Online Engagement Detection and Task Adaptation in a Virtual Reality Based Driving Simulator for Autism Intervention

Dayi Bian[1(✉)], Joshua Wade[3], Zachary Warren[4,5,6], and Nilanjan Sarkar[1,2]

[1] Department of Electrical Engineering, Vanderbilt University, Nashville, TN, USA
{dayi.bian, nilanjan.sarkar}@vanderbilt.edu
[2] Department of Mechanical Engineering, Vanderbilt University, Nashville, TN, USA
[3] Department of Computer Science, Vanderbilt University, Nashville, TN, USA
[4] Department of Pediatrics and Psychiatry, Vanderbilt University, Nashville, TN, USA
[5] Department of Special Education, Vanderbilt University, Nashville, TN, USA
[6] Vanderbilt Kennedy Center, Treatment and Research Institute of Autism Spectrum Disorders, Vanderbilt University, Nashville, TN, USA

Abstract. Individuals with Autism spectrum disorder (ASD) have difficulty functioning independently on essential tasks that require adaptive skills such as driving. Recently, computer-aided technology, such as Virtual Reality (VR), is being widely used in ASD intervention to teach basic skills to children with autism. However, most of these works either do not use feedback or solely use performance feedback from the participant for system adaptation. This paper introduces a physiology-based task adaptation mechanism in a virtual environment for driving skill training. The difficulty of the driving task was autonomously adjusted based on the participant's performance and engagement level to provide the participant with an optimal level of challenge. The engagement level was detected using an affective model which was developed based on our previous experimental data and a therapist's ratings. We believe that this physiology-based adaptive mechanism can be useful in teaching driving skills to adolescents with ASD.

Keywords: Virtual Reality (VR) · Driving simulator · ASD intervention · Dynamic difficulty adjustment (DDA) · Physiological signals · Machine learning · Affective computing · Human computer interaction

1 Introduction

Autism spectrum disorder (ASD) has a high prevalence rate of 1 in 68 children in the US [1] and is associated with high familial and societal cost [2, 3]. Individuals with ASD have difficulty functioning independently on essential tasks that require adaptive skills such as driving [3], which is crucial for independent living in developed countries.

Although there is no single accepted treatment or known cure for ASD, there is a growing consensus that skill training and educational intervention programs can significantly improve long-term outcomes for individuals with ASD and their families [4, 5]. A growing number of studies have been investigating applications of computer-aided technology [6–11], including Virtual Reality (VR) platforms, for ASD intervention. However, most of these studies do not provide feedback or solely use participants' performance or explicit user feedback as primary means of evaluation.

Task difficulty can induce a variety of cognitive workload and affective states [12]. A task that is beyond an individual's capability can be overwhelming and cause anxiety while a task that does not utilize the skill of a person might result in boredom. An individual under the states of anxiety or boredom will focus less on their tasks, learn less, be less productive, and be more prone to errors [13]. Minimizing anxiety and boredom during a task can help people maintain a high level of engagement, perform well, and learn efficiently. To provide users with an optimal level of challenge, dynamic difficulty adjustment (DDA) is used to automatically alter the task difficulty in real-time by monitoring user state. Compared to strictly performance-based feedback, physiology feedback can be more efficient in providing optimal challenge to users, increasing their engagement and improving their performance [14].

The present study introduced a VR-based adaptive driving skill training system which automatically adjusted its difficultly levels based on participants' engagement levels and performance metrics. The engagement detection model was developed by training on data that was collected from our previous study where we recorded physiological signals and engagement labels [15]. Off-the-shelf driving simulators were not suitable for this study because we needed to design new driving tasks with embedded intervention rules as well as to integrate the simulator with real-time physiological recording and engagement detection modules. To the best of our knowledge, none of the commercial driving simulators offer these flexibilities.

We believe that by utilizing physiology-based engagement detection to adjust difficulty levels in a VR-based driving system, our proposed system can be used more effectively for driving skill training of individuals with ASD.

2 System Overview

The proposed VR-based driving environment with adaptive response technology comprised of (1) a VR-based driving task module; (2) a real-time physiological data acquisition module; and (3) an individualized DDA module utilizing a physiology-based engagement prediction mechanism (Fig. 1).

2.1 VR-Based Driving Task Module

In this study, we used desktop VR applications because they are accessible, affordable and potentially minimize cyber sickness for this especially-sensitive ASD population [16]. Unity (www.unity3d.com), a commercial VR design platform, was used to design the VR environment. Within Unity, we developed a graphical user interface, created
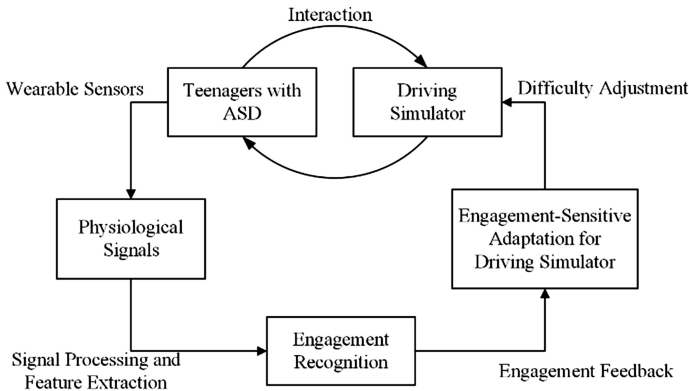
Fig. 1. System framework overview

behaviors for vehicles, pedestrians and traffic lights, designed the driving scenario and embedded customized driving rules.

The task difficulty can be adjusted by modifying several game parameters, such as responsiveness of the brake pedal, accelerator pedal, steering wheel, intensity of light in the environment etc. Detailed information of these parameters are presented in Table 1. The values of these parameters for different difficulties were chosen based on the average performance of the participants in our previous study. The parameters related to difficulty levels were grouped into two categories. One category is car controllability, which contains the responsiveness of the brake pedal, accelerator pedal and steering wheel. These are also the main factors that affect the difficulty of the game. The other category is environment, which contains speed for agent vehicles and the intensity of light in the environment. Each category has three difficulty levels from easy to hard. By combining these two categories, we have nine different difficulty states, among which the switching mechanism adjusts the difficulty.

2.2   Physiological Data Processing Module

Physiological response for an individual can be utilized to assess one's affective states [17]. The physiological data were collected using the Biopac MP150 physiological data acquisition system (www.biopac.com) with a sampling rate of 1000 Hz. Using hardware API provided by Biopac, we developed a customized physiological data acquisition program, in which we integrated socket-based communication with the Unity program to automatically record event information and time stamps. In this study, three physiological signals were investigated, which were photoplethysmogram (PPG), galvanic skin response (GSR) and respiration (RSP). These signals were measured by using light-weight, non-invasive wireless sensors. PPG and GSR were measured from toes instead of fingers in order to reduce the motion artifact from driving. Respiration was measured by using a respiration belt tied around the participant's abdomen.

The engagement model was developed offline using physiological data and therapist's labels from our previous study [15]. Therefore, it is necessary that the patterns

Table 1. Difficulty parameters

Group	Parameters	Domain	Easy difficulty	Medium difficulty	Hard difficulty
Car Controllability	Responsiveness of brake pedal	[0.35, 1]	1	0.675	0.35
	Responsiveness of the accelerator pedal	[1, 1.5]	1	1.25	1.5
	Responsiveness of the steering wheel	[1, 3.75]	1	1	3.75
Environment	Intensity of light in the environment	[0.01, 0.5]	0.5	0.226	0.01
	Speed of agent vehicles	[0.85, 1.75]	0.85	1.35	1.75

of emotional responses remain stable across participants. Although different patterns of emotional responses have been found in psychophysiological studies, Stemmler [18] argues that they are due to context deviation specificity. Since, in the current study, the engagement is elicited in the same context (the driving task), inter-participant variability should be low. Nevertheless, to further reduce this variability, three minutes' physiological data during a period of rest were acquired as a baseline. The raw physiological signals from our previous study were preprocessed to remove motion artifacts and noise. After preprocessing, the data were subsampled 10-fold. Subsampling can significantly reduce computation time for feature extraction, which is very crucial in real-time closed-loop systems. A set of features, which were highly correlated with engagement, were selected for training an engagement detection model [15]. After feature extraction, several machine learning algorithms from the Waikato Environment for Knowledge Analysis (WEKA) [19] toolkit were used to build the engagement detection model. Ten-fold cross-validation was used to validate the model. In the end, the Random Forest algorithm, which had the highest accuracy (84.72 %), was chosen to develop the engagement model.

In our proposed closed-loop driving scenario, the driving task program and engagement detection module communicated via sockets over a local area network (LAN). When defined events, like start of trial, end of trial, failure, success etc., occurred, a JSON (http://www.json.org) string containing the time stamp and event message was sent to the physiological data acquisition module. Event makers were recorded with the physiological data for future offline analysis. In the beginning of the experiment, three minutes of baseline data were recorded. After the baseline recording, baseline data were processed to get the baseline features. Every three minutes, the driving task sent an event marker to the engagement detection module to trigger detection. The engagement detection module acquire three minutes' data before this trigger. These data were preprocessed and ten selected features were extracted. The

baseline features were subtracted from the features of the three minutes' data to offset environment and subject difference. Then these features were fed into the Random Forest model to predict the engagement level. A binary label, "HIGH" or "LOW", was sent to the driving task program via the socket. Then the difficulty adjustment module took over and made the decision to switch the task difficulty (Fig. 2).
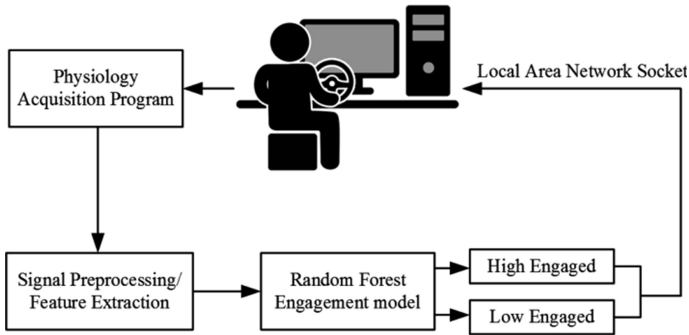

Fig. 2. Physiological data processing module

2.3   Difficulty Adjustment Module

Performance-Sensitive System (PS).   For the PS, a task-switching mechanism adjusts the difficulty states based on participants' performance metric alone. When a participant's performance is "Good" (Case 1), the task progression continues step-wise while the task difficulty level increases based on the state flow representation shown in Fig. 3. On the other hand, if a participant's performance is "Poor" (Case 2), the task progression continues step-wise while decreasing the task difficulty level. The switch mechanism continues until Case 1 reaches the most difficult level or Case 2 reaches the easiest level (Fig. 3).

Engagement-Sensitive System (ES). For the ES, however, the task switching mechanism is not only based on participants' performance but also his/her engagement level. In other words, in order to move up to higher difficulty levels, the participant must not only perform "Good" in the task, he/she must also be "High Engaged" in the task. In the ES, we fuse performance metrics and engagement levels to make the decision regarding the next step (Table 2). In some cases, the switching strategy is intuitive: If a participant is "High Engaged" and performs "Good", the system increases the difficulty level based on the finite state machine representation. On the other hand, if a participant is "Low Engaged" and performs "Poor", the system decrease the difficulty level. However, in the other two cases, where the engagement and performance metrics does not agree with each other, the switching strategy is harder to define. We give performance more priority for these cases. For case 2 in Table 2, in which engagement is "High" but performance is "Poor", the system recommends decreasing the difficulty level. For case 3, when a participant is "Low Engaged" but performs "Good", the difficulty level remains the same at least until next adjustment. At the next
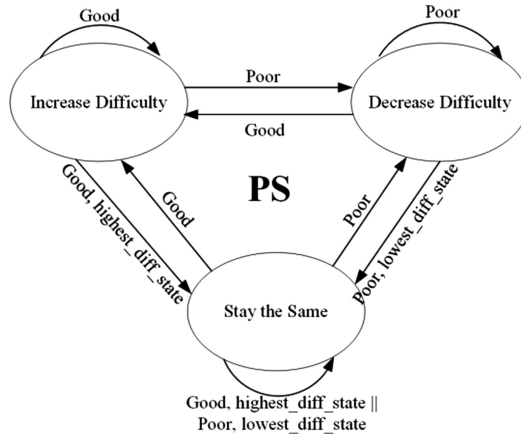
Fig. 3. Performance-Sensitive System difficulty switching state flow. (highest_diff_state: system at highest difficulty state, lowest_diff_state: system at lowest difficulty state).

Table 2. Engagement-Sensitive System cases

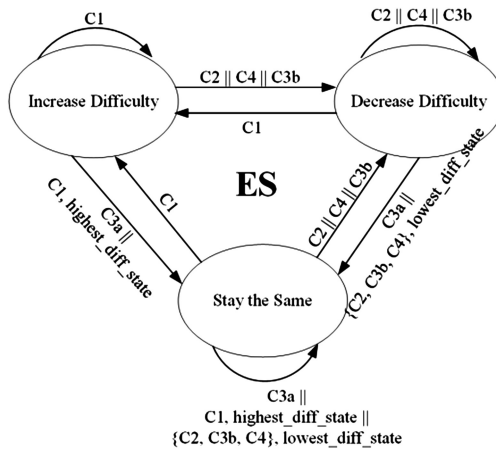|	Engagement	Performance	Action
Case 1	High	Good	Increase difficulty
Case 2	High	Poor	Decrease difficulty
Case 3a/3b	Low	Good	Same/Decrease difficulty
Case 4	Low	Poor	Decrease difficulty

Fig. 4. Engagement-Sensitive System difficulty switching state flow. (highest_diff_state: system at highest difficulty state, lowest_diff_state: system at lowest difficulty state, C1, C2, C3a, C3b, C4 are explained in Table below).

adjustment point, if the participant is still "Low Engaged" and performs "Good", the system decreases the difficulty level (Fig. 4).

3   Method and Procedure

3.1   Experimental Setup

The VR driving environment was run on a server-grade machine that could provide high quality rendering for graphics, while the peripheral physiological signals were acquired and processed in parallel on a separate. Both machines communicated over the LAN using TCP/IP. The VR driving task was presented on a 24-inch flat LCD panel (at resolution 1980 $\times$ 1080). Participants interacted with the driving environment using a Logitech G27 driving controller that was mounted on a specially-designed play-seat [20, 21]. The experiment was conducted in a laboratory with two rooms separated by one-way glass for observation. The researcher sat in the outer room.

3.2   Participants

We have recruited 4 teenagers with ASD for this phase of study. One of them is female. All participants had a clinical diagnosis of ASD from a licensed clinical psychologist as well as scores at or above clinical cutoff on the Autism Diagnostic Observation Schedule [22]. The Institutional Review Board (IRB) approval for conducting the experiment was sought and obtained.

3.3   Procedure

The participants were randomly assigned to either PS or ES. Each participant completed a 90-min session in one visit. At the start of each session, physiological sensors were placed on the participant's body by a trained experimenter. Participants watched a short video which explained basic instructions and game controls. After the tutorial, the participant was asked to remain calm and relaxed for three minutes during which physiological baseline data were collected. Participants also receive three minutes of practice driving in which there were no pedestrians and no other vehicles in the VR environment. This practice period allowed participants to familiarize themselves with the game controls and virtual environment.

After the three-minute practice, participants began the driving assignment. Through the assignment, participants followed the navigation system and tried to obey traffic rules. Disobeying any traffic rules (i.e., running a red light) caused a performance failure. Time duration for the assignment varied from 30 to 40 min depending on the participant's performance. A post-task survey was completed once the participant finished the assignment.

4   Results and Discussion

Four participants were recruited for a pilot study evaluation of the system and all of the participants completed the experiment. Two participants were assigned to the PS group and the other two were assigned to the ES group. The driving task ran smoothly according to the participants' responses; all of the participants reported that they "enjoyed the game" or "enjoyed the game a lot". Importantly, all of the participants reported noticing that the task difficulty changed during the driving task, demonstrating that the driving task difficulty can be successfully manipulated by tuning the difficulty parameters we have chosen. The performance data, such as steering wheel angle, gas and brake pedal usage, driving speed, and failure times, were recorded to hard disk in various formats (e.g., time series and event log). These data could be analyzed to evaluate participants' driving behavior.

Physiological measures including PPG, GSR and RSP were recorded and assessed during the driving task. These data were divided into three minute's windows. After preprocessing, a set of features were extracted from these windows and corrected through baseline features subtraction. These sets of features were then fed into the Random Forest engagement model to measure the average engagement level for each participant during the task. Although we cannot draw conclusion regarding differences between the PS and ES groups due to the small sample size, it is encouraging that all subjects, regardless of group affiliation, demonstrated higher engagement characteristics for a majority of the task duration ($M = 77.79$ %, $SD = 4.63$). In addition, the average engagement levels reported from the post-task survey support the acceptability of the system (Table 3).

Table 3. Engagement statistics from engagement model and post-task survey

Participant ID	Percentage of "High Engaged" output from engagement model	Engagement level from post-task survey*
PS01	77.78	5
PS02	73.84	4
ES01	84.30	5
ES02	75.25	5
Mean	77.79	4.75
Std	4.63	0.5

*The post-task survey used a 5-Likert Scale, where 1 means not engaged and 5 means highly engaged.

One advantage of our system is that it does not rely on direct input from the user; this is advantageous because the user may not be aware of his/her level of engagement during the task and may not have the time or the extra cognitive capability to provide such information. Another advantage is that the engagement detection module and DDA module are independent of the simulation software package and thus can be applied to any learning scenario—not merely driving training. The engagement level predictions are sent as JSON strings over TCP/IP, which can be read and applied

generically. Once the difficulty levels are defined for a given task, the difficulty switch logic can be applied based on the performance and engagement level of the user.

DDA is a valuable mechanism for maintaining consistent levels of engagement in people using a system. As wearable physiological sensors such as smart watches become more popular and less invasive, passive systems which make use of the user's affective states will become easier to implement, more reliable, and more common-place. In this paper, we presented our system based on the physiology-based DDA mechanism, which uses the user's engagement level and performance to dynamically adjust the difficulty level of the driving task. We demonstrated that our system is reliable and robust. Future work will focus on conducting a larger user study in order to make conclusion about the superiority of the ES system over the PS system.

References

1. CDC, Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years—Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2010 (2014)
2. Ganz, M.L.: The Costs of Autism, pp. 475–502. CRC Press, New York (2006)
3. Chasson, G.S., Harris, G.E., Neely, W.J.: Cost comparison of early intensive behavioral intervention and special education for children with autism. J. Child Fam. Stud. **16**(3), 401–413 (2007)
4. Rogers, S.J.: Empirically supported comprehensive treatments for young children with autism. J. Clin. Child Psychol. **27**(2), 168–179 (1998)
5. Palmen, A., Didden, R., Lang, R.: A systematic review of behavioral intervention research on adaptive skill building in high-functioning young adults with autism spectrum disorder. Res. Autism Spectrum Disord. **6**(2), 602–617 (2012)
6. Welch, K.C., Lahiri, U., Liu, C., Weller, R., Sarkar, N., Warren, Z.: An affect-sensitive social interaction paradigm utilizing virtual reality environments for autism intervention. In: Jacko, J.A. (ed.) HCI International 2009, Part III. LNCS, vol. 5612, pp. 703–712. Springer, Heidelberg (2009)
7. Lahiri, U., Warren, Z., Sarkar, N.: Dynamic gaze measurement with adaptive response technology in Virtual Reality based social communication for autism. In: 2011 International Conference on Virtual Rehabilitation (ICVR). IEEE (2011)
8. Lahiri, U., et al.: Design of a virtual reality based adaptive response technology for children with autism. IEEE Trans. Neural Syst. Rehabil. Eng. **21**(1), 55–64 (2013)
9. Lahiri, U., Warren, Z., Sarkar, N.: Design of a gaze-sensitive virtual social interactive system for children with autism. IEEE Trans. Neural Syst. Rehabil. Eng. **19**(4), 443–452 (2011)
10. Liu, C., et al.: Online affect detection and robot behavior adaptation for intervention of children with autism. IEEE Trans. Robotics **24**(4), 883–896 (2008)
11. Tartaro, A., Cassell, J.: Using virtual peer technology as an intervention for children with autism. Towards Universal Usability: Designing Computer Interfaces for Diverse User Populations, vol. 231, p. 62. Wiley, Chichester (2007)
12. Giakoumis, D., et al.: Identifying psychophysiological correlates of boredom and negative mood induced during HCI. In: B-Interface (2010)

13. Pekrun, R., et al.: Boredom in achievement settings: Exploring control–value antecedents and performance outcomes of a neglected emotion. J. Educ. Psychol. **102**(3), 531 (2010)
14. Rani, P., Sarkar, N., Liu, C.: Maintaining optimal challenge in computer games through real-time physiological feedback. In: Proceedings of the 11th international conference on human computer interaction (2005)
15. Dayi, B., et al.: Physiology-based affect recognition during driving in virtual environment for autism intervention. In: 2nd International Conference on Physiological Computing Systems (2015)
16. Bellani, M., et al.: Virtual reality in autism: state of the art. Epidemiol. Psychiatr. Sci. **20**(03), 235–238 (2011)
17. Sarkar, N.: Psychophysiological control architecture for human-robot coordination-concepts and initial experiments. In: IEEE International Conference on Robotics and Automation, Proceedings, ICRA 2002. IEEE (2002)
18. Stemmler, G., et al.: Constraints for emotion specificity in fear and anger: The context counts. Psychophysiology **38**(2), 275–291 (2001)
19. Hall, M., et al.: The WEKA data mining software: an update. ACM SIGKDD explorations newsletter **11**(1), 10–18 (2009)
20. Bian, D., Wade, J.W., Zhang, L., Bekele, E., Swanson, A., Crittendon, J.A., Sarkar, M., Warren, Z., Sarkar, N.: A novel virtual reality driving environment for autism intervention. In: Stephanidis, C., Antona, M. (eds.) UAHCI 2013, Part II. LNCS, vol. 8010, pp. 474–483. Springer, Heidelberg (2013)
21. Wade, J., Bian, D., Fan, J., Zhang, L., Swanson, A., Sarkar, M., Weitlauf, A., Warren, Z., Sarkar, N.: A virtual reality driving environment for training safe gaze patterns: application in individuals with ASD. In: Antona, M., Stephanidis, C. (eds.) UAHCI 2015. LNCS, vol. 9177, pp. 689–697. Springer, Heidelberg (2015)
22. Lord, C., et al.: The autism diagnostic observation schedule—generic: a standard measure of social and communication deficits associated with the spectrum of autism. J. Autism Dev. Disord. **30**(3), 205–223 (2000)