Design and Validation of a Stress Detection Model for Use with a VR Based Interview Simulator for Autistic Young Adults

Miroslava Migovich(✉), Alex Korman, Joshua Wade, and Nilanjan Sarkar

Robotics and Autonomous Systems Lab, Vanderbilt University, Nashville, TN 37212, USA
miroslava.migovich@vanderbilt.edu

Abstract. Studies show that young autistic adults are under- or unemployed, with almost half never holding a paying job in their 20's. Unemployment within this population leads to decreased personal growth and increased dependence on caregivers. Research suggests that the interview process is one of the largest barriers to employment for this population. Autistic individuals often struggle with emotion regulation, which can be exacerbated by the interview process. To address this, we propose the use of a stress detection model in conjunction with a virtual reality interview simulator. This combination will allow for the interview to adapt to the state of the participant to improve the skills and engagement of the user and positively influence their comfort level. Data regarding negative affective responses to categories of questions can also be used to inform employers on better interviewing techniques. A model was designed using data obtained from neurotypical participants completing a modified Computerized Paced Serial Addition Task (PASAT-C) and evaluated on a dataset obtained from Autistic participants who took part in a simulated interview. Agreement between the model and ground truth was compared based on Pearson correlation coefficients. It was found that was $r(289) = 0.28$, which was statistically significant ($p < .001$; CI: 0.17 to 0.38). Our preliminary results provide evidence for the validity of observer-based labeling of data captured using a wrist-worn physiological sensor.

Keywords: Emotions in HCI and design · Stress-detection · Physiological response · Machine learning

1 Introduction

It is estimated that over half of autistic[1] adults are unemployed or underemployed. This percentage is higher in comparison to adults with other developmental disabilities [1, 2]. It is also documented that one of the largest impediments to the employment of autistic individuals is the interview process and the associated stress of this process [3,

[1] In accordance with recent surveys that suggest a preference for identity-first language by autistic self-advocates, we have chosen to adopted identity-first language (*autistic person*) instead of person-first (*person with autism*) language [23].

M. Antona and C. Stephanidis (Eds.): HCII 2021, LNCS 12768, pp. 580–588, 2021.
https://doi.org/10.1007/978-3-030-78092-0_40

4] due the differences in social communication and behavior that is common in autistic individuals. Autism Spectrum Disorder is a neurodevelopmental condition that is most often characterized by difference in social communication and social behavior, and repetitive/restrictive patterns of behaviors/interests [5]. Often, the individual possesses the required experience and job skills, however they are unable to interview well by "neurotypical" standards. There is limited information regarding emotional regulation of autistic individuals, however it is known that ASD is associated with heightened emotional responses and poor emotion control, which leads to difficulties with the interview process [6]. The inability to regulate emotions effectively has been linked to all core features of ASD, including social and communication functioning [7].

The unemployment of this population adds to the growing cost of caring for autistic Americans, which was found to be $268 billion annually in 2015 [8]. Autistic self-advocates have expressed that accommodations in job interviews can reduce barriers to employment [9]. These accommodations can be integrated, in part, by the use of a virtual reality interview simulator that allows autistic adults the opportunity to practice interview skills and receive feedback in a virtual environment. The information from the virtual interview can be used to inform potential employers on how to better interview the population and provide the user with information regarding their interview performance. With the virtual interview process, it is possible to mitigate the level of stress and allow the participant to practice interview skills in a lower-stress environment by developing a real-time stress recognition module based on physiological responses and machine learning. Such a module would allow the system to dynamically recognize stress and adapt to help individuals manage their stress response and improve their performance.

Previous work has shown that physiological features are significantly correlated with one's affective state, and therefore can be used to approximate stress levels using machine learning. Autistic individuals often experience emotional stress without external expression and therefore implicit measures such as physiological signals are used to determine stress levels without external expression [10]. Physiological signals are the preferred method for emotion recognition because physical signals such as speech, facial expressions, and posture can also be manipulated by the individual and reliability cannot be guaranteed [11]. The autonomic nervous system uses sensory and motor neurons to operate between various organs and the central nervous system. Physiological signals are the body's method of response to these systems. The Cannon-Bard theory states that emotions and physiological responses are simultaneous [12]. Because the central nervous system and the autonomic nervous systems are involuntarily stimulated, physiological responses to stress cannot be as easily manipulated by the individual [11, 13]. While the relationship between individual physiological signals and emotions have been explored [14, 15], it is preferable to use multiple signals in conjunction for more in-depth correlations between stress and physiological signals [11].

There is a paucity of literature concerning the performance of stress models in autistic individuals and, to our knowledge, this is the first exploration of the use of stress detection in an interview setting for autistic individuals [16].

2 System Design

2.1 System Architecture

The proposed stress model was used in conjunction with a Virtual Reality-based interview simulator called Career Interview Readiness in Virtual Reality (CIRVR) [17] in order to inform and adjust the interview based on the individuals' perceived stress levels. CIRVR uses natural language-based interactions coupled with real-time stress measurements to create a semi-naturalistic interview experience. The CIRVR system features five key components: (1) a natural language-based communication module largely enabled by Microsoft Azure cloud services such as LUIS, Speech-to-text, and Text-to-Speech[2] (2) a real-time stress-detection module described in detail in this paper; (3) an eye-tracking module used to track visual attention patterns (e.g., using Tobii eye trackers and the Fove head-mounted display); (4) a facial expression detection module used to measure facial expressions based on universally-recognized emotions which are detailed in [18]; and (5) support for both immersive—e.g., Fove or VIVE head-mounted displays—and non-immersive interaction modalities (Fig. 1).

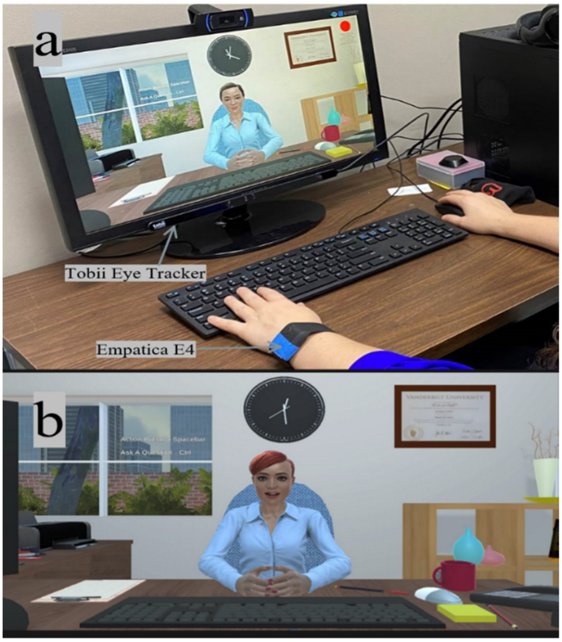

Fig. 1. CIRVR job interview simulator: (a) desktop-based interaction setup with Tobii eye tracker and the Empatica E4 physiological-sensing wristband; (b) desktop-based interaction view of the interviewer

During the interview, stress data will be collected and processed before being sent to the model to be classified. The stress values generated by the model will then be used

[2] https://azure.microsoft.com/en-us/services/cognitive-services/.

by to inform the interviewer avatar of the interviewee's stress level and allows them to adjust question types in order to reduce discomfort in the user. This closed-loop system allows for real-time adjustments to take place within the interview scenario. Adjustments include rephrasing of the question, moving onto a new question, or introduction of multiple-choice questions. The adjustments deescalate the stress level of that user so that they are able to gain valuable interview practice in a low-risk environment. The data will also be used to inform future employers on interview techniques for neurodiverse individuals (Fig. 2).

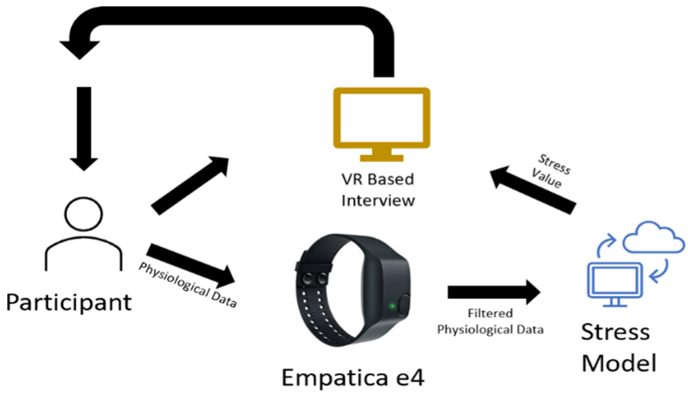

Fig. 2. Interaction of system components

2.2 Training Data Procedure

Training data were obtained from N = 20 neurotypical participants by trained research assistants. The research was approved by the Vanderbilt Institutional Review Board (IRB) and all participants provided informed written consent or assent as required by the IRB. The neurotypical participants were not compensated and training data collection took place before the COVID-19 pandemic.

In order to collect training data, a modified version of the Computerized Paced Serial Addition Task (PASAT-C) [19] was used along with an Empatica E4 sensor (www.empatica.com) to gather physiological signals. The code for the PASAT-C task was originally developed by Millisecond Software and was modified to include a baseline timer, clear distinction between the three levels, and a Likert scale self-report of perceived stress by the participant. To begin, the participant was asked demographic questions, such as age, weight, if they have exercised or drank caffeine, if they feel sick, and their dominant hand. These are characteristics that affect how the body expresses stress physiologically and should be taken into consideration. The E4 device was then placed on the non-dominant wrist. The computerized task begins with a 3-min baseline during which the participant is asked to sit still and keep their non-dominate hand steady. Following the baseline, a set of instructions is presented. The task is designed to show a series of single digit numbers that the participant is expected to sum and use the mouse to press the box which correlates with the sum. The possible answers are 1 through 18 and are shown in a clockwise manner. Once the next number is shown, the participant continues to compute the sum of only the last two digits shown, not the running total. If the participant chooses the wrong sum, an error noise is played. The task begins with 11 practice trials. This is followed by level one. Level on is three minutes long with three seconds between each digit presentation. Level two is five minutes with two seconds between digits and level three is ten minutes with 1.5 s between the digits. After each level, the participant was asked to rate their stress level in context of the game on a 1–10 Likert scale. Before level three, there is a 45 s pause in which the participant is informed that the digits will be presented faster than before, and they are given the option of a "QUIT" button during the third level to terminate the task prematurely.

After completing the PASAT-C (Fig. 3), open ended questions were asked to gather qualitative data about the participants experience and perceived stress levels.

From the E4, the heart rate, inter-beat interval, blood volume pressure, and acceleration is extracted as individual excel files. These files were then pre-processed before being used in the model to synchronize time stamps across the metrics.

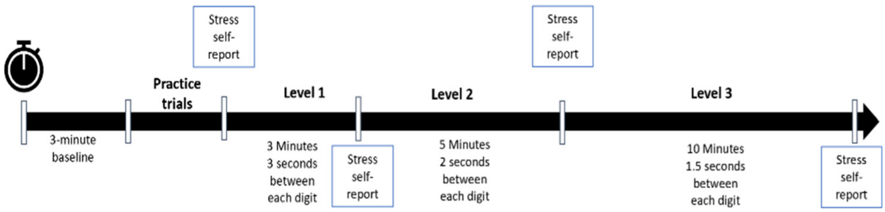

Fig. 3. Timeline of PASAT-C

2.3 Design of Models

The physiological data were processed into separate readings: heartrate, inter-beat interval, blood volume pulse, skin conductance level (SCL), skin conductance response (SCR), and acceleration in the x, y, and z direction.

The dataset was divided into training, testing, and validation sets. Each set was composed of the participant's entire dataset, as opposed to choosing random instances. Four models were initially compared, however the K-nearest neighbor classifier performed better than the support vector machine, neural network, and the random forest models. Pior research has shown this model has high accuracy classifying emotional states [20] K-nearest neighbors classification is based on classifying an unknown sample based on the known classification of neighbors [21]. The training data were segmented into five second intervals and trained according to perceived stress values, with the practice level and level one of the PASAT-C corresponding to a low stress level (labeled 0) and levels two and three corresponding to higher stress (labeled 1). The self-reported stress levels during the PASAT-C were not used due to lack of consistent grading across participants. Task data in the test set were averaged across 5 s increments with a stress value output for each segment. Hyper-parameters were finetuned on the validation set. The KNN had a Mean accuracy of 73.06 and a F1 Score of 77.26.

2.4 Testing Procedure

Five autistic participants took part in the study, leading to a sample size of n = 289 observations. Due to COVID-19 regulations and restrictions, the sample size was limited. Autistic participants were recruited to take part in one-hour sessions at the university lab. Appropriate precautions, such as symptom pre-screening 24 h before the session, follow-up prescreening day of, mandatory face covering, social distancing and disinfection of all materials, were taken to ensure safety of all involved. Written consent/assent was obtained in accordance with Vanderbilt Institutional Review Board (IRB) and then the participant completed a 20-min virtual interview using the desktop version of the CIRVR system. The interview is split into multiple sections, such as introductions, past work experience, technical experience, technical questions on a whiteboard, education, education questions on the whiteboard, and open ended/personal questions. This allows for exploration of patterns of stress during different portions of a typical interview. After the interview, the researcher led the participant through a brief qualitative interview to gain insight on the user's experience and what type of information the user would like to receive from the system. One participant was unable to complete the full interview due to system failure, however no observational data was lost. The autistic participants were compensated for their time and travel.

Video recording of the participant from the waist up were used to determine ground truth stress ratings by a researcher trained in behavioral assessment. Stress was graded on a scale of 1–10 with the participant's demeanor upon arrival set at baseline. The ground truth ratings were then paired with the outputs of the model using timestamps. As the model output a stress rating every 5 s, three model ratings were averaged and compared to the ground truth.

3 Results

Because only one rater was involved in data labeling, we decided to estimate agreement between rater observations and the E4 output using the Pearson correlation coefficient. Statistical analyses were conducted using MedCalc statistical analysis software (version 19.5.3) with default parameters [22]. In the future, other indices of agreement (e.g., kappa and concordance correlations) will be applied. Based on a sample size of n = 289, the Pearson correlation coefficient was $r(289) = 0.28$, which was statistically significant ($p < .001$; CI: 0.17 to 0.38). This result provides preliminary support for the validity of the observer-based rating approach. Next, in order to identify the optimal cutoff score for prediction of the E4 output based on rater observations, we conducted a Receiver Operating Characteristic (ROC) Area Under the Curve (AUC) analysis. As shown in Fig. 4, a statistically significant AUC of 0.604 was found ($p < .01$; CI: 0.55 to 0.66). The Youden Index was used to determine the optimal cutoff score of >5 (i.e., rater observations of 6 correspond to high stress according to the E4 output).

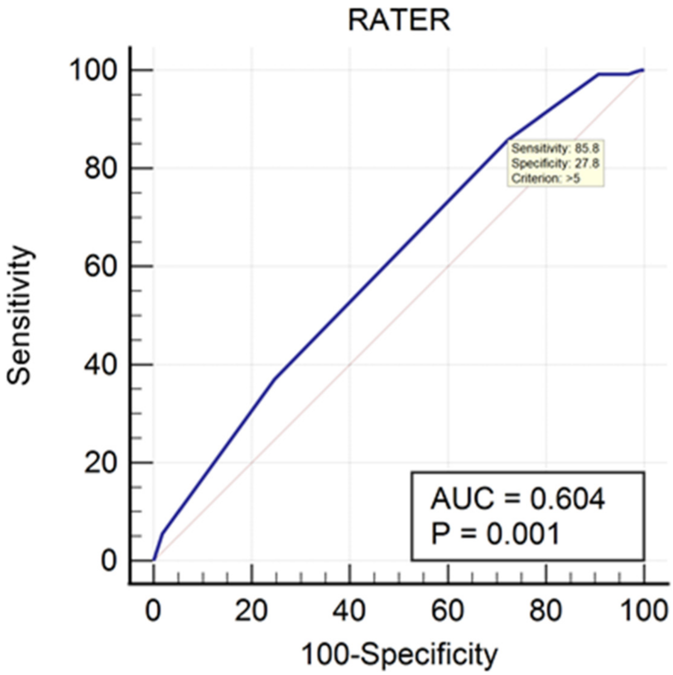

Fig. 4. Receiver operating characteristic curve analysis

4 Discussion and Conclusion

Our preliminary results provide evidence for the validity of observer-based labeling of data captured using a wrist-worn physiological sensor. While encouraging, a number of limitations must be addressed in future work. The performance of the model could be improved by a change in the labeling structure. The training data were labeled as either "stressed" or "not stressed" depending on the level of the PASAT-C. However, this assumes that the participants are less stressed during the first two levels and that their stress increases during the task. More rigorous data collection with additional tasks would help to improve the labels. While the general physiological response is predicted to be similar between the two groups, it is possible that certain physiological patterns exist within autistic physiological data that are not concurrent with neurotypical data. It is necessary to train models using autistic datasets and the observer ratings as labels for training data in order to further investigate the possible differences in physiological responses to stress between the groups.

Future work includes to integration of the model into the real-time system using Microsoft Azure's Machine Learning Studio and Azure Web Services as well as an in-depth exploration of trends of stress during the interview process with autistic participants. This information will then be used to inform both interviewees and possible employers on accommodations that can made to improve the interview experience.

Acknowledgements. This project was funded by a Microsoft AI for Accessibility grant, by the National Science Foundation under awards 1936970 and 2033413 and by the National Science Foundation Research Traineeship DGE 19-22697. The authors would like to thank the participants for their time.

References

1. Autism U.S. Department of Labor. https://www.dol.gov/agencies/odep/topics/autism. Accessed 28 Oct 2020
2. Chen, J.L., Leader, G., Sung, C., Leahy, M.: Trends in employment for individuals with autism spectrum disorder: a review of the research literature. Rev. J. Autism Dev. Disord. **2**(2), 115–127 (2014). https://doi.org/10.1007/s40489-014-0041-6
3. Burke, S.L., Li, T., Grudzien, A., Garcia, S.: Brief report: improving employment interview self-efficacy among adults with autism and other developmental disabilities using virtual interactive training agents (ViTA). J. Autism Dev. Disord. **51**(2), 741–748 (2020). https://doi.org/10.1007/s10803-020-04571-8
4. Mazefsky, C.A.: Emotion regulation and emotional distress in autism spectrum disorder: foundations and considerations for future research. J. Autism Dev. Disord. **45**(11), 3405–3408 (2015). https://doi.org/10.1007/s10803-015-2602-7
5. Diagnostic and statistical manual of mental disorders: DSM-5 (2013)
6. Mazefsky, C.A., et al.: The role of emotion regulation in autism spectrum disorder. J. Am. Acad. Child. Adolesc. Psychiatry **52**, 679–688 (2013). https://doi.org/10.1016/j.jaac.2013.05.006
7. Samson, A.C., et al.: Emotion dysregulation and the core features of autism spectrum disorder. J. Autism Dev. Disord. **44**(7), 1766–1772 (2013). https://doi.org/10.1007/s10803-013-2022-5
8. Leigh, J.P., Du, J.: Brief report: forecasting the economic burden of autism in 2015 and 2025 in the United States. J. Autism Dev. Disord. **45**(12), 4135–4139 (2015). https://doi.org/10.1007/s10803-015-2521-7
9. Booth, J., McDonnell, J.: Autism Equality in the Workplace (2016)
10. Lahiri, U., Welch, K.C., Warren, Z., Sarkar, N.: Understanding psychophysiological response to a virtual reality-based social communication system for children with ASD (2011). https://doi.org/10.1109/ICVR.2011.5971841
11. Shu, L., et al.: A review of emotion recognition using physiological signals. Sensors (Switz.) **18**(7), MDPI AG, 1 July 2018. https://doi.org/10.3390/s18072074
12. Friedman, B.H.: Feelings and the body: the Jamesian perspective on autonomic specificity of emotion. Biol. Psychol. **84**, 383–393 (2010). https://doi.org/10.1016/j.biopsycho.2009.10.006
13. Hagemann, D., Waldstein, S.R., Thayer, J.F.: Central and autonomic nervous system integration in emotion. Brain Cogn. **52**(1), 79–87 (2003). https://doi.org/10.1016/S0278-2626(03)00011-3
14. Kim, H.G., Cheon, E.J., Bai, D.S., Lee, Y.H., Koo, B.H.: Stress and heart rate variability: a meta-analysis and review of the literature. Psychiatry Invest. **15**, 235–245 (2018). https://doi.org/10.30773/pi.2017.08.17
15. Xie, J., Wen, W., Liu, G., Chen, C., Zhang, J., Liu, H.: Identifying strong stress and weak stress through blood volume pulse. In: Proceedings of the 2016 IEEE International Conference on Progress in Informatics and Computing, PIC 2016, June 2017, pp. 179–182 (2017). https://doi.org/10.1109/PIC.2016.7949490
16. Dijkhuis, R.R., Ziermans, T., van Rijn, S., Staal, W., Swaab, H.: Emotional arousal during social stress in young adults with autism: insights from heart rate, heart rate variability and self-report. J. Autism Dev. Disord. **49**(6), 2524–2535 (2019). https://doi.org/10.1007/s10803-019-04000-5

17. Wade, J.W., et al.: Career Interview Readiness in VR (CIRVR): Feasibility of an AI-Driven Platform for Employers and Neurodiverse Talent. INSAR. https://insar.confex.com/insar/2020/meetingapp.cgi/Paper/35598
18. Ekman, P.: Are there basic emotions? Psychol. Rev. **99**(3), 550–553 (1992). https://doi.org/10.1037/0033-295X.99.3.550
19. Lejuez, C., Kahler, C., Brown, R.: A modified computer version of the Paced Auditory Serial Addition Task (PASAT) as a laboratory-based stressor. Undefined (2003). Accessed 25 Oct 2020
20. Fan, J., et al.: A step towards EEG-based brain computer interface for autism intervention. In: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, November 2015, pp. 3767–3770 (2015). https://doi.org/10.1109/EMBC.2015.7319213
21. Cover, T., Hart, P.: Nearest neighbor pattern classification, pp. 21–27, January 1967. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1053964. Accessed 27 Jan 2021
22. Schoonjans, F., Zalata, A., Depuydt, C.E., Comhaire, F.H.: MedCalc: a new computer program for medical statistics. Comput. Methods Programs Biomed. **48**(3), 257–262 (1995). https://doi.org/10.1016/0169-2607(95)01703-8
23. Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., Pellicano, E.: Which terms should be used to describe autism? Perspectives from the UK autism community. Autism **20**(4), 442–462 (2016). https://doi.org/10.1177/1362361315588200