sensors and Mobile Phones. In *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*. IEEE, 671–676.

[58] Sarah A Schoen, Lucy Jane Miller, Barbara Brett-Green, and Susan L Hepburn. 2008. Psychophysiology of Children with Autism Spectrum Disorder. *Research in Autism Spectrum Disorders* 2, 3 (2008), 417–429.

[59] Immanuel Schweizer, Roman Bärtl, Axel Schulz, Florian Probst, and Max Mühläuser. 2011. NoiseMap-real-time Participatory Noise Maps. In *Second international workshop on sensing applications on mobile phones*. Citeseer, 1–5.

[60] 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.

[61] Marco Simões, Susana Mouga, Andreia C Pereira, Paulo de Carvalho, Guiomar Oliveira, and Miguel Castelo-Branco. 2020. Virtual Reality Immersion Rescales Regulation of Interpersonal Distance in Controls But Not in Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders* 50, 12 (2020), 4317–4328.

[62] 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.

[63] 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.

[64] 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.

[65] 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.

[66] 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.

[67] Joshua Wade, Dayi Bian, Lian Zhang, Amy Swanson, Medha Sarkar, Zachary Warren, and Nilanjan Sarkar. 2014. Design of a Virtual Reality Driving Environment to Assess Performance of Teenagers with ASD. In *International conference on universal access in human-computer interaction*. Springer, 466–474.

[68] Jamie A Ward, Daniel Richardson, Guido Orgs, Kelly Hunter, and Antonia Hamilton. 2018. Sensing Interpersonal Synchrony between Actors and Autistic Children in Theatre Using Wrist-worn Accelerometers. In *Proceedings of the 2018 ACM International Symposium on Wearable Computers*. 148–155.

[69] Megha Yadav, Md Nazmus Sakib, Kexin Feng, Theodora Chaspari, and Amir Behzadan. 2019. Virtual Reality Interfaces and Population-specific Models to Mitigate Public Speaking Anxiety. In *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 1–7.

[70] Jacky Yang, Michael Chan, Alvaro Uribe-Quevedo, Bill Kapralos, Norman Jaimes, and Adam Dubrowski. 2020. Prototyping Virtual Reality Interactions in Medical Simulation Employing Speech Recognition. In *2020 22nd Symposium on Virtual and Augmented Reality (SVR)*. IEEE, 351–355.

[71] 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 Korean Version of the PEERS® Parent-Assisted Social Skills Training Program for Teens With ASD. *Autism Research* 7, 1 (2014), 145–161.

[72] 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.

[73] Lian Zhang, Amy S Weitlauf, Ashwaq Zaini Amat, Amy Swanson, Zachary E Warren, and Nilanjan Sarkar. 2020. Assessing Social Communication and Collaboration in Autism Spectrum Disorder Using Intelligent Collaborative Virtual Environments. *Journal of Autism and Developmental Disorders* 50, 1 (2020), 199–211.

[74] Huan Zhao, Amy R Swanson, Amy S Weitlauf, Zachary E Warren, and Nilanjan Sarkar. 2018. Hand-in-hand: A Communication-enhancement Collaborative Virtual Reality System for Promoting Social Interaction in Children with Autism Spectrum Disorder. *IEEE transactions on human-machine systems* 48, 2 (2018), 136–148.

[75] Hang Zhao, Yujing Wang, Juanyong Duan, Congrui Huang, Defu Cao, Yunhai Tong, Bixiong Xu, Jing Bai, Jie Tong, and Qi Zhang. 2020. Multivariate Time-series Anomaly Detection via Graph Attention Network. In *2020 IEEE International Conference on Data Mining (ICDM)*. IEEE, 841–850.