A Gaze-Contingent Adaptive Virtual Reality Driving Environment for Intervention in Individuals with Autism Spectrum Disorders

JOSHUA WADE, LIAN ZHANG, DAYI BIAN, and JING FAN, Department of Electrical Engineering and Computer Science, Vanderbilt University
AMY SWANSON and AMY WEITLAUF, Treatment and Research Institute for Autism Spectrum Disorders, Kennedy Center, Vanderbilt University
MEDHA SARKAR, Department of Computer Science, Middle Tennessee State University
ZACHARY WARREN, Department of Pediatrics and Psychiatry, Vanderbilt University
NILANJAN SARKAR, Department of Mechanical Engineering, Vanderbilt University

3

In addition to social and behavioral deficits, individuals with Autism Spectrum Disorder (ASD) often struggle to develop the adaptive skills necessary to achieve independence. Driving intervention in individuals with ASD is a growing area of study, but it is still widely under-researched. We present the development and preliminary assessment of a gaze-contingent adaptive virtual reality driving simulator that uses real-time gaze information to adapt the driving environment with the aim of providing a more individualized method of driving intervention. We conducted a small pilot study of 20 adolescents with ASD using our system: 10 with the adaptive gaze-contingent version of the system and 10 in a purely performance-based version. Preliminary results suggest that the novel intervention system may be beneficial in teaching driving skills to individuals with ASD.

---

1. INTRODUCTION

The last decade has witnessed a rapid increase in the rate of Autism Spectrum Disorder (ASD), with diagnoses in American children increasing an estimated 30% from 2012 to 2014 alone [Wingate et al. 2014]. ASD is a lifelong neurodevelopmental disorder wherein children demonstrate behavioral characteristics that differ significantly from their typically developing (TD) peers. Such characteristics may include reduced eye-contact during social interactions, repetition of activities, a decreased ability to process and understand others' emotions based on facial expressions, difficulty with motor coordination and planning, atypical gaze processing, and the inability to communicate verbally or nonverbally [American Psychiatric Association 2013; Senju et al. 2005]. In addition to social and behavioral symptoms, individuals with ASD often struggle to develop the adaptive skills necessary to achieve independence. Everyday tasks such as driving a vehicle or having a conversation are essential parts of independent living that may prove very challenging to individuals on the autism spectrum. Due in part to its lifelong course, ASD intervention and care results in substantial costs to the families of diagnosed individuals as well as to society at large. Recent conservative estimates suggest a national cost in the United States of $61 billion for children and $175 billion for adults annually [Buescher et al. 2014]. The sources of these costs include—but are not limited to—medical treatment, special education, individual and parental productivity loss, and paid staff to care for these individuals outside of a medical setting.

Although at present there is no single accepted cure or treatment for ASD, there is a growing consensus that appropriately targeted individualized behavioral and educational intervention programs have the potential to positively impact the lives of individuals and their families [Weitlauf et al. 2014]. This has very clearly been the case with regards to early intervention programs for preschool children, which have shown fairly robust impacts on cognitive and language skills. However, the intervention literature regarding functional gains in real-world and adaptive skills at later ages is much more equivocal. Specifically, interventions designed to address higher-level adaptive impairments (i.e., processing and integrating information from the environment, participating in new situations, and learning skills related to functional independence) have most often demonstrated short-term improvements in targeted skill areas but much more limited generalization of such skills to real-world environments. More simply, traditional skill instruction of social and adaptive interventions typically show low treatment effects and generalization [Bellini et al. 2007]. In this regard, with a lack of widely available efficacious treatments for addressing important complex life skills at later points in childhood, adolescence, and early adulthood, it is not surprising that the majority of individuals with ASD often fail to achieve many traditional markers of functional independence as adults [Shattuck et al. 2012].

Interestingly, many adolescents with ASD display an affinity for technology. Several recent studies have shown that adolescents with ASD demonstrate a higher level of engagement in certain tasks when interacting with a robot than with a human [Zheng et al. 2014]. Virtual reality (VR) environments have also proven to be engaging to these individuals, often in the form of video games [Tanaka et al. 2010]. As a result, researchers have developed a wide range of intervention programs based on these engaging paradigms. These intervention programs have included systems for teaching social behaviors such as emotion recognition [Bekele et al. 2013] and joint-attention [Bekele et al. 2011], as well as systems for teaching skills for adaptive independence such as operating a vehicle [Bian et al. 2013] or preparing meals [Alqahtani and Schoenfeld 2014].

A growing body of recent research has begun to focus on the driving behaviors of individuals diagnosed with ASD [Classen and Monahan 2013; Classen et al. 2013; Cox

et al. 2012; Daly et al. 2014; Reimer et al. 2013; Sheppard et al. 2010; Wade et al. 2014]. This work seems generally to suggest that (a) driving is more challenging for people with ASD, (b) these individuals may experience a greater risk to personal safety while driving than their TD peers, and (c) driving as it relates to people with ASD is highly underresearched. Driving intervention programs often cannot be realistically implemented in real-world driving environments due to the high risk of injury and obvious associated costs. The VR paradigm, however, provides an opportunity to create ideal environments that are safe, controlled, and capable of collecting a wide variety of data from both drivers and their environments. Such environments are not only capable of operating off of simple performance paradigms but also hold the potential for integrating other aspects of information processing into the environment in order to optimize learning (e.g., eye gaze and physiology).

Specialized real-world and simulated driving systems have been designed for applications unrelated to ASD which offer unique approaches to driving intervention. Fletcher and Zelinsky point out that drivers from all backgrounds display moments of inattention while driving. They, therefore, developed a gaze-sensitive driver assistance system to augment driver awareness in real-world driving environments. They sought to minimize inattention by alerting drivers through audio feedback to objects determined not to have been seen by the drivers. They demonstrated the feasibility of the system to identify road objects (such as signs and pedestrians) using eye-tracking technology coupled with sophisticated image processing software [Fletcher and Zelinsky 2009]. Rezaei and Klette used two cameras in real-world driving scenarios to obtain several features of the environment and the driver including distance from other vehicles, head orientation, yawning, and head-nodding—the latter two being indicators of drowsiness. They showed that this system could reliably detect whether a driver was looking at relevant vehicles in the environment [Rezaei and Klette 2014]. Ho et al. were also interested in reducing driver inattentiveness but employed a warning mechanism based on vibrotactile feedback rather than audio or video. Using commercial simulation software, they designed a system that activated a vibrotactile belt worn by drivers when the VR vehicle was either too close to a leading vehicle or too far behind. Vibration was spatially applied to the front (i.e., the stomach) when drivers were too close and in the rear (i.e., the lower back) when drivers were too far behind. They found drivers using this system braked earlier when a leading vehicle decelerated and left more space between themselves and the leading vehicle when stopped [Ho et al. 2006].

Relatively little research specifically investigates driving in the population with ASD [Classen and Monahan 2013]. However, since many individuals with ASD display affinity for nonbiological motion [Klin et al. 2009] and may find it difficult to attend to multiple stimuli during driving tasks [Cox et al. 2012], understanding their driving behaviors and aiding them in their driving skills require further research. Much of the work that has been done in this area has not utilized technology effectively to obtain objective, quantitative information about these individuals' driving behaviors. Sheppard et al. recruited 23 participants with ASD and 21 TD controls—all nondriving adult males—and showed them videos of hazardous driving scenarios that contained either a social hazard (e.g., a pedestrian) or a nonsocial hazard (e.g., an automobile). They found that individuals with ASD identified social hazards significantly less often than the controls and were slower than controls overall in identifying hazards [Sheppard et al. 2010]. Cox et al. surveyed the parents of adolescents and young adults with ASD about their experiences with their children learning to drive. The results of their survey indicated that a majority of parents were generally very concerned about the safety of their child while driving because of their child's ASD and parents tended to feel that multitasking (e.g., managing speed while entering a highway) was, for their child, a particularly problematic skill [Cox et al. 2012]. Another survey-based study conducted

by Daly et al. questioned licensed driving adults with and without ASD about their driving histories. They found that individuals with ASD reported being older at the age of licensure, spending less time driving, feeling less confident about their driving abilities, and experiencing a higher number of traffic violations than their TD peers [Daly et al. 2014].

Very few empirical studies have investigated a driving simulation paradigm to specifically assess the driving performance of individuals with ASD—one of which is the previous work of the present authors [Classen et al. 2013; Reimer et al. 2013; Wade et al. 2014]. Classen et al. used proprietary driving simulation software to observe the between group distinctions of seven adolescents with ASD and 22 TD controls [Classen et al. 2013]. They found that subjects with ASD demonstrated significantly more driving errors than controls. Reimer et al. also used proprietary driving simulation software to analyze the performance of young adult males diagnosed with higher-functioning ASD (HF-ASD) as compared to controls [Reimer et al. 2013]. In addition, this study observed physiological signals (i.e., heart rate and skin conductance) as well as eye gaze from participants. They found heart rate and skin conductance levels were nominally higher in the HF-ASD group, suggesting higher levels of anxiety while driving. Furthermore, compared to controls, the pattern of gaze in the HF-ASD group tended to be higher vertically and more to the right on the screen on which the VR driving environment was projected. Similar patterns of gaze and increased anxiety were reproduced in our previous study [Wade et al. 2014]. In this pilot study, we assessed the driving performance of individuals with ASD compared to TD controls using a novel driving simulator. The results showed that individuals with ASD displayed significant physiological differences from the TD group with respect to skin conductance, likely indicating higher levels of anxiety while driving. The same pattern of gaze reported by the Reimer et al. study was present, and, additionally, the ASD group demonstrated a significantly higher number of driving errors compared to their TD peers.

In summary, a growing body of work suggests that driving is a challenging skill for individuals with ASD and this challenge appears to be a direct result of how individuals with ASD process information in the driving environment. If this is the case, then effective interventions must incorporate methods for altering information processing, not merely focus on repeated exposure without processing support (e.g., performance systems). A primary reason why many current intervention approaches show limited improvements in functional adaptive skills may be that traditional skill-based methodologies often fail to systematically match intervention strategies to specific underlying processing deficits associated with targeted skills. We therefore hypothesize that in order to address the driving skill deficit of individuals with ASD, one would need to design an intelligent driving simulator that can (a) have embedded rules geared specifically towards ASD intervention, (b) provide individualized tasks and feedback to improve driving outcomes, and (c) be integrated with a host of sensors such as an eye tracker and physiological and EEG sensors to measure features of the driver and create dynamic closed-loop interaction. We further hypothesize that realization of an autonomous system capable of providing real-time gaze-contingent feedback would contribute to both enhanced performance within the driving environment (e.g., fewer driving errors) as well as shape how individuals with ASD were scanning the relevant objects in the environment (e.g., alterations in gaze). Although a range of high-quality, off-the-shelf driving simulation software exists that can be used to assess driving behaviors, these tools do not provide necessary access to the source code in order to design closed-loop systems that utilize information from both driving performance *and* internal driver state. They thus do not satisfy the aforementioned requirements for ASD intervention. As a result, commercial driving simulators were not deemed appropriate for this work.

In this article, we discuss the design and application of a VR-based, real-time gaze-contingent driving simulator capable of providing individualized feedback
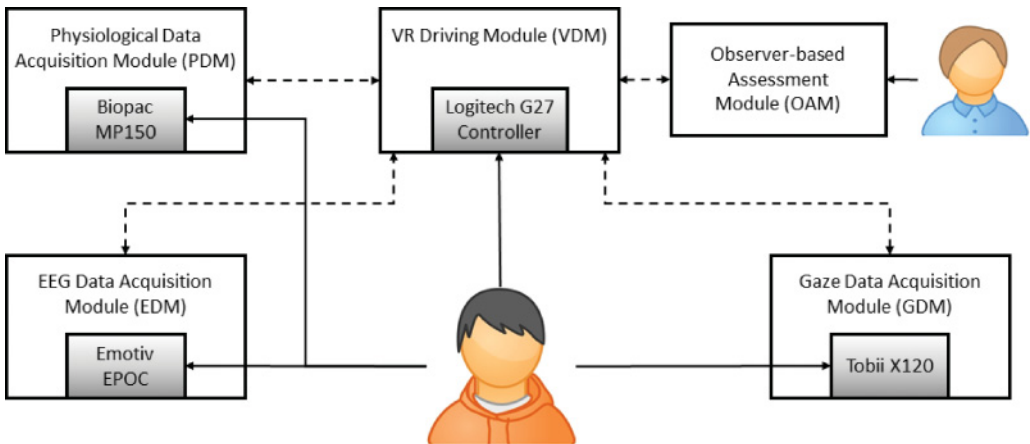
Fig. 1. VADIA block diagram. The solid lines indicate human input and the dashed lines represent network communication channels.

about how drivers scan their visual environment while driving. This system was designed to evaluate participant responses across two feedback modalities: strictly performance-dependent feedback and both performance- and gaze-contingent feedback. Although performance is an important factor in developing driving skills, it is just as important that drivers demonstrate appropriate gaze patterns while behind the wheel. The contributions of this work are threefold:

—The design and implementation of a novel adaptive VR driving simulation platform for ASD intervention that combines real-time gaze monitoring and optional gaze-contingent feedback to the driver in the virtual environment.
—The presented architecture supports integration of sensor feedback in a closed loop manner, which has been utilized to design an intervention program based on eye gaze information.
—The intervention system has been validated through a preliminary user study and results suggest that the system is functional and robust.

The development of this VR Adaptive Driving Intervention Architecture (henceforth referred to as VADIA) is described in the next section. Note that this article is a significantly expanded version, in terms of both technical content and results, of our earlier conference publications [Bian et al. 2013; Wade et al. 2014; Wade et al. 2015]. This article is organized as follows: Section 2 examines the structure of VADIA and the design of its component modules; Section 3 details the driving task design under the gaze-contingent mode of operation as well as the experimental procedure; Section 4 discusses the results from a validation of the system and the results of a user study involving 20 subjects with ASD; and Section 5 concludes the article with a review of the contributions as well as current limitations of VADIA and future work.

2. VADIA DEVELOPMENT

VADIA is composed of a set of interacting modules in which the central member is the VR driving module (VDM). In addition to the VDM, VADIA contains the following modules: the gaze data acquisition module (GDM), the physiological data acquisition module (PDM), the electroencephalography (EEG) data acquisition module (EDM), and the observer-based assessment module (OAM). Figure 1 gives a block diagram of the architecture. VADIA can be configured to operate with various combinations of modules depending on the intervention objectives. In addition, some modules may be disabled

under special circumstances. For example, the EDM can be disabled without impacting the rest of the system in cases where the participant refuses to wear the EEG device. This section discusses in detail VADIA's current modules identifying each module's role within the overall architecture as well as important features of the module.

2.1. Models of Computation

Formal models of computation (MoC) are tools used in the development process to design software systems in a way that compartmentalizes the system's behaviors and responsibilities. There are many reasons to suggest modeling systems using a formal MoC. Because a formal MoC has established syntax and semantics, it provides a framework for modeling that can be followed by anyone. Software often requires maintenance in order to fix a bug or extension to perform some new task. For very complex systems, these can prove to be major challenges. A model-centric approach to design may reduce the difficulty of these tasks by allowing developers to more easily identify the components of the system that require change as well as to see how the changes will impact the overall system. Additionally, the graphical representation of a MoC may serve as both documentation of a system's behavior and a useful tool for explaining how a system works to others. Finally, a formal MoC is amenable to analysis which allows developers to make guarantees about how a system will execute in a target environment. Quantitative analyses may be carried out to assess the worst-case execution time of a piece of code or the throughput of a network component. Formal verification methods such as model checking have also been developed that can be used to prove that a system will exhibit certain properties under all possible circumstances.

MoCs have been primarily applied in the context of Embedded Systems development, but recently they have begun to gain interest in other domains [Buntins et al. 2013; Schwarze et al. 2013]. Many different types of MoCs exist and the choice of which to use depends on the intended behavior of the system. Some examples include finite state machines (FSM), statecharts or hierarchical state machines (HSM), discrete events systems, hybrid automata (HA), petri nets, and dataflow networks [Lee and Seshia 2011]. Within each of these types of MoCs, there are many recognized variations, each with its own formal syntax and semantics. In this work, a mixture of FSMs, HSMs, and HAs were used to model various components of VADIA's modules. We discuss these in more detail in the next subsection.

2.2. Implementation

Here we discuss each of VADIA's modules in depth. For each module, we describe its role in the overall architecture, associated hardware interface (if applicable), software modeling approach taken, and implementation details of note.

2.2.1. VR Driving Module (VDM). The fundamental module of VADIA is the VDM. The VDM consists of both VR driving simulation software and a hardware driving interface that allow users to perform driving tasks in a controlled environment. The simulation system was designed such that users could engage in a variety of meaningful driving tasks and that parameters of the tasks such as complexity, difficulty, and length could be explicitly controlled. Drivers could interact with traffic lights, pedestrians, and other vehicles while completing tasks and a built-in navigation system aided drivers in reaching their respective destinations. Behaviors were defined for autonomous vehicles and pedestrians in order to make the tasks more naturalistic. The system was designed in such a way that tasks could be completed or failed and either outcome resulted in some appropriate feedback to the drivers.

The first step towards developing the virtual driving environment was to create a three-dimensional (3D) model city. Using CityEngine software (www.esri.com/software/
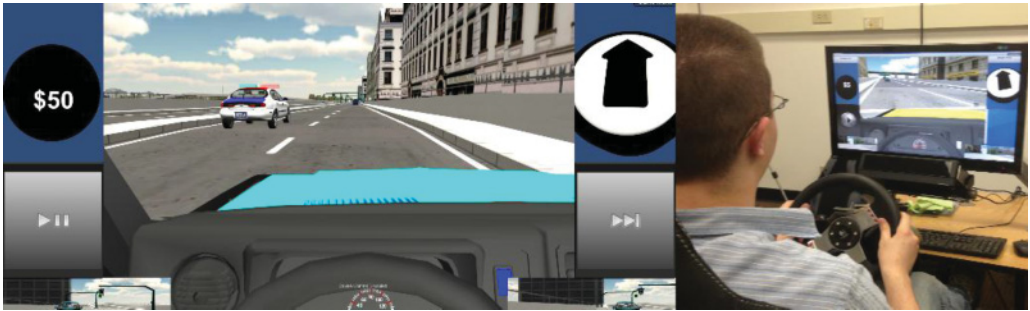
Fig. 2.   The VDM interface: driver perspective (left) and Logitech G27 controller (right).

cityengine), we constructed a model city sufficiently large to allow the creation of driving tasks lasting up to a few hours. This model consisted of diverse regions including a large downtown area with skyscrapers, a residential community, and an industrial park. A wide variety of roads were represented, including narrow one-way streets, multilane highways, and sharp turns. Next, the city was appropriately populated with typical objects such as vehicles, pedestrians, traffic lights, and signs. Vehicle and pedestrian objects were obtained from free online repositories of 3D models. The rest of the required models were created using Autodesk Maya software (www.autodesk.com/products/maya).

The Unity3D game development platform (or simply "Unity"; www.unity3d.com) was selected to implement the VR driving application. Unity was a good option for development for a number of reasons: It supports a variety of high-level programming languages, has a robust physics application programming interface (API), is available in both free and paid versions, and generates executables on all major operating systems. Additionally, Unity natively interfaces with many input devices including the Logitech G27 driving controller, which was used in the presented work. The G27 controller includes a steering wheel with several customizable buttons, a pedal board with three independent pedals, and a gear shifter which was not used in this study. Figure 2 shows the G27 controller in use as well as a screenshot from the driving perspective.

The software of the VDM is made up of several major components (see Table I for a description of some of these). At the low level, there are components such as *TurnSignalControl* and *TrafficLightDisplay* that define the behaviors of simple elements of the driving environment. More complex components dictate the behavior of semiautonomous objects such as *AIPedestrian* and *AIVehicle*. There are other components like *EEGServer* and *QServer* dedicated to communicating over the network with VADIA's other modules. At the highest level, the supervisory controller *GameMonitor* manages the lower-level components, synchronizes them, broadcasts events, and effectively dictates the flow of the system. Figure 3 gives a view of the hierarchy among the VDM's major components. Having control over all of these elements of the driving environment was key in developing the embedded rules of the system. For example, traffic light signals and autonomous vehicle behaviors were manipulated to create a range of task complexity that was suitable to drivers across a spectrum. Additionally, feedback presentation mechanisms suitable to individuals with ASD were incorporated, such as object-highlighting and audio/text presentation. A full description of each of the VDM's components is well beyond the scope of this article and also not necessary. Instead, the discussion is focused on the supervisory *GameMonitor* component in order to demonstrate the application of model-centric design and to show that it is appropriate for software design outside the context of Embedded Systems.

Table I. Major Components of the VDM Software

Name	Purpose
GameMonitor	Supervisory controller manages highest level of application logic such as menu navigation (e.g., vehicle and assignment selection), initiating network communication, and beginning driving tasks.
TrialMonitor	Initializes trials and monitors failure and success events.
PathMonitor	Manages placement of driver's vehicle at trial start and reset of a trial after a failure.
GazeMonitor	Maintains hashtable of fixation duration information and reports gaze-related failures in the gaze-contingent version of the system.
SoundMonitor	Schedules various audio clips used in the simulator (e.g., congratulatory sounds for successful trial completion).
DataManager	Logs user performance data as well as fixation duration data and metadata.
FeedbackModule	Manages the presentation of feedback within the system.
EyeTrackerModule	Handles network communication with the GDM and provides real-time gaze information to the GazeMonitor component.
PhysioModule	Handles network communication with the PDM.
EEGServer	Handles network communication with the EDM.
Qserver	Handles network communication with the OAM.
VehicleManager	Controls the driver's vehicle state based on the driver's input via the G27 controller.
AIVehicle	Controls the autonomous vehicles in the driving tasks.
AIPedestrian	Controls the autonomous pedestrians in the driving tasks.
TrafficLightDisplay	Manages the scheduling of traffic light signals.
ExcessiveSpeedFSM	Detects driving failures due to exceeding the speed limit.
OffRoadDetectionFSM	Detects driving failures due to driving off the road (i.e., driving too far onto the shoulder or onto the grass).
IncorrectLaneFSM	Detects driving failures due to driving in the wrong lane.
RidingSidewalkFSM	Detects driving failures due to driving onto the sidewalk.
PassingFSM	Detects driving failures that result from a driver either not passing a vehicle when it is required or passing in an improper way.
RunStopSignFSM	Detects driving failures related to failing to stop appropriately at a stop sign.
RunRedFSM	Detects driving failures related to running a red light at a traffic light intersection.
GpsController	Manages the navigation system used by the driver to complete tasks.

For the automata presented here, discrete modes or states are represented as rounded rectangles and transitions between states are given as arcs. Transition arcs are typically labeled using the format *guard/actions* where *guard* indicates the set of conditions under which the transition is valid and *actions* represents a set of tasks to be performed when the transition takes place. Such actions may include setting the value of a variable or generating an event. States are labeled with a descriptive name and, in the case of HA, include a set of ordinary differential equations that describe the behavior of the continuous variables in that state.

The following desired behavior was specified for the *GameMonitor* component: (a) handle the decision to enter practice mode or assignment mode, (b) manage the selection of vehicles by users, (c) manage the choice of level and assignment by users, (d) initiate the start of an assignment when users are ready to begin, (e) provide the
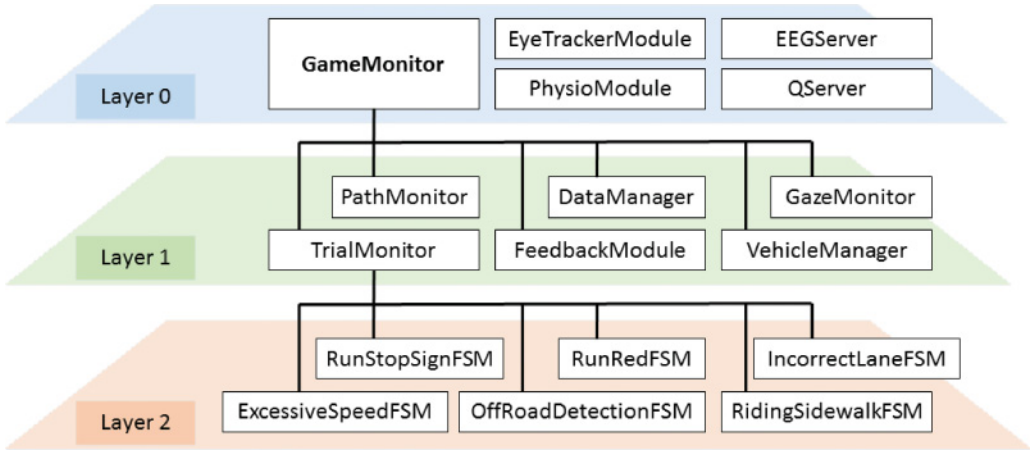
Fig. 3. Hierarchical view of some of the major components of the VDM software. Layer 0 is populated by concurrent, independent components such as the supervisory controller and network components while layers 1 and 2 are subordinate to the layers above them.
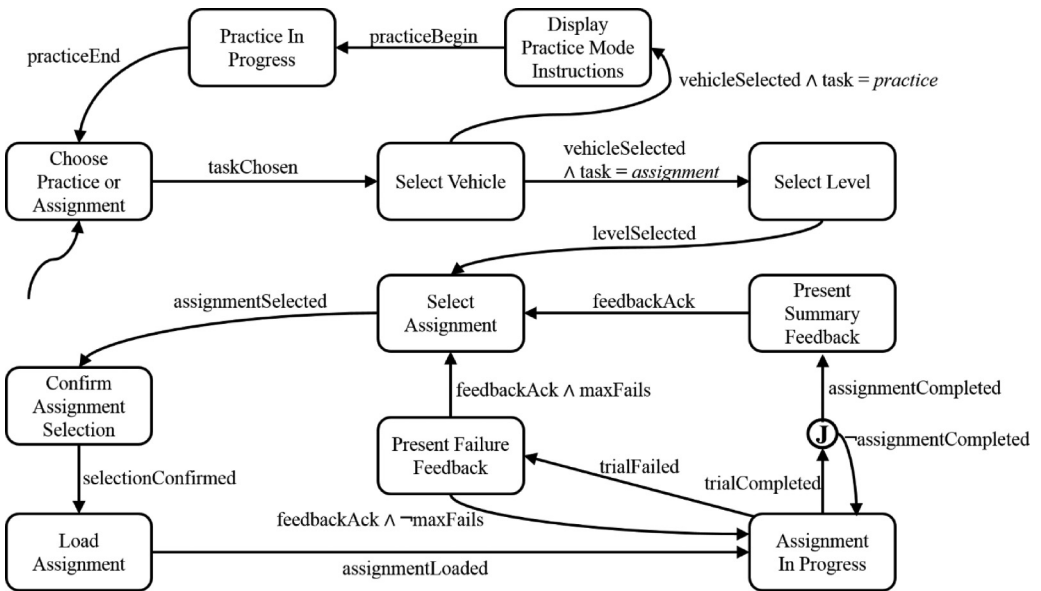
Fig. 4. FSM model of the *GameMonitor* component of the VDM. Note that self-loops, transition actions, and arcs for go-back options are excluded for simplicity.

option to return to the previous state of menu selection in case users make a mistake in selecting options, (f) keep track of trial successes and failures and assignment successes and failures, and (g) schedule the presentation of feedback. Since the behavior described in this specification deals exclusively with discrete events and variables with finite domains, we chose to model this behavior as a FSM shown in Figure 4. Keep in mind that other components such as *EyeTrackerModule* and *EEGServer* were in the same hierarchical layer as *GameMonitor* and so network-related logic is not the responsibility of this particular FSM.

The initial state of this FSM is labeled *Choose Practice or Assignment* and offers users the choice to run the system in either practice mode or regular assignment mode before allowing them to select a vehicle to operate in the *Select Vehicle* state. Depending on users' choices in the initial state, the FSM next transitions to either the state *Display Practice Mode Instructions* (followed by *Practice in Progress*) where users confirm the start of practice mode or the state *Select Level* where users choose the difficulty level to play. When users choose the assignment path (i.e., *Select Level → Select Assignment → Confirm Assignment Selection*), the appropriate assignment is loaded and executed in the state labeled *Assignment In Progress*. It is in this state that the *TrialMonitor* component is enabled and actively presents driving tasks to the user. Because *TrialMonitor* is only active when the *GameMonitor* is in this state, the *TrialMonitor* component is said to be subordinate to the *GameMonitor* component, hence the "hierarchical" state machine (Figure 3). While the assignment is in progress, the *TrialMonitor* may generate events indicating that a driving task has been either failed or completed successfully. When either of these events occur, the *GameMonitor* transitions to the appropriate next state: to the *Present Failure Feedback* state when a task is failed or to the junction **J** when the trial is completed. From the junction **J**, the assignment may end if it has been completed and will thus be followed by congratulatory feedback in the state *Present Summary Feedback* or cycle back to the *Assignment In Progress* state otherwise. We have opted not to include some details of the model in Figure 4 such as self-loops and transition actions, but this model accurately represents the specification outlined in the previous paragraph. Each MoC produced for components of the VDM system was transformed to software code using object-oriented, automata-based programming techniques. In all, over 30 formal models were mapped to source code for the VDM's implementation.

2.2.2. Gaze Data Acquisition Module (GDM). The GDM allowed VADIA to obtain eye gaze information from the driver. This information may be logged for offline analysis or utilized in a closed-loop system that uses gaze data to generate individualized driving feedback as we shall describe in Section 3.1.2.

The Tobii X120 remote eye-tracking device (www.tobii.com) was used to extract eye gaze data from participants at a sampling rate of 120 Hz. The X120 can achieve high accuracy (up to $0.5°$) and allows the participant some movement of the head in the range of $30 \times 22 \times 30$ cm (w × h × d) [Tobii Technology 2011]. The Tobii software API provides developers access to a variety of information about a subject's gaze including pupil diameter, blink rate, and gaze position on a surface (e.g., a monitor or projector screen). Pupil diameter and blink rate have been shown to be physiological indicators of a person's internal state and can reveal an individual's level of anxiety or engagement while gaze position may yield information about an individual's attention [Anderson et al. 2006; Jensen 2009].

An existing device interface application was modified to allow communication between the GDM and VDM. A simple FSM was created to model the pairwise network communication logic with the VDM and was implemented within the application. This software performed device calibration onto a $24''$ monitor (1920×1080 resolution), logged the subject's gaze data locally, and transmitted relevant data to the VDM over the network. The subject's gaze position information was used by the VDM for online processing while pupil diameter and blink rate were logged solely for offline analysis.

2.2.3. Physiology and EEG Data Acquisition Modules (PDM and EDM). The PDM and EDM acquired physiological and EEG data, respectively, from the driver for the purpose of offline analysis. Their primary roles in the presented work were to develop models of drivers' affect for a series of concurrent studies [Bian et al. 2015; Fan et al. 2015; Zhang et al. 2015]. Both modules were connected to the VDM via a local area network (LAN) and received event messages from the VDM for the purpose of aligning signals

with respect to important driving events. These modules are discussed here as part of the presentation of VADIA, but this article is not concerned with the development of a driver affect model.

The Biopac MP150 was used for physiological data acquisition (www.biopac.com). This system wirelessly sampled several physiological signals from the driver at 1000Hz. Signals of interest included electromyogram (EMG), electrocardiogram (ECG), galvanic skin response (GSR), photoplethysmogram (PPG), skin temperature, and respiration. These signals were selected because of their ability to predict the affective state of the subject [Cacioppo et al. 2007].

For EEG data acquisition, the Emotiv EPOC headset was used (www.emotiv.com). This device was used to sample from 14 locations on the wearer's head at 128 Hz. The data were streamed wirelessly to a custom PC-side application via a proprietary dongle. This application logged the EEG readings as well as signal quality and rotational velocity information collected from a gyroscope within the device.

2.2.4. Observer-Based Assessment Module (OAM). As mentioned in Section 2.2.3, the roles of the PDM and EDM were simply to log data regarding the internal state of subjects in order to derive a model of subject affect. In order to achieve this, the OAM was developed to provide a means for specifying the ground truth of subject affect by allowing an observer to label epochs of data in a meaningful way. As with the PDM and EDM, the OAM is described here for completeness, but the results of the affect model training are not within the scope of this article.

The OAM was designed for both online and offline use. In online mode, an observer could attend a driving session and make assessments about a driver's state in real time. These assessments were labeled with timestamps that could be mapped to data collected from the PDM and/or EDM. Video recordings of the subjects' faces and the driving environment were collected during driving sessions. Using this video, which included a timestamp overlay, observers were able to use OAM in offline mode as well.

The OAM prompted observers for five categories of input. Four measures of subject affect—engagement, enjoyment, boredom, and frustration—and an additional measure of perceived difficulty of the task for the subject were collected from the observer. Each category could be rated on a continuous scale in the range [0, 9] (where 0 indicated lowest intensity) by dragging a slider with a mouse. An observer was permitted to make an assessment at any time during the driving session—these were referred to as *periodic* assessments. There were also *summary* assessments that were made by the observer at the end of driving assignments. Assignments are described in detail in Section 3.1.

2.2.5. Network Communication. VADIA uses a star network topology in which the VDM is the central node. This is not strictly necessary but was convenient for the conducted study. All modules required a connection with the VDM but not to other modules. For example, the PDM had no need to interact directly with the OAM, nor did the GDM haven any reason to communicate with the EDM. All nodes in the network communicated via TCP/IP and used object serialization to send objects as messages between nodes. Each node was connected to a common router to create a LAN. The physical nodes consisted of three computers: the first executing the VDM and GDM software and interfacing with the G27 controller and eye tracker, the second running the PDM and EDM software and interfacing with the Biopac system and EPOC headset, and the third running the OAM software. Other assignments of processes to hardware were possible, but this configuration yielded the best performance.

3. STUDY DESIGN

In this section, we describe the design of the driving study. Although some of the details of the performance-based task design may be found in our previous work [Bian et al.

Table II. Level Difficulty Parameters

Name	Description	Domain
A_s	Speed of autonomous vehicles, a scalar	$\mathbb{R}$
A_a	Aggressiveness of autonomous vehicles, a scalar	$\mathbb{R}$
H_s	Traffic light alert sound	{enabled, disabled}
R_b	Responsiveness of the brake pedal, a scalar	$\mathbb{R}$
R_a	Responsiveness of the accelerator pedal, a scalar	$\mathbb{R}$
R_s	Responsiveness of the steering wheel, a scalar	$\mathbb{R}$
W	Weather condition state	{sunny, overcast, rainy}
L	Light intensity value, a scalar	$\mathbb{R}$
N_v	Number of vehicles at intersections/hwy entrances	$\{1, 2, \ldots, 5\}$
S_d	Duration of time to permit driving on sidewalk, in seconds	$\mathbb{R}$

2013; Wade et al. 2014; Wade et al. 2015], we resummarize major points here that are
key to understanding the new system.

3.1. Task Design

Here we describe the structure of the driving tasks, the development of the closed-loop
gaze-contingent operating mode, and important choices made in designing both the
performance-based and gaze-contingent modes of operation.

3.1.1. Task Structure. The driving scenarios used to evaluate driver performance were
called *trials*. Four classes of trials were defined: turning, merging, speed-maintenance,
and laws. Turning trials consisted of all those trials in which the participant made a left
or right turn at an intersection; this implies a street change and does not include driving
on sharply curved roads. Merging trials were characterized by any scenario in which
drivers either passed another vehicle or entered/exited a highway. Speed-maintenance
trials were those in which drivers were required to modify their speed to comply with
the changing environment. Active school zones and areas of road construction are
examples of scenarios in which the speed limit may change and drivers should adjust
their speed accordingly. The last category, laws, deals with an assortment of driving
scenarios such as waiting for a school bus to unload and stopping at stop signs, which
require drivers to know specific road laws.

A set of eight trials were assembled together, one after the other, into larger tasks
referred to as *assignments*. The number eight was chosen because it produced assign-
ments of desirable duration (roughly 5min). Three assignments were grouped together
into *levels*. In all, six levels were developed (i.e., 18 assignments or 144 trials). The
levels increased in difficulty from level one (the easiest) to level six (the hardest). In
order to implement six difficulty levels, we defined a set of difficulty parameters that
could be appropriately tuned to fit the desired difficulty settings. Table II describes
the parameters chosen. The functions of most of the parameters should be evident by
their descriptions, but we shall mention a few in passing. When the parameter H_s is
enabled, drivers are alerted via audio when relevant traffic lights change from red to
green. The parameter L controls the intensity of lighting in the environment (i.e., it
can be brighter or darker affecting the driver's visibility). The parameter S_d indicates
how long drivers can drive on a sidewalk without a penalty; sometimes drivers acci-
dentally drive onto the sidewalk (e.g., when turning a sharp corner) and so a kind of
"forgiveness" factor was quantified.

3.1.2. Operating Modes. Two distinct modes of operation were defined for VADIA: a
performance-based mode and a gaze-contingent mode, henceforth referred to as PB
and GC modes, respectively. A configuration file is used to specify the mode of opera-
tion employed by VADIA at runtime. The PB mode of operation presents users with
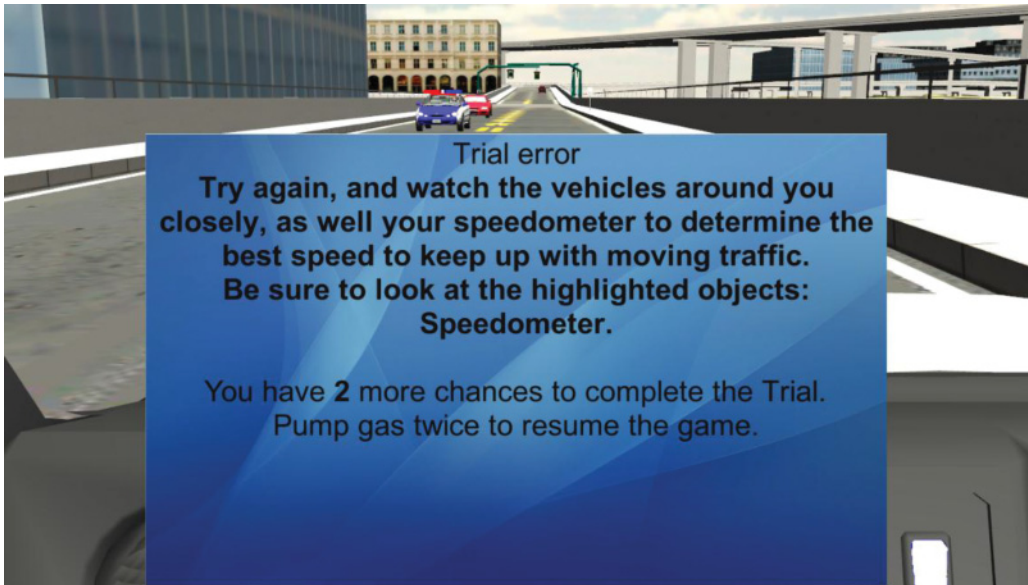
Fig. 5. A feedback message appears after a driver fails to look at the vehicle's speedometer during a speed-maintenance trial.

driving tasks in which progression depends entirely on their performance within the system. The GC mode of operation also enforces proper driving performance, but very importantly, it also requires that drivers demonstrate particular patterns of gaze while driving. In both modes, there exist the notions of *trial failures* and a *trial successes*. A trial failure is an event that occurs when the driver performs some kind of error during a trial. Similarly, a trial success is an event that occurs when the driver completes a trial without error. In the PB mode, trial failures can occur only due to performance-related issues such as failing to stop at a stop sign or exceeding the speed limit. In GC mode, trial failures could occur for performance-related errors as well as gaze-related problems. We will explain more precisely what this means in Sections 3.1.3 and 3.1.4. In essence, GC mode requires that drivers look at specified regions of interest (ROI) during trials in order to progress through tasks. Failing to look at these ROIs results in a *gaze failure* as opposed to a *performance failure*.

Trial failures and successes are met with feedback from the system. When drivers successfully completed a trial, a money counter (an arbitrary scoring system) increased by $5 and a congratulatory audio clip played (the obligatory cash register sound). When an entire assignment was completed successfully, drivers were presented with a congratulatory feedback message (e.g., *"Great job driving! Get ready for the next assignment!"*). On the other hand, when trial failures occurred, a feedback window would appear on the screen with a text message and corresponding audio explaining what drivers did wrong and how to correct it moving forward (see Figure 5). Additionally, for trial failures that occurred in GC mode, the feedback message would include the names of ROIs that the driver should be looking at (e.g., traffic light, stop sign, left side view mirror, etc.).

Following trial failure events, trials were reinitialized so drivers could make subsequent attempts at completing the failed trials. The number of reattempts permitted per assignment depended on the operating mode. In PB mode, three trial failures per assignment were allowed and the fourth failure resulted in the termination of the

Fig. 6. Example of ROI highlighting after a gaze-related trial failure. Both the male pedestrian and the pedestrian crossing sign have become highlighted with a green light to draw the attention of the driver.

assignment without a chance to reattempt. In GC mode, however, three trial failures were granted for both categories (i.e., three performance failures and three gaze failures) and a fourth trial failure in either category resulted in the termination of the assignment without a chance to reattempt. In GC mode only, when a trial was reset after a gaze failure, all of the ROIs relevant to that trial were highlighted with a green light to draw the driver's attention. Figure 6 shows an example of a scenario in which the ROIs become highlighted. In this case, the driver did not look at one or both of the male pedestrian and pedestrian crossing sign and thus are repeating the trial with the ROIs highlighted.

3.1.3. Selection of ROIs. The set of relevant ROIs selected for each trial were purposefully chosen on a trial-by-trial basis. Each of the 144 trials were exhaustively evaluated and a list of only the most crucial ROIs were identified as being key to the successful completion of each particular trial. The basic inclusion criterion for an object to be selected as a ROI was that the act of observing the object must be essential to safe driving. That being said, most trials in the same category (i.e., turning, merging, speed-maintenance, and laws) had very similar sets of relevant ROIs. In every speed-maintenance trial, for example, the relevant set of ROIs included at a minimum the vehicle's speedometer and the nearby speed limit signs. Without observing these two objects, drivers cannot be demonstrating truly safe driving behavior because they either do not know the speed at which they are travelling or do not know the current speed limit. Similarly, for trials requiring drivers to stop at pedestrian crossings, drivers must always be aware of both the crossing zone (i.e., the pedestrian crossing signs) and the pedestrian. Failing to notice one or both of these objects does not reflect optimally safe driving.

A comprehensive listing of the ROIs selected for each trial is not provided due to space consideration. Instead, Table III identifies the typical choice of ROIs by trial category. In a few cases, multiple virtual objects constituted the same ROI. For example, yield signs were located on both sides of highway entrance ramps. Rather than requiring

Table III. Typical Selection of ROIs by Trial Category

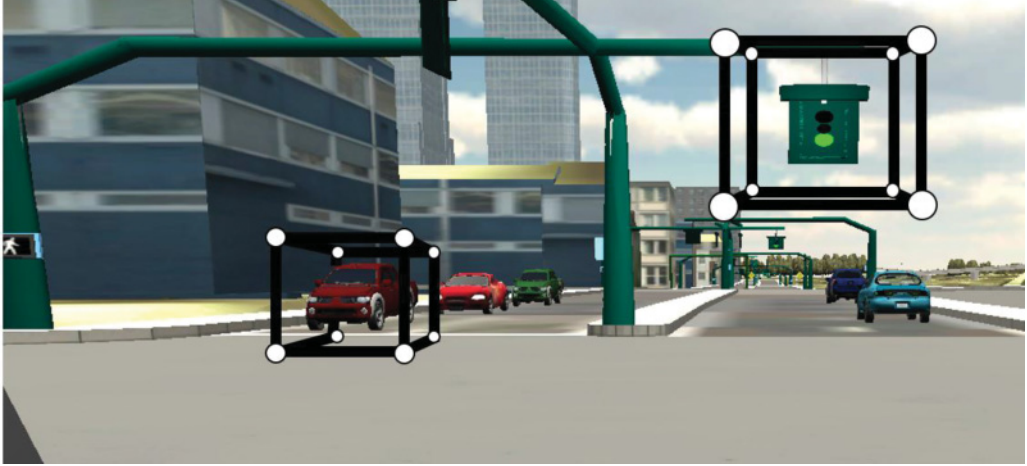| Category	Typical Regions of Interest (ROIs)
Turning	oncoming traffic, traffic lights, flashing lights, stop signs
Merging	side view mirrors, yield signs
Speed-maintenance	speedometer, speed limit signs, road work signs
Laws	pedestrian crossing signs, pedestrians, school bus stop sign, stop signs

Fig. 7. Bounding boxes around two ROIs in a turning-related trial: an oncoming vehicle and a traffic light.

drivers to observe both of these objects, one or the other was sufficient. Similarly, if multiple speed limit signs were present along a stretch of road, then drivers need only look at one of them to satisfy the gaze requirement.

3.1.4. Online Gaze Monitoring. Tracking drivers' eye gaze in order to know what objects they are looking at while driving is an essential function of the GC operating mode. As such, an algorithm was needed that would accept drivers' gaze as input and in turn report the amount of time drivers spent looking at various ROIs. The term *fixation duration* (FD) is used to indicate the amount of time that a driver spends looking at a particular ROI during a driving trial.

We developed an algorithm that performed the desired FD calculation (Algorithm 1). This algorithm utilized a hashtable data structure to store key-value pairs in which the keys were identifiers for specific ROIs, and the values were the FDs for the associated ROIs in seconds. The algorithm had two parts: an initialization part and an update part. Initialization was performed at the start (or restart) of each trial in order to clear the hashtable of old key-value pairs. The update part performed the actual FD calculations. The basic intuition of the algorithm was to check for intersection of the user's gaze $g \in \mathbb{R}^2$, with a circle of radius $r \in \mathbb{R}$, around each relevant ROI's centers of mass $c \in \mathbb{R}^3$. If the gaze position fell within the circle, then the driver was determined to be looking at the ROI.

Formally, a ROI was defined as a tuple $p = (c, E)$, where $c \in \mathbb{R}^3$ was the ROI's center of mass and $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_8\}$, $\forall_i \mathbf{e}_i \in \mathbb{R}^3$ was the set of extents of a bounding box around the ROI (see Figure 7). A transformation $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ was required to transform points in the virtual environment to coordinates on the screen; Unity's API provided such a transformation. If the algorithm determined that a gaze intersection for a particular ROI did in fact occur, then the FD time for that ROI was incremented by the elapsed time in seconds between the current and previous video frame. This

value was also available via the Unity API and is denoted in Algorithm 1 as Δ_t. In line 5 of Algorithm 1, the furthest extent e^* from the ROI's center of mass c_i (in screen space) is obtained. Then, in line 6, the radius r of a screen space circle is computed as the largest of either the screen space distance between c_i and e^* multiplied by scalar α, or the constant k, where k is the minimum allowable radius length. The value of k was equal to 1cm in screen space (roughly $0.82°$) and was selected based on a validation of the eye-tracking system (see Section 4.1 for a justification of the values of α and k). Line 7 of the algorithm determines if an intersection occurred and, if so, then line 8 increments the FD time for the relevant ROI.

ALGORITHM 1: Fixation Duration Calculation for Regions of Interest

---

Input: A hashtable H of key-value pairs where a key is a ROI name and the value is the fixation duration in seconds, the driver's gaze position g, the set R of ROIs relevant to the current trial, and the flag *initFlag* indicating when to (re)initialize the H.

Output: The updated hashtable H.
INITIALIZATION OF HASHTABLE
1: **if** *initFlag* **then**
2:     **for** $i = 1$ to n **do**
3:         $H[i] = 0$
UPDATE HASHTABLE WITH NEW FIXATION DURATIONS
4: **for** $i = 1$ to n **do** {Iterate over each of the ROIs}
5:     $e^* = \arg\max_{e \in E_i} \parallel T[c_i] - T[e] \parallel$ {Find the furthest extent from the center of mass}
6:     $r = \max(a * \parallel T[c_i] - T[e^*] \parallel, \quad k)$ {Compute the screen space ROI radius}
7:     **if** $\parallel g - T[c_i] \parallel \leq r$ **then**
8:         $H[i] = H[i] + \Delta_t$ {Increment by elapsed time between video frames}
9: **return** H

3.2. Participants

Twenty adolescents aged 13 to 18 years ($M = 15.29, SD = 1.65$) and diagnosed with ASD by our team of collaborating psychologists via administration of gold-standard research diagnostic instruments (e.g., Autism Diagnostic Observation Schedule [ADOS; Lord et al. 2000]) were recruited for a small study to evaluate the reliability of VADIA's PB and GC modes. Because ASD is much more common in males as females (3–5 times), and although recruitment was open to both sexes, the majority of our participants (19 of 20) were male. Randomized group placement assigned 10 participants to the PB group and the remaining 10 participants to the GC group. Of the 20 participants, 3 had a learner's permit and only 1 had a driver's license. Subjects' parents completed the Social Responsiveness Scale, second edition [SRS-2; Constantino and Gruber 2012], to quantify the severity of their child's ASD symptoms. The randomized group placement resulted in two individuals with learner's permits being assigned to the PB group and one individual with a driver's license and another with a learner's permit being assigned to the GC group. Table IV gives a detailed comparison of group characteristics. Informed consent and assent was appropriately obtained for each participant and participants were compensated at each visit for their time. This study was approved by the Institutional Review Board at Vanderbilt University.

3.3. Session Structure

Participants came to our lab facilities to complete six sessions, each lasting approximately 75min. In most cases, sessions were performed on separate days. The first and last of these sessions consisted of a pre- and posttest, respectively. The pre- and posttests were identical in terms of the task and served as the basis of comparison for

Table IV. Participant Characteristics (n = 20)

Group M (SD)			
	PB (n = 10)	GC (n = 8)	Failed to complete (n = 2)
Gender (% male)	100%	87.50%	100%
Chronological age (years)	15.1 (1.58)	15.76 (1.86)	14.37 (1.05)
SRS-2 total raw score	96.1 (31.66)	96.5 (21.05)	112 (35.36)
SRS-2 Tscore	74.3 (10.03)	75.63 (9.53)	80.5 (13.44)
Permit-holders (%)	20%	12.50%	0%
License-holders (%)	0%	12.50%	0%

Table V. Difficulty Level and Assignment Number per Session Given
as a Pair (Level, Assignment)

	1 (Pre-test)	2	3	4	5	6 (Post-test)
1st Assignment	2, 2	1, 1	3, 1	4, 1	6, 1	2, 2
2nd Assignment	5, 1	1, 2	3, 2	4, 2	6, 2	5, 1
3rd Assignment	5, 2	1, 3	3, 3	4, 3	6, 3	5, 2

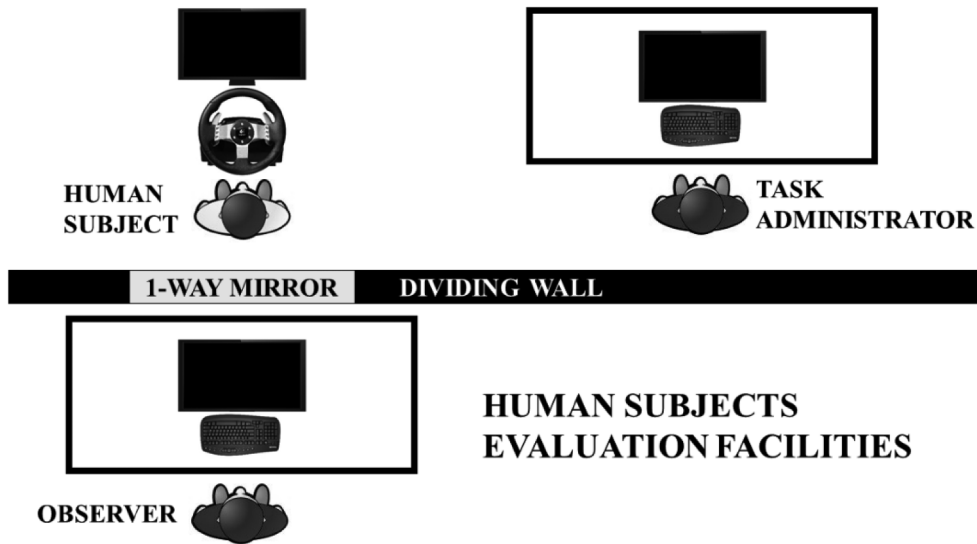| License-holders (%)	0%	12.50%	0%

Fig. 8. Experimental evaluation room schematic.

performance outcomes. The pre- and posttest session tasks consisted of three assignments selected from levels two and five in order to observe driver performance under a variety of difficulty settings. Sessions 2–5 were each unique in terms of both task and difficulty. Task difficulty was the easiest during session two, the hardest during session five, and medium-difficulty during sessions three and four (see Table V). Upon arriving for each session, physiological sensors were applied to the participant's body. Next, the EPOC headset was fitted to the participant's head. Following this, the Tobii eye-tracking device was calibrated following the calibration procedures recommended by the manufacturer [Tobii Technology 2011]. Figure 8 gives a schematic of the experimental facilities. Participants were seated in the playseat such that their eyes were approximately 70cm from the eye-tracking device. Each subjects' eyes were individually calibrated to a 24″ monitor (1920 × 1080 resolution) on which the VR driving

environment was displayed. A researcher administering the evaluation sat to the right of the subject at a PC running the simulation application. An observer sat behind a one-way mirror in a separate room at a PC running the observer-based assessment application.

At the first session, a brief tutorial was presented to the participants. This tutorial outlined the control of the G27 device, the basics of road safety, meanings of common signs, and how to complete driving tasks. Participants were informed only of the rules of the mode of the group in which they were placed (i.e., PB or GC group/mode). Following the tutorial, a silent baseline data collection period of 3min was performed to obtain a reference model of affect for each subject. When the baseline was complete, subjects began a 3min practice driving session to gain some experience with the controls before beginning the actual driving tasks (first session only). Finally, subjects attempted the core driving task which consisted of three assignments. As mentioned before, assignments could be attempted only once and when all three assignments had been attempted, the session was complete.

4. RESULTS

We present system validation results as well as the results of the user study. The purpose of the system validation tests was to show that the closed-loop gaze system was both functional and robust. The user study was conducted to demonstrate the acceptability and usability of both the PB and GC modes of the system as well as the potential of the designed intelligent system for future ASD intervention.

4.1. System Validation

Key elements of the system were quantitatively analyzed to ensure that they behaved according to specification. Specifically, we analyzed the network communication performance between the VDM and GDM modules and the measurement uncertainty of ROI detection using our proposed method. The developed system is far too complex for formal model checking in its entirety. Instead, we discuss a few important invariant properties of the system and show that they held for every observed execution of the system. Specifically, we show that feedback is always presented to the driver after a trial failure has occurred, and after every trial failure, another trial failure will not occur until at least 5s after the next trial begins.

Network communication throughput between the VDM and GDM is defined as the number of times per second that the VDM receives a new driver gaze coordinate from the GDM. The communication speed should be sufficiently fast to reliably represent drivers' gaze within the driving environment. We analyzed 32 assignments (256 trials) and found that the rate of transmission was approximately 11.6Hz or about every 86ms. Note that the GDM software logged drivers' gaze data at 120Hz irrespective of the network throughput. The observed frequency was more than adequate for real-time gaze monitoring because it has been shown that fixation durations less than about 300ms are not long enough to indicate actual fixation in a dynamic scene but rather saccadic movement [Rayner 1998].

The measurement uncertainty of the eye tracker in detecting users' gaze was also validated. Seven volunteers were enlisted to participate in a short data collection task. The task consisted of a small white circle appearing at nine known locations in a random order for 2s, and subjects were asked to look at the center of the circle as it moved along the screen. The same type of monitor was used for this task as in the driving task (24″ monitor with 1920×1080 resolution) and subjects were appropriately seated 70cm from the monitor [Tobii Technology 2011]. A mean error of $0.88°$ (1.08cm) was found with the error appearing to be less severe at the bottom of the monitor ($0.36°$, 0.44cm) than at the top ($1.57°$, 1.92cm). Because this error appeared to be linear, the

computed radii for ROIs were scaled based on their vertical position in screen space (α in line 6 of Algorithm 1) so ROIs near the top of the screen had slightly larger radii than those at the bottom of the screen. Additionally, the mean error provided the basis for selecting the value of the minimum ROI radius $k = 1$ cm in Algorithm 1.

For this intervention system to be effective, feedback must be presented to the driver after every single driving error. Over the course of the 120 driving sessions, 912 instances of driving errors were registered by the system. The event logs produced by the *DataManager* component of the VDM recorded the time of the failure, followed by the time of the feedback presentation as well as the time that the driver acknowledged the feedback message. In every instance of a trial failure, the feedback presentation and acknowledgment events were present in the log. Basic statistics were computed for the duration of time that the feedback was presented on screen and the mean duration time was 7.44s ($SD = 8.25$s, $Median = 4.55$s). Note that feedback was presented as both text and audio. Therefore, for those subjects who may have quickly acknowledged the feedback text in order to return to gameplay sooner, they were still presented with the audio version of feedback.

Since it is possible that multiple driving errors could occur at the same time (e.g., a driver makes a wrong turn *and* drives onto the sidewalk), it is necessary that multiple failure events are not reported at or near the same time. A timed automaton was designed specifically to guarantee that no trial failure event could be registered between the times that a trial failure event occurred until 5s after the next trial started. Of the exactly 600 instances of time intervals between consecutive trial failures measured during the driving sessions, the shortest time duration was 5.08s ($M = 55.39$s, $SD = 48.93$s).

4.2. User Study Results

Pre- and posttest outcomes were evaluated using two-sided Wilcoxon signed rank tests so medians are given along with Z and p statistics. This choice was due to the nonnormally distributed variables and the nature of the study (i.e., sampling the same population at two different points in time). In addition, Cohen's d is reported to give an estimate of effect size. The significance cutoffs for Cohen's d are as follows: $d \geq 0.8$ represents a large effect, $d \geq 0.5$ a medium effect, and $d \geq 0.2$ a small effect.

All of the subjects visited the lab for all six sessions and attempted the assignments in each task. We formally defined an inclusion criterion wherein subjects must complete at least one assignment over the course of all six visits for consideration in data analysis. This criterion resulted in the elimination of subjects from the data set. Removing these subjects changed the group balance to 10 participants in the PB group and only 8 in the GC group.

We analyzed two metrics of performance from the pre- and posttests of both groups to assess whether the intervention programs were effective. Median trial duration was computed for each trial completed by each subject and gives an indication of how difficult a task is to complete. Shorter median trial durations in the posttest are preferable because they indicate a higher proficiency in task completion. We also analyzed the median number of trial failures per assignment. This metric indicates the level of improvement within the task and should be smaller in the posttest than the pretest.

A Wilcoxon signed rank test indicated that, for subjects in the PB group, median trial duration was significantly longer during the pretest ($Mdn = 43.81$s) than during the posttest ($Mdn = 23.62$s) with large effect, $Z = 3.26$, $p < 0.01$, $d = 0.82$. Additionally, the PB group showed a significant decrease in the number of trial failures during the pretest visit ($Mdn = 7$) compared to the posttest visit (Mdn = 3) with very large effect, $Z = 2.37$, $p < 0.05$, $d = 1.43$.

GC group subjects also demonstrated a significant decrease in median trial duration from pretest ($Mdn = 36.05s$) to posttest ($Mdn = 28.82s$), $Z = 2.42$, $p < 0.05$, $d = 0.48$. GC group subjects also showed a significant decrease in total trial failures per visit from pretest ($Mdn = 11$) to posttest ($Mdn = 6$) also with a large effect, $Z = 2.37$, $p < 0.05$, $d = 1.12$.

5. DISCUSSION AND CONCLUSION

Both the PB and GC groups demonstrated statistically significant improvements in performance within the system over time. In both groups, the time taken to complete trials decreased and the number of driving errors resulting in trial failures decreased from pretest to posttest. It remains to be shown, of course, whether these improvements in the simulated environment translate to the real world. Thus this is an area for future study. There is, however, evidence of a correspondence between performance in the simulator and performance in the real world, which may indicate that performance improvements in the simulated environment also transfer to the real world [Bédard et al. 2010; Shechtman et al. 2010]. While the driving simulation tasks were similar across groups, the metrics for success differed for these groups. Specifically, the GC group had to *both* perform tasks correctly *and* look successfully at salient aspects of the environment. The PB group could achieve success just based on performance. With this caveat in place, the improvement in the GC group may be indicative of success in both scanning the driving environment as well as in performing the task. However, to determine whether this is the case, future work will warrant evaluating two groups of subjects on a common performance-based task who have received training from either the PB or GC methods. If in this task subjects from the GC group perform better than their PB group counterparts, then we may conclude that the GC version of the system is the more effective intervention method. Even without these experiments, however, both of the present intervention methods show very promising results for improving driving outcomes for individuals with ASD.

An overwhelming majority of individuals with ASD often fail to achieve many traditional markers of functional independence as adults [Shattuck et al. 2012]. Although there have been substantial advances in interventions demonstrating improvement for young children in targeted areas, effective interventions designed to address higher-level adaptive impairments (i.e., processing and integrating information from the environment, participating in new environments, and learning skills related to functional independence) with meaningful generalization to real-world environments are much more limited. One reason why many current intervention approaches show limited improvements in functional adaptive abilities may be that traditional skill-based methodologies often fail to systematically match intervention strategies to specific underlying processing deficits associated with targeted skills. Additionally, these intervention approaches may struggle to create opportunities for addressing such skills and deficits within and across naturalistic settings in appropriately intensive dosages, such that generalization of these skills is more challenging [Goodwin 2008].

Although promising, the present study is preliminary in nature with several clear key areas of needed future work. Our study's sample size, although characteristic of initial user studies in general, weakens the statistical power of the results to a certain extent. Readers are thus cautioned against making broad inferences at a population level from these data. A larger randomized clinical trial is required to verify the translation of these results to the general ASD population. Further, although we tested enhanced performance across systems, we did not embed a methodologically equivalent metric for calculating true performance and processing differences across the two conditions. Despite these limitations, our results represent meaningful steps toward developing dynamic VR driving environments linked with gaze technology for possible

ASD intervention. In future work, we plan to refine our simulator and further develop our VR program to test the ability of the system to advance learning beyond the intervention environment itself. We anticipate that providing a safe environment in which to practice driving skills, with ongoing monitoring of performance, engagement, and processing will lead to individuals with ASD having more confidence and successful navigation of driving tasks in the real world.

Cumulatively, our findings support the potential of developing technological tools such as VR driving simulators with embedded gaze-contingent feedback. Such integrated systems may be able to dynamically display important aspects of functional tasks, potentially guiding and altering gaze processing and attention, thereby enhancing processing and performance within these environments over time. Specifically, such a system could become a valuable intervention tool by enacting changes not just based on performance but also on how participants recognize and process environmental cues. Further, addressing underlying performance vulnerabilities on a processing level may result in changes that generalize more powerfully than current approaches for teaching functional skills, as real-world driving often requires fast and accurate interpretation of, and response to, others' behavior within environments that can be unsafe and allow little room for error. Finally, it is important to note that this difference in processing of dynamic information is likely not circumscribed to driving but may be related to many other challenges and vulnerabilities associated with ASD. As such, work addressing differences in processing (rather than just performance) may be important to designing intervention paradigms across other areas of skill vulnerability.

ACKNOWLEDGMENTS

REFERENCES

Hanadi H. Alqahtani and Naomi A. Schoenfeld. 2014. Teaching cooking skills to young women with mild intellectual disability: The effectiveness of internet websites. *Curr. Issues Educ.* 17, 2 (2014).

Christa J. Anderson, John Colombo, and Jill D. Shaddy. 2006. Visual scanning and pupillary responses in young children with autism spectrum disorder. *J. Clin. Exp. Neuropsychol.* 28, 7 (2006), 1238–1256.

American Psychiatric Association. 2013. *DSM-V Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). American Psychiatric Association.

Michel Bédard, Marie Parkkari, Bruce Weaver, Julie Riendeau, and Mike Dahlquist. 2010. Assessment of driving performance using a simulator protocol: Validity and reproducibility. *Am. J. Occupat. Ther.* 64, 2 (2010), 336–340.

Esubalew Bekele, Uttama Lahiri, Julie Davidson, Zachary Warren, and Nilanjan Sarkar. 2011. Development of a novel robot-mediated adaptive response system for joint attention task for children with autism. In *Proceedings of the International Symposium on Robot and Human Interactive Communication (RO-MAN), 2011 IEEE*. IEEE, Washington, DC, 276–281.

Esubalew Bekele, Zhi Zheng, Amy Swanson, Julie Crittendon, Zachary Warren, and Nilanjan Sarkar. 2013. Understanding how adolescents with autism respond to facial expressions in virtual reality environments. *IEEE Trans. Visual. Comput. Graphics* 19, 4 (2013), 711–720.

Scott Bellini, Jessica K. Peters, Lauren Benner, and Andrea Hopf. 2007. A meta-analysis of school-based social skills interventions for children with autism spectrum disorders. *Remed. Spec. Educ.* 28, 3 (2007), 153–162.

Dayi Bian, Joshua Wade, Amy Swanson, Zachary Warren, and Nilanjan Sarkar. 2015. Physiology-based affect recognition during driving in virtual environment for autism intervention. In *Proceedings of the 2nd International Conference on Physiological Computing Systems (PhyCS-2015)*. 137–145.

Dayi Bian, Joshua W. Wade, Lian Zhang, Esubalew Bekele, Amy Swanson, Julie Ana Crittendon, Medha Sarkar, Zachary Warren, and Nilanjan Sarkar. 2013. A novel virtual reality driving environment for autism intervention. *Universal Access in Human-Computer Interaction: User and Context Diversity*. Springer, Berlin 474–483.

Matthias Buntins, Jens W. Schicke, Frank Eggert, and Ursula Goltz. 2013. Hybrid Automata as a Modelling Approach in the Behavioural Sciences. *Electron. Not. Theor. Comput. Sci.* 297 (2013), 47–59.

Ariane V. S. Buescher, Zuleyha Cidav, Martin Knapp, and David S. Mandell. 2014. Costs of autism spectrum disorders in the United Kingdom and the United States. *J. Am. Med. Assoc. Ped.* 168, 8 (2014), 721–728.

John T. Cacioppo, Louis G. Tassinary, and Gary Berntson. 2007. *Handbook of Psychophysiology*. Cambridge University Press, Cambridge.

Sherrilene Classen and Miriam Monahan. 2013. Evidence-based review on interventions and determinants of driving performance in teens with attention deficit hyperactivity disorder or autism spectrum disorder. *Traffic Injury Prevent.* 14, 2 (2013), 188–193.

Sherrilene Classen, Miriam Monahan, and Stephanie Hernandez. 2013. Indicators of Simulated Driving Skills in Adolescents with Autism Spectrum Disorder. *Open J. Occupat. Ther.* 1, 4 (2013), 2:1–13.

J. Constantino and C. Gruber. 2012. *Social Responsiveness Scale, Second Edition*. Western Psychological Services, Los Angeles, CA.

Neill B. Cox, Ronald E. Reeve, Stephany M. Cox, and Daniel J. Cox. 2012. Brief Report: Driving and young adults with ASD: Parents' experiences. *J. Autism Dev. Disord* 42, 10 (2012), 2257–2262.

Brian P. Daly, Elizabeth G. Nicholls, Kristina E. Patrick, Danielle D. Brinckman, and Maria T. Schultheis. 2014. Driving behaviors in adults with Autism Spectrum Disorders. *J. Autism Dev. Disord.* 44, 12 (2014), 3119–3128.

Jing Fan, Joshua W. Wade, Dayi Bian, Alexandra P. Key, Zachary E. Warren, Lorraine C. Mion, and Nilanjan Sarkar. 2015. A step towards EEG-based Brain computer interface for autism intervention. In *Proceedings of the 37th Annual International Conference of the IEEE, Engineering in Medicine and Biology Society (EMBC)*.

Luke Fletcher and Alexander Zelinsky. 2009. Driver inattention detection based on eye gaze—road event correlation. *The Int. J. Robot. Res.* 28, 6 (2009), 774–801.

Matthew S. Goodwin. 2008. Enhancing and accelerating the pace of autism research and treatment. *Focus Autism Dev. Disabil.* 23, 2 (2008), 125–128.

Cristy Ho, Nick Reed, and Charles Spence. 2006. Assessing the effectiveness of "intuitive" vibrotactile warning signals in preventing front-to-rear-end collisions in a driving simulator. *Accident Anal. Prevent.* 38, 5 (2006), 988–996.

B. Jensen, B. Keehn, L. Brenner, S. P. Marshall, A. J. Lincoln, and R. A. Müller. 2009. Increased eye-blink rate in autism spectrum disorder may reflect dopaminergic abnormalities. In *Proceedings of the International Meeting for Autism Research (IMFAR)*.

Amy Klin, David J. Lin, Phillip Gorrindo, Gordon Ramsay, and Warren Jones. 2009. Two-year-olds with autism orient to non-social contingencies rather than biological motion. *Nature* 459, 7244 (2009), 257–261.

Edward A. Lee and Sanjit A. Seshia. 2011. *Introduction to Embedded Systems: A Cyber-Physical Systems Approach*. Lee & Seshia.

Catherine Lord, Susan Risi, Linda Lambrecht, Edwin H. Cook Jr., Bennett L. Leventhal, Pamela C. DiLavore, Andrew Pickles, and Michael Rutter. 2000. 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 (2000), 205–223.

Keith Rayner. 1998. Eye movements in reading and information processing: 20 years of research. *Psychol. Bull.* 124, 3 (1998), 372–422.

Bryan Reimer, Ronna Fried, Bruce Mehler, Gagan Joshi, Anela Bolfek, Kathryn M. Godfrey, Nan Zhao, Rachel Goldin, and Joseph Biederman. 2013. Brief report: Examining driving behavior in young adults with high functioning autism spectrum disorders: A pilot study using a driving simulation paradigm. *J. Autism Dev. Disord.* 43, 9 (2013), 2211–2217.

Mahdi Rezaei and Reinhard Klette. 2014. Look at the driver, look at the road: No distraction! No accident! In *Proceedings of the IEEE Conference Computer Vision and Pattern Recognition (CVPR)*. IEEE, 129–136.

Anke Schwarze, Matthias Buntins, Jens Schicke-Uffmann, Ursula Goltz, and Frank Eggert. 2013. Modelling driving behaviour using hybrid automata. *IET Intell. Transport Syst.* 7, 2 (2013), 251–256.

Atsushi Senju, Yshikuni Tojo, Kiyoshi Yaguchi, and Toshikazu Hasegawa. 2005. Deviant gaze processing in children with autism: an ERP study. *Neuropsychologia* 43, 9 (2005), 1297–1306.

Paul T. Shattuck, Sarah C. Narendorf, Benjamin Cooper, Paul R. Sterzing, Mary Wagner, and Julie L. Taylor. 2012. Postsecondary education and employment among youth with an autism spectrum disorder. *Pediatrics* 129, 6 (2012), 1042–1049.

Orit Shechtman, Kezia D. Awadzi, Sherrilene Classen, Desiree N. Lanford, and Yongsung Joo. 2010. Validity and critical driving errors of on-road assessment for older drivers. *Am. J. Occupat. Ther.* 64, 2 (2010), 242–251.

Elizabeth Sheppard, Danielle Ropar, Geoffrey Underwood, and Editha van Loon. 2010. Brief report: Driving hazard perception in autism. *J. Autism Dev. Disord.* 40, 4 (2010), 504–508.

James W. Tanaka, Julie M. Wolf, Cheryl Klaiman, Kathleen Koenig, Jeffrey Cockburn, Lauren Herlihy, Carla Brown, Sherin Stahl, Martha D. Kaiser, and Robert T. Schultz. 2010. Using computerized games to teach face recognition skills to children with autism spectrum disorder: The Let's Face It! Program. *J. Child Psychol. Psychiat.* 51, 8 (2010), 944–952.

Tobii Technology. 2011. *Accuracy and Precision Test Method for Remote Eye Trackers* (Edition 2.1.1). Tobii.

Joshua Wade, Dayi Bian, Lian Zhang, Amy Swanson, Medha Sarkar, Zachary Warren, and Nilanjan Sarkar. 2014. Design of a Virtual Reality Driving Environment to Assess Performance of Teenagers with ASD. In *Universal Access in Human-Computer Interaction. Universal Access to Information and Knowledge*. Springer, Berlin, 466–474.

Joshua Wade, Dayi Bian, Jing Fan, Lian Zhang, Amy Swanson, Medha Sarkar, Amy Weitlauf, Zachary Warren, and Nilanjan Sarkar. 2015. A virtual reality driving environment for training safe gaze patterns: Application in individuals with ASD. In *Universal Access in Human-Computer Interaction: Access to Learning, Health and Well-Being*. Springer, Berlin, 689–697.

Amy S. Weitlauf, Melissa L. McPheeters, Brittany Peters, Nila Sathe, Rebekah Travis, Rachel Aiello, Edwin Williamson, Jeremy Veenstra-VanderWeele, Shanthi Krishnaswami, and Rebecca Jerome. 2014. Therapies for children with autism spectrum disorder. *Comp. Effectiv. Rev.* 137 (2014).

Martha Wingate, Russell S. Kirby, Sydney Pettygrove, Chris Cunniff, Eldon Schulz, Tista Ghosh, Cordelia Robinson, Li-Ching Lee, Rebecca Landa, and John Constantino. 2014. Prevalence of autism spectrum disorder among children aged 8 years-autism and developmental disabilities monitoring network, 11 sites, United States, 2010. *Morbid. Mortal. Wkly. Rep. Surveill. Summ.* 63, 2 (2014).

Lian Zhang, Joshua Wade, Dayi Bian, Jing Fan, Amy Swanson, Amy Weitlauf, Zachary Warren, and Nilanjan Sarkar. 2015. Multimodal fusion for cognitive load measurement in an adaptive virtual reality driving task for autism intervention. In *Universal Access in Human-Computer Interaction. Access to Learning, Health and Well-Being*. Springer, Berlin, 709–720.

Zhi Zheng, Shuvajit Das, Eric M. Young, Amy Swanson, Zachary Warren, and Nilanjan Sarkar. 2014. Autonomous robot-mediated imitation learning for children with autism. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2707–2712.