computers

MDPI

Article

Multimodal Analysis of Hazard Perception Learning in Novice Drivers with Autism Using a Simulation-Based Training Environment

Erik J. Sand, Matthew T. Marino and Charles E. Hughes *

School of Modeling, Simulation and Training, University of Central Florida, Orlando, FL 32816, USA; erik.sand@ucf.edu (E.J.S.); matthew.marino@ucf.edu (M.T.M.)
* Correspondence: charles.hughes@ucf.edu

Abstract

Simulation-based driver training has shown promise for improving hazard perception in novice drivers; however, how learners with autism adapt behaviorally, visually, and physiologically during such training remains poorly understood. This study examined the effects of a game-based, hazard-focused driving simulation on hazard detection accuracy, gaze behavior, and heart rate in novice drivers with autism using a single-case, multi-phase design. Five participants completed repeated trials across baseline, treatment, and withdrawal phases while behavioral performance, eye movements, and physiological response were recorded. Across outcome domains, participants demonstrated highly individualized learning trajectories with substantial variability in both the direction and magnitude of change. Improvements in hazard detection accuracy were not consistently accompanied by changes in gaze organization or physiological response. While one participant exhibited a canonical pattern of coordinated improvement across behavioral, visual, and physiological measures, others showed dissociation between modalities, including reduced physiological arousal without performance gains or modest accuracy improvements despite sustained physiological engagement. Exploratory peri-hazard analyses further revealed participant-specific heart rate responses aligned to hazard detection, with no uniform temporal signature associated with learning. These findings suggest that hazard perception learning in drivers with autism does not follow a single pathway and cannot be inferred from any single performance or physiological metric. Instead, multimodal, within-participant analysis is critical for capturing meaningful individual adaptation during simulation-based training. The results have implications for the design and evaluation of driver training systems, supporting flexible, learner-specific assessment frameworks and adaptive approaches that accommodate diverse patterns of engagement and learning.

Keywords: autism spectrum disorder; driver training; hazard perception; eye tracking; road surveillance; game-based learning; driving simulation

Check for updates

Academic Editor: Amit Kumar Mishra

Received: 23 January 2026
Revised: 20 February 2026
Accepted: 24 February 2026
Published: 27 February 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

1.1. Teen Crash Risk and Road Surveillance

Motor vehicle crashes remain one of the leading causes of injury and death among adolescents and young adults worldwide [1]. Newly licensed teen drivers are disproportionately represented in crash statistics due to developmental, cognitive, and experiential factors that limit hazard detection and situational awareness. Compared with experienced

drivers, novices demonstrate reduced visual scanning breadth, delayed hazard recognition, and diminished anticipatory attention, particularly in complex traffic environments [2,3].

These risks are further amplified for adolescents on the autism spectrum. Autism Spectrum Disorder (ASD) is associated with differences in attentional allocation, executive functioning, sensory processing, and information integration, all of which are critical for safe driving. Prior research has shown that drivers with autism may exhibit narrower scanning patterns, delayed orienting to hazards, and elevated cognitive workload during driving tasks, potentially increasing susceptibility to missed or delayed responses.

Effective road surveillance requires continuous distribution of visual attention across multiple elements, including forward roadway context, peripheral motion, traffic control devices, and emerging hazards. This ongoing prioritization of visual information imposes substantial cognitive demands, particularly for novice drivers and those with ASD [4,5]. When attentional resources are overtaxed, hazard detection reliability declines, increasing crash risk [2,4,6].

Despite the importance of hazard perception as a foundational driving skill, traditional driver education provides limited direct training in visual scanning strategy and hazard anticipation. Instruction typically relies on static materials or limited on-road exposure, offering restricted opportunities for systematic repetition, performance measurement, and adaptive feedback. As a result, many novice drivers enter real traffic environments without having developed robust road surveillance strategies.

Improving hazard detection and visual attention efficiency in novice drivers on the autism spectrum therefore represents both a public safety priority and a training systems challenge, requiring approaches capable of safely targeting the perceptual–cognitive mechanisms underlying effective road surveillance.

1.2. Why Traditional Driver Training Falls Short

Conventional driver education relies on classroom instruction, video examples, and limited on-road practice. While these approaches introduce traffic laws and vehicle operation, they provide relatively little direct training in systematic visual scanning, hazard anticipation, and attentional prioritization—the perceptual–cognitive skills most closely associated with crash risk in novice drivers [2,3,7].

On-road instruction presents inherent limitations for hazard perception training. Real-world environments provide unpredictable and infrequent exposure to critical hazards, making repeated practice difficult to ensure [8,9]. Instructors must prioritize vehicle control and safety, limiting opportunities for targeted feedback on visual attention strategy. For adolescent drivers with ASD, these challenges may be compounded by elevated cognitive load, anxiety, and sensory sensitivity during live traffic exposure [10–12].

Classroom and video-based methods improve safety and allow repeated exposure but lack performance contingency. Learners can observe hazards without demonstrating real-time detection or sustained visual surveillance under dynamic conditions. Consequently, improvements in conceptual understanding do not always translate into improved visual scanning behavior [7,8].

Importantly, traditional instruction lacks direct instrumentation of attentional state. Instructors infer hazard perception indirectly through steering, braking, or verbal report, which do not provide precise, time-resolved measures of visual attention. This limitation makes it difficult to determine whether hazards were detected visually in time or whether delayed responses reflect attentional, perceptual, or motor factors [7,13].

These limitations highlight the need for training environments that are safe, repeatable, hazard-dense, and capable of directly measuring and shaping the perceptual processes

underlying hazard detection, particularly for specialized populations such as novice drivers with ASD [4,10,14].

1.3. Why Gaze and Physiology Matter

Hazard perception fundamentally depends on visual attention [7,8]. Eye tracking provides a direct, time-resolved measure of where attention is allocated and how visual information is sampled from the driving environment. Unlike steering input or braking response, gaze data reveal the spatial and temporal structure of visual scanning behavior.

Effective road surveillance requires continuous allocation of attention between forward roadway context, peripheral cues, signage, and emerging hazards. This dynamic prioritization is cognitively demanding, particularly for novice drivers and individuals with ASD, whose attentional allocation and sensory integration may differ from typical patterns [10,15]. Gaze behavior therefore provides a critical window into how visual attention supports hazard perception.

Physiological measures offer complementary insight by indexing cognitive and emotional workload [13,16]. Heart rate, in particular, has been widely used as an indicator of mental effort, stress, and task demand during driving. Elevated or highly variable heart rate may reflect increased workload or inefficient information processing. Combined with gaze tracking, physiological monitoring enables differentiation between attentional limitations and workload-related performance constraints.

Together, behavioral, gaze, and physiological measures support a multimodal assessment framework in which learning can be evaluated not only by performance outcomes but also by underlying attentional and workload dynamics. This multimodal instrumentation transforms simulation-based training into a measurable human–machine interaction system, enabling detailed analysis of how attention, performance, and workload evolve during hazard perception learning.

1.4. Game-Based Simulation as the Intervention

Simulation offers a uniquely powerful medium for hazard perception training because it enables safe, repeatable, and tightly controlled exposure to critical driving events. Unlike real-world instruction, simulated environments can guarantee repeated presentation of dangerous scenarios without placing the learner, instructor, or public at risk. From a training systems perspective, simulation also supports precise control over stimulus timing, difficulty scaling, and performance feedback—features that are difficult or impossible to achieve consistently on the road.

Game-based simulation extends these advantages by embedding training tasks within continuous, engaging interaction, sustaining learner attention while providing high volumes of practice. When coupled with real-time instrumentation, game-based systems can function not only as training tools, but also as measurement platforms that capture perceptual, behavioral, and physiological responses with millisecond-level precision [17,18].

The present study employs a customized version of SpeedLimit, a Unity-based driving simulation designed specifically to train and evaluate hazard detection and road surveillance in novice drivers. The game presents participants with a continuously advancing roadway populated by scripted hazards embedded within rural, suburban, and urban scenes. Hazards emerge from multiple spatial locations and at variable time scales, requiring persistent visual scanning and rapid attentional shifting rather than simple reaction to a single forward focal point.

Gameplay difficulty increases systematically through manipulation of vehicle speed and hazard density, enabling progressive challenge across repeated trials. Performance feedback can be enabled or disabled in real time, allowing the same simulation environment

to support baseline measurement, reinforced training, and withdrawal testing without altering the underlying perceptual structure of the task.

Crucially, SpeedLimit is fully instrumented for multimodal telemetry. Gaze location, hazard events, performance scoring, heart rate, and synchronized video output are captured concurrently during all trials. This integration transforms the simulation from a conventional training game into a measurement-rich experimental platform capable of supporting fine-grained analysis of how visual attention, workload, and hazard detection evolve with training exposure.

By uniting repeatable hazard delivery, adaptive difficulty, and multimodal instrumentation, the simulation provides a controlled yet ecologically meaningful environment for examining whether targeted game-based exposure can reshape the perceptual–cognitive processes that underlie effective road surveillance in novice drivers on the autism spectrum.

1.5. Study Purpose and Contributions

The purpose of this study is to evaluate whether a fully instrumented, game-based driving simulation could influence hazard detection, visual scanning efficiency, physiological workload, and driving-related attitudes in novice drivers on the autism spectrum. Rather than focusing solely on behavioral performance outcomes, the study adopts a multimodal evaluation perspective that examines how attention, performance, and physiological response evolve together across repeated training exposure.

Specifically, the study examined whether repeated interaction with a hazard-dense simulation environment would:

1. Improve the rate and consistency of hazard detection during dynamic driving scenes.
2. Promote more efficient visual scanning behavior, reflected in reduced gaze offset and more stable gaze organization.
3. Reduce physiological workload, as indexed by heart rate during gameplay.
4. Support more positive attitudes toward driving, as reported by both participants and parents or caregivers.

These outcomes were evaluated using a single-case, multi-phase experimental design with repeated measures across baseline, treatment, and withdrawal conditions. This structure enabled fine-grained within-participant analysis of performance, gaze behavior, and physiological response across progressive levels of task difficulty.

From a computational and systems perspective, this work makes several key contributions:

- It demonstrates the design and deployment of a multimodal, measurement-rich driving simulation platform integrating real-time gameplay, eye tracking, physiological monitoring, and synchronized video capture.
- It presents a reproducible evaluation pipeline for studying the coupled evolution of attention, performance, and workload in interactive safety training systems.
- It provides empirical evidence that targeted, game-based hazard exposure can reshape visual scanning behavior and hazard detection performance in some novice drivers on the autism spectrum, highlighting both responsive and non-responsive learning trajectories.
- It illustrates how instrumented serious games can function simultaneously as training tools and experimental testbeds for human–machine interaction research.

Together, these contributions position the present study at the intersection of simulation, human–computer interaction, and computational behavioral measurement, extending the role of game-based systems from engagement-oriented training platforms to quantitative tools for studying perceptual–cognitive learning under dynamic task conditions.

2. Related Work

2.1. Gaze and Hazard Detection in ASD

Effective hazard perception depends fundamentally on visual attention. A driving hazard cannot be anticipated or avoided unless it is first visually detected, making gaze behavior a primary mechanism underlying safe road surveillance. Across both simulated and real-world driving contexts, novice drivers demonstrate systematic limitations in visual scanning, including narrowed search patterns, delayed attention shifts, and reduced anticipation of emerging hazards [7,8]. These limitations are particularly consequential in complex, dynamic traffic environments where hazards may arise from peripheral or partially occluded locations.

In drivers on the autism spectrum, these novice-related limitations appear to persist and, in some cases, are amplified by differences in attentional allocation and cognitive control. Eye-tracking studies of adolescents and adults with autism have reported atypical visual search strategies during hazard perception tasks, including delayed fixation on hazards, reduced anticipatory scanning, and greater reliance on focal rather than distributed gaze patterns [11,19]. Under conditions of increased task demand, these attentional differences are further influenced by executive function and cognitive workload, with greater variability in gaze organization and hazard detection timing observed as cognitive load increases [14,20]. Together, this work suggests that hazard perception difficulties in drivers with autism are not solely attributable to inexperience but reflect underlying differences in how visual attention and cognitive resources are coordinated during dynamic driving tasks. These differences are not readily observable through steering or braking behavior alone.

Importantly, several studies indicate that gaze-related differences in drivers with autism are not necessarily reflected in gross driving outcomes alone. Population-level analyses using licensing, crash, and violation records have shown that adolescents and young adults on the autism spectrum exhibit crash rates and traffic outcomes that are comparable to, or in some cases lower than, neurotypical peers once licensed [21,22]. At the same time, behavioral performance measures such as lane keeping, speed maintenance, or collision counts may appear broadly comparable at an aggregate level, even as more fine-grained simulator and eye-tracking measures reveal meaningful differences in tactical driving behavior and visual scanning strategy [14,19,20]. This dissociation suggests that performance outcomes alone may mask inefficient or fragile perceptual–cognitive strategies, particularly in novice drivers who rely heavily on effortful attentional control under dynamic task demands. Despite the recognized importance of visual attention in hazard perception, much of the existing ASD driving literature relies on indirect indicators of attention, such as reaction time or post hoc verbal report. Relatively few studies incorporate continuous eye-tracking to directly verify whether hazards were visually perceived, when attention was allocated, and how gaze organization evolves across repeated exposure. When eye tracking is employed, it is often used descriptively rather than as a core determinant of detection performance.

Together, these findings underscore the importance of directly measuring gaze behavior when evaluating hazard detection in drivers on the autism spectrum. They also highlight a methodological gap in existing work: the lack of training and evaluation environments that combine repeated hazard exposure with continuous, time-resolved measurement of visual attention. Addressing this gap requires systems that can both elicit sustained road surveillance demands and instrument the perceptual processes that support hazard detection, rather than inferring attention solely from behavioral outcomes.

2.2. Simulation- and Game-Based Driver Training

Driving simulation has long been used as a training and assessment tool for novice drivers due to its ability to provide safe, repeatable exposure to hazardous scenarios that would be impractical or unethical to stage in real traffic. Simulator-based training allows precise control over scenario timing, environmental complexity, and task demands, enabling systematic evaluation of driver behavior under conditions designed to elicit safety-critical responses [8]. For novice drivers in particular, simulation has been shown to support improvements in hazard anticipation and visual scanning behaviors that generalize beyond specific trained scenarios.

Within autism-focused driving research, simulation has played a central role in examining perceptual, cognitive, and motor aspects of driving performance. Multiple studies have employed driving simulators to assess lane keeping, speed regulation, hazard response, and executive function under controlled conditions, revealing subtle but meaningful differences in how drivers with autism allocate attention and respond to dynamic traffic demands [5,10,23]. Importantly, these simulator-based differences are often more pronounced under increased workload or complex environmental conditions, suggesting that simulation is particularly well suited for probing the limits of attentional control and hazard processing.

More recent work has extended simulation beyond assessment toward intervention, using structured simulator-based training programs to improve driving-related skills in adolescents and young adults with autism. These interventions typically emphasize repeated practice, graduated difficulty, and targeted feedback, with outcomes measured through simulator performance, off-road assessments, or self- and parent-reported driving confidence [24–26]. While results have generally been promising, prior systems often rely on discrete scenarios, scripted trials, or instructor-mediated feedback, limiting the continuity of attentional demands and the density of hazard exposure.

Game-based simulation represents a further evolution of this approach by embedding hazard perception training within continuous, interactive gameplay. By leveraging mechanics drawn from commercial games such as sustained motion, escalating difficulty, and immediate feedback, game-based systems can maintain engagement while delivering high volumes of perceptual practice. Several researchers have argued that such designs may be particularly advantageous for learners with autism, who often benefit from structured, predictable environments paired with clear performance contingencies [17,18].

Despite these advances, many existing game-based or simulator-based training systems emphasize behavioral outcomes such as collision avoidance or task completion without directly measuring the attentional processes that support those outcomes. As a result, it remains unclear whether observed performance improvements reflect genuine changes in visual scanning strategy or reflect compensatory adaptations that may be fragile under real-world conditions.

The present study builds on prior simulation-based and game-based driver training research by integrating continuous hazard exposure with multimodal instrumentation, enabling simultaneous measurement of performance, gaze behavior, and physiological workload. This approach allows evaluation not only of whether training improves outcomes, but also of how underlying perceptual–cognitive strategies evolve during repeated exposure to hazard-rich driving tasks.

2.3. Multimodal Monitoring in Training Systems

Complex skills such as driving involve tightly coupled perceptual, cognitive, and motor processes that unfold continuously over time. As a result, single-channel performance measures such as task accuracy, collision counts, or completion time provide only

a partial view of learner behavior and may obscure the mechanisms underlying improvement or failure. In response, a growing body of research has emphasized the value of multimodal monitoring, integrating behavioral, attentional, and physiological signals to better characterize learner state and task engagement during training.

Within driving research, eye tracking has emerged as a particularly informative modality for assessing attentional allocation during dynamic tasks. Gaze measures provide time-resolved information about where visual attention is directed, how frequently it shifts, and whether hazards are visually sampled prior to behavioral response. Studies using eye tracking in simulated driving contexts demonstrate that gaze behavior can reveal differences in hazard anticipation, search efficiency, and attentional strategy that are not apparent from driving performance metrics alone, particularly under conditions of increased cognitive demand or learner variability [11,18,19,25].

Physiological measures provide a complementary window into driving performance by indexing cognitive workload and task-related stress. Heart rate and heart rate variability have been widely used to reflect changes in mental effort, attentional demand, and resource allocation during complex tasks, including simulated and real-world driving. From a theoretical perspective, increases in physiological activation are understood as responses to heightened task demands and competition for limited cognitive resources [13,16]. When interpreted alongside behavioral performance and gaze behavior, physiological measures help distinguish conditions in which performance decrements arise from attentional misallocation versus those driven by overload, stress, or inefficient resource management.

Recent work in adaptive training systems has increasingly sought to combine behavioral performance, gaze data, and physiological signals to support richer assessment and personalization. Multimodal approaches have been explored in virtual reality-based driving environments for learners with autism, where synchronized telemetry enables analysis of how attentional allocation and workload fluctuate in response to task demands [23,27]. These systems demonstrate the feasibility of integrating multiple sensing modalities within interactive simulations, though they are often deployed primarily for assessment rather than sustained training or longitudinal learning analysis.

Despite these advances, multimodal instrumentation remains relatively underutilized in driver training interventions, particularly those targeting hazard perception and road surveillance. Many systems collect gaze or physiological data in isolation or use these measures solely for post hoc analysis rather than as integrated components of the training and evaluation framework. Consequently, few studies have examined how visual attention, workload, and performance evolve together across repeated training exposure.

The present study builds on this emerging multimodal perspective by embedding synchronized gaze tracking, physiological monitoring, and performance logging within a continuous, game-based driving simulation. By capturing attention, workload, and behavior concurrently across baseline, treatment, and withdrawal phases, the system enables a fine-grained examination of how perceptual–cognitive processes change with training, moving beyond outcome-only evaluation toward a mechanistic understanding of learning in complex interactive environments.

3. Game-Based Training System

3.1. System Overview

The game-based training system used in this study was designed to function simultaneously as an interactive learning environment and as a multimodal experimental measurement platform. Rather than treating gameplay and data collection as separate components, the system integrates real-time task execution, performance logging,

eye tracking, physiological monitoring, and synchronized event timing within a unified simulation architecture.

The training environment was implemented using the Unity version 2017.3.0f3, which provides precise control over scene rendering, event scheduling, and input/output synchronization. Unity's component-based architecture enabled modular integration of gameplay logic, hazard scripting, telemetry capture, and external sensor streams while maintaining a consistent temporal reference across all data modalities. This design ensured that gaze location, hazard events, scoring outcomes, and physiological signals could be aligned at fine temporal resolution for post hoc analysis.

The system was configured to support repeated experimental trials across multiple phases (baseline, treatment, and withdrawal) while preserving a consistent underlying task structure. Rather than dynamically adjusting gameplay parameters within trials, the simulation employed a fixed set of predefined trial configurations, each associated with a specific vehicle speed and hazard presentation profile. These trial configurations were held constant across experimental phases, ensuring that participants encountered the same pacing and hazard structure whenever a given trial level was repeated.

This design choice emphasized experimental consistency over adaptive pacing. By locking trial parameters programmatically, the system ensured that observed changes in hazard detection performance, gaze behavior, or physiological response across phases could be attributed to learning and repeated exposure rather than to variation in task mechanics or environmental structure. At the same time, the fixed pacing and hazard rate-imposed constraints on how finely task difficulty could be tuned. This is an issue that will be revisited in the discussion of system limitations and future design improvements.

Importantly, the system was intentionally designed to minimize extraneous motor and interface demands. Steering and speed control were simplified to reduce confounding influences from complex vehicle handling, allowing the task to emphasize visual scanning, hazard anticipation, and attentional prioritization. By constraining motor complexity, the simulation isolates perceptual–cognitive processes central to road surveillance while remaining representative of core driving demands.

Together, these design choices position the system not merely as a serious game, but as a controlled experimental platform capable of capturing how attention, performance, and workload evolve during repeated exposure to hazard-rich driving scenarios (see Figure 1).

3.2. Driving Task and Hazard Presentation

The primary driving task consists of a continuously advancing forward-motion scenario in which participants navigate a straight roadway while monitoring for emerging hazards. Within game development, this design pattern is commonly referred to as an *infinite runner* mechanic. The visual environment is segmented into rural, suburban, and urban scenes, each designed to present distinct visual clutter, motion density, and hazard emergence patterns commonly encountered in real-world driving.

Hazards are scripted events that appear dynamically from multiple spatial locations, including the forward roadway, roadside periphery, and occluded regions such as behind parked vehicles or roadside structures. This design requires participants to distribute visual attention broadly rather than fixating on a single forward focal point. Hazards vary in timing, spatial origin, and visual salience, preventing anticipatory responses based solely on pattern memorization.

Participants were instructed to detect hazards as quickly and accurately as possible using a single response input. Correct detections were defined by gaze fixation within a predefined spatial window around each hazard and were time-stamped relative to hazard

onset, enabling precise measurement of detection latency and missed events. Importantly, hazards did not pause gameplay; instead, they unfolded within a continuous flow of motion, preserving attentional demands associated with real-world driving.



Figure 1. Representative in-game view from the SpeedLimit training environment, showing the participant's first-person driving perspective within a stylized suburban scene. The display includes forward roadway context, mirrors, and interface elements visible during hazard-perception trials. This view illustrates the visual field available to participants during training and measurement, within which hazards emerged and gaze behavior, performance, and physiological responses were recorded. No participant-identifying information is shown.

Task difficulty is manipulated through two primary parameters: vehicle speed and hazard density. As speed increases, the temporal window for hazard detection narrows, increasing perceptual load and response urgency. Higher hazard density increases visual competition and attentional switching demands, particularly in urban scenes where multiple non-hazard stimuli are present. These manipulations enable systematic scaling of cognitive demand across trials while maintaining identical hazard logic.

Feedback presentation can be enabled or disabled depending on experimental phase. During training trials, performance feedback reinforces successful hazard detection and missed events. During baseline and withdrawal phases, feedback is withheld to assess unassisted performance. This approach allows the same task environment to support both learning and evaluation without altering core perceptual demands.

By embedding hazard detection within a continuous, hazard-dense driving flow, the task emphasizes sustained road surveillance rather than isolated reaction time. This structure aligns closely with real-world driving demands while remaining sufficiently controlled to support detailed analysis of visual attention and workload under varying levels of difficulty.

The simplified control scheme, which relies on continuous forward motion without full vehicle steering dynamics, was intentionally designed to isolate perceptual–cognitive aspects of hazard perception while minimizing variability associated with fine motor control and vehicle handling skill. This approach improves experimental control by allowing hazard detection performance, gaze organization, and physiological response to be evaluated independently of steering proficiency. However, this design represents a tradeoff with ecological validity, as real-world driving requires continuous integration of perceptual, cognitive, and motor processes. Accordingly, the present study focuses specifically on

attentional allocation and hazard detection within a controlled simulation context, and findings should be interpreted within this scope.

3.3. Eye-Tracking Integration

Eye-tracking was integrated into the simulation to provide continuous, objective measurement of visual attention during gameplay. Gaze data were collected using a desktop-mounted Tobii EyeX eye tracker purchased from Amazon.com positioned below the display monitor. The system sampled binocular gaze position at 20 Hz, which provides sufficient temporal resolution for characterizing visual scanning behavior and hazard-directed attention in simulated driving tasks.

Prior to each session, a standard calibration procedure was performed to ensure accurate mapping of gaze coordinates to on-screen locations. Calibration quality was visually verified before trials commenced, and sessions with inadequate tracking quality were excluded from analysis. Participants were seated at a fixed distance from the display to maintain consistent viewing geometry across trials.

Raw gaze coordinates were recorded continuously throughout gameplay and synchronized with game events, including hazard onset, scoring events, and scene transitions. Gaze data were mapped into screen-centered coordinate space, enabling calculation of gaze offset (distance from screen center) and directional gaze bias over time. These measures provide complementary indicators of visual scanning efficiency and attentional organization during dynamic driving scenes.

For hazard detection analysis, participant responses were registered via a single paddle-shift input on the game controller. Following each response, gaze fixation within a predefined spatial window surrounding the active hazard was evaluated to determine whether the hazard had been visually detected. Fixation timing relative to hazard onset enabled precise estimation of detection latency, while responses occurring without gaze fixation within the hazard window were classified as missed or incorrect detections. Responses occurring outside the hazard's active temporal window were ignored for detection analysis. This response-gated gaze criterion ensured that detection metrics reflected verified perceptual attention rather than response timing or motor execution alone.

By embedding eye tracking directly into the simulation loop and synchronizing gaze telemetry with task events, the system enabled fine-grained, time-resolved analysis of how visual attention evolved across trials and experimental phases. This integration was essential for evaluating whether training-related changes in performance were accompanied by measurable shifts in underlying visual scanning strategy.

3.4. Heart Rate Monitoring

Heart rate was recorded during all gameplay sessions as a physiological indicator of cognitive and emotional workload. Cardiac data were collected using a heart rate sensor worn by participants throughout each trial. Heart rate was sampled at a fixed frequency and synchronized with gameplay events, including hazard onset, participant responses, and scene transitions.

Heart rate measures were analyzed as time series aligned to trial progression rather than as isolated summary values. This approach enabled examination of both overall workload trends across trials and transient physiological responses associated with hazard-rich segments of gameplay. Elevated or highly variable heart rate during task performance was interpreted as reflecting increased cognitive demand, stress, or inefficient information processing rather than physical exertion, as gameplay involved minimal physical movement.

To facilitate integration with gaze and performance data, heart rate recordings were temporally aligned to the same reference clock used for eye tracking and game telemetry. This synchronization enabled concurrent analysis of attentional allocation, behavioral performance, and physiological response within a unified multimodal framework.

By incorporating heart rate monitoring alongside gaze and performance metrics, the system supports differentiation between conditions in which degraded performance reflects attentional misallocation versus those characterized by elevated workload or stress. This distinction is particularly relevant for evaluating training-related changes in efficiency, where improved performance may be accompanied by reduced physiological demand even in the absence of overt behavioral change.

3.5. Performance Feedback and Scoring

The simulation incorporated an internal scoring and health system to provide structured, performance-contingent consequences during gameplay. Participants accumulated points for hazards that were correctly detected, with scoring weighted by response timing relative to hazard onset (i.e., faster verified detections yielded higher values). In addition to points, the system tracked a health variable that decreased when hazards were missed. Because the driving task continued as an uninterrupted run, reductions in health could shorten trial duration by terminating the run early when health was depleted.

Feedback presentation varied by experimental phase. During treatment trials, participants received performance feedback associated with scoring and health changes, providing reinforcement for timely hazard detection and clear consequences for misses. During baseline and withdrawal phases, feedback was withheld while telemetry continued to be recorded, enabling measurement of hazard detection and attention allocation without overt reinforcement cues.

Because missed hazards could shorten trial duration, total score alone did not provide a complete representation of performance. Outcome analyses therefore accounted for variable exposure by emphasizing time-resolved detection measures, including verified detections and latency relative to hazard onset, in conjunction with gaze behavior and physiological workload indices. This approach ensured that performance differences reflected perceptual and attentional processes rather than differences in trial length.

By maintaining consistent scoring and health logic across phases while manipulating feedback availability, the system supported evaluation of whether improvements in hazard detection and scanning efficiency generalized beyond the presence of explicit reinforcement.

4. Materials and Methods

4.1. Participants

Five adolescents and young adults identified by their parents or caregivers as being on the Autism Spectrum participated in this study. Participants ranged in age from 15 to 18 years old and were all novice drivers who were either preparing to begin formal driving instruction or were in the early stages of learning to drive. All participants had sufficient cognitive and motor functioning to operate a desktop-based driving simulation and to comply with study procedures.

Participants were recruited as a convenience sample through the University of Central Florida's Center for Autism and Related Disabilities (CARD). Inclusion criteria required (1) parent or caregiver identification of an Autism Spectrum diagnosis, (2) novice driving status, and (3) normal or corrected-to-normal vision sufficient for eye-tracking. Exclusion criteria included any medical, neurological, or visual condition that would interfere with safe participation in simulated driving tasks.

All procedures were reviewed and approved by the University of Central Florida Institutional Review Board (IRB). Written informed consent was obtained from participants and/or their legal guardians prior to participation. Participants were assigned anonymized subject identifiers (Subject01–Subject05) for all data collection, analysis, and reporting to ensure confidentiality.

Although the sample size was necessarily small due to the intensive, multi-phase nature of the experimental protocol, the use of a single-subject experimental design enabled detailed within-participant analysis across baseline, treatment, and withdrawal phases and is consistent with methodological standards for intervention research in specialized populations.

4.2. Setting and Apparatus

4.2.1. Study Setting

Study sessions were conducted in one of two locations selected for parental convenience. Four participants completed sessions in a conference room located at the Florida Interactive Entertainment Academy (FIEA) facility on the University of Central Florida's downtown Orlando campus. The fifth participant completed sessions at the Toni Jennings Exceptional Education Institute (TJEEI) on UCF's main campus. Both rooms were comparable in size, lighting, and seating configuration to ensure consistency across testing environments.

All sessions took place during standard business hours. Each room contained a conference table and seating for the participant, researcher, and a parent or guardian, all of whom remained present throughout each session. The participant was seated approximately 24 inches from the display monitor, with the physical game controller mounted on the table between the participant and the screen. Overhead fluorescent lighting was used for all sessions.

Each participant completed three study sessions on separate days (baseline, treatment, and withdrawal), with each session lasting approximately 60–90 min.

4.2.2. Apparatus Overview

The experimental apparatus consisted of a desktop-based driving simulation system integrating real-time gameplay, eye-tracking, physiological monitoring, and video recording. The SpeedLimit game is a Unity-based driving simulation originally developed by a student team in the FIEA GameLab program and adapted for research use in this study. The game employs an infinite-runner style forward-motion roadway environment in which players identify and respond to scripted roadway hazards through visual scanning and manual input.

Eye position was recorded continuously using a Tobii EyeX eye-tracking device operating at approximately 20 Hz. The system logged time-stamped gaze coordinates and continuously compared gaze location to the on-screen position of each hazard. When gaze was directed at a hazard within the target tolerance window during the relevant time interval, the event was classified as a successful detection. For each trial, the game automatically exported a comma-separated value (CSV) file containing time-stamped hazard events, continuous gaze location data, and summary performance metrics.

Physiological data were collected using a Polar Verity Sense purchased from Amazon.com optical heart rate sensor worn on the participant's inner forearm and sampled at 1 Hz. Heart rate data were recorded using the Polar Sense mobile application downloaded from the Apple's App Store and later synchronized with game event data during post-processing.

Full gameplay sessions were recorded using OBS Studio version 31.0.2 to capture continuous visual output and to assist with time-alignment verification across systems.

4.2.3. Hardware Configuration

A standardized hardware configuration was used across all sessions. The primary computing platform was a Dell G-Series laptop (Intel i7 processor, 32 GB RAM, Windows 10) purchased by UCF from Dell Direct, which hosted the SpeedLimit game, Tobii Eye-Tracking Core Software, version 2.16.8 and OBS Studio. Visual output was presented on a 24-inch Dell external monitor positioned at a fixed distance to preserve stable gaze geometry.

Participants interacted with the simulation using a Logitech G25 steering wheel from Amazon.com with paddle shifters mounted securely to the table. Although steering input was not used as a behavioral outcome measure, the controller provided a realistic tactile posture anchor during gameplay. Seating consisted of a standard conference-room chair positioned to maintain consistent viewing geometry. A parent or guardian was seated nearby throughout each session.

4.2.4. Data Outputs

Multiple synchronized data streams were generated during each trial, including:

- Game logs (CSV): Time-stamped hazard events, continuous gaze coordinates, and end-of-trial performance summaries.
- Heart rate logs (CSV): One-hertz beats-per-minute values downloaded from the Polar cloud platform.
- Session video recordings (MP4): Continuous screen recordings captured via OBS Studio.
- Survey responses: Paper-based Driver Attitude Survey forms completed pre-baseline and post-withdrawal by both participants and parents.

Clock synchronization across systems was verified daily. A verbal trial-start cue recorded into the gameplay video stream provided an additional temporal alignment reference to ensure accurate merging of gaze, heart rate, and game-event data during later processing.

4.3. Experimental Design

The study employed a single-case, within-subject reversal design with an increasing trials-to-criterion structure to evaluate whether repeated gameplay with progressively higher difficulty improves participants' hazard-detection and visual-scanning performance. This design was selected to support fine-grained observation of changes in individual behavior over time and is well suited for intervention research in specialized populations.

Each participant progressed through three primary phases: Baseline, Treatment, and Withdrawal.

During the Baseline phase, participants played the SpeedLimit game with no visible scoring and no performance feedback. Although hazards were presented in the same manner as later phases, all detections were logged silently in the background. This phase established each participant's initial hazard-detection and gaze-allocation patterns in the absence of reinforcement.

During the Treatment phase, real-time performance feedback and a running score were activated to reinforce early and accurate hazard detection. Feedback was presented categorically as "Perfect," "Great," "Good," or "Missed," based on hazard response accuracy and timing. Treatment was administered across three sequential difficulty levels, each characterized by increased driving speed and greater hazard density. This escalation provided systematically increasing perceptual and cognitive challenge across trials.

During the Withdrawal phase, visible scoring and feedback were again removed, returning the game to a non-reinforced condition. This phase was included to assess whether improvements in hazard detection and gaze behavior were maintained in the absence of external reinforcement.

This within-participant design allowed direct comparison of performance across phases while emphasizing observed behavioral change rather than group-level averages. The reversal structure provides evidence regarding both acquisition and retention of trained behavior within each participant.

4.4. Procedures

Each participant completed three study sessions scheduled on separate days corresponding to the baseline, treatment, and withdrawal phases. Each session lasted approximately 60–90 min.

4.4.1. Session 1 (Baseline)

Session 1 included informed consent, pre-study surveys, system calibration, and baseline gameplay trials. Upon arrival, the researcher reviewed the informed consent document with the parent or guardian and then with the participant, answering all questions prior to initiating the study. Written consent was obtained from the parent or guardian, and participant assent was obtained before any study activities began.

Following consent, both the participant and the parent or guardian completed the Driver Attitude Survey. Surveys were labeled using anonymized participant codes and stored securely for later digitization and analysis.

Participants then received a brief tutorial on the SpeedLimit game. Because the game employed a familiar infinite-runner-style mechanic, only minimal training was required. The Tobii EyeX eye tracker was calibrated using Tobii Eye-Tracking Core Software, version 2.16.8, in which the participant fixates on a series of visual targets presented on-screen. Calibration settings were stored and reused for subsequent sessions.

After eye-tracking calibration, the Polar Verity Sense optical heart rate monitor was secured to the participant's inner forearm and paired via Bluetooth to the Polar Sense mobile application. A one-minute resting heart rate recording was collected prior to gameplay.

Baseline trials were then conducted with no visible feedback or scoring displayed to the participant. All hazard detections, gaze data, and physiological measurements were recorded silently in the background.

4.4.2. Session 2 (Treatment)

Session 2 consisted of all treatment trials across three escalating difficulty levels. At the beginning of the session, participants were fitted with the Polar Verity Sense device and loaded with their saved Tobii eye-tracking calibration settings. Resting heart rate was again recorded prior to gameplay.

Visible performance feedback and scoring were enabled throughout the treatment session. Each hazard was scored as "Perfect," "Great," "Good," or "Missed" based on the timing and accuracy of gaze-directed detection. Driving speed and hazard density increased systematically across the three treatment levels to provide progressively greater perceptual and cognitive challenge.

4.4.3. Session 3 (Withdrawal)

Session 3 consisted entirely of withdrawal trials. As in prior sessions, participants were fitted with the heart rate sensor, eye-tracking calibration was loaded, and a one-minute resting heart rate was recorded. Visible feedback and scoring were disabled during this session to evaluate maintenance of learned behavior in the absence of reinforcement.

At the conclusion of Session 3, both participants and parents or guardians completed the post-study Driver Attitude Survey.

4.4.4. Trial Initiation and Temporal Synchronization

Each trial required near-simultaneous initiation of three processes: gameplay, physiological recording, and video capture. To ensure temporal alignment, continuous screen recording was initiated using OBS Studio prior to the start of trials. At the beginning of each trial, the researcher verbally announced the phase and trial number (e.g., "Treatment 2, Trial 3, starting now"). On the spoken cue "now," the researcher-initiated gameplay and immediately started the heart rate recording timer.

Initial system testing confirmed that timestamps generated by the game engine and the Polar heart rate system were almost perfectly aligned. The verbal cue captured within the video stream provided an additional synchronization reference during later data processing.

4.4.5. Post-Session Handling

Upon completion of each study session, all game output files, heart rate recordings, session videos, and survey materials were securely transferred to University of Central Florida cloud storage and organized within participant-specific directories. Raw output files were removed from local device storage once verified as successfully uploaded. Paper-based surveys and consent documents were scanned and stored in secure access-controlled directories.

At the conclusion of the final session, participants were thanked for their participation and parents were provided with instructions to redeem a $25 Amazon gift card as compensation.

4.5. Measures

Four primary outcome measures were used to evaluate the effects of game-based training across study phases: hazard detection accuracy, gaze location, heart rate, and driver attitude.

4.5.1. Measures of Hazard Detection Accuracy

Hazard detection accuracy was defined as whether the participant successfully identified an oncoming roadway hazard within the timing threshold dictated by the current game speed. Each scripted hazard event within SpeedLimit generated a detection window during which gaze was required to intersect the on-screen hazard position.

A hazard was classified as a successful detection when the participant's gaze intersected the hazard location within the tolerance window during the defined response interval. Detection events were time-stamped and logged automatically by the game engine for every trial. This measure served as the primary outcome variable for Hypothesis 1, evaluating whether participants demonstrated improved hazard detection during treatment and withdrawal relative to baseline.

4.5.2. Measures of Gaze Offset and Gaze Bias

Eye position was recorded continuously at approximately 20 Hz as a two-dimensional (x, y) coordinate pair. Gaze data were collected relative to the fixed screen coordinate system, with (0, 0) representing the lower-left corner of the game display and the horizon center of gameplay corresponding to approximately (840, 525).

Gaze data served two purposes:

- Verification of hazard detection—ensuring that hazard detections were associated with true visual fixation rather than arbitrary input.

- Quantification of gaze behavior change—evaluating how average gaze location shifted across phases and trials.

Two complementary gaze metrics were derived:

- Gaze offset—the mean Euclidean distance of gaze coordinates from the central roadway hazard origin point, representing the magnitude of deviation from centered visual scanning.
- Gaze bias—the directional tendency of gaze deviations relative to screen center (e.g., systematic high-right or low-left bias).

These measures served as the primary outcomes for Hypothesis 2, evaluating whether gaze became more efficiently organized with gameplay exposure.

Gaze offset and gaze bias were selected as primary measures because they directly quantify the spatial organization of visual attention relative to the roadway hazard region. These metrics operationalize gaze organization in terms of both magnitude and directional distribution of visual attention, which are critical for hazard anticipation in dynamic driving environments. While alternative eye movement measures such as fixation duration and saccade rate provide valuable information about temporal characteristics of visual scanning, they do not directly capture the spatial alignment of gaze with hazard-relevant regions. Because the present study focused on changes in attentional allocation to hazard emergence zones, spatial measures of gaze organization provided the most direct and interpretable indicators of perceptual–cognitive adaptation across training phases.

4.5.3. Measures of Heart Rate

Physiological arousal was measured using heart rate recordings sampled at 1 Hz via the Polar Verity Sense optical heart rate monitor worn on the participant's inner forearm. At the start of each session, approximately one minute of resting heart rate was recorded prior to gameplay to establish a baseline reference.

Heart rate data were used to evaluate changes in physiological stress across gameplay exposure, serving as the primary outcome for Hypothesis 3. Reduced heart rate across later phases relative to baseline was interpreted as evidence of decreasing physiological workload associated with improved visual scanning efficiency.

Because baseline physiological levels vary substantially across individuals, heart rate analysis focused on within-participant changes relative to each participant's own baseline phase rather than direct cross-participant comparison. This within-subject reference framework provides an implicit normalization of physiological response and is consistent with single-case experimental design methodology, allowing physiological adaptation to be interpreted at the individual level.

The Polar Verity Sense device has been validated against ECG-based reference instrumentation (Polar H10 chest strap), demonstrating near-perfect agreement for beat-to-beat intervals, with correlation coefficients exceeding 0.95 during light-to-vigorous activity [28]. This provides strong evidence of measurement reliability for the present application.

4.5.4. Measures of Driver Attitude

Driver attitudes were measured using the Driver Attitude Survey, administered to both participants and their parents or caregivers at the beginning of baseline and at the conclusion of the withdrawal phase. The survey captures both positive and negative emotional perceptions of driving across three contexts:

- Talking about driving
- Getting ready to drive
- Sitting behind the wheel

This measure served as the primary outcome for Hypothesis 4, evaluating whether gameplay exposure was associated with increased positive driving attitudes and reduced anxiety.

The Driver Attitude Survey demonstrates strong internal consistency and sensitivity to pre–post changes in attitudes toward driving among adolescents with autism and their parents, with reported Cronbach's α values exceeding 0.80 [29]. This provides evidence of adequate reliability for detecting attitudinal change in the present study.

4.6. Data Analysis

Data analysis focused on evaluating within-participant change across experimental phases using a combination of visual inspection, descriptive statistics, and non-overlap-based effect size estimation. This multi-pronged approach is consistent with best practices for single-case experimental designs, where emphasis is placed on pattern detection, level change, trend, and consistency rather than population-level inference.

4.6.1. Hazard Detection Accuracy

Hazard detection accuracy was analyzed at the trial level for each participant across baseline, treatment, and withdrawal phases. For each trial, the proportion of successfully detected hazards was computed from the automatically generated game log files.

Primary evaluation relied on:

- Visual analysis of phase-level accuracy trajectories
- Across-phase level comparisons
- Consistency of change across participants

To supplement visual inspection, non-overlap-based effect size estimates were computed to quantify the magnitude of change between baseline and subsequent phases. These effect estimates served as a descriptive index of intervention impact rather than as a formal inferential test.

4.6.2. Gaze Offset and Gaze Bias

Continuous gaze location data were analyzed using custom R-based processing scripts developed in RStudio v2025.09.1. For each trial, gaze coordinates were transformed into two derived measures:

- Gaze offset: the mean Euclidean distance of gaze from the central roadway hazard region
- Gaze bias: the directional distribution of gaze deviations relative to screen center

For each participant, gaze offset and bias were summarized across trials within each experimental phase. Phase-to-phase changes were evaluated through:

- Visual inspection of trial-wise distributions
- Phase-wise boxplot comparisons
- Heatmap-based visualization of gaze density patterns

These analyses were used to assess changes in both the magnitude and directional organization of visual scanning behavior with gameplay exposure.

4.6.3. Heart Rate

Heart rate data were processed at a 1 Hz resolution and synchronized with corresponding gameplay trials using timestamp alignment. For each trial, mean heart rate was computed after excluding the pre-trial resting baseline. Trial-level values were then grouped by experimental phase.

Physiological change was evaluated through:

- Visual inspection of heart rate time series

- Phase-level distribution comparisons
- Within-participant trend assessment across sessions

This analysis strategy allowed assessment of whether physiological arousal decreased as participants became more efficient in hazard detection and visual scanning.

4.6.4. Driver Attitude

Driver Attitude Survey responses were scored according to published scoring guidelines. Pre-baseline and post-withdrawal survey scores were compared descriptively at the individual participant level for both parent- and self-report forms. Because of the small sample size and the single-case design structure, attitudinal outcomes were interpreted using:

- Direction of change
- Consistency across respondents
- Convergence between parent and participant ratings

Rather than formal inferential testing.

4.6.5. Visualization Strategy

Given the central role of visual analysis in single-case methodology, multiple complementary visualization formats were generated, including:

- Trial-wise phase plots of hazard detection accuracy
- Box-and-whisker plots comparing phase-level performance
- Two-dimensional gaze heatmaps with overlaid mean and dispersion ellipses
- Time-aligned multi-modal plots combining gaze, hazard events, and heart rate

All plots were generated using custom R scripts using RStudio version 2025.09.1 to ensure reproducibility and consistent scaling across participants.

5. Results

5.1. Hazard Detection Accuracy

Hazard detection accuracy was evaluated at the trial level for each participant across baseline, treatment, and withdrawal phases as shown in Figure 2. For each trial, accuracy was computed as the proportion of successfully detected hazards relative to the total number of hazard events presented. Performance trajectories for all five participants are shown in [24,25]. Phase-level patterns were evaluated using within-participant visual analysis with attention to level, trend, and consistency of change.

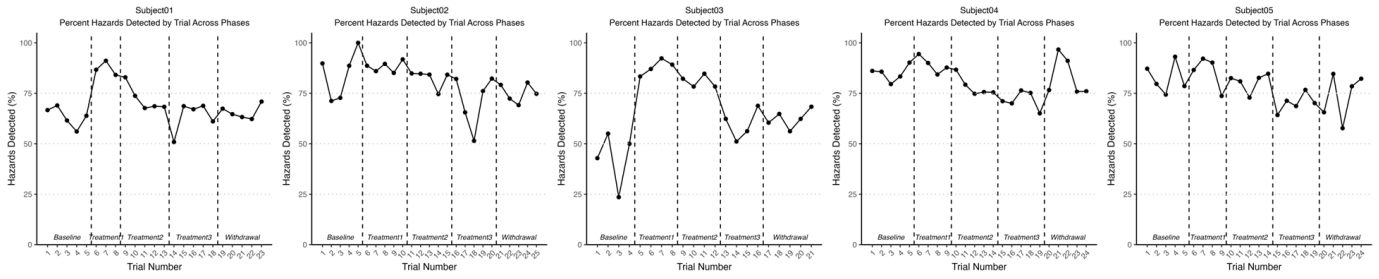


Figure 2. Trial-by-trial hazard detection accuracy for each of the five participants across baseline, treatment, and withdrawal phases. Each panel represents one participant, with hazard detection accuracy plotted as a function of trial sequence. Vertical phase transitions boundaries indicate changes in experimental condition. These trajectories illustrate substantial inter-individual variability in baseline performance, treatment response, and retention during withdrawal, consistent with the individualized learning patterns characteristic of single-case experimental designs.

5.1.1. Baseline Performance

Baseline performance varied substantially across participants. Subject01, 02, 04, and 05 entered baseline with relatively higher hazard detection accuracy, generally ranging between approximately 60–90% across trials, though with notable trial-to-trial variability. In contrast, Subject03 exhibited lower and more unstable baseline performance, with multiple trials falling below 50% detection, establishing a markedly weaker initial performance level.

This heterogeneity in baseline performance underscores the importance of within-subject, phase-based evaluation rather than reliance on group-averaged comparisons.

5.1.2. Treatment Effects

Treatment-related effects differed sharply across participants.

Subject03 demonstrated a clear and sustained improvement in hazard detection accuracy following the introduction of treatment. Accuracy increased rapidly from low and unstable baseline levels to consistently moderate and high levels during Treatment 1 and remained elevated across Treatments 2 and most of 3. This pattern reflects a robust training-related acquisition effect, with accuracy stabilizing well above baseline variability.

In contrast, Subject01, 02, 04, and 05 did not exhibit treatment effects of comparable magnitude or stability. While some participants showed modest or transient improvements during portions of the treatment phases, these changes were smaller in magnitude, less phase-linked, and less consistently sustained than the acquisition pattern observed in Subject03.

5.1.3. Withdrawal Phase

During the withdrawal phase, Subject03 maintained accuracy levels above baseline, indicating partial retention of treatment-related gains following removal of performance feedback.

For the remaining four participants, withdrawal performance either remained similar to late-treatment levels or showed continued modest decline, with no evidence of post-treatment rebound above baseline. These patterns suggest that, for most participants, treatment exposure did not result in clearly durable enhancements in hazard detection accuracy beyond the withdrawal phase.

5.1.4. Individual Variability

Considerable inter-individual variability was observed in both baseline performance and treatment response. Subject03 demonstrated the most pronounced and sustained improvement in hazard detection accuracy following treatment introduction, whereas the remaining participants exhibited more modest, variable, or unstable performance patterns across phases despite equivalent exposure to the simulation environment and feedback structure. This divergence highlights the heterogeneous nature of hazard perception learning within novice drivers with autism and reinforces the value of a single-case analytic framework for resolving individual change processes that would be obscured under group averaging.

5.1.5. Summary of Accuracy Outcomes

In summary:

- One participant (Subject03) exhibited the most pronounced and sustained improvement in hazard detection accuracy following treatment introduction.
- The remaining participants (Subject01, 02, 04, and 05) demonstrated more modest, variable, or unstable performance patterns during treatment, with no participant showing a comparably large and phase-linked acquisition effect.

- During withdrawal, accuracy for Subject03 remained above baseline levels, whereas performance for the other participants either stabilized near late-treatment levels or continued modest decline, with no evidence of rebound above baseline.

Collectively, these findings indicate that, within this cohort, game-based hazard exposure was associated with highly variable learning outcomes. While one participant demonstrated the largest observed and most stable acquisition pattern, most participants showed limited or inconsistent improvements in hazard detection accuracy.

5.2. Gaze Offset and Bias

Gaze behavior was evaluated using two complementary metrics: gaze bias, defined as the directional displacement of gaze relative to the roadway centerline (signed), and gaze offset, defined as the magnitude of deviation from center regardless of direction (unsigned). Together, these measures characterize both the directional tendency and the spatial dispersion of visual attention during dynamic driving scenes.

Trial-by-trial gaze bias and gaze offset for all five participants across baseline, treatment, and withdrawal phases are shown in Figure 3. Phase-level trends were evaluated using within-participant visual analysis, with complementary summary representations provided by box plots, heatmaps, and Tau-U effect size estimates.

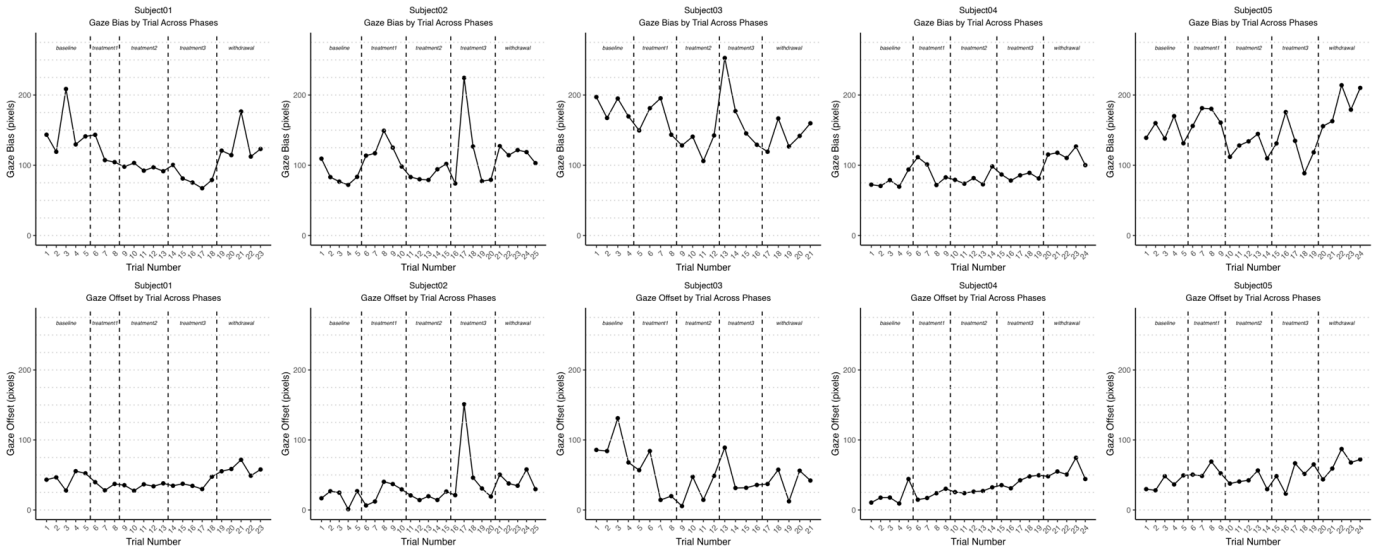


Figure 3. Trial-by-trial gaze organization measures for each of the five participants across baseline, treatment, and withdrawal phases. The top row shows gaze bias, representing directional displacement of gaze relative to the roadway centerline, while the bottom row shows gaze offset, representing the magnitude of gaze deviation regardless of direction. Each column corresponds to one participant, with values plotted across sequential trials. Vertical phase transitions boundaries indicate changes in experimental condition. Lower gaze offset and reduced bias magnitude reflect more centered and stable visual attention, consistent with more organized road surveillance.

5.2.1. Baseline Gaze Characteristics

Baseline gaze characteristics varied substantially across participants, with distinct patterns observed for gaze bias and gaze offset. Subject01, 03, and 05 demonstrated relatively elevated gaze bias magnitudes during baseline, indicating directional gaze tendencies, whereas Subject02 and 04 exhibited lower directional bias and less pronounced baseline asymmetry.

Baseline gaze offset magnitudes were broadly similar across Subject01, 02, 04, and 05, while Subject03 exhibited comparatively larger initial dispersion. Together, these baseline differences indicate that participants entered the study with distinct patterns of visual sampling that were not uniform across the cohort.

Importantly, baseline gaze organization did not map cleanly onto baseline hazard detection accuracy, reinforcing that similar behavioral detection rates may arise from different attentional strategies.

5.2.2. Gaze Changes During Treatment

Across treatment phases, changes in gaze behavior were observed for several participants, though patterns differed substantially by individual and by metric (gaze bias values versus gaze offset).

- Subject01 showed a progressive reduction in gaze bias values across treatment, reaching a local minimum during Treatment 3, followed by partial rebound during withdrawal. Gaze offset for this participant remained relatively stable at low magnitudes, with a modest upward drift late in the protocol.
- Subject02 exhibited large but transient spikes in both gaze bias values and gaze offset during Treatment 3, driven by a single trial in which both metrics increased concurrently. Outside of this excursion, gaze measures showed partial stabilization during withdrawal, consistent with episodic spatial variability rather than a sustained shift in gaze organization.
- Subject03 demonstrated an overall reduction in gaze offset from baseline through treatment, with offset magnitudes becoming smaller and more stable after Treatment 1. Gaze bias values for this participant were also reduced relative to baseline, remaining below baseline levels across treatment phases with a single bounded excursion that did not exceed the baseline range. Unlike offset, bias values did not show a monotonic decline but exhibited constrained variability following treatment onset.
- Subject04 showed gradual increases in gaze offset across treatment phases, with small but systematic growth from baseline through withdrawal. Gaze bias values remained comparatively low throughout the protocol, though a slow upward drift was observed late in treatment.
- Subject05 exhibited increasing gaze offset across later treatment phases and withdrawal, accompanied by elevated and variable gaze bias values during Treatment 3 and withdrawal. These patterns are consistent with increased visual dispersion rather than progressive tightening of attentional organization.

5.2.3. Relationship Between Gaze and Accuracy

Subject03, who exhibited the most pronounced and sustained improvement in hazard detection accuracy (Section 5.1), also demonstrated a sustained reduction in gaze offset magnitude during treatment, with offset values becoming smaller and more stable across phases. This convergence represents the clearest observed alignment between improvements in detection performance and changes in visual sampling behavior within the cohort.

For the remaining participants, changes in gaze behavior did not consistently align with accuracy outcomes. Several participants showed gaze broadening or increased variability despite stable or declining detection performance. These dissociations indicate that gaze reorganization and behavioral performance did not evolve in a uniform or tightly coupled manner across individuals.

5.2.4. Withdrawal Phase

During withdrawal, maintenance of treatment-related changes in gaze organization was limited and participant-specific. Subject03 retained relatively low gaze offset magnitudes compared with baseline and exhibited gaze bias values that remained at or below baseline levels throughout withdrawal. In contrast, other participants either shifted toward baseline offset distributions (e.g., Subject01) or continued patterns of increased dispersion observed during late treatment (e.g., Subject04 and 05). Across the cohort, gaze bias values during withdrawal remained variable, with no consistent pattern of maintenance evident beyond Subject03.

5.2.5. Summary of Gaze Outcomes

In summary:

- The most pronounced and sustained reduction in gaze offset magnitude was observed for Subject03, coinciding with that participant's improvement in hazard detection accuracy.
- Other participants exhibited heterogeneous gaze responses during treatment, including transient stabilization, progressive broadening, and episodic spikes in both gaze bias values and offset.
- Changes in gaze bias values and gaze offset were partially dissociated, indicating that directional tendencies and dispersion magnitude capture distinct aspects of visual attention organization.
- Across participants, changes in gaze behavior did not consistently align with changes in detection accuracy, indicating that gaze measures provided complementary rather than redundant information about training effects.

Taken together, these results indicate that game-based hazard exposure was associated with highly individualized patterns of visual attention change, with the clearest alignment between gaze organization and behavioral learning observed for Subject03.

5.3. Physiological Engagement (Heart Rate)

Physiological engagement during gameplay was assessed using heart rate (HR) as an index of cognitive and emotional workload across repeated trials. Heart rate data were synchronized with gameplay events and analyzed across baseline, treatment, and withdrawal phases for each participant.

Figure 4 shows average heart rate by trial, phase, and participant. Across participants, heart rate responses exhibited substantial variability both within and between experimental phases. Some participants showed elevated HR during early trials that attenuated over repeated exposure, while others demonstrated relatively stable or inconsistent HR patterns across sessions. Notably, changes in heart rate did not consistently parallel changes in hazard detection performance or gaze-based measures.

At the individual level, reductions in mean heart rate across trials were observed for some participants during treatment phases; however, these reductions were not systematically associated with improved hazard detection accuracy. In several cases, participants who exhibited decreased physiological arousal across trials showed stable or declining performance, while participants with relatively elevated or variable HR demonstrated modest performance gains. One participant (Subject03), however, exhibited a canonical pattern in which reductions in mean heart rate across treatment phases co-occurred with improvements in hazard detection performance, consistent with a more efficient physiological response during task performance.

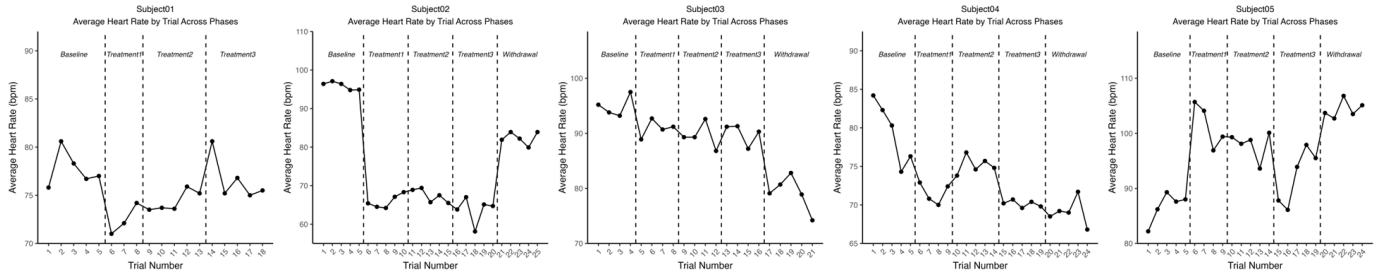


Figure 4. Trial-level average heart rate for each of the five participants across baseline, treatment, and withdrawal phases. Each panel represents one participant, with mean heart rate (beats per minute) plotted for sequential trials. Dashed vertical lines indicate phase transitions between experimental conditions. Heart rate provides an index of physiological workload and arousal during task performance, enabling examination of how physiological response changed across training and withdrawal phases on a within-participant basis.

Heart rate variability across trials also differed by participant, suggesting individualized physiological responses to task demands. These patterns indicate that physiological engagement during the simulation was sensitive to repeated exposure and task demands, but not uniformly predictive of behavioral performance outcomes.

Taken together, the heart rate results suggest that physiological adaptation occurred for some participants during repeated gameplay, but that changes in arousal alone.

5.4. Driving Attitudes and Survey Responses

Driving attitudes were assessed using participant self-report and caregiver-reported survey instruments administered at baseline and withdrawal. These measures were intended to capture global perceptions of driving confidence, comfort, and concern rather than trial-level learning dynamics.

Participant self-reports showed minimal change from baseline to withdrawal across both positive and negative attitude ratings Figure 5. Across participants, summed positive and negative scores remained largely stable, indicating that repeated exposure to the simulation did not substantially alter participants' self-perceived driving attitudes within the study timeframe.

In contrast, caregiver reports exhibited more pronounced change. Caregivers reported a reduction in negative driving-related attitudes at withdrawal relative to baseline, accompanied by modest increases in positive ratings. This pattern suggests that caregivers perceived improvements in participants' driving readiness or comfort following study participation, even when participant self-reports remained unchanged.

Importantly, changes in caregiver-reported attitudes did not consistently align with objective performance measures. Participants whose caregivers reported more favorable post-study attitudes did not uniformly demonstrate improvements in hazard detection accuracy, gaze organization, or physiological response. Conversely, participants who showed objective performance gains did not consistently report increased confidence or reduced concern. These dissociations indicate that subjective perceptions of driving readiness captured by survey instruments reflected individualized interpretations of experience rather than direct indices of learning or skill acquisition.

Taken together, the survey results highlight divergence between participant and caregiver perspectives and underscore the complexity of interpreting attitudinal change in the context of simulation-based driver training. Subjective perceptions of progress and readiness evolved differently from objective behavioral, visual, and physiological measures, motivating an integrated examination of individual variability across outcome domains.

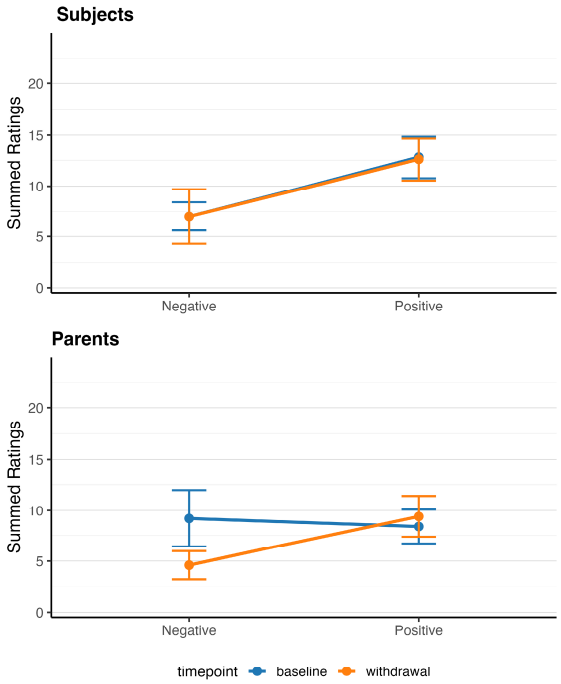


Figure 5. Participant- and caregiver-reported driving attitude ratings at baseline and withdrawal phases. The top panel shows participant self-report scores, and the bottom panel shows caregiver-reported scores. Positive and negative attitude subscales are plotted separately, with summed scores reflecting overall driving-related confidence and concern. These measures provide a subjective complement to behavioral, gaze, and physiological metrics, enabling assessment of perceived changes in driving readiness and confidence following simulation-based training.

5.5. Individual Variability

Considerable individual variability was observed across all outcome measures, including hazard detection accuracy, gaze behavior, physiological response, and self- and caregiver-reported driving attitudes. Although all participants completed the same experimental protocol, patterns of change across experimental phases differed substantially between individuals, with no single trajectory characterizing the cohort.

One participant (Subject03) demonstrated the most pronounced improvement in hazard detection performance across treatment phases, accompanied by corresponding changes in gaze organization and modest reductions in physiological arousal. Other participants exhibited stable performance, fluctuating patterns, or gradual declines in accuracy despite repeated exposure to the simulation environment. In several cases, reductions in physiological arousal occurred without concurrent improvements in detection performance or gaze organization, while some participants showed modest behavioral gains in the presence of elevated or variable heart rate.

Gaze behavior likewise followed participant-specific trajectories. Some individuals exhibited reductions in gaze offset magnitude or increased stability across trials, whereas others demonstrated minimal change, progressive broadening, or episodic variability. Changes in gaze bias values and gaze offset were partially dissociated, and gaze reorganization did not consistently align with improvements in hazard detection accuracy.

Survey responses further reflected this heterogeneity. Participant self-reports showed minimal attitudinal change across phases, whereas caregiver reports indicated reduced negative perceptions and modest increases in positive attitudes following study participation. These subjective shifts did not systematically correspond to objective behavioral, visual, or physiological outcomes, reinforcing the dissociation between perceived driving readiness and measured performance.

Collectively, these findings highlight substantial heterogeneity in learning trajectories, attentional organization, physiological engagement, and subjective experience among novice drivers with autism. The single-case, multi-phase design was essential for capturing these individualized patterns, which would have been obscured under group-level or aggregate analyses.

6. Discussion

This section interprets and integrates the findings reported in Section 5, with the goal of understanding how repeated exposure to a hazard-rich driving simulation influenced behavioral performance, visual scanning behavior, and physiological responses among novice drivers with autism. Importantly, treatment phases were designed to increase systematically in difficulty, such that declines in raw performance across successive treatment phases were anticipated and reflect escalating task demands rather than training failure. Rather than evaluating outcomes in isolation, the discussion considers how these modalities relate to one another over time and across experimental phases. Particular attention is given to individual learning trajectories, as the results revealed substantial heterogeneity in how participants responded to training across behavioral, visual, and physiological dimensions.

The discussion is organized to first summarize key findings by hypothesis and outcome domain, followed by a deeper examination of cross-modal relationships and dissociations observed at the individual level. Consistent with the single-case, multi-phase design, emphasis is placed on within-participant patterns rather than group-level averages, as aggregate trends often obscured meaningful individual differences. This structure allows consideration of why some participants demonstrated coordinated changes across measures, while others showed partial, delayed, or divergent responses, and provides a foundation for interpreting the implications of these patterns for training design and future research.

6.1. Summary of Findings

This study examined the effects of a game-based, hazard-focused driving intervention on behavioral performance, visual scanning behavior, and physiological response in novice drivers with autism using a single-case, multi-phase design. Across outcome domains, results revealed substantial individual variability in both the direction and magnitude of change, with no single pattern characterizing all participants. Changes in one modality did not consistently co-occur with changes in others, underscoring the importance of within-participant analysis when evaluating training effects in this population.

With respect to hazard detection accuracy (H1), no participant demonstrated monotonic improvement across successive treatment phases, as raw accuracy declined from Treatment 1 to Treatment 3 for all participants. These declines were expected given the systematic increase in scenario difficulty across treatment phases. Despite this pattern, one participant (Subject03) exhibited higher detection accuracy during the withdrawal phase than during baseline, suggesting a delayed or post-treatment performance benefit rather than incremental gains during treatment itself. Overall, these findings indicate that repeated exposure to hazard-rich scenarios did not produce uniform improvements in raw detection accuracy during treatment, and that learning effects, when present, were expressed in individualized and temporally displaced ways.

Regarding visual scanning behavior (H2), changes in gaze offset and gaze bias were similarly heterogeneous. Some participants showed reductions in gaze offset or shifts toward more centralized scanning patterns across treatment, whereas others demonstrated minimal or inconsistent change. Improvements in gaze-based measures did not consistently

align with changes in hazard detection accuracy, indicating a dissociation between visual attentional organization and overt behavioral performance at the individual level.

For physiological response (H3), trial-level and phase-level analyses showed that reductions in average heart rate during treatment occurred for some participants but were not reliably associated with improved hazard detection or gaze outcomes. One participant exhibited a canonical pattern in which reduced heart rate co-occurred with improved performance, whereas others showed decreased physiological arousal without corresponding behavioral gains or sustained physiological variability despite modest performance improvements. These findings suggest that heart rate indexed individualized engagement or adaptation processes rather than serving as a uniform marker of learning.

Finally, driving attitude measures (H4) demonstrated limited and inconsistent change across participants and phases. While some participants and caregivers reported modest shifts in perceived confidence or comfort with driving, these changes did not systematically track with behavioral, visual, or physiological outcomes. Taken together, the findings highlight the complex and individualized nature of learning, engagement, and adaptation in response to simulation-based hazard training among novice drivers with autism.

These multimodal findings underscore the importance of distinguishing between observable performance gains and the underlying perceptual–cognitive mechanisms that support hazard detection. In several cases, improvements in hazard detection accuracy occurred without corresponding reductions in gaze offset or physiological workload. This dissociation suggests that learning may involve strategic adaptation, improved sensitivity to peripheral visual cues, or cognitive anticipation that does not immediately manifest as more centered gaze or reduced physiological demand. Conversely, reductions in heart rate without accompanying behavioral improvement may reflect familiarization with the simulation environment rather than functional enhancement of hazard perception. Together, these patterns highlight the value of multimodal instrumentation for revealing how attention, performance, and physiological response evolve along partially independent trajectories during training.

6.2. Relation to Prior Work

The behavioral and visual findings indicate that improvements in hazard detection were not uniformly driven by changes in visual scanning organization. While prior hazard perception research has linked more anticipatory and stable visual scanning patterns to improved detection performance under certain conditions [7], the present results demonstrate that this relationship did not manifest consistently under the conditions of this study. Some participants exhibited modest gains in detection accuracy alongside shifts toward more centralized or stable gaze patterns, whereas others showed changes in one domain without corresponding change in the other. This dissociation indicates that visual scanning behavior, as indexed by gaze offset and bias, was neither a necessary nor sufficient condition for improved hazard detection across participants.

For the subset of participants who exhibited coordinated changes in gaze organization and detection accuracy, the observed patterns are consistent with prior work suggesting that efficient allocation of visual attention may support hazard perception [7]. In these cases, reduced gaze dispersion and increased stability may reflect more anticipatory scanning strategies developed through repeated exposure to hazard-rich scenarios. However, as in previous studies involving neurodiverse populations, this pattern was observed in only a minority of participants and did not generalize across individuals [6,30].

In contrast, several participants demonstrated stable or declining detection accuracy despite measurable changes in gaze behavior, indicating that alterations in visual scanning alone did not reliably translate into improved performance. Conversely, modest

performance gains were sometimes observed in the absence of corresponding changes in gaze offset or bias, suggesting that alternative mechanisms such as task familiarity, heuristic learning, or non-visual strategies may have supported hazard detection under increasing task demands. These findings align with prior work emphasizing that multiple perceptual–cognitive pathways can contribute to hazard perception, particularly in complex or progressively challenging environments [7,11].

Consistent with earlier cautions against treating eye-tracking metrics as direct proxies for driving competence, the present findings suggest that gaze measures are best interpreted as indicators of attentional organization rather than as direct predictors of behavioral performance [7]. In novice drivers with autism, this interpretation implies that gaze patterns may reflect individual differences in perceptual strategy, sensory processing, or interaction with the simulation environment rather than progression toward a single optimal scanning pattern.

Overall, the behavioral and visual results refine prior models of hazard perception by demonstrating that the relationship between visual scanning organization and detection performance is conditional, individualized, and sensitive to task structure. Rather than converging toward a uniform learning trajectory, participants exhibited distinct combinations of gaze behavior and detection accuracy over time. This variability reinforces the value of single-case, multi-phase designs and motivates further examination of how physiological engagement may interact with behavioral and visual adaptation, as discussed in the following section.

The withdrawal phase provides additional insight into the persistence and stability of learning following removal of structured feedback. For some participants, hazard detection performance remained elevated relative to baseline despite the absence of performance-contingent reinforcement, suggesting that at least some perceptual–cognitive adaptations generalized beyond the immediate training context. In other cases, performance returned toward baseline levels, indicating that improvements may have depended in part on ongoing feedback or task familiarity. This variability in retention patterns is consistent with prior findings emphasizing individualized learning trajectories in neurodiverse populations and highlights the importance of evaluating both acquisition and persistence when assessing training effectiveness [6,30].

6.3. Design Implications

From a training and design perspective, physiological response, indexed by heart rate, showed substantial variability across participants and experimental phases and did not consistently align with changes in hazard detection accuracy or visual scanning behavior. Although some participants exhibited reductions in average heart rate during treatment phases, these changes were not uniformly associated with improvements in detection performance or gaze organization. Overall, heart rate patterns appeared to reflect individualized engagement or adaptation processes rather than serving as a direct indicator of learning or performance efficiency.

One participant (Subject03) demonstrated a coordinated pattern in which reductions in mean heart rate across treatment phases co-occurred with improvements in hazard detection accuracy and more stable gaze behavior, consistent with increased task familiarity and more efficient behavioral and physiological regulation. However, this pattern was not representative of the broader sample. Other participants showed decreased heart rate without corresponding behavioral or visual gains, indicating that reduced physiological arousal alone was insufficient to infer learning or improved hazard perception.

In contrast, one or more participants exhibited markedly lower average heart rate during treatment sessions compared to both baseline and withdrawal phases, despite no

corresponding advantage in hazard detection performance. This pattern suggests that reduced physiological arousal during training may reflect increased comfort with the simulation environment, habituation to task demands, or reduced engagement rather than improved perceptual efficiency. Such cases illustrate that lower heart rate during training cannot be assumed to index learning or performance gains and underscore the importance of interpreting physiological measures within the broader behavioral and contextual profile of each participant.

Conversely, several participants exhibited elevated or variable heart rate throughout treatment phases despite modest improvements in detection accuracy. In these cases, physiological arousal may have reflected sustained effort, stress, or cognitive demand rather than inefficiency or disengagement. Such patterns highlight the complexity of interpreting heart rate measures in isolation, particularly in autistic populations where autonomic responses may be influenced by sensory sensitivity, task novelty, or individual differences in regulation strategies.

Taken together, these findings indicate that heart rate should not be treated as a standalone proxy for cognitive efficiency or training success in simulation-based hazard learning. Instead, physiological measures appear to index individualized patterns of engagement and adaptation that interact with—but do not deterministically predict—behavioral and visual outcomes. From a design perspective, these results caution against over-reliance on physiological signals as indicators of learning and underscore the need for evaluation frameworks that integrate physiological, behavioral, and visual measures when assessing training effectiveness.

These findings also have implications for the design of future adaptive training and assistive systems. Multimodal instrumentation that integrates behavioral performance, gaze behavior, and physiological response provides a richer and more reliable basis for assessing learner state than any single modality alone. Emerging computational approaches, including Automated Machine Learning techniques, may further enhance the ability to identify meaningful patterns within complex multimodal datasets and support development of intelligent training systems that dynamically adjust difficulty, feedback, and pacing based on individual learning trajectories. Such adaptive systems may be particularly valuable for neurodiverse learners, whose perceptual–cognitive adaptation patterns may differ substantially from neurotypical populations and may benefit from individualized training strategies [31]. Future work may also leverage high-resolution gaze time-series data to identify temporal and spatial attention patterns not captured by summary metrics such as gaze offset and bias, further refining multimodal models of hazard perception learning.

6.4. Training and Clinical Implications

The findings of this study have important implications for the delivery and evaluation of simulation-based driver training for novice drivers with autism. Across behavioral, visual, and physiological measures, participants exhibited highly individualized learning trajectories, with limited convergence toward a single pattern of improvement. These results indicate that training approaches based on uniform progression or aggregate performance benchmarks may fail to capture meaningful changes occurring at the individual level.

The dissociation observed between hazard detection accuracy, gaze behavior, and physiological response suggests that improvements in one domain should not be assumed to reflect broader gains in driving competence. Training systems that rely on a single performance metric such as accuracy scores or physiological thresholds risk mischaracterizing learner progress. Multi-dimensional assessment frameworks that integrate behavioral outcomes with visual and physiological indicators may therefore provide a more accurate representation of learner engagement and adaptation over time.

Variability in response patterns also highlights the need for flexibility in training implementation. While some participants appeared to benefit from repeated exposure to hazard-rich scenarios, others showed limited or inconsistent gains despite continued practice. These findings support the potential value of adaptive training approaches that adjust task difficulty, feedback, or pacing based on individual response patterns rather than relying on fixed-sequence interventions.

Importantly, the results caution against interpreting reductions in physiological arousal as a universal indicator of learning or efficiency. Although decreased heart rate coincided with improved performance for one participant, others exhibited reduced arousal without behavioral gains or sustained arousal despite modest improvements. Training programs that incorporate physiological monitoring should therefore treat such signals as indicators of engagement or regulation rather than definitive markers of learning success.

Finally, the observed heterogeneity underscores the importance of individualized progress monitoring in both research and applied contexts. Single-case, multi-phase approaches provide a framework for identifying meaningful within-participant change that may be obscured by group-level analyses. For clinicians, researchers, and system designers, this perspective supports shifting emphasis from normative performance targets toward learner-specific trajectories, enabling more responsive, interpretable, and inclusive approaches to driver training for individuals with autism.

7. Limitations and Future Work

While the present study provides insight into individualized learning trajectories during simulation-based hazard training, several limitations should be considered when interpreting the findings. These limitations define the scope of the current work and inform directions for future research aimed at refining assessment methods, training design, and real-world applicability.

7.1. Limitations

First, the small sample size and single-case, multi-phase design limit the generalizability of results to broader populations of drivers with autism. While this design enabled detailed examination of within-participant learning trajectories and cross-modal dissociations, the findings should be interpreted as illustrative of individual variability rather than representative of population-level effects.

Second, the study was conducted within a desktop-based driving simulation, which constrains inferences about real-world driving behavior. Although the simulation allowed for controlled exposure to hazard-rich scenarios and precise measurement of behavioral, visual, and physiological responses, transfer of training effects to on-road performance was not assessed. Future work will be needed to determine how observed changes in hazard detection, gaze behavior, and physiological response relate to real-world driving outcomes.

Third, eye-tracking measures were limited by the spatial and temporal resolution of the sensing hardware. While gaze offset and bias provided useful indices of visual attention organization, these measures do not capture all aspects of visual search behavior, such as peripheral awareness or head movements. As a result, changes in gaze metrics should be interpreted as partial indicators of attentional strategy rather than comprehensive representations of visual processing.

Fourth, heart rate was used as a coarse index of physiological response and engagement. Heart rate is influenced by multiple factors, including effort, stress, sensory load, and individual differences in autonomic regulation, particularly in autistic populations [11]. Consequently, heart rate changes cannot be uniquely attributed to learning efficiency or cognitive load.

Finally, physiological data were incomplete for some participants and phases due to sensor connectivity issues and the exclusion of resting segments that could not be reliably matched to trial onset. Although these constraints did not affect the primary behavioral and visual analyses, they limited the completeness of longitudinal physiological comparisons across all participants and phases.

7.2. Future Work

Future research should build on the individualized learning trajectories observed in this study by exploring training approaches that are responsive to learner-specific patterns of performance, visual attention, and physiological engagement. Rather than assuming uniform progression, adaptive simulation-based systems that dynamically adjust task difficulty, pacing, or feedback based on individual response profiles may better support diverse learners and warrant systematic investigation.

Additional work is also needed to refine multimodal assessment strategies for evaluating hazard perception training. Integrating behavioral performance measures with higher-resolution eye-tracking and richer physiological sensing—such as heart rate variability or additional autonomic indicators—may provide more sensitive indices of engagement and adaptation. Logging hazard onset times within the simulation would further enable more precise peri-event analyses, allowing future studies to distinguish between perceptual, decisional, and response-related components of hazard detection.

Longitudinal and transfer-focused studies represent another important direction for future research. Examining whether changes observed during simulation-based training generalize to on-road driving performance, persist over time, or influence real-world safety outcomes would strengthen the applied relevance of these findings. Such studies may also clarify how visual and physiological adaptations interact with experience and environmental complexity beyond the training context.

Finally, future work should consider expanding participant samples to capture a broader range of autistic learner profiles, including variation in sensory sensitivity, attentional style, and prior driving experience. Larger samples would enable more systematic characterization of individual response patterns while preserving the within-participant analytic approach demonstrated here. Together, these directions offer a pathway toward more personalized, effective, and inclusive driver training interventions.

8. Conclusions

This study examined behavioral, visual, and physiological responses to a game-based, hazard-focused driving simulation in novice drivers with autism using a single-case, multiphase design. Across outcome domains, participants demonstrated highly individualized learning trajectories, with limited convergence toward a uniform pattern of improvement. Changes in hazard detection accuracy, gaze behavior, and heart rate did not consistently co-occur, underscoring substantial dissociation across modalities at the individual level.

The findings indicate that improvements in hazard detection performance cannot be inferred from changes in visual scanning organization or physiological response alone. While a canonical pattern of coordinated behavioral, visual, and physiological change was observed in one participant, this pattern did not generalize across the sample. Instead, participants exhibited distinct combinations of engagement, adaptation, and performance over repeated exposure to hazard-rich scenarios.

Collectively, these results highlight the importance of multimodal, within-participant analysis for evaluating simulation-based driver training in autistic populations. Training effectiveness may be better characterized by individualized response patterns than by aggregate performance metrics or single-domain indicators alone. From both research

and applied perspectives, these findings support the development of flexible assessment frameworks and adaptive training approaches that account for diverse learner profiles and pathways toward skill acquisition.

Author Contributions: Conceptualization, E.J.S. and C.E.H.; methodology, E.J.S. and M.T.M.; software, E.J.S. and C.E.H.; validation, E.J.S., M.T.M. and C.E.H.; formal analysis, E.J.S.; investigation, E.J.S.; resources, E.J.S. and C.E.H.; data curation, E.J.S.; writing—original draft preparation, E.J.S.; writing—review and editing, M.T.M. and C.E.H.; visualization, E.J.S.; supervision, M.T.M. and C.E.H.; project administration, M.T.M. and C.E.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The studies involving human participants were reviewed and approved by the University of Central Florida Institutional Review Board. Further, the study was conducted in accordance with institutional guidelines and adhered to the principles of the Declaration of Helsinki and Title 45 of the US Code of Federal Regulations (Part 46, Protection of Human Subjects). The protocol was approved by the Institutional Review Board of University of Central Florida FWA00000351 IRB00001138, IRB00012110. The IRB ID for our project is STUDY00007597. The IRB Date of Approval is 18 March 2025.

Informed Consent Statement: The parents of the participants provided written consent, and the participants provided verbal assent with the option to stop participating at any time.

Data Availability Statement: The datasets generated in this study will not be made public, as the IRB requires explicit permission for access to each individual's data.

Acknowledgments: The authors gratefully acknowledge the original student development team at the Florida Interactive Entertainment Academy (University of Central Florida) for the initial creation of the SpeedLimit simulation platform, and in particular Christian Tuttle, who served as the original project lead. The authors further extend sincere thanks to Joseph DeMartini (Embry-Riddle Aeronautical University) for his invaluable assistance in adapting and extending the SpeedLimit software for experimental use in the present study. His technical expertise, responsiveness, and collaborative spirit were instrumental to the successful execution of this research. Article processing charges were provided in part by the UCF College of Graduate Studies Open Access Publishing Fund. Generative artificial intelligence tools were used to assist with manuscript drafting and editing, including refinement of text clarity and organization. No generative AI tools were used for data generation, data analysis, or figure creation. All interpretations and conclusions were determined by the authors.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Teen Drivers and Passengers: Get the Facts | Transportation Safety | Injury Center | CDC. Available online: https://www.cdc.gov/teen-drivers/about/ (accessed on 13 June 2023).
2. Curry, A.E.; Hafetz, J.; Kallan, M.J.; Winston, F.K.; Durbin, D.R. Prevalence of teen driver errors leading to serious motor vehicle crashes. *Accid. Anal. Prev.* **2011**, *43*, 1285–1290. [CrossRef]
3. McKnight, A.J.; McKnight, A.S. Young novice drivers: Careless or clueless? *Accid. Anal. Prev.* **2003**, *35*, 921–925. [CrossRef]
4. Bednarz, H.M.; Kana, R.K.; Svancara, A.M.; Sherrod, G.M.; Stavrinos, D. Neuropsychological predictors of driving hazard detection in autism spectrum disorder and ADHD. *Child Neuropsychol. A J. Norm. Abnorm. Dev. Child. Adolesc.* **2021**, *27*, 857–887. [CrossRef]
5. Cox, S.M.; Cox, D.J.; Kofler, M.J.; Moncrief, M.A.; Johnson, R.J.; Lambert, A.E.; Cain, S.A.; Reeve, R.E. Driving Simulator Performance in Novice Drivers with Autism Spectrum Disorder: The Role of Executive Functions and Basic Motor Skills. *J. Autism Dev. Disord.* **2016**, *46*, 1379–1391. [CrossRef]
6. Sheppard, E.; Ropar, D.; Underwood, G.; van Loon, E. Brief report: Driving hazard perception in autism. *J. Autism Dev. Disord.* **2010**, *40*, 504–508. [CrossRef]

7. Underwood, G. Visual attention and the transition from novice to advanced driver. *Ergonomics* **2007**, *50*, 1235–1249. [CrossRef]
8. Pradhan, A.K.; Pollatsek, A.; Knodler, M.; Fisher, D.L. Can younger drivers be trained to scan for information that will reduce their risk in roadway traffic scenarios that are hard to identify as hazardous? *Ergonomics* **2009**, *52*, 657–673. [CrossRef] [PubMed]
9. Thomas, F.D.; Korbelak, K.T.; Divekar, G.; Blomberg, R.D.; Romoser, M.R.E.; Fisher, D.L. *Evaluation of an Updated Version of the Risk Awareness and Perception Training Program for Young Drivers*; National Highway Traffic Safety Administration: Washington, DC, USA, 2017.
10. Chee, D.Y.T.; Lee, H.Y.; Patomella, A.-H.; Falkmer, T. The visual search patterns of drivers with Autism Spectrum Disorders in complex driving scenarios. *J. Transp. Health* **2019**, *14*, 100597. [CrossRef]
11. Reimer, B.; Fried, R.; Mehler, B.; Joshi, G.; Bolfek, A.; Godfrey, K.M.; Zhao, N.; Goldin, R.; Biederman, J. 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.* **2013**, *43*, 2211. [CrossRef]
12. Vindin, P.; Wilson, N.J.; Lee, H.; Cordier, R. The Experience of Learning to Drive for People with Autism Spectrum Disorder. *Focus Autism Other Dev. Disabil.* **2021**, *36*, 225–236. [CrossRef]
13. Wickens, C.D. Multiple Resources and Mental Workload. *Hum. Factors* **2008**, *50*, 449–455. [CrossRef]
14. Patrick, K.E.; Schultheis, M.T.; Agate, F.T.; McCurdy, M.D.; Daly, B.P.; Tarazi, R.A.; Chute, D.L.; Hurewitz, F. Executive function "drives" differences in simulated driving performance between young adults with and without autism spectrum disorder. *Child Neuropsychol. A J. Norm. Abnorm. Dev. Child. Adolesc.* **2020**, *26*, 649–665. [CrossRef] [PubMed]
15. Chee, D.Y.; Lee, H.C.; Patomella, A.-H.; Falkmer, T. Driving Behaviour Profile of Drivers with Autism Spectrum Disorder (ASD). *J. Autism Dev. Disord.* **2017**, *47*, 2658–2670. [CrossRef] [PubMed]
16. Sweller, J. Measuring cognitive load. *Perspect. Med. Educ.* **2018**, *7*, 1–2. [CrossRef]
17. Mak, G.; Zhao, L. A systematic review: The application of virtual reality on the skill-specific performance in people with ASD. *Interact. Learn. Environ.* **2020**, *31*, 804–817. [CrossRef]
18. 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 *Universal Access in Human-Computer Interaction*; Springer International Publishing: New York, NY, USA, 2015. [CrossRef]
19. Sheppard, E.; van Loon, E.; Underwood, G.; Ropar, D. Attentional Differences in a Driving Hazard Perception Task in Adults with Autism Spectrum Disorders. *J. Autism Dev. Disord.* **2017**, *47*, 405–414. [CrossRef]
20. Patrick, K.E.; Hurewitz, F.; McCurdy, M.D.; Agate, F.T.; Daly, B.P.; Tarazi, R.A.; Chute, D.L.; Schultheis, M.T. Driving Comparisons Between Young Adults with Autism Spectrum Disorder and Typical Development. *J. Dev. Behav. Pediatr.* **2018**, *39*, 451–460. [CrossRef]
21. Curry, A.E.; Yerys, B.E.; Huang, P.; Metzger, K.B. Longitudinal study of driver licensing rates among adolescents and young adults with autism spectrum disorder. *Autism Int. J. Res. Pract.* **2018**, *22*, 479–488. [CrossRef]
22. Curry, A.E.; Metzger, K.B.; Carey, M.E.; Sartin, E.B.; Huang, P.; Yerys, B.E. Comparison of Motor Vehicle Crashes, Traffic Violations, and License Suspensions Between Autistic and Non-Autistic Adolescent and Young Adult Drivers. *J. Am. Acad. Child Adolesc. Psychiatry* **2021**, *60*, 913–923. [CrossRef]
23. Wade, J.; Weitlauf, A.; Broderick, N.; Swanson, A.; Zhang, L.; Bian, D.; Sarkar, M.; Warren, Z.; Sarkar, N. A Pilot Study Assessing Performance and Visual Attention of Teenagers with ASD in a Novel Adaptive Driving Simulator. *J. Autism Dev. Disord.* **2017**, *47*, 3405–3417. [CrossRef]
24. Baker-Ericzén, M.J.; Smith, L.; Tran, A.; Scarvie, K. A Cognitive Behavioral Intervention for Driving for Autistic Teens and Adults: A Pilot Study. *Autism Adulthood* **2021**, *3*, 168–178. [CrossRef]
25. Cox, D.J.; Brown, T.; Ross, V.; Moncrief, M.; Schmitt, R.; Gaffney, G.; Reeve, R. Can Youth with Autism Spectrum Disorder Use Virtual Reality Driving Simulation Training to Evaluate and Improve Driving Performance? An Exploratory Study. *J. Autism Dev. Disord.* **2017**, *47*, 2544–2555. [CrossRef]
26. Vindin, P.; Cordier, R.; Wilson, N.J.; Parsons, L.; Lee, H. Development and feasibility of a driving training program for Autistic student drivers. *PLoS ONE* **2025**, *20*, e0324934. [CrossRef]
27. Fan, J.; Wade, J.W.; Key, A.P.; Warren, Z.E.; Sarkar, N. EEG-Based Affect and Workload Recognition in a Virtual Driving Environment for ASD Intervention. *IEEE Trans. Biomed. Eng.* **2018**, *65*, 43–51. [CrossRef] [PubMed]
28. Schweizer, T.; Gilgen-Ammann, R. Wrist-Worn and Arm-Worn Wearables for Monitoring Heart Rate During Sedentary and Light-to-Vigorous Physical Activities: Device Validation Study. *JMIR Cardio* **2025**, *9*, e67110. [CrossRef] [PubMed]
29. Ross, V.; Cox, D.J.; Reeve, R.; Brown, T.; Moncrief, M.; Schmitt, R.; Gaffney, G. Measuring the attitudes of novice drivers with autism spectrum disorder as an indication of apprehensive driving: Going beyond basic abilities. *Autism Int. J. Res. Pract.* **2018**, *22*, 62–69. [CrossRef]

30. Ross, V.; Jongen, E.; Van Vlierden, K.; Brijs, K.; Brijs, T.; Hens, R.; Wets, G.; Vanvuchelen, M. Process of learning to drive by young persons with autism: Experiences of the young persons themselves, parents, and driving instructors. *Trans. Transp. Sci.* **2018**, *9*, 42–56. [CrossRef]
31. Addula, S.R.; Sajja, G.S. Automated Machine Learning to Streamline Data-Driven Industrial Application Development. In *2024 Second International Conference Computational and Characterization Techniques in Engineering & Sciences (IC3TES)*; IEEE: New York, NY, USA, 2024; pp. 1–4. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.