Assistive Technology
The Official Journal of RESNA

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/uaty20

Multimodal job interview simulator for training of autistic individuals

Deeksha Adiani, Michael Breen, Miroslava Migovich, Joshua Wade, Spencer Hunt, Mahrukh Tauseef, Nibraas Khan, Kelley Colopietro, Megan Lanthier, Amy Swanson, Timothy J. Vogus & Nilanjan Sarkar

To cite this article: Deeksha Adiani, Michael Breen, Miroslava Migovich, Joshua Wade, Spencer Hunt, Mahrukh Tauseef, Nibraas Khan, Kelley Colopietro, Megan Lanthier, Amy Swanson, Timothy J. Vogus & Nilanjan Sarkar (2023): Multimodal job interview simulator for training of autistic individuals, Assistive Technology, DOI: 10.1080/10400435.2023.2188907

To link to this article: https://doi.org/10.1080/10400435.2023.2188907

Published online: 24 Apr 2023.

ASSISTIVE TECHNOLOGY
https://doi.org/10.1080/10400435.2023.2188907

Multimodal job interview simulator for training of autistic individuals

Deeksha Adiani, BS[a], Michael Breen, BS[b], Miroslava Migovich, MS[b], Joshua Wade, MS[b], Spencer Hunt, BS[b], Mahrukh Tauseef, BS[c], Nibraas Khan, BS[a], Kelley Colopietro, BS[b], Megan Lanthier, BASc[d], Amy Swanson, MA[d], Timothy J. Vogus, PhD[e], and Nilanjan Sarkar, PhD[a,b,c]

[a]Computer Science, Vanderbilt University, Nashville, Tennessee, USA; [b]Mechanical Engineering, Vanderbilt University, Nashville, Tennessee, USA; [c]Electrical and Computer Engineering, Vanderbilt University, Nashville, Tennessee, USA; [d]TRIAD, Vanderbilt University Medical Center, Nashville, Tennessee, USA; [e]Owen School of Management, Vanderbilt University, Nashville, Tennessee, USA

ABSTRACT

Autistic individuals face difficulties in finding and maintaining employment, and studies have shown that the job interview is often a significant barrier to obtaining employment. Prior computer-based job interview training interventions for autistic individuals have been associated with better interview outcomes. These previous interventions, however, do not leverage the use of multimodal data that could give insight into the emotional underpinnings of autistic individuals' challenges in job interviews. In this article, the authors present the design of a novel multimodal job interview training platform called CIRVR that simulates job interviews through spoken interaction and collects eye gaze, facial expressions, and physiological responses of the participants to understand their stress response and their affective state. Results from a feasibility study with 23 autistic participants who interacted with CIRVR are presented. In addition, qualitative feedback was gathered from stakeholders on visualizations of data on CIRVR's visualization tool called the Dashboard. The data gathered indicate the potential of CIRVR along with the Dashboard to be used in the creation of individualized job interview training of autistic individuals.

ARTICLE HISTORY
Accepted 2 March 2023

KEYWORDS
autism; employment; job interview; multimodal data; virtual reality

Introduction

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder, with a prevalence of 1 in 44 children in the US (Maenner et al., 2021), characterized by persistent deficits in social communication and interaction (e.g., reduced facial expressions, differences in perceived eye-contact), as well as presence of restrictive patterns of behavior, interests, or activities (e.g., the need for routine and discomfort with unexpected events), where individuals range in their level of support needs – from requiring support to very substantial support needs in daily activities (American Psychiatric Association [APA], 2013). Employment outcomes for autistic[1] adults are often poor (Black et al., 2020; Nicholas & Klag, 2020; Sarrett, 2017; Wehman et al., 2018). A report indicated that among autistic individuals who received developmental disability services in a community, only 14% had paying jobs (Roux et al., 2017), and those employed are often overqualified for the work (Baldwin et al., 2014). Autistic individuals cite anxiety, stress, and discrimination in the workplace as challenges to obtaining employment (Booth, 2016; Nicholas et al., 2017; Taylor et al., 2019). Among those, the job interview process poses a significant barrier to employment where autistic individuals are often expected to communicate and act like their neurotypical (NT) peers (Taylor et al., 2019), with some communication styles (e.g., blunt) associated with autism seen as deviating from the employer's image of a "good employee." Social Anxiety is a common phenomenon in autistic individuals (Spain et al., 2018), and can be amplified in uncertain contexts or where they feel they may need to "camouflage" autistic traits to be perceived well by NT individuals (APA, 2013; Senju & Johnson, 2009). For example, autistic job candidates may be especially aware of the need to make eye contact and engage in it even when they find it aversive. These behaviors and emotional experiences often lead to employers erroneously perceiving this population as undesirable candidates despite them having the required skills and talents (Maras et al., 2020). Moreover, when autistic job candidates disclose their diagnosis without further elaboration, interviewers tend to make "thin-slice" judgments that tend to correspond with a global, negative view of autism often rooted in misperceptions of ASD (Bublitz et al., 2017; Sarrett, 2017; Whelpley & May, 2022). In addition, autistic individuals are often subject to social judgments in the workplace by their NT peers, regardless of presence of differences in behavior (Sasson et al., 2017).

Prior research on interview simulators has demonstrated the importance of practice and training for improving interview performance (Smith et al., 2021). The development of CIRVR was motivated by several works in the field, all of which have shown some benefit for interviewees articulating their strengths, interview performance, and employment outcomes for adults and transition-age youth (TAY) (ages 16–25). Burke et al. (2018) developed Virtual Interactive Training Agents (ViTA) in simulated job interviews that capture video of the sessions as well as determine facial expressions from camera images. An efficacy study with ViTA was associated with improvement in

CONTACT Deeksha Adiani deeksha.m.adiani@vanderbilt.edu Department of Computer Science, Vanderbilt University, Olin Hall 518, 2400 Highland Ave, Nashville, TN 37013, USA.

[1]Autistic individuals have shown preference for identity-first language ("autistic") rather than person-first ("person with autism") (Kenny et al., 2016). Thus, the authors have used this terminology throughout the article.

participants' ability to identify and promote their strengths and respond to situational and behavioral interview questions (Burke et al., 2018). Training using JobTIPS, a video modeling-based job interview training platform for autistic youth, was associated with improvement in autistic participants' interview performance (Strickland et al., 2013). Virtual Reality Job Interview Training (VR-JIT) uses over 1,000 prerecorded interview questions and multiple-choice style responses (Smith et al., 2014). Training with VR-JIT was associated with subsequent positive employment outcomes for autistic job candidates (Smith et al., 2015). Recently, the authors of VR-JIT expanded their work to adapt the system to fit the needs of autistic TAY (Smith et al., 2020), and their Virtual Interview Training for Transition Age Youth (VIT-TAY) demonstrated positive vocational outcomes for autistic TAY (Smith et al., 2021). Two other systems have used multimodal data (DeVault et al., 2014; Xu et al., 2015) to assess interviewee distress and conscious and unconscious nonverbal behavior (e.g., eye contact and speech volume).

Motivated by the above-mentioned works, a novel multimodal job interview simulator was developed called Career Interview Readiness in VR or CIRVR. CIRVR offers a more holistic set of approaches (naturalistic dialog-based conversation, stress monitoring via affective computing, real-time gaze monitoring, and real time emotion recognition via facial expressions) to capture richer data (speech, physiology, eye gaze, and facial expressions). Note that gaze data is collected, not with the expectation that autistic individuals will adopt the gaze patterns of their NT peers, but to help job coaches and employers understand observed differences and calibrate their expectations. Such an approach is valuable as prior research finds that multimodal data provide insight into users' behaviors and learning experiences (Giannakos et al., 2019), that can help employers understand their autistic employees, and can help enhance or predict learning outcomes (Wang et al., 2018). CIRVR conveys this information via a Dashboard of visualizations to facilitate customized interview feedback for autistic individuals and job coaches, while providing employers with a more sophisticated understanding of the outcomes and experiences produced with their interview processes.

The scope of this article is to present a feasibility study of CIRVR to determine whether (1) the interviews ran smoothly and largely autonomously without human intervention, (2) multimodal data were captured in the intended format without appreciable data loss, and (3) the data could be autonomously processed and displayed on the Dashboard for analysis. The authors further discuss how these data can potentially contribute to meaningful improvement of job interview experiences and outcomes for autistic individuals. Once feasibility is established, subsequent research will longitudinally assess the interview performance improvement of the users.

Methods

Stakeholder input

The design and implementation of CIRVR involved several stakeholders—10 autistic self-advocates, 10 job coaches and career counselors, and 3 employers. A combination of qualitative and survey data were used to understand typical barriers to successful job interviews for autistic candidates and to shape the development of CIRVR. Interviews were conducted with 23 stakeholders, individually, or as part of a focus group, and survey data was further collected from 27 autistic individuals, to learn what features might be helpful for improving interview performance and user interaction. The autistic interviewees ranged from recent college graduates to seasoned professionals with experience in multiple industries (e.g., consulting, information technology, sales and [retail] service) and reported an average of 6.56 years of work experience. The 10 job coaches and career counselors in post-secondary educational organizations, all provided direct support to autistic job candidates and reported an average of 10.44 years of work experience. The 3 employers were managers or human resource professionals with an average of 9.67 years of work experience, and currently working for large companies (i.e., more than 500 employees) with established autism hiring initiatives. Each of the employers had more than 5 years of experience in hiring and working with autistic employees.

Data collection from individuals and focus groups

Individual stakeholder interviews ranged from 45 to 60 minutes and focus groups from 90 to 120 minutes based on stakeholder preferences. The semi-structured interview protocol was tailored to each stakeholder group and included additional prompts and follow-up questions to be used as needed. Interviews with focus groups explored the factors that influence autistic individuals' job interview experiences (e.g., physical environment, interview questions, interviewer interpersonal style).

The authors first developed the interview questions to explore employment interview experiences (see Figure B1, Figure B2, and Figure B3 in Appendix B). Before implementing these protocols, the interview protocols were shared with individuals who are part of the stakeholder groups (i.e., key informants, including autistic individuals) for feedback and protocol refinement to ensure it used clear and understandable language. This produced wording changes and additional probing questions in our protocols. From feedback, it was learned that some open-ended interview questions might be too broad and vague for participants to answer. Hence, the wording was clarified and questions were added. For example, when autistic individuals were asked to broadly discuss their interview preparation experience with the researchers, a question was added on the impact of receiving interview questions prior to the job interview (see question 3c in Figure B1 in Appendix B).

Data analysis of qualitative data from focus groups and individual interviews

The authors used qualitative content analysis using an inductive category application approach to identify themes (Mayring, 2019). Three authors were involved in the data analysis process and used *Dedoose (v4.12)*[2] qualitative coding software. The authors began by reading the transcripts, session notes, and written memos. The coding process first involved the authors' understanding of text content and the application of codes to

[2]https://www.dedoose.com/.

units of meaning (i.e., words, sentences, or paragraphs that convey a specific idea or perspective) (Campbell et al., 2013). The three authors independently coded the same transcript and then compared their respective coding units, codes, and code labels to assess consistency of the approach. After this review, they decided that a coding unit may include sentences of paragraphs instead of words to capture the contextual meaning of the content. In addition, the code labeling structure was also discussed; for example, all three authors agreed that one of the first-level codes was *Barriers_hypothetical questions*, which depicted the challenges autistic job seekers had in responding to interview questions based on hypothetical situations. All other transcripts were coded and reviewed by at least two out of three authors who were involved in the qualitative analysis process. One author read and coded transcripts with first-level coding, while the other looked for consistency in first-level codes. The next step was to reduce the number of codes and determine categories, i.e., codes with similar ideas were sorted into the same category. The categories generated were then reviewed and refined. The authors then searched for common categories among employers, autistic individuals, and service providers, and examined consistencies and differences in concepts across the three stakeholder groups, which followed by making connections between different categories. Themes were identified by analyzing and interpreting how categories, codes, and narratives were related to each other and how the relationship could help understand the nuances of the employment interview process for the neurodiverse workforce. Data saturation was reached when there was no more new information and the researchers identified similar concepts from participants resulting in the same themes. Inter-coder reliability was performed to ensure coding accuracy, and inter-coder agreements for main categories were calculated using Cohen's Kappa coefficient (Cohen, 1960). The Cohen's Kappa inter-coder agreement ranged from 0.54 to 1.00 for all categories, yielding a fair-to-excellent agreement (Miles & Huberman, 1994), and all discrepancies were resolved through discussion. The results from the thematic analysis above have been discussed in the Results section.

Design of CIRVR

CIRVR was created on the Unity 3D Game Engine (version 2019.3) (Unity 3D 2019.3, 2022), and consists of four components: 1) a virtual environment with a virtual interactive interviewer (see Figure 1); 2) a conversation management system (CMS) that executes the interview script and processes interviewee responses to output consecutive interview questions; 3) multimodal data capture module that allows real-time collection of several multimodal data as shown in Tables 1 and 4) a Dashboard for displaying visualizations of the captured data. Components 1, 2, and 3 (see Figure 2) are part of the Unity application, and component 4 is a web application developed using React, a library for building web interfaces (React -A JavaScript library for building user interfaces, 2022) (see Figures 3 and 4). We used Azure's AI services (Microsoft Azure Cognitive Services, 2022) to facilitate

Figure 1. The avatar in a virtual office.

development that includes Speech-to-Text and Text-to-Speech conversion, Language Understanding (LU) and Text Analytics for intent and entity recognition, Face API for classification of facial expressions, an Azure Function that is used to serve the response retrieval program for interviewee-initiated questions, and the Azure Machine Learning (ML) service to host a stress detection model developed as part of prior work (Adiani et al. 2022) to allow for stress detection in real-time in the future from any location where CIRVR is running. Once the interview has concluded, the data collected is stored in the Cosmos database (Microsoft Azure Cosmos DB, 2022) and are then retrieved via the Dashboard, that uses this data to create initial visualizations as seen in Figures 3 and 4. This Dashboard can be used by a job coach to analyze the data and develop strategies to help the interviewee improve performance, and to help them understand their own performance during a simulated interview. Previous work provides further details regarding the system (Adiani et al. 2022).

Experiment protocol

For the feasibility study of CIRVR, 23 autistic individuals (mean age = 18.55, SD = 2.31, 18 males and 5 females, 21 White out of which 4 were Hispanic, and 2 African American), were recruited for a feasibility study to capture data from the four modes of input while interacting with CIRVR. All participants had a clinical diagnosis of ASD by licensed clinical psychologists. The inclusion criteria were – participants needed to be verbal and fluent in English, clinically diagnosed with ASD, and were at least the age for full-time employment in the state where the research was conducted (i.e., 16 years or older). The participants had postsecondary education and did not have co-occurring intellectual disabilities. A job interview for a Data Entry position was administered which was structured into segments – Greetings, Previous Work Experience, Technical/Educational/Personal question-answers, and interviewee-initiated Questions – similar to the approach of Baur et al. (2013). The study was approved by the Institutional Review Board (IRB)[3] of the lead author's university and informed consent/assent were obtained from the participants before administering CIRVR. To capture post-

[2]https://www.dedoose.com/.
[3]IRB Approval #191277.

Table 1. The four modes of data captured by CIRVR and what they are used to infer.

Mode	Device	Data Captured	To Infer
Verbal	Headphones with microphone	Interviewee utterances, their speech	Entities and intents, and speech acts to understand how the interviewee responds to questions.
Physiological	Empatica E4 wristband sensor	Photoplethysmography (PPG) and Electrodermal activity (EDA) (see Table 3)	Stress predicted by a machine learning model trained on the sensor data.
Gaze	Tobii Eye X or Tobii Eye Tracker 5	Eye gaze points in the form of *(x, y)* coordinates and regions of interest (ROI), and fixation points where fixation is a gaze point maintained for a period of time	Where the participant looks during the interview, and the duration of each gaze point (fixation duration[4])
Facial Expressions	Logitech C920 HD Pro web camera	Facial images	One of the eight Ekman emotions.

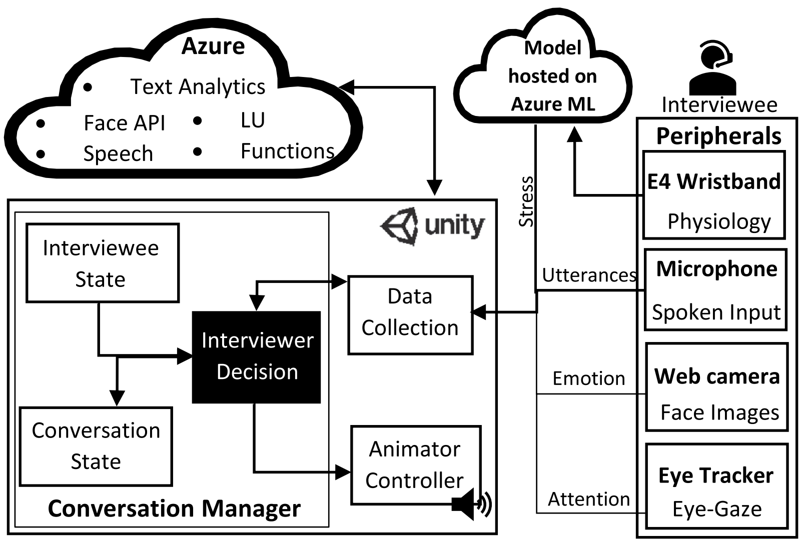

Figure 2. Components of CIRVR.

session user experience data, a questionnaire was administered to capture comfort and confidence scores and additional feedback on CIRVR as reported in the Results section in Table 4.

The qualitative data described above informed the design of CIRVR and the Dashboard. Once the Dashboard was developed, unstructured feedback was solicited regarding its usefulness, following a demonstration of the Dashboard (i.e., a display and description of its components and features) collectively from 4 autistic employees and 1 job coach of a local company. This informal feedback was aimed to seek a general reaction to the system and to ask questions about the system as they arose. Data were collected in a more structured way through a brief questionnaire to fill out by hand, that included questions regarding demographics and background of the participants as well as requesting open-ended comments regarding their perceptions of the Dashboard. A screenshot of the questionnaire is provided in Appendix A (see Figure A1). Qualitative feedback from the four autistic employees and one job coach were analyzed for shared reactions to the Dashboard. Given the written responses to each open-ended question in the questionnaire were limited in length (usually one to two sentences), they were not analyzed in the same manner as the lengthier stakeholder interviews. Instead, it was concluded that a shared reaction exists when the majority of all responses (i.e., 3 of 5 participants' responses to each question) contained a particular valence (positive/negative) or detail (e.g., examples of how CIRVR can aid empathy), and if both employees and job coach shared the reaction. A summary of the responses is presented in the Results section.

Multimodal data capture

Verbal mode: utterance classification

Spoken interaction during interviews in CIRVR were labeled and classified to provide insights into an autistic individual's interaction with the interviewer. Understanding how an autistic interviewee responds to questions during an interview will potentially help a job coach identify where an individual needs coaching and create an individualized training plan for them. The utterances were classified using a dialogue act annotation scheme in an interview scenario described in (Chakravarty et al., 2019; Farzana et al., 2020), with a few modifications. In addition to the

[4]Fixation duration is the time duration of the eyes gazing upon an object in its surroundings, where a fixation lasts approximately 250 milliseconds, but may vary, either shorter or longer (Galley et al., 2015).

Table 2. Answer dialogue acts.

Label	Description	Example
y	Yes answer or a variation.	"Yes," "yeah," "yep," etc.
y-d	Yes answer with explanation.	"Yes, I have experience in Java."
n	No answer or a variation.	"No," "nah," "not at all," etc.
n-d	No answer with explanation.	"No, I do not have experience with Unity."
sno	Statements of non-opinion.	"I worked as a someone who used to work in maintenance back in 2016 and prior to that I was a childcare assistant for New Hope Academy."
so	Statements of opinion.	"But I think that I've done a pretty good job so far."
ack	Acknowledgments.	"Okay," "uh huh," "I see," etc.
dno	Other answers.	"I don't know."
xx	No Response (interviewee did not respond or took too long to answer and time-limit expired, or the user was interrupted by avatar which led to incomplete responses).	"User didn't respond, "The user did not ask anything," "the . . . ," "I . . . ," etc.
query	Interviewee-initiated question.	"What is my salary?"
ft	Thanking.	"Thanks," "thank you."
fa	Apologies.	"Sorry"
fp	Conventional, such as greetings and farewell.	"Hi," "hello," "nice to meet you," "bye," etc.
fe	Exclamations or curses.	"geez," "gosh," "goodness"

Table 3. Physiological features.

Physiological Signal	Feature	Unit of Measurement
Photoplethysmography (PPG)	The mean value of the heart rate	Beats per min (bpm)
	Standard Deviation of Heart Rate	No unit
Electrodermal Activity (EDA)	Mean of the skin conductance level (SCL), which gives a moving baseline of the EDA	microSiemns (µS)
	Standard Deviation of SCL	No unit
	Skin Conductance Response (SCR) Rate gives the stimuli sensitive peaks found from the EDA data	Response Peaks/s
	Mean of the Skin Conductance Response (SCR)	Response Peaks/s
	Standard Deviation of SCR	No unit
	SCR Max	Response Peaks/s

Table 4. Results from post-session interview.

Independent Variable	Participants Mean (SD)
Age (N = 23)	18.55 (2.31)
Post Session Review (N = 20[5]**)**	
I think that I would like to use this system frequently[a]	3.30 (1.10)
I found the system unnecessarily complex	2.05 (1.13)
I found various functions in the system to be well integrated	3.80 (0.93)
I would imagine that most people would learn to use the system very quickly	3.70 (1.15)
I think that I would need the support of a technical person to be able to use this system	2.80 (1.50)
I needed to learn a lot of things before I could get going with the system	2.55 (1.32)
How confident[b] did you feel during the interview?	7.00 (2.10)
How comfortable did you feel . . . overall during the interview?	7.15 (2.08)
. . . during your conversation with the interviewer?	6.63 (2.18)
. . . during the whiteboard section?	6.95 (2.40)
. . . during the knock at the door interruption?	6.00 (3.03)
. . . with the wristband and the headset?	8.65 (1.65)
How much do you agree[c] with the following? Practicing with this system would help me get a job in the future.	7.30 (2.08)
How much do you agree with the following? If it was available, I would use this system to prepare for a job interview.	7.70 (2.26)
In what format might you prefer to use this system? VR, desktop, laptop, mobile?	10 (50%) VR 5 (25%) Mobile 10 (50%) Desktop 4 (20%) Laptop
In what setting might you prefer to use this system? At school, at home, at a vocational support office?	6 (30%) School 12 (60%) Home 7 (35%) Support Office

[a]Reported on a 5-point Likert scale (1=Strongly Disagree, 5=Strongly Agree), [b]Self-reported scores on a 10-point rating scale (1=Not confident at all or Not Comfortable at all, 10=Very Confident or Very Comfortable), [c]Reported on a 10-point rating scale (1=Completely disagree, 10=Completely agree).

labels for answers given in (Chakravarty et al., 2019), six more labels have been included (see last six entries in Table 2). The last four labels have been adapted from (Jurafsky & Shriberg, 1997; Stolcke et al., 2000). Two annotators were trained to classify the utterances into dialog acts using the Question dialogue act labeling scheme in Chakravarty et al. (2019), and the Answers dialogue act labeling scheme in Table 2. The inter-rater reliability was 86% and was calculated by taking the percentage of the number of utterances labeled the same dialogue act by both annotators, out of the total number of utterances.

[5]Some sessions were conducted during the COVID-19 pandemic and the consent forms and post-session questionnaires were administered electronically and post-session scores were lost for 3 participants due to technical errors (e.g., corrupt file, unsaved file lost).

Figure 3. Visualizations in the dashboard.[6] (a) the stress graph showing the user's overall stress values captured throughout the interview (b) a radial graph showing the user's overall emotion values captured during the interview and the overall average emotion of all users. (c) bar graph of 7 emotions displaying the data of one user relative to the average emotion levels of all participants in the study to date. (d) a scatter chart showing the user's eye gaze-based fixation points plotted on a static image of the interviewer in the virtual environment. (e) the heat graph displaying fixations accrued during the entire interview where intensity of the color in the individual cells of the grid represent the number of fixation points located in the area, i.e., the longer the interviewee gazed at the area.

Physiological mode: physiology and stress detection

Understanding and managing stress and related emotions (e.g., anxiety) during interviews could help improve interview performance and, potentially, employment outcomes (McCarthy & Goffin, 2004). Ground truth stress ratings on a level of 1–10 in 15 second intervals were obtained from videos of the participants undergoing the interview by a behavior analyst trained in rating autistic behavior. These labels were then converted to binary labels with a rating below 4 equaling 0 and a rating of 5 or higher equaling 1. The ground truth stress ratings allow for initial analysis of stress trends. Stimuli which are sensitive to physiological changes in the cardiovascular system and electrodermal activity have been linked with stress (Bian et al., 2015; Cacioppo et al., 2007; Kim et al., 2018). Therefore, physiological data were gathered from the wearable sensor. Specific physiological and statistical features, shown in Table 3, were derived after a baseline was

[6]Created using JavaScript-based data visualization package, *Nivo* (Nivo, 2022).

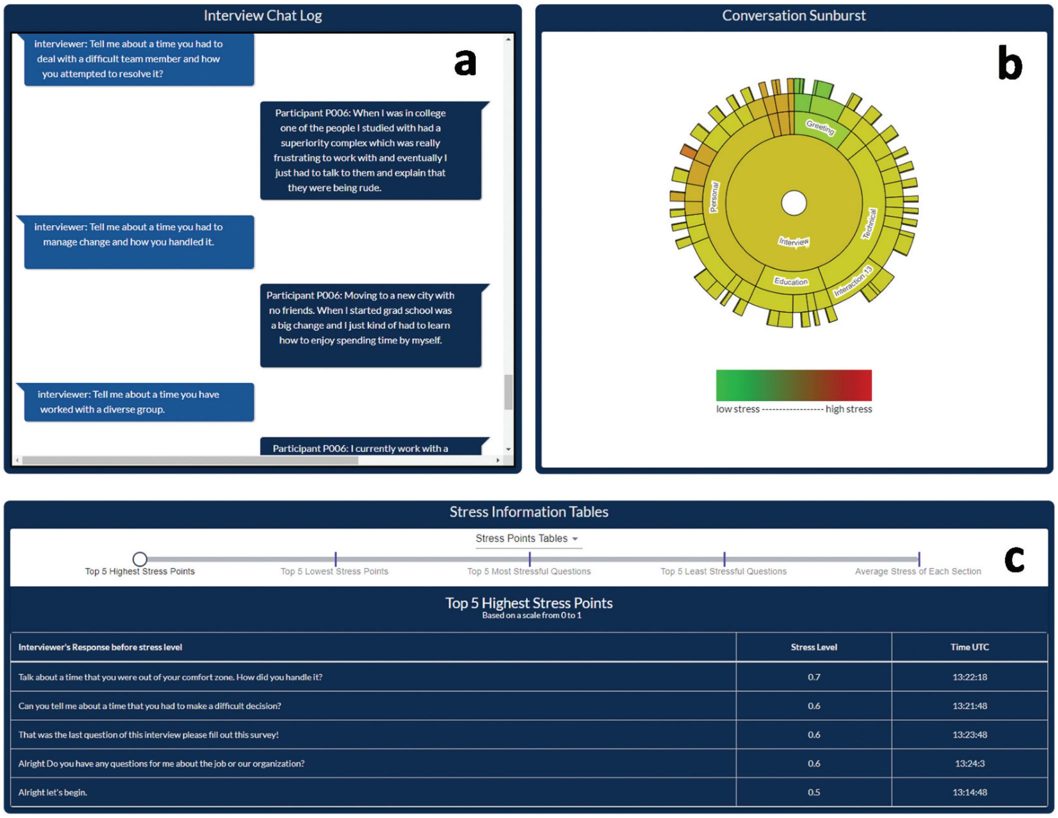

Figure 4. More visualizations in the Dashboard. (a) the interview chat log[7] of one session, with interviewer questions and interviewee questions presented in chat bubbles similar to a messaging/texting platform. (b) the conversation sunburst[8] graph that represents the conversation split into different levels of rings where the colors signify stress levels. (c) stress information table[9] displaying the top 5 highest stress points before questions were asked, and their timestamps.

established based on recommendations from prior research (Bian et al., 2015; Migovich et al., 2021).

Gaze mode: eye gaze detection

Eye gaze is an important aspect of social interaction and can be used to perceive information from others and to communicate attention (Ca˜nigueral & Hamilton, 2019). Studies have shown that percentage of time an interviewee gazes at the interviewer influences likelihood of being hired (Nguyen et al., 2014) where less eye-contact is viewed unfavorably by NT interviewers (McGovern & Tinsley, 1978). As such, autistic candidates are more likely to be viewed unfavorably irrespective of their competency and skillset (Maras et al., 2020). Therefore, we capture this metric to show that gaze patterns of autistic candidates appear unrelated to quality of responses to interview questions. To capture eye gaze, data from the eye-tracker were recorded as (x, y) coordinates where the bottom left of the screen is $(0, 0)$ and the top right of the screen is $(1, 1)$. CIRVR creates several regions of interest (ROI) within the virtual environment and keeps track of where the participant is looking (e.g., the eyes, the face, the surroundings like the clock on the wall). For this study, the focus was specifically on the interviewer. To understand the range of eye gaze made toward the interviewer when compared to the overall gaze throughout the session for each participant, the percentage of gaze points that fell within the interviewer ROI out of the total number of gaze points in the session is calculated, where percentages were grouped as ranges, i.e., 0–3%, 4–6%, 6–9%, etc.

Facial expression mode: emotion detection from facial expressions

Facial expressions can be used to assess the emotional state of an individual. They also can arouse certain emotions in an individual who is observing them, which in turn, can influence one's response during a social interaction (Frith, 2009). Autistic individuals often struggle with understanding and describing their own emotions (Silani et al., 2008). Stress data and facial expressions have been previously used to help autistic children and their caregivers understand their emotional state (Gay et al., 2013). Therefore, facial expressions are an important mode that can be used along with verbal, physiological, and eye gaze responses to provide an overall understanding of an interviewee's interaction and performance during job interview training. The webcam was used to capture one's facial image every 5.5 seconds. The Face API (Microsoft Azure Cognitive Services, 2022) called from the Unity application processed each image and returned confidence scores for each of the eight universally recognized emotions: joy,

[7]Created using the package *react-chat-elements* (React Chat Elements, 2022).
[8]Created using the package *sunburst-chart* (Sunburst Chart, 2022).
[9]Create using the package *react-bootstrap-table* (React Bootstrap Table, 2022).

surprise, fear, anger, disgust, contempt, neutral, and sadness (Ekman, 1993). The one with the highest confidence score was assigned as the emotion for the image with a facial expression. The percentage of the number of each detected facial expression per image by the total number of images captured for all participants was calculated. The results are reported in the next section.

Results

From analysis of qualitative data before the development of CIRVR

The thematic analysis on the qualitative data from interviews with stakeholders identified 5 major themes as participants discussed barriers and facilitators to successful interviews. These overarching themes were identified by multiple interviews among three stakeholder groups as described in the Stakeholder Input section in Methods. Where there were differing emphases within a theme, subthemes were created. The 5 major themes were 1) navigating unpredictability (with subthemes of knowing what to expect in advance, interpreting questions and comments), 2) introducing flexibility and modifications, 3) relationship building strategies (with the subtheme of masking as a strategy), 4) importance of self-awareness and self-advocacy, and 5) nuances of self-disclosure. Autistic employees described barriers experienced and strategies used during an employment interview. Employers described characteristics that they observed from autistic candidates and employees who were successful at the job interview, as well as strategies adopted by their companies to support autistic job seekers. Service providers described their role and strategies for supporting autistic candidates and employers.

Some representative examples of the "navigating unpredictability" theme derived from the interviews and focus groups included – autistic interviewees often struggle with unstructured questions and unexpected events (e.g., interruptions) in part because they trigger negative emotion (e.g., feeling overwhelmed). Insights obtained motivated the incorporation of features such as open-ended questions that are unstructured and unexpected, and interruptions during the Interview to provide practice opportunities to address these issues. Additionally, from the subtheme of "navigating unpredictability" of "interpreting questions and comments," all stakeholders expressed need for more coaching in answering interview questions (e.g., being able to structure answers) and practice with emotion regulation and stress management. A fuller analysis of the wider qualitative data collection and analysis, not directly tailored to CIRVR is documented elsewhere Chang et al. (In Press).

From the CIRVR data collection study

The CIRVR interviews ran smoothly 22 out of 23 times with only two instances of the researcher intervening when CIRVR hung during one of the sessions. It was an issue with the speech recognition which hung when the participant pressed the key multiple times to submit one answer. The issue was fixed for subsequent studies via error handling within CIRVR. Multimodal data were captured successfully—22/23 times for speech, 23/23 times for gaze, 21/23 times for facial expressions, and 17/23 times for physiology – in the required format except for physiological data loss for 6 participants due to recording errors where participants may have been wearing the sensor loosely. This was corrected for subsequent participants and no participant mentioned any discomfort with the wristband as we can see from the results in Table 4. Facial expression data were lost for the initial two participants due to errors in saving within the system and one speech log was considered invalid due to incorrect response submission. All data were gathered in-situ and in the desired format in Comma Separated Values (CSV) files for autonomous processing and displaying on the Dashboard. These data can be uploaded to the database using the Dashboard interface, which is immediately processed and displayed for viewing. The data are reported as descriptive statistics and visualizations. The results in Table 4 show the mean and standard deviation of the scores from all the participants' responses from each question of the post-session /post-experiment questionnaire.

From feedback on the dashboard

Positive feedback was received on the layout of the Dashboard (see Figures 3 and 4). All five participants reported positive reactions to the Dashboard layout. Some comments were – was "easy to follow" (as noted specifically by an employee and job coach) and looked professional (four of the five participants) with other specific examples of the Dashboard being positively highlighted including the "variety of visualizations" and "readability" of the dashboard. The one participant who disagreed noted that some "additional interpretation/narrative" of how to interpret the Dashboard, that could be "toggled on/off," would be useful. Four of the five participants described how the Dashboard presented information in a way that would help them empathize with job candidates. The job coach specifically noted that the information in the dashboard would help them empathize with autistic job candidates and allow them "to better understand perspectives that are not conventional." All four autistic employees felt the Dashboard would improve their interview performance by helping "analyze stress levels and eye contact," "pinpoint which questions . . . to prepare for more," and "help prepare for interviews and become comfortable before an interview." Two autistic employee participants expressed that the fixations scatter point visualization (see Figure 3d) was useful and would further help interview performance. Another participant felt that knowing which stress inducers enhance or do not enhance one's job performance, might be informative to autistic individuals and/or job coaches. Overall, their insights provided initial reinforcement of the utility of the Dashboard as a feature of CIRVR.

From compiling collected multimodal data

Multimodal data are presented in Figures 5–9. Figure 5 shows the overall percentage of each of the types of dialogue acts for the utterances from 22 participants (e.g., approximately 59% of the responses were *Statement Non-Opinion*, i.e., the participants spoke about themselves). Physiological data were recorded from participant seven onwards due to technical problems with the wristband sensor and/or due to the

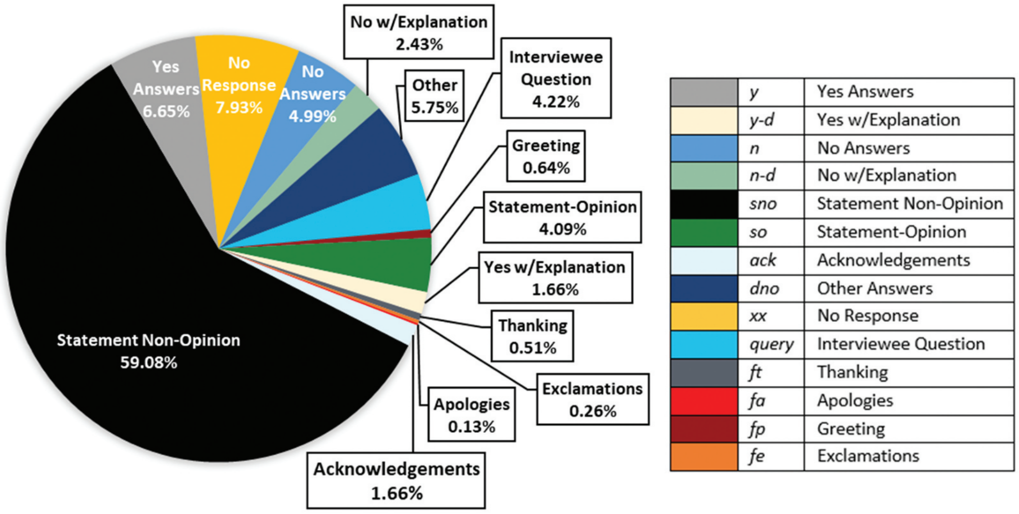

Figure 5. Overall distribution of dialogue acts of utterances from $N = 22$ participants.

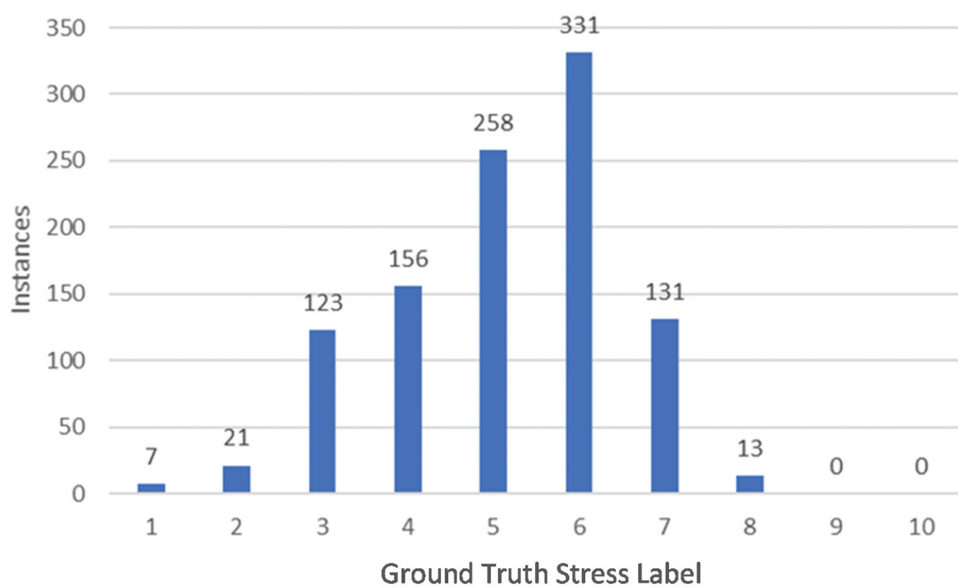

Figure 6. Distribution of ground truth labels ($N = 17$).

wristband being worn loosely. The distribution of ground truth stress labels from the 17 participants are presented in Figure 6, where 1 and 10 correspond to the lowest and highest stresses, respectively. Table 5 displays the average stress level for each segment of the interview (e.g., the average annotated stress during the Education segment of the interview was 5.57 out of 10). Figure 7 shows how often the participants looked at the interviewer avatar's face, which varied from 0.00% to approximately 18% (e.g., 12 participants were within the 0–3% range). The emotion ratings received from Face API were used to understand the emotional state of the interviewee during the interview. Figure 8 shows the overall distribution of emotions during the interview for all the participants (e.g., happiness was detected in 23% of the participants). Figure 9 shows, as an example, the distribution of emotions over time for one of the participants where we can see the fluctuation in emotions detected throughout the interview at each timestep.

Discussion

Insights from stakeholders data

The design and development of CIRVR and the Dashboard were informed by the insights of three sets of stakeholders – autistic job candidates, employers, and service providers. Overall, they illuminated challenges to successfully navigating employment interviews and obtaining employment as well as experiences, information, and strategies that might help surmount these challenges. Specifically, the "navigating unpredictability" theme led to CIRVR incorporating open-ended questions that are both unstructured and unexpected as well as interruptions during the interview, all of which provide experiences to better understand and address specifically difficult situations. Emotion recognition and stress detection modules as well as the Dashboard including displays of the emotional experience of the interviewee throughout the interview (e.g., when stress and negative emotions are

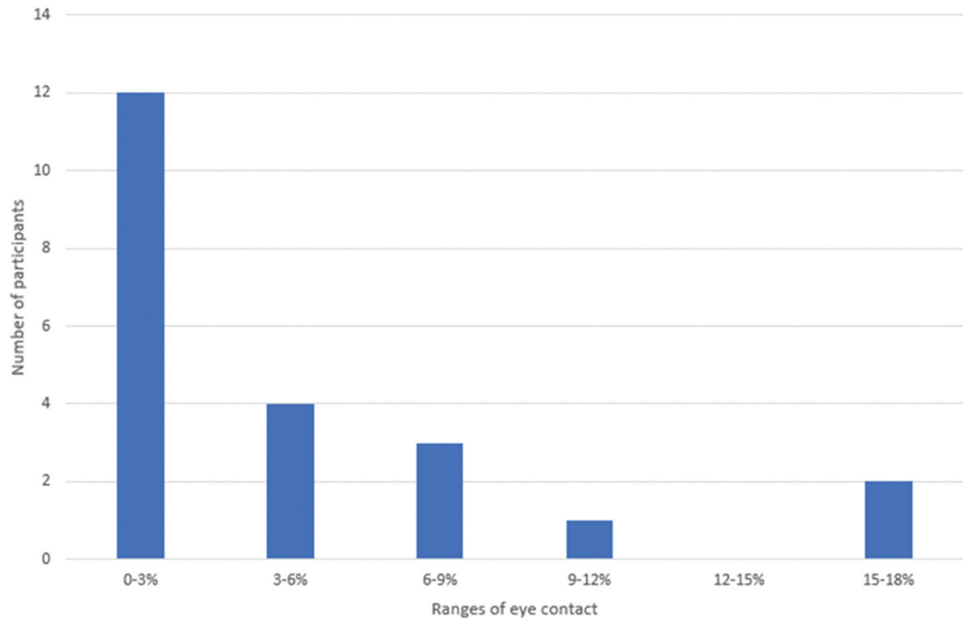

Figure 7. Overall range of eye-contact made toward the interviewer during the interview ($N = 23$).

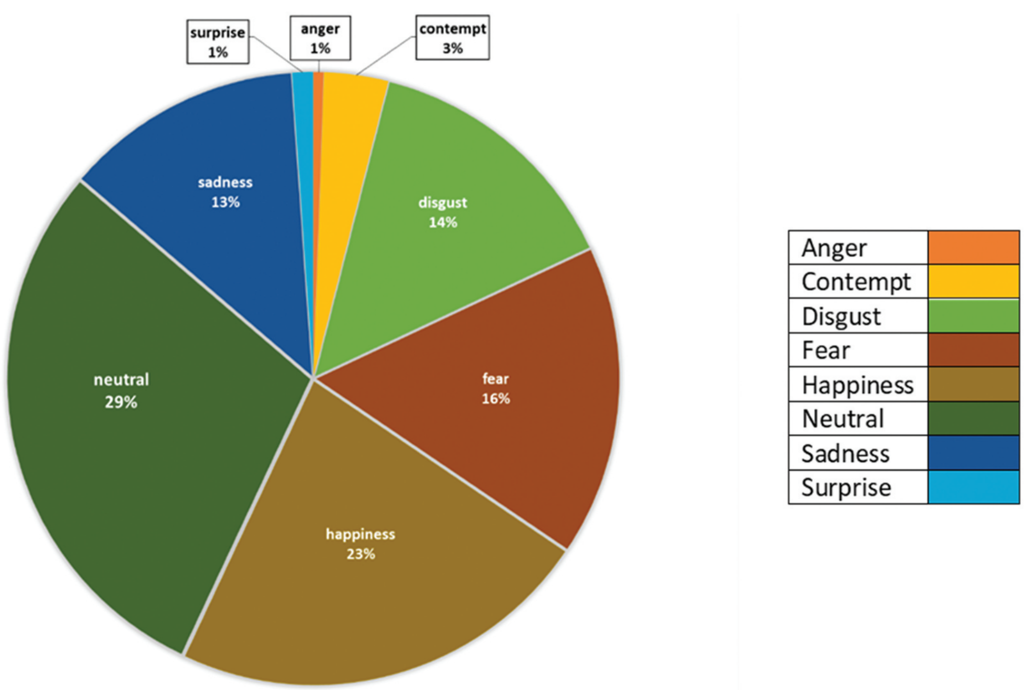

Figure 8. Overall distribution of emotions during the interview as predicted by Face API ($N = 21$).

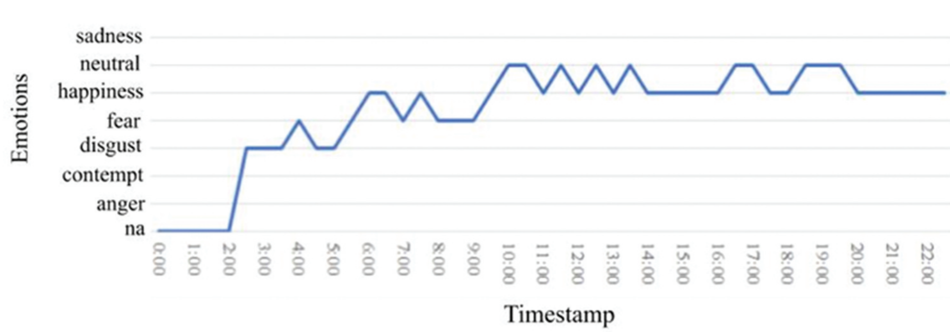

Figure 9. Emotions detected from facial expressions of participant 5 over the span of the interview (in minutes).

Table 5. The overall stress values per segment.

Interview Segment	Average Stress (out of 10)
Greeting	4.09
Previous Work Experience	4.57
Technical	4.69
Education	5.57
Personal	5.52
Questions	4.92

experienced) and how emotion corresponds to each interview question, resulted from feedback about the importance of navigating unpredictability and understanding emotional response to the unpredictability.

Insights from multimodal data

Utterance classification

CIRVR aims to build upon the strengths of current job interview training technologies by integrating multimodal data to provide richer insight into the emotional experience and verbal and behavioral response patterns of autistic individuals during interviews. In the study, it was observed (see Figure 5) that the *sno* type dialogue act, which indicates personal statements where the interviewee says something about themselves, has the highest percentage during an interview, similar to earlier work (Stolcke et al., 2000). The next highest percentage is of *xx* labels, which was both unexpected and new. The label *xx* accounts for the times when the participant either did not respond to a question because time-limit (1–2 minutes) for that question ran out, or they were interrupted by simulated interruptions. Out of the 23, 4 participants had several non-response *xx* tagged utterances. On viewing the video and file logs, the authors observed that one participant was responding to the questions, however, they were not hitting the correct submit key despite there being instructions on screen on which key control was for submission, as the conversation log had "USER DID NOT RESPOND" for all of the questions. As for the other three participants, the logs revealed that in the beginning they were still figuring out the system or needed more time to respond to the questions and they had to be repeated for them, and hence they were labeled as *xx*. The non-response resulting from time pressure also reveals to employers the importance of recalibrating expectations regarding timeliness of response and allowing for pauses or other opportunities to think through answers longer. This also emerged in the qualitative interviews and focus groups in the "introducing flexibility and modifications" theme, as a barrier and facilitator (when given more time, performance was reported to improve). CIRVR's ability to categorize participant responses to questions in a meaningful way through dialogue act classification could provide insight to job coaches and support professionals for offline analysis of spoken interaction to create an effective and individualized training program.

Stress detection

Figure 6 indicates that CIRVR did not cause extreme stress (e.g., rating of 9 or 10) in any of the participants. Most of the ratings were a 6, at 31.8% of the total ratings, suggesting that the virtual interview elicited a moderate stress response, as expected of the simulated task. The stress data can also be used to explore trends in stress during specific segments of the interview and for particular questions. Table 5 shows that the most stressful segments for participants, on average, were the Personal and Education question segments. This information can be used to create interview scripts that target these segments to provide more exposure to questions that may be asked. Informing interviewers that personal questions and educational questions tend to be the most stressful for this population may also lead to modifying the way these questions are asked to reduce stress (e.g., with greater prompting, more structure). The ability to collect and derive physiological features that are known to be correlated to stress opens the door for the creation of machine learning models that can detect stress without the additional labor needed from a trained behavior expert. Overall, it is important to be able to detect stress so that the system can adapt to the participant and increase their comfort.

Eye gaze

According to Shellenbarger (2013), NT adults make eye-contact between 30–60% of the time in human-human social interaction. While, overall, the autistic participants did not maintain eye contact during the interview at the same level as expected from NT adults (see Figure 7), this did not indicate that they were not paying attention, as each of them completed their individual interview, despite taking more time to think about some questions. Such information is useful to educate job interviewers who may not have experience interviewing autistic individuals. The eye gaze data can also be used to inform autistic individuals of their own gaze patterns.

Facial expression

As for the facial expression data, neutral and happiness were the most frequent emotions exhibited by all the participants (i.e., over 50% of the total number of instances) whereas there were few instances of anger, surprise, and contempt recorded (see Figure 8). Since the facial images were recorded every 5.5 seconds, the system can capture the variability of emotions over the entirety of the interview. The emotions marked by timestamps can be mapped to different questions asked during the interview. For the first nine minutes of the interview (see Figure 9), disgust and fear were the most detected emotions with few instances of happiness for Participant 5. As the interview progressed, there were more instances of neutrality and happiness. This might be an indication of the interviewee feeling more nervous in the beginning and getting more comfortable with time. Such data can be insightful to understand what aspects of the interview aroused negative or positive emotions. This can lead to discussions that can help job coaches use a more targeted approach to train the interviewees and interviewers (i.e., provide strategies for managing stress initially). It can also help employers by signaling the importance of starting with questions that are especially clear and likely to build comfort and confidence. Moreover, it might also suggest that some of the initial trepidation could be managed by providing job candidates with the interview questions in advance so they know what to expect. One important aspect of facial emotion recognition is that the training dataset must appropriately represent autistic faces to reduce bias. Currently there are no such dataset available publicly and CIRVR used a NT face dataset. Although CIRVR is able to use any dataset for facial emotion recognition for training, current results are based on an NT dataset and thus needs to be interpreted with caution.

Insights from post-experiment feedback

From the post-session feedback (see Table 4), the authors found that participants felt CIRVR's features were well integrated, and they would use the system if available to prepare for job interviews. The overall confidence and comfort scores varied across participants and the overall comfort score was higher than the confidence score, which aligns with previous work that autistic individuals often have lower confidence during interviews than their NT peers (Maras et al., 2020). Overall, the participants felt comfortable using CIRVR and thought that the data obtained from CIRVR would be useful with a job coach to guide them through the interview process.

Overall insights

CIRVR can classify utterances with dialogue act labels to provide insight about conversation patterns during an interview. Stress ratings obtained from the system can be analyzed in order to provide employers and job coaches with data to help identify specifically challenging interview questions or segments and adapt interviews to be less aversive and more effective at assessing the skills of the job candidates. CIRVR is able to record physiological data that will give further insight into stress levels in future versions. CIRVR can capture fine-grained eye gaze data; initial results demonstrate that although autistic individuals made much less eye contact relative to NT-based expectations, they were able to complete the interview process, despite having to take more time to think about their response, including saying "no" or "I don't know," to some questions. The authors believe data resulting from these capabilities can be provided to employers and job coaches to help reduce bias faced by autistic interviewees. CIRVR also captures facial expressions experienced during the interview process. These data can provide important insights into how a participant feels during an interview. A de-identified group data of facial expressions can be useful in designing more inclusive interviews by helping interviewers recognize signs of distress and altering the implementation and design of interview scripts. Overall, data visualized on the Dashboard, can help highlight sources of strength in the interview that may be applied to other questions seen as more challenging and stressful. This information can be used to understand which questions might require more preparation substantively and emotionally, for which the autistic candidate may desire to prepare more with the help of a job coach. It also informs a future interviewer, which questions are likely to be uncomfortable or difficult for autistic individuals and may need to be restructured/rephrased or removed from the interview altogether.

Limitations and future work

CIRVR, despite its potential, does have limitations. Although CIRVR's functionality allows the interviewee to ask the interviewer questions they have about the job (e.g., work environment, daily tasks, about the job, and more), the domain is still limited to the job interview and thus limits truly free-flowing conversation. Stress ratings are currently obtained from a trained behavior expert which is a laborious process. This version of CIRVR does not include automated stress detection such as via machine learning as more research is needed to understand how to avoid bias. CIRVR, currently uses commercially available Facial Emotion Recognition (FER) models that are not trained on autistic faces. Studies have shown that autistic individuals produce facial expressions differently from the NT population (Trevisan et al., 2017) and thus it is important to develop FER models trained on autistic datasets in the future. Finally, there were some limitations of the Dashboard interface. Participants mentioned that they would need some explanation on the purpose of the visualizations, as it was unclear what each one represented (e.g., the conversation sunburst in Figure 4b). The participants mentioned that they would like to see key performance metrics pertaining to answering questions, not currently part of the Dashboard, to provide clearer guidance and feedback on how to improve. For example, highlighting questions where they performed well, and how the candidate can replicate that result across the interview.

Future work includes developing a feedback mechanism which provides suggestions for how to improve based on recommendations from job coaches as well as other support professionals. CIRVR will be upgraded into a closed-loop system with an enhanced dialogue framework where the interviewer's questions will be based on information gathered from not only spoken input, but also eye gaze, stress, and the interviewee's affective state. Now that more physiological data has been collected from the interview sessions, the authors will further improve the initial Random Forest stress detection model (Adiani et al. 2022) which will be part of future work as a standalone project. The authors will also strive to create a test dataset of facial images of autistic individuals since NT facial expression datasets do not capture autistic facial expressions well. Based on the feedback received on the Dashboard interface, an information tab needs to be added for each visualization that explains the purpose of the data and what insights can be obtained via observation, and a few enhancements need to be made to the general interface (e.g., font size, color schemes, clear axis labels, etc). Hence, a manual will be made to accompany the Dashboard to aid the job coaches in using this tool. For the CIRVR interface, a tutorial is needed to help familiarize the participants with the keyboard controls before the main interview session begins, as the pre-session and on-screen instructions did not seem enough since some participants would press the wrong key for submission. This tutorial will be part of future work. Finally, a feasibility study of CIRVR and the Dashboard with a broader set of job coaches and employers is underway to further aid the successful real-world deployment and implementation of CIRVR.

Conclusion

The current work shows feasibility of CIRVR in collecting multiple modes of data which can be used to improve interview experience and potentially interview outcomes of autistic individuals. The interviews ran smoothly, with the exception of two instances that required intervention. Aside from initial physiological data loss due to errors of the Empatica E4 wristband sensor, data were successfully collected during the study in the desired format which can be processed via an automated pipeline and displayed immediately post-session on the Dashboard for analysis and coaching. CIRVR provides useful multimodal information regarding an interviewee's emotional experience and interaction during job interviews and thus will help create customizable job interview training protocols for autistic candidates and more targeted interventions by job coaches and support professionals. It also holds promise for guiding employers toward interviews that are more inclusive of neurodiversity by identifying which questions or segments of the interview elicit high stress and/or lower quality responses.

Acknowledgements

The authors gratefully acknowledge the support provided by Microsoft AI for Accessibility Program and by the National Science Foundation under grant numbers 1936970 and 2033413 for this project. They would like to thank the reviewers and the Editor-in-Chief of *Assistive Technology Journal* for their constructive and insightful comments which helped clarify the contributions of the work and in turn, strengthen the manuscript. The authors are solely responsible for the content of this article. None of the authors have any financial conflict of interest.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work was supported by the National Science Foundation [1936970, 2033413]; AI for Accessibility Grant from Microsoft [No number].

References

Adiani, D., Itzkovitz, A., Bian, D., Katz, H., Breen, M., Hunt, S., Swanson, A., Vogus, T. J., Wade, J., & Sarkar, N. (2022, March). Career Interview Readiness in Virtual Reality (CIRVR): A platform for simulated interview training for autistic individuals and their employers. *ACM Transactions on Accessible Computing*, *15*(1), 28. https://doi.org/10.1145/3505560

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of mental disorders: DSM-5* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Baldwin, S., Costley, D., & Warren, A. (2014). Employment activities and experiences of adults with high-functioning autism and asperger's disorder. *Journal of Autism and Developmental Disorders*, *44*(10), 2440–2449. https://doi.org/10.1007/s10803-014-2112-z

Baur, T., Damian, I., Gebhard, P., Porayska-Pomsta, K., & Andr´e, E. (2013). A job interview simulation: social cue-based interaction with a virtual character. *2013 International Conference on Social Computing* (pp. 220–227). https://doi.org/10.1109/SocialCom.2013.39

Bian, D., Wade, J., Swanson, A., Warren, Z., & Sarkar, N. (2015). Physiology-based affect recognition during driving in virtual environment for autism intervention. *Proceedings of the 2nd International Conference on Physiological Computing Systems* (pp. 137–145). https://doi.org/10.5220/0005331301370145

Black, M. H., Mahdi, S., Milbourn, B., Scott, M., Gerber, A., Esposito, C., Falkmer, M., Lerner, M. D., Halladay, A., Ström, E., D'Angelo, A., Falkmer, T., Bölte, S., & Girdler, S. (2020). Multi-informant international perspectives on the facilitators and barriers to employment for autistic adults. *Autism Research*, *13*(7), 1195–1214. https://doi.org/10.1002/aur.2288

Booth, J. (2016). *Autism equality in the workplace. removing barriers and challenging discrimination*. Jessica Kingsley Publishers.

Bublitz, D. J., Fitzgerald, K., Alarcon, M., D'onofrio, J., & Gillespie-Lynch, K. (2017). Verbal behaviors during employment interviews of college students with and without ASD. *Journal of Vocational Rehabilitation*, *47*(1), 79–92. https://doi.org/10.3233/JVR-170884

Burke, S. L., Bresnahan, T. L., Li, T., Epnere, K., Rizzo, A., Partin, M. T., Ahlness, R. M., & Trimmer, M. (2018). Using virtual interactive training agents (ViTA) with adults with autism and other developmental disabilities. *Journal of Autism and Developmental Disorders*, *48*(3), 905–912. https://doi.org/10.1007/s10803-017-3374-z

Ca˜nigueral, R., & Hamilton, A. F. D. C. (2019). The role of eye gaze during natural social interactions in typical and autistic people. *Frontiers in Psychology*, *10*, 10. https://doi.org/10.3389/fpsyg.2019.00560

Cacioppo, J. T., Tassinary, L. G., & Berntson, G. (2007). *Handbook of Psychophysiology*. Cambridge university press.

Campbell, J. L., Quincy, C., Osserman, J., & Pedersen, O. K. (2013). Coding in-depth semistructured interviews: Problems of unitization and intercoder reliability and agreement. *Sociological Methods & Research*, *42*(3), 294–320. https://doi.org/10.1177/0049124113500475

Chakravarty, S., Chava, R. V. S. P., & Fox, E. A. (2019). *Dialog acts classification for question-answer corpora*. ASAIL@ICAIL.

Chang, H. Y., Saleh, M. C., Bruyere, S. M., & Vogus, T. J. (In Press). Making the employment interview work for a neurodiverse workforce: Perspectives of individuals on the autism spectrum, employers, and service providers. *Journal of Vocational Rehabilitation*.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46. hdl:1942/28116. S2CID 15926286. https://doi.org/10.1177/001316446002000104

DeVault, D., Artstein, R., Benn, G., Dey, T., Fast, E., Gainer, A., Georgila, K., Gratch, J., Hartholt, A., Lhommet, M., Lucas, G., Marsella, S., Morbini, F., Nazarian, A., Scherer, S., Stratou, G., Suri, A., Traum, D., Wood, R., … Morency, L.-P. (2014). SimSensei kiosk: A virtual human interviewer for healthcare decision support. *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-Agent Systems* (pp. 1061–1068).

Ekman, P. (1993). Facial expression and emotion. *The American Psychologist*, *48*(4), 384–392. https://doi.org/10.1037/0003-066X.48.4.384

Farzana, S., Valizadeh, M., & Parde, N. (2020). Modeling dialogue in conversational cognitive health screening interviews. *Proceedings of the 12th Language Resources and Evaluation Conference* (pp. 1167–1177).

Frith, C. (2009). Role of facial expressions in social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3453–3458. https://doi.org/10.1098/rstb.2009.0142

Galley, N., Betz, D., & Biniossek, C. (2015). Fixation durations-why are they so highly variable? Das ende von rational choice? *Zur Leistungsf ¨ahigkeit der Rational-Choice-Theorie*, 93. https://doi.org/10.13140/RG.2.1.3128.1769

Gay, V., Leijdekkers, P., Agcanas, J., Wong, F., & Wu, Q. (2013). *CaptureMyemotion: Helping autistic children understand their emotions using facial expression recognition and mobile technologies*. Bled eConference.

Giannakos, M. N., Sharma, K., Pappas, I. O., Kostakos, V., & Velloso, E. (2019). Multimodal data as a means to understand the learning experience. *International Journal of Information Management*, *48*, 108–119. https://doi.org/10.1016/j.ijinfomgt.2019.02.003

Jurafsky, D., & Shriberg, E. (1997). Switchboard SWBD-DAMSL shallow-discourse-function annotation coders manual.

Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., & Pellicano, E. (2016). Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism*, *20*(4), 442–462. https://doi.org/10.1177/1362361315588200

Kim, H. G., Cheon, E. J., Bai, D. S., Lee, Y. H., & Koo, B. H. (2018). Stress and heart rate variability: A meta-analysis and review of the literature.

Psychiatry Investigation, 15(3), 235–245. https://doi.org/10.30773/pi.2017.08.17

Maenner, M. J., Shaw, K. A., Bakian, A. V., Bilder, D. A., Durkin, M. S., Esler, A., Furnier, S. M., Hallas, L., Hall Lande, J., Hudson, A., Hughes, M. M., Patrick, M., Pierce, K., Poynter, J. N., Salinas, A., Shenouda, J., Vehorn, A., Warren, Z., Constantino, J. N., . . . Cogswell, M. E. (2021). Prevalence and characteristics of autism spectrum disorder among children aged 8 years — autism and developmental disabilities monitoring network, 11 sites, United States, 2018. *MMWR Surveillance Summaries*, *70*(11), 1. https://doi.org/10.15585/mmwr.ss7011a1

Maras, K. L., Norris, J. E., Nicholson, J., Heasman, B., Remington, A., & Crane, L. (2020). Ameliorating the disadvantage for autistic job seekers: An initial evaluation of adapted employment interview questions. *Autism*, *25*(4), 1060–1075. https://doi.org/10.1177/1362361320981319

Mayring, P. (2019). Qualitative content analysis: Demarcation, varieties, developments. *Forum: Qualitative Social Research*, *20*, 3. https://doi.org/10.17169/fqs-20.3.3343

McCarthy, J., & Goffin, R. (2004). Measuring job interview anxiety: Beyond weak knees and sweaty palms. *Personnel Psychology*, *57*(3), 607–637. https://doi.org/10.1111/j.1744-6570.2004.00002.x

McGovern, T. V., & Tinsley, H. E. (1978). Interviewer evaluations of interviewee nonverbal behavior. *Journal of Vocational Behavior*, *13*(2), 163–171. https://doi.org/10.1016/0001-8791(78)90041-6

Microsoft Azure Cognitive Services [Computer Software]. (2022). https://azure.microsoft.com/en-us/services/cognitive-services/

Microsoft Azure Cosmos DB [Computer Software]. (2022). https://azure.microsoft.com/en-us/services/cosmos-db/

Migovich, M., Korman, A., Wade, J., & Sarkar, N. (2021). Design and validation of a stress detection model for use with a VR based interview simulator for autistic young adults. *International Conference on Human-Computer Interaction* (pp. 580–588). https://doi.org/10.1007/978-3-030-78092-0_40

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Sage.

Nguyen, L. S., Frauendorfer, D., Mast, M. S., & Gatica-Perez, D. (2014). Hire me: Computational inference of hirability in employment interviews based on nonverbal behavior. *IEEE Transactions on Multimedia*, *16*(4), 1018–1031. https://doi.org/10.1109/TMM.2014.2307169

Nicholas, D. B., Hodgetts, S., Zwaigenbaum, L., Smith, L. E., Shattuck, P., Parr, J. R., Conlon, O., Germani, T., Mitchell, W., Sacrey, L., & Stothers, M. E. (2017). Research needs and priorities for transition and employment in autism: Considerations reflected in a “special interest group” at the international meeting for autism research. *Autism Research*, *10*(1), 15–24. https://doi.org/10.1002/aur.1683

Nicholas, D. B., & Klag, M. (2020). Critical reflections on employment among autistic adults. *Autism in Adulthood*, *2*(4), 289–295. https://doi.org/10.1089/aut.2020.0006

Nivo [Computer Software]. (2022). https://nivo.rocks/

React-A JavaScript library for building user interfaces [Computer Software]. (2022). https://reactjs.org/

React Bootstrap Table [Computer Software]. (2022). http://allenfang.github.io/react-bootstrap-table/

React Chat Elements [Computer Software]. (2022). https://detaysoft.github.io/docs-react-chat-elements

Roux, A. M., Rast, J. E., Anderson, K. A., & Shattuck, P. T. (2017). *National autism indicators report: Developmental disability services and outcomes in adulthood*. Life Course Outcomes Program, AJ Drexel Autism Institute, Drexel University.

Sarrett, J. C. (2017). Interviews, disclosures, and misperceptions: autistic adults’ perspectives on employment related challenges. *Disability Studies Quarterly*, *37*(2). https://doi.org/10.18061/DSQ.V37I2.5524

Sasson, N. J., Faso, D. J., Nugent, J., Lovell, S., Kennedy, D. P., & Grossman, R. B. (2017). Neurotypical peers are less willing to interact with those with autism based on thin slice judgments. *Scientific Reports*, *7*(1), 1–10. https://doi.org/10.1038/srep40700

Senju, A., & Johnson, M. H. (2009). Atypical eye contact in autism: Models, mechanisms and development. *Neuroscience & Biobehavioral Reviews*, *33*(8), 1204–1214. https://doi.org/10.1016/j.neubiorev.2009.06.001

Shellenbarger, S. (2013). *Just look me in the eye already*. The Wall Street Journal. Retrieved March 4, 2022, from https://www.wsj.com/articles/SB10001424127887324809804578511290822228174

Silani, G., Bird, G., Brindley, R., Singer, T., Frith, C., & Frith, U. (2008). Levels of emotional awareness and autism: An fMRI study. *Social Neuroscience*, *3*(2), 97–112. https://doi.org/10.1080/17470910701577020

Smith, M. J., Fleming, M. F., Wright, M. A., Losh, M., Humm, L. B., Olsen, D., & Bell, M. D. (2015). Brief report: Vocational outcomes for young adults with autism spectrum disorders at six months after virtual reality job interview training. *Journal of Autism and Developmental Disorders*, *45*(10), 3364–3369. https://doi.org/10.1007/s10803-015-2470-1

Smith, M. J., Ginger, E. J., Wright, K., Wright, M. A., Taylor, J. L., Humm, L. B., Olsen, D. E., Bell, M. D., & Fleming, M. F. (2014). Virtual reality job interview training in adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *44*(10), 2450–2463. https://doi.org/10.1007/s10803-014-2113-y

Smith, M. J., Pinto, R. M., Dawalt, L., Smith, J. D., Sherwood, K., Miles, R., Taylor, J., Hume, K., Dawkins, T., Baker-Ericzén, M., Frazier, T., Humm, L., & Steacy, C. (2020). Using community-engaged methods to adapt virtual reality job-interview training for transition-age youth on the autism spectrum. *Research in Autism Spectrum Disorders*, *71*, 101498. https://doi.org/10.1016/j.rasd.2019.101498

Smith, M. J., Sherwood, K., Ross, B., Smith, J. D., DaWalt, L., Bishop, L., Humm, L., Elkins, J., & Steacy, C. (2021). Virtual interview training for autistic transition age youth: A randomized controlled feasibility and effectiveness trial. *Autism*, *25*(6), 1536–1552. https://doi.org/10.1177/1362361321989928

Spain, D., Sin, J., Linder, K. B., McMahon, J., & Happé, F. (2018). Social anxiety in autism spectrum disorder: A systematic review. *Research in Autism Spectrum Disorders*, *52*, 51–68. https://doi.org/10.1016/j.rasd.2018.04.007

Stolcke, A., Coccaro, N., Bates, R., Taylor, P., Van Ess-Dykema, C., Ries, K., Shriberg, E., Jurafsky, D., Martin, R., & Meteer, M. (2000). Dialogue act modeling for automatic tagging and recognition of conversational speech. *Computational Linguistics*, *26*(3), 339–373. https://doi.org/10.1162/089120100561737

Strickland, D. C., Coles, C. D., & Southern, L. B. (2013). JobTIPS: A transition to employment program for individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *43*(10), 2472–2483. https://doi.org/10.1007/s10803-013-1800-4

Sunburst Chart [Computer Software]. (2022). https://www.npmjs.com/package/sunburst-chart

Taylor, J. L., DaWalt, L. S., Marvin, A. R., Law, J. K., & Lipkin, P. (2019). Sex differences in employment and supports for adults with autism spectrum disorder. *Autism*, *23*(7), 1711–1719. https://doi.org/10.1177/1362361319827417

Trevisan, D. A., Roberts, N., Lin, C.Y. -H., Birmingham, E., & Hadjikhani, N. (2017). How do adults and teens with self-declared autism spectrum disorder experience eye contact? A qualitative analysis of first-hand accounts. *Plos One*, *12*(11), 1–22. https://doi.org/10.1371/journal.pone.0188446

Unity 3D 2019.3 [Computer Software]. (2022). https://unity.com/releases/2019- 3

Wang, G., Zhang, J., & Fu, G. -S. (2018). Predicting student behaviors and performance in online learning using decision tree. *2018 Seventh*

International Conference of Educational Innovation through Technology (EITT) (pp. 214–219). https://doi.org/10.1109/EITT.2018.00050

Wehman, P., Taylor, J., Brooke, V., Avellone, L., Whittenburg, H., Ham, W., Brooke, A. M., & Carr, S. (2018). Toward competitive employment for persons with intellectual and developmental disabilities: What progress have we made and where do we need to go. *Research and Practice for Persons with Severe Disabilities*, *43*(3), 131–144. https://doi.org/10.1177/1540796918777730

Whelpley, C. E., & May, C. P. (2022). Seeing is disliking: Evidence of bias against individuals with autism spectrum disorder in traditional dfjob interviews. *Journal of Autism and Developmental Disorders*, 1–12. https://doi.org/10.1007/s10803-022-05432-2

Xu, Q., Cheung, S. S., & Soares, N. (2015). LittleHelper: An augmented reality glass application to assist individuals with autism in job interview. *2015 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)* (pp. 1276–1279). https://doi.org/10.1109/APSIPA.2015.7415480

Appendix A. Post-Session Questionnaire

Participant ID: Date:

Time:

Dashboard Post-Session Qualitative Feedback Questionnaire

Background:

1. What is your **professional title?**
2. In your current role, are you involved in **making hiring decisions** at your organization?
3. In your current role, do you **provide support** to autistic or neurodiverse individuals on the job or during the job search? For instance, either as a job coach or career counselor?
4. Does your organization have an **existing neurodiverse hiring initiative**?

Experience with Webpage

1. What did you **like** about the webpage's **visualizations** and **metrics**?
2. What did you **dislike** or what did you find **confusing** or **unclear**?
3. **FOR PROFESSIONALS INVOLVED IN HIRING:** Which (if any) of the dashboard metrics that were presented do you feel might influence your decision to **hire** a candidate?
4. **FOR PROFESSIONALS INVOLVED IN JOB COACHING and/or CAREER SUPPORT ONLY:** Which (if any) of the dashboard metrics that were presented do you feel might help you with **coaching/supporting** a candidate?
5. Are there metrics that are not included in the current dashboard that you would be interested in seeing?
6. If any, which metrics do you think might be informative **to autistic individuals and/or job coaches**, and why?
7. Talk about whether or not you agree with the following statements:
   a. The **layout** of the web page was well organized.
   b. The **charts and visualizations** presented information clearly.
   c. The **information contained in the dashboard** would help me **empathize** with job candidates.
   d. The **information contained in the dashboard** would help me **select** candidates.
   e. The **visual presentation of the web page** looked professional.
8. As development of the dashboard continues, what would you say is the **single most important** area in need of further development?
9. How might you use the dashboard to improve your interview performance?
10. Do you have any additional thoughts about the dashboard or its application?

Figure A1 . Screenshot of the dashboard post-session qualitative feedback questionnaire.

Appendix B. Stakeholders Interview Questions

Employee/Candidate Focus Group Questions

General Questions

1. To start, let's learn a little more about you. Can you please tell us if you are currently working? What type of job do you have? How many years have you been working?
2. Think about previous jobs you've had, what helped you get the job? Think about previous jobs you've applied to. What have been the biggest barriers to getting a job?
    a. Lacking the right skills for the job?
    b. Specific training?
    c. Certifications; licenses?
    d. Specific educational preparation/degrees?

Interview Process

3. We are interested in learning about your experience during an employment interview, can you describe the process of an interview that you have done? For example,
    a. What types of questions did they ask?
    b. What did they ask you to do during the interview (work in teams, do work tasks, etc.)?
    c. Did you receive the questions you were going to be asked in advance? Could that have helped?
4. Do you disclose your diagnosis when applying for a job?
    a. Can you explain what helps you to make the decision?
    b. Have your experiences differed when you have disclosed versus when you haven't?
5. What were some challenges that you have experienced during an employment interview?
    a. What kinds of questions make you feel the most nervous?
    b. How do you respond to those questions?
    c. What are some interviewer's traits/approach make you most nervous?
    d. What interview situation makes you feel most nervous?
    e. Did you have adequate time to form and deliver your answer?
6. Have you ever received specific feedback on why you didn't get a job?
    a. [If YES] What was that feedback?
7. What do you think could make the hiring process (and interviews) more effective for autistic people/people with autism?
    a. Is there a specific training that could help you? What kind of training?
    b. Is there specific training that could help employers? What else could help employers?
8. Have the interviews differed in any way from the time you get the job or not?
9. Do you have gaps in your employment history?
    a. Did the employer ask questions about that?
    b. How do you think that might have influenced your interview?

Figure B1. Screenshot of employee/candidate focus group questions.

Employer Focus Group and Interview Questions

General Questions

1. To start, let's learn a little more about you. In particular, please tell us what industry is your business in, your position within the organization, and how long have you been with your organization.
   a. What industry is your business in?
   b. What is your position title?
   c. How long have you been with your organization?
2. Have you previously hired people on the autism spectrum?
   a. If YES, how many? Did you know when you hired them that they were on the spectrum?
3. For what types of roles have you hired autistic people?
   a. Did they work by themselves?
   b. Did they work with other autistic people?
   c. Did they work with other non-autistic people?
   d. Did they work with clients/customers?

Interview Process

Thank you for letting us know about your experience in supporting people with ASD. For the rest of this interview, we want to learn more about your experience working with autistic people. In particular, we are interested in the interview process, and things that would support employees with ASD, once hired. We will begin with questions to learn more about interviewing people with ASD.

4. How do you go about identifying and hiring autistic talent?
5. What is your hiring process?
   a. Do you use interviews? If so, how?
      i. What types of questions do you ask?
         - Behavioral? Brainteasers? Open-ended? Cases?
      ii. Do individuals who interview autistic candidates get any specific guidance or training on how to do so?
         - If not, how do you select interviewers?
   b. Do you use task-oriented assessments (e.g., assessment centers, inbox exercises, team-based work)?
   c. Who conducts these exercises? What is their background? Do they have experience/knowledge/training with disability/autism?
   d. How are the results interpreted? Any other measures were taken into consideration when reviewing the assessment results?

Figure B2. Screenshot of employer focus group and interview questions.

Service Professional/Job Coach Focus Group and Interview Questions

General Questions

1. To start, let's learn a little more about you. Please tell us a little bit about your work role and organization. Have you worked with people on the autism spectrum? In what capacity?
2. What are the work roles, educational backgrounds for autistic people that you support?
   a. Did they work by themselves?
   b. Did they work with other autistic people?
   c. Did they work with other non-autistic people?
   d. Did they work with clients/customers?

Interview Process

Thank you for letting us know your about experience in supporting people with ASD. For the rest of this interview, we want to learn more about your experience supporting autistic people as they prepare for interviews, and strategies you have used to help them to succeed at work/school.

3. What have you seen as effective processes for educating, identifying, and hiring autistic talent?
4. What types of interview process, or interview questions, have you seen or used that you have found effective?
   a. Behavioral (e.g., tell me about a time …)?
   b. Brain teasers?
   c. Cases studies and associated questions?
   d. Open-ended?
5. Have you seen organizations use task-oriented assessments (e.g., assessment centers, inbox exercises, team-based work)?
   a. How effective do you think these assessments are? What makes the assessment effective or not effective?
   b. In what ways were they tailored to autism? In what ways could they be more tailored?
   c. Who conducted these assessments? What is their background? Do they have experience/knowledge/training with disability/autism?
6. Who *should* conduct these assessments?
7. Have you provided employers who interview autistic candidates any specific guidance or training on how to interview effectively?
   a. What do you think employers and other organizations could do better?
8. Have you provided autistic people with any specific guidance or training on how to interview effectively?
   a. Where do you think autistic people need the most help?

Figure B3. Screenshot of service professional/job coach focus group and interview questions.