Accident Analysis and Prevention 166 (2022) 106565

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Driver distraction and its effects on partially automated driving performance: A driving simulator study among young-experienced drivers

Noa Zangi, Rawan Srour-Zreik, Dana Ridel, Hadas Chassidim, Avinoam Borowsky *

Ben-Gurion University of the Negev, Israel

A R T I C L E I N F O

Keywords:
Partially automated driving
Driving simulator
Driver distraction
Non-driving related task
Eye movements
Heart rate variability

A B S T R A C T

Drivers of partially automated vehicles (PAVs) are relieved from parts of the driving tasks allocated to the automated driver. Ironically, these drivers are obligated to continuously monitor the driving task at all times and keep their attention on the roadway. This reduction in the driving task's demands and cognitive workload may encourage drivers to engage with non-driving related tasks (NDRT), which may impair drivers' awareness of the road environment and, as a result, compromise safety. This study examined how engagement with a visual-manual NDRT affects the driving performance of PAV drivers. Thirty-seven participants were randomly assigned to one of two experimental conditions in a driving simulator. Each consisted of two simulated drives. The first experimental condition included one drive under manual driving conditions and another under partially automated driving conditions (i.e., L2). Both drives had no NDRT involved. The second experimental condition included one drive under L2 without an NDRT and one drive under L2, including engagement with an NDRT. Participants' eye movements and heart rate were recorded throughout the experiment. Across various measures, the findings showed that under L2 driving conditions, engagement with an NDRT impairs driving performance in two primary aspects: (1) drivers were less aware of road hazards, and (2) their mental workload was higher when they engaged with an NDRT. In addition, the findings reveal that for drivers engaged with an NDRT, the attentional time-sharing strategy between the NDRT and the roadway monitoring task affected the probability of identifying a hazard. This study shows the adverse effects of engagement with an NDRT under L2 driving conditions on driving performance. Future studies should examine different interventions to mitigate these effects, assuring that drivers are constantly aware of the roadway environment.

1. Introduction

Driver inattention, including distraction, contributes to around 65% of safety–critical events (Cunningham & Regan, 2018). According to Peters & Stavrinos (2017), 1.6 million distracted-driving-related crashes and 330,000 distracted-driving-related injuries occur each year worldwide. Young et al. (2007) have argued that there are four types of distractions: cognitive, visual, auditory, and physical. Visual distraction occurs when the driver allocates visual attention to off-road targets for extended periods. Engagement with non-driving related tasks (NDRTs) while driving may affect driving performance and safety, depending on the task's type (Young et al., 2007). Since the driving task is primarily visual, thus requiring visual attention, visual NDRT competes over the limited visual attention resources. Therefore, visual NDRT are considered distractions that adversely impact awareness of road hazards (e.g., Borowsky et al., 2016), speed management (Ortiz-Peregrina et al., 2020), and a higher number of lane excursions and a higher standard deviation of lane position (Ortiz et al., 2018). Therefore, to maintain an appropriate level of driving performance and concurrently engage with a visual NDRT, drivers must regulate their NDRT engagement per the driving context and driving requirements. For example, drivers can decide to engage with an NDRT if the current driving situation allows it and not engage if they perceive it as complex and mentally demanding. A recent study by Nilsson et al. (2020) investigated the visual behavior among manual car drivers during the execution of a cognitive task in two simulated traffic scenarios. They showed that the number of glances towards potential threats decreased as the cognitive load increased. Schömig and Metz (2013) proposed a three hierarchical levels model to describe how drivers regulate their engagement with an NDRT while driving. The levels are planning, decision and control. On the planning level, drivers plan a priori (before starting their drive) where and when it would be most appropriate to engage with an NDRT along their route.

* Corresponding author.
E-mail address: borowsky@bgu.ac.il (A. Borowsky).

https://doi.org/10.1016/j.aap.2022.106565
Received 12 September 2021; Received in revised form 29 December 2021; Accepted 3 January 2022
Available online 13 January 2022

On the decision level, drivers decide in real-time whether to begin engagement with an NDRT or postpone it based on the current situation. Finally, on the control level, drivers who have already initiated engagement with an NDRT will shift their attention between the driving and the NDRTs by applying various divided attention strategies.

Some evidence of the various strategies drivers may apply during the control level comes from multitasking and manual driving literature (Schömig and Metz 2013). Driving and concurrently engaging with an NDRT can be performed as *concurrent multitasking* or *sequential multitasking* along the multitasking continuum (Salvucci & Taatgen, 2009). An example of concurrent multitasking is when the driver is driving and talking over the phone simultaneously. An example of sequential multitasking is when the driver might spend a while focusing on reading a text message before stopping and moving his gaze back to the road (Salvucci & Taatgen, 2009). In sequential multitasking, the operator must prioritize one task over the other since concurrent task performance is impossible in overload situations (Wickens et al., 2015). People can *self-interrupt*—stop performing one task and switch to a potentially more critical one. A person can estimate the time spent on a task by setting an internal clock to run for a set period, and when the timer reaches its time, switch to another task (Salvucci & Taatgen, 2009).

Recently, Ahlstrom, Kircher and Kircher (2013) developed a real-time distraction detection algorithm, termed AttenD, that can potentially capture different time-sharing strategies between an NDRT and the roadway. Its fundamental assumption is that drivers' attention follows the same object as their gaze. Everything inside the vehicle, except for the mirrors and the speedometer, is irrelevant for the driving task. The algorithm follows the principle that a sign for driver distraction is either a single glance longer than 2 s (Horrey and Wickens, 2007) or frequent short glances away from the forward roadway. The general idea behind the AttenD algorithm is that the driver has a time buffer of 2 s, which depletes in real-time when the driver looks away from the forward roadway. The buffer will start to fill up after a latency period of 0.1 s, only after the driver's gaze is redirected toward the forward roadway. If the buffer runs empty, the driver is classified as distracted (Ahlstrom et al., 2013).

Under these circumstances of driver distraction, drivers must continuously anticipate hazardous situations and perform evasive actions to avoid a crash. Drivers' ability to anticipate hazards is a skill known as hazard perception (HP; Borowsky et al., 2010). HP, i.e., hazard anticipation or hazard awareness, may be defined as drivers' ability to read the road and identify hazardous traffic situations (Yamani et al., 2016). Of the many driving skills a driver possesses, HP is the only driving skill that correlates with traffic crashes (Horswill & McKenna, 2004; Horswill, Hill, & Wetton, 2015). Measuring HP performance is typically done by analyzing participants' eye movements and response behavior when driving in a driving simulator, orobserving short video clips of real-world hazardous situations. Driver distraction impairs the detection of road hazards. For example, drivers engaged with texting while driving were less likely to identify and respond to hazards than drivers who drove without texting (Burge & Chaparro, 2018). Likewise, in a study conducted by Savage et al. (2013), participants observed video clips of hazardous situations. They had to solve various thinking puzzles, which preceded the hazard perception clips. The findings revealed that when the participants were preoccupied with an NDRT (puzzle solving), they responded to approximately 30% more non-hazardous stimuli (i.e., false alarms) and were slower by 3 s in responding to hazards. These findings imply increased visual attentional demands imposed by specific NDRTs. The lack of attentional resources due to engagement with an NDRT was likely to impair hazard perception performance. Similarly, Borowsky et al. (2016) showed that drivers who drove in a driving simulator and engaged with a visual-manual task for two seconds, adjacent to the beginning of a hazardous situation, were less likely to identify a pre-cued hazard than drivers not engaged with an NDRT. They explained that hazard perception is dependent on working memory, and when it contains information of a visual-manual task, hazard perception performance deteriorates.

Despite the plethora of knowledge on the effects of NDRT on manual driving performance, there is insufficient knowledge of its impact on partially automated driving. Due to the rapid penetration of partially automated vehicles into the market, studying NDRT effects on partially automated driving becomes essential. These topics are discussed next.

Automation in vehicles is developing rapidly and brings many benefits such as improved traffic flow, increased road safety, and advantages to special populations such as younger and older drivers and mobility impaired drivers (Fisher et al., 2016). In addition, it can improve the detection of a driver's state in situations where the driver is under the influence of alcohol, fatigue, or drugs (Fisher et al., 2016). The Society of Automotive Engineers (SAE) defined six levels of driving automation (L0 to L5; SAE, 2018). L0 represents manual driving, and L5 represents full automation. L2 is defined as partial automation where the automated system is responsible for both the lateral and longitudinal control of the vehicle. Still, the human driver is the sole driving task supervisor. Thus, the driver should monitor the roadway environment continuously and take over control of the vehicle in hazardous situations (e.g., overtaking a lead vehicle, obstacle avoidance) or when the automation system reaches its operational limits (De Winter et al., 2016). This requirement of continuous roadway monitoring implies that the driver must remain in the driving loop at all times. Under these conditions, the primary driving task demands are remaining vigilant, monitoring the roadway environment, and searching for hazards (Fisher et al., 2016; Lin et al., 2019).

Nevertheless, drivers under L2 driving conditions are ironically relieved from some driving tasks allocated to the automated system. This reduction in the driving task demands, and possibly in cognitive workload imposed on drivers, results in a higher tendency to engage with non-driving related tasks (NDRT), higher frequency of interaction with NDRT, and a higher average duration of glancing away from the forward roadway compared to manual driving (Solís-Marcos et al., 2018). In addition, over trust and complacency in automation may lead to similar effects such as fewer glances toward the road and increased engagement with NDRT, preventing drivers from noticing automation failures or degraded ability to take over control when required (Wilson et al., 2020). For example, Carsten et al. (2012) allowed drivers to choose whether and when to engage with NDRT from a variety of options during three levels of automation: manual, semiautomated (longitudinal/lateral control), and highly automated driving (both longitudinal and lateral control). Results showed that higher levels of automation led to an increased willingness to engage with NDRT, especially listening to the radio and watching a DVD.

Therefore, a crucial aspect that requires exploration is how PAV drivers distribute their attention between the forward roadway and performing NDRT (Naujoks et al., 2016) and how these divided attention strategies affect their hazard perception performance. Naujoks et al. (2016) conducted a field study and found that under L2 driving conditions, drivers adjusted their level of engagement with NDRT to the traffic situation; as their vehicle's velocity increased, drivers focused their attention more on the forward roadway and less on the NDRT. This decrement in NDRT engagement due to traffic conditions can be considered situation-adaptive behavior since it may reduce the perceived safety or increase the perceived workload during partially automated driving. According to De Winter et al. (2014), drivers using adaptive cruise control (ACC) performed about 12% more in-vehicle visual tasks than manual driving. This excessive engagement with NDRT followed by prolonged in-vehicle glances, and at a higher rate, may simultaneously compete with the process of hazard perception over the limited resource of visual attention. It may thus interfere with drivers' ability to identify hazards and take over control safely. To the best of our knowledge, the ways engagement with NDRT affects hazard perception under L2 driving conditions has not been studied yet, and this is one of the primary goals of this study.

A higher tendency to engage with NDRT during L2 driving also implies that drivers might experience high levels of mental workload during their engagement with NDRT. Since high levels of mental workload can impair driving performance (Paxion et al., 2014) and hazard perception (Recarte & Nunes, 2003), it is essential to measure mental workload and assess its effects. Evaluating mental workload can be done subjectively using standard questionnaires such as NASA R-TLX (Rubio et al., 2004). For example, recent evidence shows that cognitively distracted participants by variations of n-back tasks report lower levels of mental workload when driving in a driving simulator under varying levels of driving automation than manual driving (Lu et al., 2021). Nevertheless, in this study, drivers' control takeover performance (when a lane change was required or braking due to a decelerating lead vehicle) was inferior under higher levels of automation than manual driving. This contradictory evidence may imply objective mental overload when drivers had to takeover control. Similarly, a recent on-road study revealed that participants who drove in three different L2 vehicle models in urban and highway environments reported a higher mental workload and lower acceptance when driving under partially automated mode than manual driving (Kim et al., 2021).

Mental workload can also be assessed objectively using physiological measures such as heart rate variability (Hoover et al., 2012) or evaluating the performance of the NDRT (e.g., Olsson & Burns, 2000, Martens & Van Winsum, 2000). Commonly used physiological measures for assessing changes in mental workload in general, and specifically due to changes in the driving demands, is cardiac monitoring (Pohlmeyer & Coughlin, 2008, Stapel et al., 2019; Lohani et al., 2019). Heart rate variability (HRV), which is one form of cardiac monitoring, was found to decrease under high mental effort (Lohani et al., 2019; Stapel et al., 2019) and is sensitive to workload increases due to vigilance and situational awareness demands of the task (Lohani et al., 2019). HRV metrics can be categorized as time-domain parameters evaluating the variations in heartbeat intervals. One of them is the RMSSD - root means square of successive R-R intervals (Lohani et al., 2019). In a recent on-road study involving four types of partially automated vehicles, Lohani et al. (2021) used RMSSD to examine whether partially automated driving increases cognitive demands compared to manual driving. The authors have found no difference in RMSSD between the two driving modes leading to the conclusion that partially automation without an additional secondary task does not lead to an increase in cognitive demands. In this study, we will use both subjective and objective measures of mental workload levels.

To summarize, while the prevalence of L2 partially automated vehicles is increasing (Rosenzweig & Bartl, 2015), there is a lack of literature on how engagement with NDRT affects automated vehicle deactivation, hazard perception, and drivers' attentional time-sharing strategies. Thus, this study had two main goals: (1) examine how engagement with a visual-manual NDRT under L2 driving conditions affects driving performance. The driving performance measures included hazard perception and automated vehicle deactivation. (2) Examine drivers' attentional time-sharing strategies between monitoring the road and engagement with an NDRT under L2 driving conditions.

Participants in this study used a driving simulator to navigate through two virtual drives. Each drive included the same seven hazardous situations embedded in a pseudo-randomized order. Participants were randomly assigned to a baseline condition or experimental condition. The baseline condition consisted of two drives, one under manual and one under L2 driving conditions. Both drives had no engagement with NDRT. The experimental condition included two drives, one under L2 driving conditions without an NDRT and one under L2 driving conditions with an NDRT. The order of the two drives in both conditions was counterbalanced among participants. Participants' eye movements and heart rate were monitored and recorded throughout the experiment in addition to questionnaires and vehicle dynamics data.

2. Hypotheses

We elicited three hypotheses to investigate the effects of drivers' engagement with NDRT on their driving performance under L2 driving conditions.

H1: Under L2 driving conditions, drivers' awareness of hazardous situations (i.e., hazard perception performance) will deteriorate when engaging with an NDRT such that:

H1.1: They will be less likely to identify hazards than drivers who do not engage with an NDRT.
H1.2: They will have fewer glances at a hazard than drivers who do not engage with an NDRT.
H1.3: When encountering potential hazards, they will be less likely to deactivate the automation mode than drivers who do not engage with an NDRT.

H2: Under L2 driving conditions, drivers' mental workload will increase when engaging with an NDRT such that:

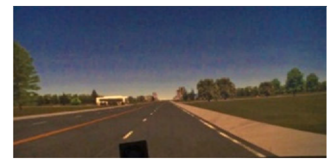
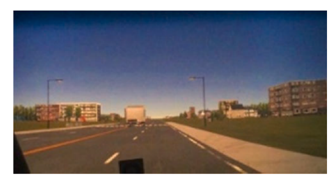
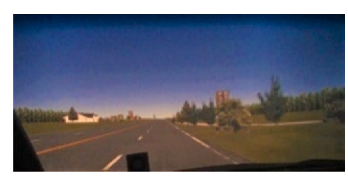
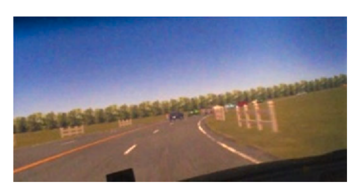
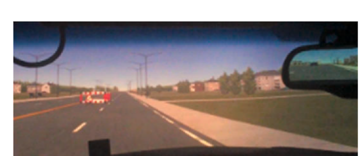
H2.1: They will report higher ratings of mental workload (higher score in NASA- TLX questionnaire) than drivers who do not engage with an NDRT.
H2.2: Their HRV will be smaller when engaged with an NDRT than when they do not engage with an NDRT.

H3: Different attentional time-sharing strategies between monitoring the road and engaging with an NDRT will differently affect the ability to identify hazards:

H3.1: As the number of glance shifts between the road and the NDRT increases, drivers will be more likely to identify hazards.
H3.2: As the total time of the NDRT increases, drivers will be more likely to identify hazards.
H3.3: As the cumulative glances' durations on the NDRT increase, drivers will be less likely to identify hazards.

Fig. 1. Real-time technologies, high fidelity driving simulator (left), an in-vehicle touch screen display (center), and rear-projector (right).

Table 1
A detailed description of the seven hazardous scenarios.

#	Scenario Name	Description	Scenario picture
1	Pulling out from a line of parked vehicles	**Description**: The participant's vehicle drives in an industrial area (the same road type as the urban environment). There are vehicles parked on the right in parallel to the sidewalk. One vehicle uses its left signal indicator to indicate its intention to merge into the driver's lane. A truck is on a left side street waiting to enter traffic, distracting the driver who approaches the parked vehicles. **Hazard**: The vehicle intending to pull out and the truck on a left side street. **Road type**: Urban, one lane in each travel direction. **Speed limit**: 60 km/h **AOI**: The parked vehicle with the signal light. **Cue**: Parked vehicle's left signal indicator. **Expected response**: The driver should glance at the truck on the left and the car on the right that intends to pull out from parallel parking.	
2	A car approaches an intersection from the left-hand side	**Description:** The participant's vehicle is driving on the right lane in an urban environment. Approximately 500 m before the participant approaches a stop-controlled intersection (the participant has the right of way), a car is approaching on the left lane from the left-hand side of the intersection. A bus stopped on the right lane on the left-hand side of the intersection obscures the car when it reaches the stop line in front of the intersection. **Hazard**: The obscured car on the left-hand side of the intersection can turn left and merge onto the participant's path. **Road type:** Highway, two lanes in each travel direction. **Speed limit**: 90 km/h **AOI**: Area in front of the bus. **Cue**: Car approaching from the left side of the intersection. **Expected response**: When the participant is ready to cross the intersection, he should glance towards the front of the bus and notice the obscured car to make sure it stopped and does not burst into the intersection.	
3	SUV obscures the view of a midblock crosswalk	**Description**: The participant's vehicle drives on the right lane of a road. An SUV is stopping on the left lane right in front of the crosswalk. **Hazard**: The front edge of the SUV obscures possible pedestrians crossing the road in front of the SUV. **Road type:** Urban, two lanes in each travel direction. **Speed limit:** 60 km/h **AOI:** Area in front of the van at the crosswalk. **Cue**: crosswalk, stopping SUV. **Expected response:** The driver should notice the crosswalk and slow down. Before passing the crosswalk, the participant should look left ahead of the SUV and search for a possible pedestrian. Note that the automation will do not slow down in this scenario.	
4	Exit from a gas station	**Description**: The participant's vehicle drives on the right lane. The participant passes near a gas station with an entrance and an exit from and to the main road. Approximately 500 m before the gas station, there is a vehicle that enters the gas station. When the participant's PAV passes, a vehicle at the exit is waiting to merge into the main road. **Hazard**: The vehicle on the right is waiting to merge onto traffic from the exit area of the gas station. **Road type:** Highway, two lanes in each travel direction. **Speed limit:** 90 km/h **AOI**: The vehicle at the gas station's exit. **Cue**: The vehicle ahead enters the gas station, indicating the possibility of other vehicles exiting the gas station further down the road. **Expected response:** The driver should look at the car at the gas station's exit.	
5	Sudden traffic slowing cascade	**Description**: The participant's vehicle drives on the right lane of a 4-lane unidirectional curved road in a highway environment and approaches a traffic jam consisted of several lines of slowing vehicles. **Hazard**: Decelerating vehicles in front. The driver may collide with these vehicles in case the automation suddenly fails. **Road type**: Highway, two lanes in each travel direction. **Speed limit**: 90 km/h **AOI:** Braking lights of vehicles in front of the participant. Cascade at the end of the curve. **Cue**: The participant can make glances across the curve's apex and notice slowing car lines. **Expected response:** The participant's car will automatically slow down because of the ACC, but we would like to see if the driver notices and looks at the braking lights of the lead vehicles. Other drivers might decide to disengage from the automation and apply their brake to prevent a crash.	
6	Vehicle from behind approaches fast near a construction zone	**Description**: The participant's vehicle drives on the right lane of a two-lane bidirectional road in an urban environment. A construction zone ahead on the left lane is approximately 500 m in front of the participant's car. While the participant is driving on the right lane, a car on the left lane behind the participant approaches the construction zone relatively fast. **Hazard**: The approaching vehicle may pass the ego car and then cut in the	

(*continued on next page*)

Table 1 (*continued*)

#	Scenario Name	Description	Scenario picture
		simulator car to not enter the construction zone. **Road type**: Urban, two lanes in each travel direction. **Speed limit:** 60 km/h **AOI**: Rear and left mirrors. **Cue**: Working zone. **Expected response:** The driver should notice to the vehicle in the mirrors and the work zone, disengage from the automation, and slow down.	
7	Give right of way to the vehicle on the right-hand side of a four-way intersection	**Description**: The participant's vehicle is driving on the right lane and approaches a four-way intersection. A lead vehicle in front of the participant slows down before entering the intersection to ensure no crossing traffic is present. Then, it accelerates and clears the intersection. Then the participant's PAV approaches the intersection. Concurrently, a third vehicle, which has the right of way, approaches from the right-hand side of the intersection. **Hazard**: The car on the right-hand side of the intersection. **Road type:** Urban, two lanes in each travel direction. **Speed limit:** 60 km/h **AOI**: The right side of the intersection. **Cue**: Give right of way sign and the vehicle approaching from the right side of the intersection. **Expected response**: The driver should stop and give way to the car coming from the right.	

3. Method

3.1. Participants

Thirty-seven participants, 19 females (mean age = 25.73, SD = 1.91) and 18 males (mean age = 26.11, SD = 1.74), were randomly assigned into one of two experimental groups. All participants were undergraduate students from Ben-Gurion University of the Negev (BGU) in Beer Sheva city in Israel and received monetary compensation for their participation. All participants were right-handed and had normal or corrected-to-normal visual acuity of 6/9 or better (eye contact lenses were allowed), normal contrast sensitivity (Ginsburg, 2003), normal color vision, and no background of heart problems. Participants had a valid driver's license for at least five years and reported no previous experience with a driving assistance system providing lateral and longitudinal support. The average driving experience was 8.5 years (SD = 2.05) and the weekly average number of driving hours was 4.67 (SD = 3.85). Four participants were involved in minor car crashes. Twenty participants were car owners. The BGU IRB ethical committee approved the study.

3.2. Apparatus

Driving Simulator. An RTI Driving Simulator (Realtime Technologies, Inc.) was used for the study. The driving simulator consisted of an engineless Cadillac-STS sedan and a 7 m diameter curved screen (2.4 m × 6.1 m), creating a visual angle of 165 degrees of the virtual world. The curved screen is located at about 1 m in front of the Cadillac (Fig. 1 left). Three laser projectors displayed the virtual world on the curved screen, and a designated software (Wrapalizer, Inc.) did the edge blending. A rear projector and a screen at the back of the simulator (Fig. 1 right) presented the virtual environment through the in-vehicle rear-view mirror. In addition, each physical side mirror included a 7″ LCD showing the respective views of the virtual environment. An in-vehicle 10″ touchscreen display was used to display the NDRT. It was located on the central stack on the dashboard (see Fig. 1 center). The virtual environment, traffic scenarios, and NDRT were generated with RTI SimCreator, SimVista, and Altia Designer software (Altia Designer, Inc.). The driving simulator provides various vehicle dynamics measures such as driving speed, steering angles, acceleration, and braking at a rate of 30 Hz.

The participant's Simulated Vehicle. The simulated vehicle was designed in two modes: manual driving mode and L2 driving mode. In the manual driving mode, the participant was controlling all aspects of the driving task. In the L2 driving mode, the PAV controlled the car's speed, acceleration, time headways from lead vehicles, and lane position (i.e., longitudinal and lateral control). Activating the PAV was done by the driver using a designated button located inside the car. The participant could disengage from the PAV and switch back to manual driving at any time by pressing the brake pedal.

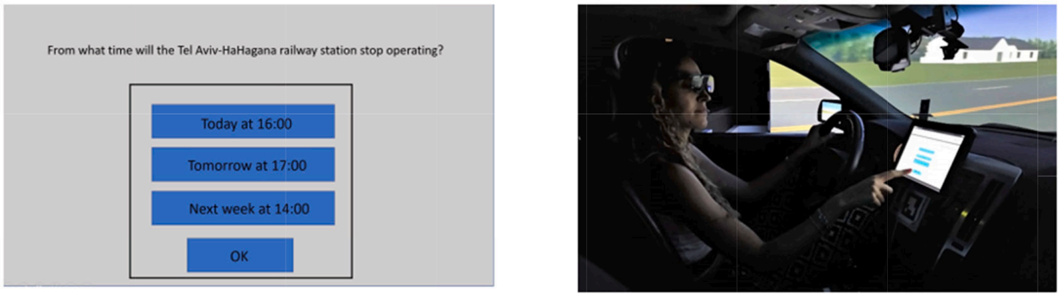
Eye tracker. Participants' eye movements were recorded via Tobii Pro Glasses 2 head-mounted Eye Tracking System throughout the experiment. The glasses weigh 45gr and include four eye cameras, a full HD wide-angle scene camera, a gyroscope, and accelerometer sensors (Tobii, 2018). The eye tracker samples gaze position at a rate of 50 Hz with an accuracy of 0.5 degrees of visual angle. The Tobii Pro Lab software (ver. 1.142) was utilized to analyze glance patterns. The main output of the data includes videos with gaze positions superimposed.

Electrocardiogram (ECG). A BioPac ECG system (MP150) was utilized to measure participants' heart rate at 2000 Hz. The system consists of a matched transmitter and receiver module, which emulated a "wired" connection from the subject to the computer. Connecting the electrodes to the participant was done by pasting three stickers on the participant's chest. These electrodes interface with two parts: the platform of data - acquisition and analysis and the AcqKnowledge software. The ECG and the driving simulator were synchronized.

Questionnaires. The participants completed three different questionnaires administered in Google forms: (1) A demographic questionnaire included seven questions about the participant's background and experience with driving assistance systems (DAS). (2) A Multidimensional Driving Style Inventory questionnaire (MDSI; Taubman-Ben-Ari et al., 2004) included 44 questions related to the participant's driving style. Each question contained a 6-points Likert scale ("1″ represents no agreement; "6″ complete agreement). This questionnaire distinguishes eight driving styles: Dissociative, Anxious, Risky, Angry, High-velocity, distress-reduction, Patient and Careful. (3) The user's adoption and trust aspects of autonomous vehicles' questionnaire consists of 12 questions with a 7-point Likert scale ("1″ represents strongly disagree; "7″ represents strongly agree) (Choi & Ji, 2015). (4) After each experimental drive, participants completed a NASA-TLX questionnaire (Hart & Staveland, 1988) for mental load assessment. This questionnaire consists of 6 questions with a 9-point Likert scale ("1″ represents low and "9″ represents high). The six items of mental demand, physical demand, temporal demand, effort, performance, and frustration provide a comprehensive measure of cognitive workload (Hart, 2006).

3.3. Driving scenarios

The virtual drive included both highways environments and urban



(a) Task's interface (b) Engagement with the Secondary task

Fig. 2. The experimental setup of the in-vehicle NDRT.

environments in daylight and took about 24 min of driving. The roads in the urban environment included 2-lanes bi-directional roads except for one scenario that had one lane in each travel direction. Participants navigated through six different latent (un-materialized) hazard scenarios and one materialized hazard during the drive. The seven hazardous scenarios were distributed throughout the virtual drive at an average time interval of 3.5 min (average distance of 4 km) between two successive hazardous scenarios. The materialized hazard scenario always appeared at the end of the drive. Five different combinations of the six hazardous scenarios, and the seventh hazard at the end of each combination, were built and pseudo-randomized among the participants to reduce learning effects (see Table 1 for a description of the seven hazardous scenarios). The "cue" field describes the hazard-related visual cues that preceded each hazardous situation and cued the participant to the upcoming hazard.

For each of the six latent hazards, we defined a time window consisting of the period during which the participant should identify the visual cue(s) indicating the developing hazard and the period during which the driver should glance at the hazard instigator to avoid a potential crash. The start of the second period is called a launch zone, a predefined area of each driving environment where the hazard becomes visible, and the driver must begin glancing toward the target zone. A Target zone is a visual area of interest (AOI) of a latent hazard (Krishnan et al., 2019). The time window ended when the detection of the latent hazard would have been too late to avert a crash (Vlakveld et al., 2018). However, none of the six latent hazards ever materialized. The six-time windows ranged from 18 to 38 s and the seventh hazard time window was 22 s. The second period (launch zone) lasted between 1 and 3 s.

3.4. The NDRT

The participants performed a distracting visual task while driving (see Fig. 2- panel (a) for the task's interface). The self-developed NDRT included 14 text messages that popped up in the in-vehicle display for 10 s each. Following each text message, participants answered two multiple-choice questions (28 questions altogether) (see Fig. 2- panel (b)). The perceived difficulty level of each question was examined in a preliminary validation study to ensure that all the 28 questions are perceived at a similar level of difficulty. The topics of the text messages were diverse and focused on informative content (e.g., public transportation, traveling, shows, etc.). This task simulated a typical interaction with SMS messages that drivers may engage with while driving. The NDRT triggers were located along the drive at the same locations for all participants. Seven of the NDRT instances were initiated at the beginning of the launch zone of hazardous situations (one instance per scenario). The other seven NDRT instances were located at other drive parts that do not embed hazardous scenarios. The purpose of embedding these latter seven NDRT instances was to prevent drivers from associating between NDRT triggers and hazardous situations. The messages and the order of the questions were randomized for each participant.

3.5. Experimental design and variables

The experimental design was a mixed within-between-subjects design. The study included two, between-subjects, experimental conditions: (1) manual driving without an NDRT vs. L2 driving without an NDRT (i.e., baseline). This condition included 15 participants and (2) L2 driving without engagement with an NDRT vs. L2 driving with engagement with an NDRT. This condition included 22 participants. Participants allocated to the first experimental condition completed a virtual drive (containing a pseudo-randomized order of the seven scenarios) twice. In one drive, the participants drove under manual driving conditions, and in the other drive under L2 driving conditions. Participants allocated to the second experimental condition drove two virtual drives under L2 driving conditions, where one of these drives also included engagement with an NDRT. In both experimental conditions, the order of the virtual drives was counterbalanced across participants and one week separated between the two drives.

Independent variables. The independent variables were: (1) the **level of automation** which was a within-subjects variable (but only for those

Table 2
A summary of the dependent variables and their meanings.

Group	Measurement	Definition	Source
Hazard perception performances	Hazard identification	Whether a participant had at least one glance of at least 100 ms at the hazards	Eye tracker
	Total number of glances at the hazards	The total number of glances that a participant made at the hazards	Eye tracker
Automated vehicle deactivation	Did the participant deactivate the automation mode?	Whether a participant deactivated the automation mode in response to a hazard and drove manually	Driving simulator
Subjective ratings of mental workload	Mental workload	The average workload rate for each drive	Questionnaire NASA-TLX
NDRT scanning patterns	Total glance duration on NDRT (msec)	The total glances duration on the NDRT during the scenario	Eye tracker
	Total time of NDRT (msec)	The total duration of engagement in the NDRT in a scenario	Eye tracker
	Number of glance shifts	Number of glances shifts from the road to the NDRT and vice versa	Eye tracker
HRV-RMSSD	log(QuotientRMSSD)	$\log\left(\frac{\text{Recorded RMSSD}}{\text{Recorded Rest RMSSD Baseline}}\right)$	ECG
	log(QuotientSubRMSSD)	$\log\left(\frac{\text{Recorded RMSSD}}{\lvert\text{Recorded Rest RMSSD Baseline} - \text{Recorded RMSSD}\rvert}\right)$	ECG

Fig. 3. Schematic description of the experimental procedure; Black- Relevant for both drives, Blue- Relevant for the first drive, Green - Relevant for the second drive.

who were assigned to experimental condition 1) and included two levels: manual (L0), L2. (2) **NDRT existence** which included two levels: yes, no. This variable was relevant only for those who drove under L2 conditions in both drives (experimental condition 2). (3**) Drive number** included two levels: 1st drive, 2nd drive. (4) The **road type** of the specific scenario included two levels: Highway, Urban. These two environments were selected to provide a variety of typical hazardous situations that PAV may encounter.

Dependent variables. The dependent variables are presented in Table 2.

3.6. Procedure

At first, participants were requested to digitally sign an online informed consent form and filled in all questionnaires online except for the NASA-TLX. Upon their arrival at the human performance evaluation lab (HPEL), the participants underwent a visual acuity test (Snellen Chart) and a functional acuity contrast test (FACT) (Ginsburg, 2003). Participants who qualified were given a written explanation about the simulator vehicle, the devices they are about to use, and the experimental procedure following the driving condition to which they were assigned. They were instructed to follow the Israeli traffic laws and drive as they would have typically done in similar real-world situations. Participants who began their first drive under L2 driving conditions learned about the automated vehicle's capabilities (Adaptive cruise control; ACC, and Lane-keeping system; LKS). Specifically, participants were informed that they are fully responsible for their driving safety even when automation is active (i.e., they should keep monitoring the driving environment continuously at all times). Switching between the PAV and manual driving modes was possible at all times. Then, participants answered a short quiz about the automated system capabilities to verify they understand how it works. For the virtual drive under L2 driving conditions where engagement with NDRTs was included, participants received an explanation about the NDRT and how to interact with it. Participants were asked to engage with the NDRT as much as possible as long as they felt they are not compromising their safety and the safety of other road users. These instructions ensured that drivers would engage with an NDRT as much as possible, allowing us to further investigate drivers' behavior during the "control" level of Schömig and Metz's (2013) model.

Next, the ECG electrodes were positioned on the participants' chests. In their first drive, they were given an IKEA magazine to browse while resting for 5 min. Their heart rate and R-R- interval baseline was measured to serve as a baseline measure. Before each experimental drive, participants drove a 4-minute training session that matched the relevant driving conditions (manual / L2 with NDRT / L2 without NDRT) to become familiarized with the simulated driving conditions, the steering wheel sensitivity, and the car pedals. In this training session, participants experienced driving on straight roads, curves, intersections, and a transition from a two-lane to a four-lane road. Following the training session, the participant wore the eye-tracker glasses, and the experimenter calibrated his gaze position. Then, the first experimental drive began. Once the participants finished driving their first virtual drive, they completed the NASA-TLX questionnaire. A week later, each participant returned to the lab approximately at the same hour as in the first session for the second session of the experiment, where he underwent the second experimental drive. On their second session, participants underwent the same procedure as in the first session except for the 5 min rest period measured only once in the first session that each participant experienced. At the end of the two sessions, the experimenter thanked the participant and compensated them (100 NIS) for their participation. Fig. 3 presents a schematic description of the experimental procedure.

Table 3
Descriptive statistics of the experiment's variables – experimental condition 2.

Scenario	Hazard identification = Yes		Deactivate the automation mode	
Driving condition	L2	L2 + NDRT	L2	L2 + NDRT
Pulling out from a line of parked vehicles	18/20 90%	16/20 80%	9/20 45%	6/20 30%
A truck approaches an intersection from the left-hand side.	18/20 90%	12/20 60%	1/20 5%	0/20 0%
SUV obscures the view of a midblock crosswalk	19/20 95%	14/20 70%	9/20 45%	6/20 30%
Exit from a gas station	18/20 90%	7/20 35%	5/20 25%	3/20 15%
Vehicle from behind approaches fast near a construction zone	8/20 40%	4/20 20%	1/20 5%	2/20 10%
Give right of way to the vehicle on the right-hand side of a four-way intersection	16/20 80%	9/20 45%	15/20 75%	11/20 55%

3.7. Statistical analysis procedure

All statistical analyses were carried out at a significance level of 0.05. For dependent variables that are binary distributed, we used a logistic regression model with a logit link function within the Generalized Linear Mixed Models (GLMM) framework. These variables included hazard identification and deactivation of the automation mode. Hazard identification is a binary variable that describes whether a participant had at least one glance of at least 100 msec at the hazard ("1″) or not ("0″). For the dependent variable that is Poisson distributed, we used a Poisson regression model with a log link function within the GLMM framework. This variable included the total number of glances at the hazard. Next, we used a linear regression model within the Linear Mixed Models (LMM) framework for dependent variables with a normal or log-normal distribution. These variables included the mental workload score, logQuotientRMSSD, and logQuotientSubRMSSD.

The initial model in each analysis included the following fixed effects: drive number (1/2), NDRT existence (yes/no; appropriate only for the second experimental condition), road type (highway/urban), automation level (0/2; suitable only for the first experimental condition) and the first and second-order interactions between the different variables. In addition, the random effects of the initial model included scenario and subject. The final model of each analysis was achieved via a backward elimination procedure where all non-significant interaction effects were removed from the model. For significant fixed effects with more than two levels, post hoc pairwise contrasts comparisons analysis was applied. The Tukey HSD procedure was used to correct alpha for multiple comparisons (all statistical analyses were conducted in R software).

4. Results

The analyses reported in this section include all six un-materialized

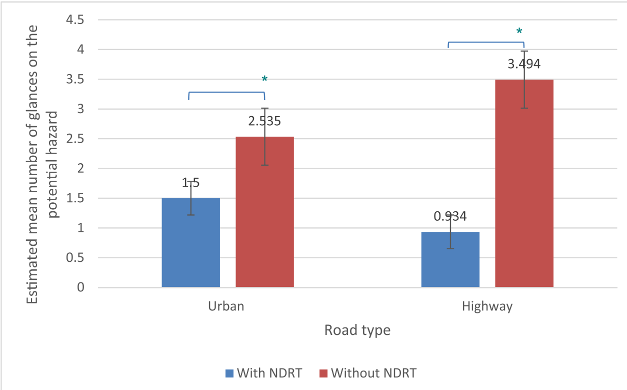


Fig. 4. The interaction between NDRT existence and road type, ** - P adjusted < 0.01, *- P adjusted < 0.05.

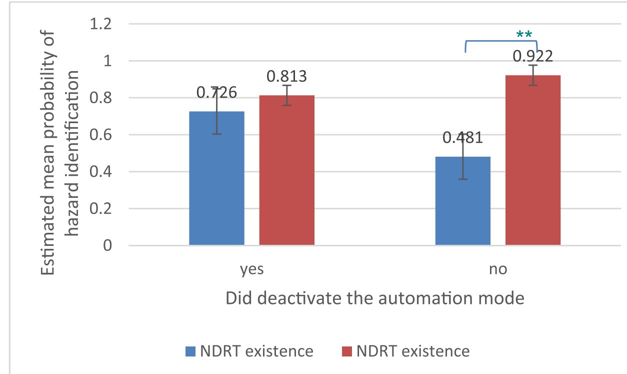


Fig. 5. The interaction between NDRT existence and did deactivate the automation mode ** - P adjusted < 0.01, *- P adjusted < 0.05.

hazardous scenarios, excluding the seventh scenario containing a materialized hazard. The seventh scenario examined how drivers mitigate a materialized hazard, but it is out of the scope of this paper. The data of 5 drives (a manual drive, two L2 drives without engagement with NDRT, and two L2 drives including engagement with NDRT) were invalid due to technical issues and were removed from all analyses. Thus, the final database that was used for all analyses consisted of 35 drives under L2 without engagement with NDRT (20 under experimental condition 2, 15 under experimental condition 1), 20 drives under L2 including engagement with NDRT, and 15 manual drives. Table 3 presents a summary of the descriptive statistics corresponding to the second experimental condition, with 20 drives analyzed for each condition (with or without NDRT).

The first experimental condition's primary purpose was to verify the absence of significant differences between manual and L2 driving conditions where an NDRT is not involved. Indeed, the analysis yielded no significant effects, which allowed us to focus on the second experimental condition where driving under L2 without NDRT served as a baseline. The following analyses rely on the second experimental condition (one driver under L2 with an NDRT and one drive under L2 without an NDRT). Participants in the first experimental condition (i.e., one drive manual and one drive L2 without an NDRT) were excluded from further analyses. In addition, only the NASA-TLX analysis resulted in significant differences between the experimental conditions of all the questionnaires reported in the method section. There were no significant differences between the groups for the other questionnaires, and most participants provided the same score regardless of the driving condition. Thus, in terms of subjective questionnaires, the results section presents only the NASA-TLX results.

4.1. Hazard perception performance

Hazard identification. The final logistic regression model included two significant main effects of NDRT existence ($X_1^2 = 21.837$, $p < 0.01$) and drive number ($X_1^2 = 3.852$, $p = 0.049$). First, for the main effect of NDRT existence, in the presence of the NDRT, the estimated mean probability that participants will have at least one glance at the hazard was smaller (Estimated Mean = 0.610, SE = 0.146) than in the absence of the NDRT (EM = 0.942, SE = 0.033). Second, concerning the main effect of drive number, the estimated mean probability that participants will have at least one glance at the hazard was higher in their first drive (Estimated Mean = 0.817, SE = 0.105) than in their second drive (EM = 0.675, SE = 0.153).

Total number of glances at a hazard. The final Poisson regression model included one significant main effect of NDRT existence ($X_1^2 = 64.011$, $p < 0.01$) and one significant interaction between NDRT existence and road type ($X_1^2 = 16.282$, $p < 0.01$). First, for the main effect of NDRT existence, participants had fewer glances at the hazard in the presence of an NDRT (Estimated Mean = 1.180, SE = 0.224) than in its absence (EM = 2.980, SE = 0.521). Second, for the interaction between NDRT existence and road type, Fig. 4 presents the estimated mean number of glances at a hazard for each level of the interaction.

As Fig. 4 presents, while driving in a highway environment, in the presence of an NDRT, participants had a fewer number of glances at the hazard (Estimated Mean = 0.934, SE = 0.253) than in its absence (EM = 3.494, SE = 0.833; P adjusted < 0.01). Likewise, while driving in an urban environment, in the presence of an NDRT, participants had a fewer number of glances at the hazard (Estimated Mean = 1.500, SE = 0.299) than in its absence (EM = 2.535, SE = 0.487; P adjusted < 0.01). Furthermore, there was no significant difference between highway and urban environments (P adjusted = 0.101) when an NDRT was present. Similarly, there was no significant difference between driving in highway and urban environments (P adjusted = 0.207) when an NDRT was absent.

4.2. Automated vehicle deactivation

The initial Logistic regression model included a binary dependent variable that describes whether a participant deactivates the automation mode during a scenario ("1″) or not ("0″). The final logistic regression model included three significant main effects of NDRT existence ($X_1^2 = 48418$, $p < 0.01$), drive number ($X_1^2 = 417255$, $p < 0.01$) and road type ($X_1^2 = 441295$, $p < 0.01$), and two significant interactions between drive number and road type ($X_1^2 = 606516$, $p < 0.01$) and between NDRT existence and drive number ($X_1^2 = 133072$, $p < 0.01$). First, for the main effect of NDRT existence, in the presence of an NDRT, the estimated probability that the participants will deactivate the automation was smaller (Estimated Mean = 0.084, SE = 0.0002) than in its absence (EM = 0.173, SE = 0.0003). Second, concerning the main effect of drive number, the estimated mean probability that the participants will deactivate the automation was higher in their first drive (Estimated Mean = 0.178, SE = 0.0003) than in their second drive (EM = 0.081, SE = 0.0002). Third, for the main effect of road type, the estimated mean probability that the participants will deactivate the automation was higher in an urban environment (Estimated Mean = 0.295, SE = 0.0007) than in a highway environment (EM = 0.043,SE = 0.0001). Since the two significant interactions are not the main focus of this study, post hoc pairwise contrasts comparisons analysis is not reported.

4.3. Hazard identification as a mediating variable

The initial Logistic regression model included an additional fixed effect "did deactivate the automation mode." The final logistic regression model yielded one significant main effect of NDRT existence ($X_1^2 = 18.394$, $p < 0.01$) and one significant interaction between NDRT existence and did deactivate the automation mode ($X_1^2 = 5.133$, $p = 0.023$). First, concerning the main effect of NDRT existence, in the presence of

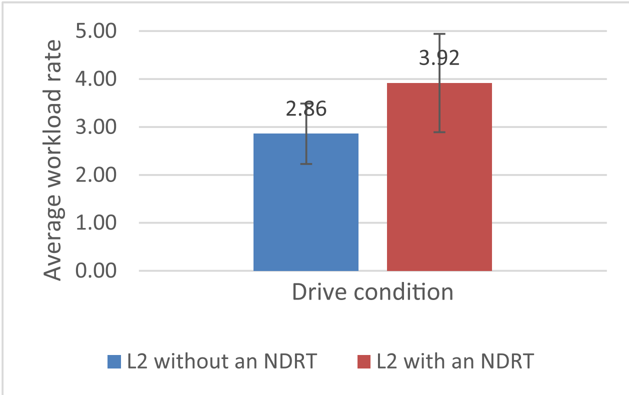


Fig. 6. The average reported workload rate for each condition- L2 driving with and without an NDRT.

an NDRT, the estimated mean probability that participants will have at least one glance at the hazard was smaller (Estimated Mean = 0.611, SE = 0.178) than in its absence (EM = 0.877, SE = 0.081). Second, concerning the interaction between NDRT existence and did deactivate the automation mode, Fig. 5 presents the estimated mean probabilities that participants will have at least one glance at the hazard for each level of the interaction.

As shown in Fig. 5, when the driver did not deactivate the automation mode, the estimated mean probability that participants will have at least one glance at a hazard was higher when an NDRT was absent (Estimated Mean = 0.922, SE = 0.057) than when it was present (EM = 0.481, SE = 0.181; P adjusted < 0.01). When the driver deactivated the automation mode, there was no significant difference between the conditions of L2 without an NDRT and the L2 with an NDRT (Estimated Mean = 0.769, SE = 0.157; P adjusted = 0.501). Post hoc pairwise contrasts comparisons analysis did not reveal any other significant difference. In the presence of an NDRT, there was no significant difference in the estimated mean probabilities between deactivating the automation or not (P adjusted = 0.122). Likewise, in the absence of an NDRT, there was no significant difference in the estimated mean probabilities between deactivating the automation or not (P adjusted = 0.166).

4.4. Mental workload

Descriptive statistics of the average mental workload ratings are presented in Fig. 6. Drivers who drove under L2 with an NDRT rated their workload higher (Mean = 3.920, SD = 1.120) than when they drove without an NDRT (Mean = 2.860, SD = 0.690). The statistical analysis of the mental workload rating scores included a linear regression model. The dependent variable was the mental workload ratings reported at the end of each drive. The final model had two significant main effects of NDRT existence ($X_1^2 = 240.361$, $p < 0.01$) and drive number ($X_1^2 = 15.501$, $p < 0.01$). First, for the main effect of NDRT existence, participants provided higher mental workload ratings on average in the presence of an NDRT (Estimated Mean = 3.810, SE = 0.191) than in its absence (EM = 2.910, SE = 0.191). Second, for the main effect of drive number, participants rated their mental workload on average as higher in their first drive (Estimated Mean = 3.470, SE = 0.191) than in their second drive (EM = 3.250, SE = 0.192).

4.5. NDRT scanning patterns

Hazard identification performance as a function of dual-task strategies. This analysis included only the data under the condition of L2 with NDRTs across the six scenarios. The initial Logistic regression model included a binary dependent variable representing hazard identification performance. This variable describes whether a participant had at least one glance of at least 100 msec at the hazard ("1″) or not ("0″). The model included three fixed effects. The first was the total duration of all glances directed at the NDRT (msec) during the engagement period. The second was the total duration of engagement with the task (msec), thus including the durations of on-road glances during engagement with the NDRT. The third was the total number of glance shifts between the roadway and the NDRT display. The final Logistic regression model included three significant main effects: the total glances duration on an NDRT ($X_1^2 = 8.434$, $p < 0.01$), the total duration of an NDRT ($X_1^2 = 4.248$, $p = 0.039$), and the number of glance shifts ($X_1^2 = 4.629$, $p = 0.031$). First, concerning the main effect of the total glances durations on an NDRT, the estimated coefficient was −0.039. This negative coefficient means that when the total duration of all glances at the in-vehicle display (NDRT engagement) increases, the probability of having at least one glance at the hazard decreases. Second, for the main effect of the total time of the NDRT, the estimated coefficient was 0.063, which means that as the total time of engagement with the NDRT increases, the probability of having at least one glance at the hazard will increase. Third, for the main effect of the number of glance shifts, the estimated coefficient was 0.417, which means that as the total number of glance shifts increases, the probability of having at least one glance at the hazard will increase. For additional analysis of scanning patterns, see Appendix A.

4.6. Heart rate variability (HRV) as a measure of mental workload

log*(QuotientRMSSD).* The final Linear Mixed model included one significant main effect of NDRT existence ($X_1^2 = 5.885$, $p = 0.015$). In the presence of an NDRT the estimated mean of log(QuotientRMSSD) measure was smaller (Estimated Mean = -0.327, SE = 0.137) than in its absence (EM = -0.052, SE = 0.137).

log*(QuotientSubRMSSD).* The final Linear Mixed model included two significant main effects of NDRT existence ($X_1^2 = 4.940$, $p = 0.026$) and drive number ($X_1^2 = 9.550$, $p < 0.01$). First, concerning the main effect of NDRT existence, in the presence of an NDRT, the estimated mean of log(QuotientSubRMSSD) measure was smaller (Estimated Mean = 0.364, SE = 0.160) than in its absence (EM = 0.700,SE = 0.160). Second, for the main effect of drive number, in the second drive, the estimated mean of log(QuotientSubRMSSD) measure was smaller (Estimated Mean = 0.299, SE = 0.162) than in the first drive (EM = 0.766,SE = 0.157).

5. Discussion

This study investigated the effects of engagement with an NDRT under L2 automated driving conditions on awareness of road hazards (i.e., hazard perception) and automated vehicle deactivation. As shown by Solís-Marcos et al. (2018), relieving the driver from the vehicle's lateral and longitudinal motion control results in a higher tendency to engage with NDRT, albeit the driver is required to monitor the driving task and automation continuously and intervene when necessary (SAE, 2018). Therefore, this study focused on drivers already engaged with a visual-manual NDRT (the "control" level in the model of Schömig and Metz (2013)). Recall that our experimental design included two experimental conditions. The first condition served as a baseline to show no significant differences between manual and L2 simulated driving when drivers are not engaged with an NDRT and continuously monitor the road environment. The second experimental condition examined the effects of engagement with an NDRT on L2 driving conditions. Thus, both drivers included L2 driving, whereas only one drive included engagement with an NDRT.

The first hypothesis (H1) stated that drivers' awareness of hazardous situations would deteriorate under L2 driving conditions when drivers engage with an NDRT. We analyzed three dependent variables to test this hypothesis: whether the driver had at least one glance at a hazard (yes/no; H1.1), the number of glances at a hazard given hazard identification (H1.2), and whether the driver deactivated the automation (yes/no; H1.3). All three analyses consistently confirmed the hypothesis. Concerning the first analysis, the findings showed that under L2 driving conditions, drivers engaged with an NDRT were substantial (35%) less likely to identify hazards than drivers not engaged with an NDRT. This considerable decrement in the probability of identifying a hazard can partially be due to a higher number of prolonged off-road glances that drivers make when they engage with an NDRT than when they are not (Noble et al., 2021). This substantial decrement can also result from shorter and intermittent on-road glances that drivers engaged with NDRT make. These on-road glances constitute 20% to 50% of the total task duration (Schömig & Metz, 2013). Consistently, for the second analysis, the findings showed that under L2 driving conditions, drivers engaged with an NDRT made substantially fewer glances at a hazard than drivers not engaged with an NDRT (1.18 vs. 2.98 respectively). This result complies with the argument that when the task's complexity is relatively high, as in this study, the number of glances at the road decreases, whereas the number of glances at the task increases compared to the baseline driving without an NDRT (Victor, 2005). It is also consistent with Nilsson et al. study, showing that as the cognitive load due to engagement with a secondary task while driving increases, the number of glances at potential threats decreases (Nilsson et al., 2020). Finally, the third analysis revealed that under L2 driving conditions, drivers engaged with an NDRT deactivated the automation mode half as much less often in response to potential hazards than drivers not engaged with an NDRT (probability of automation deactivation = 0.084 vs. 0.173, respectively). This pattern complies with the findings of Dogan et al. (2017), showing that half of their drivers engaged with a visual-manual NDRT when they had the option, which resulted in longer reaction times to a take over control request.

The automation deactivation analysis also revealed a relationship between hazard detection and automation deactivation. We found that hazard identification performance mediates between engagement with an NDRT and automation deactivation. That is, in the presence of an NDRT, drivers tended to deactivate the automation mode less often and had inferior hazard perception performance than in the absence of an NDRT. In other words, we found that drivers deactivated the automation mode only after they perceived the upcoming hazard. Our findings support this argument showing that when drivers deactivated the automation, they were equally likely to identify a hazard (based on glance position) regardless of the absence or presence of an NDRT. On the other hand, drivers who did not deactivate the automation when an NDRT was present were less likely to identify a hazard than when an NDRT was absent. This pattern suggests that these drivers engaged with an NDRT to the extent they were unaware of the hazard and thus did not deactivate the automation.

Next, we found that the road environment affected how drivers regulate their attention allocation at the road and the in-vehicle interface. First, participants had fewer glances in the presence of an NDRT than in its absence, under both highway and urban environments. Nevertheless, the differences found were significantly more considerable for the highway environment than for the urban environment (estimated mean difference was 2.56 vs. 1.03, respectively). Second, the driving context played a role in the way drivers deactivated the automation mode. The findings revealed that drivers halted the automation in response to a hazard more often during urban driving than highway driving. These results imply that drivers are aware of the driving context, to some extent, and are probably more attentive to the road during driving in urban environments. These results are in line with the findings of Ortiz-Peregrina et al.(2020) indicating drivers are aware of the driving context and adjusting their behavior accordingly. Ortiz-Peregrina et al. (2020) found that drivers reduce their speed when faced with more demanding driving conditions such as curved road segments, with more traffic intersections and when parked cars are present. Further support to our findings comes from a study by Oviedo-Trespalacios et al. (2017), who investigated how driving task demands influence speed adaptation of distracted drivers under various road infrastructure and traffic complexity conditions. Their results indicated that complex road environments (such as urbanization, car-following situations along suburban roads, and curved road alignment) significantly influenced speed adaptation behavior. In their study, distracted drivers selected a lower speed while driving along a curved road or during car-following situations.

Furthermore, the findings revealed that also experience with the system affected drivers' performance. On their 2nd drive, as participants accumulated experience with the partially automated driver and the hazardous scenarios, they were less likely to identify hazards and deactivate the automation than in the first drive regardless of NDRT's existence. Two possible reasons may explain this. First, in their second drive, participants who were already familiar with the scenarios to some extent from their first drive learned that the hazards never materialize. Second, during their second drive, drivers might have developed higher trust in the automated system and became complacent. Overtrust in an automated system may lead to adverse behavioral adaptation effects. Therefore, consistent with Hergeth et al.'s (2016) findings, they were less likely to identify hazards and decided not to deactivate the automation and let the automated system handle the situation.

Thus far, the findings reveal a complex picture regarding drivers' awareness of the roadway environment under L2 driving conditions in the presence of an NDRT. Following the control level of Schömig and Metz (2013), drivers in this study engaged with an NDRT whenever they felt safe. When drivers decided to engage with an NDRT, they were less aware of the driving context and the upcoming hazard and thus, took active control less often than when not deciding to engage with it. However, when drivers did not engage with an NDRT (regardless of the experimental condition), they could better distinguish between different traffic environments and allocate their attention differentially. There was also an effect of accumulated experience with the automated system and the hazardous scenarios. Nevertheless, the current study cannot determine whether these differences result from the experience with the automation system or familiarity with the hazardous scenarios in the second drive.

Next, consistent with our second hypothesis, the NASA-TLX analysis results show that drivers provided significantly higher NASA-TLX ratings in the presence of an NDRT than in its absence. Hence, we argue that engagement with a visual-manual NDRT while driving under L2 driving conditions causes a higher mental workload because both tasks require the same limited visual attention resources. This argument is further supported by the HRV analysis showing that both measurements of drivers' HRV (log(QuotientRMSSD) and log(QuotientSubRMSSD)) were smaller in the presence of an NDRT than in its absence. These results align with the HRV literature reporting lower HRV values under high mental effort (Lohani et al., 2019; Stapel et al., 2019) and HRV sensitivity to increased mental workload due to excess vigilance and situational awareness demands of the task (Lohani et al., 2019).

Finally, consistent with the third hypothesis, the hazard identification analysis showed that the time-sharing strategy between the NDRT and the driving monitoring task affected the probability of identifying hazards (see section 4.4). We defined several time-sharing strategies based on the analyses of the following variables (1) total duration of glances at the NDRT, (2) number of glance shifts from the road to the NDRT, and (3) total duration of engagement with the NDRT. The first time-sharing strategy is the "**unsafe driver**." Drivers who applied this time-sharing strategy focused on the task and continuously glanced at it for long periods. As the cumulative dwell time on the NDRT increased, the driver was more engaged with the task, and it consumed more attentional resources, thus leading to poor hazard perception performance. The second time-sharing strategy is the "**safe driver**." Drivers who applied this strategy tended to repeatedly shift their glances between the road and the NDRT without increasing the total engagement time with the NDRT. The extent to which the driver shifted his attention between the NDRT and the road reflects the driver's involvement in the driving task and the probability of identifying hazards. The larger the number of glance shifts, the higher involvement in the driving task, the better hazard perception performance. The third time-sharing strategy is also defined as the "**safe driver**". Drivers who applied this strategy tended to extend their total engagement time with the NDRT, parsing the interaction with the NDRT into small action units, allowing them to shift their glances between the NDRT and the road. As the total time of engagement with the NDRT increased, the driver tended to allocate more attention to the roadway, and therefore, was more likely to identify hazards.

6. Conclusions, limitations, and future work

The arguments raised in the discussion emphasize that although automation brings many benefits, there is a consistent pattern that engagement with a visual-manual NDRT under L2 driving conditions negatively affects driving performance and increases mental workload. This study focused on the "control" level in the model of Schömig and Metz (2013). The study showed that when young-experienced drivers engaged with an NDRT during L2 driving, they had to divide their attention between engagement with the NDRT and monitoring the driving task and automation performance. This mentally demanding dual-tasking work increased mental workload and impaired hazard awareness. We postulate that engagement with an NDRT while driving under L2 conditions competes simultaneously with various driving-related processes, such as monitoring the automation performance and perceiving road hazards, over limited visual attention resources. This study concludes that car manufacturers and in-vehicle interface designers should focus on countermeasures that can mitigate the adverse effects of driver distraction on driving safety, especially during L2 driving, where driver distraction is becoming prevalent. Some countermeasures should focus on monitoring drivers' attention and alerting them if they are distracted. Other countermeasures should focus on preventing engagement with distracting tasks in the first place, and especially during mentally demanding driving situations. Next, we discuss some limitations of this study.

First, the study was conducted in a driving simulator which is different from a real-world driving environment. Therefore, future studies should further conduct field studies to support the simulation results. Secondly, the study's sample was limited to young-experienced and relatively homogeneous drivers (undergraduate students). Thus, the participants' population in future studies should be more heterogeneous and include other age groups (e.g., older drivers) and different levels of education and ages. Third, this study consisted of two similar drives with the same hazardous scenarios in each drive. Although the scenarios' order was randomized between drives, there might have been learning effects that influenced drivers' performance in the second drive. Future studies should include a larger variety of scenarios that will be different from one drive to the next to understand behavioral adaptation effects better.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study was performed within the MEDIATOR project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement 814735. The first author is supported by a scholarship from the Israeli Smart Transportation Research Center (ISTRC). The authors are solely responsible for the study's contents, which do not necessarily represent the official views of the Mediator EU-project or ISTRC

Appendix A Additional analysis of scanning patterns

One way of detecting visual distraction is the AttendD algorithm. This algorithm considers single long off-road glances separated by short-duration glances back at the forward roadway, which do not allow enough time to restore full awareness of the road. When the driver is focused on the road environment and is not distracted by a visual task, the AttendD buffer is full and equals 2 s. This buffer will deplete when the driver glances away from the forward roadway. When this buffer reaches 0 s, the driver is classified as visually distracted. We applied the algorithm to the experimental data as follows. First, for each scenario where the driver engaged with an NDRT, the recorded duration of glances at the task and the forward roadway were inserted into a time vector, including an indication of their position (forward roadway or NDRT). Whenever a glance shifted from the NDRT to the forward roadway, a cost (delay) of 0.1 s was subtracted from the on-road glance duration (See Ahlström et al., 2021). We used another vector for keeping the AttenD buffer's values. At the beginning of each engagement with an NDRT, the buffer's value was 2 s. The buffer's value was re-calculated and inserted into the vector for each glance shift between the roadway and the task. When the driver shifted his glance to the NDRT, the glance duration was subtracted from the buffer's value. When the driver moved his glance back at the road, the glance duration was added to the buffer's value (the 0.1 s delay is already embedded in the glance duration vector). Finally, at the end of this procedure, values above 2 s were set at 2 s, and negative values at zero. By implementing this algorithm, three different scanning patterns emerged (Fig. A.1).. The first one is a pattern of a "Risky driver" (Fig. A.1- panel (a)), which is characterized by a long period (~12 sec) that the attention buffer equals 0. That means that the driver had long glances toward the task that depleted the attention buffer, and he did not pay enough attention at the road. The second pattern is that of a "Safe driver" (Fig. A.1- panel (b)), characterized by multiple glance shifts and positive values of attention buffer, which mostly does not reach the maximum value of 2 s. This second pattern means that the driver properly divided his attention between the task and the road with extra safety that never depleted the allowed buffer. The third pattern is that of an "Extra safe driver" (Fig. A.1- panel (c)), characterized by positive values of the attention buffer, which mostly equals 2 s and never reaches zero. This pattern means that the driver divided his attention properly between the task and the road and never exceeded the allowed buffer. All three drivers achieved a task score of 2/2 in this example. We implemented the AttendD algorithm on three drivers driving the same driving scenario under L2 with an NDRT. Three different scanning patterns emerged: risky, safe, and extra safe driver. The risky driver had long glances toward the task that depleted the attention buffer, and he did not pay enough attention toward the road. On the contrary, the safe driver divided his attention properly between the task and the road and never depleted the allowed buffer. The extra safe driver divided his attention properly between the task and the road with extra safety that never exceeded the permitted buffer. These strategies showed that engagement with an NDRT differently affects each driver and might cause adverse effects. Also, drivers' personal driving styles should be taken into account. Finally, the AttendD algorithm was applied on three drivers only in one specific scenario to show its implementation potential in our study to reveal different scanning patterns. We suggest that future work should implement this algorithm on a larger sample of drives under different scenarios. This implementation may reveal other scanning patterns that could have been missed here. Future work could extend the usage of this algorithm by using it in real-time in a way where the autonomous vehicle could classify the driver's scanning pattern and accordingly if it is necessary to alert the driver when the buffer gets empty.

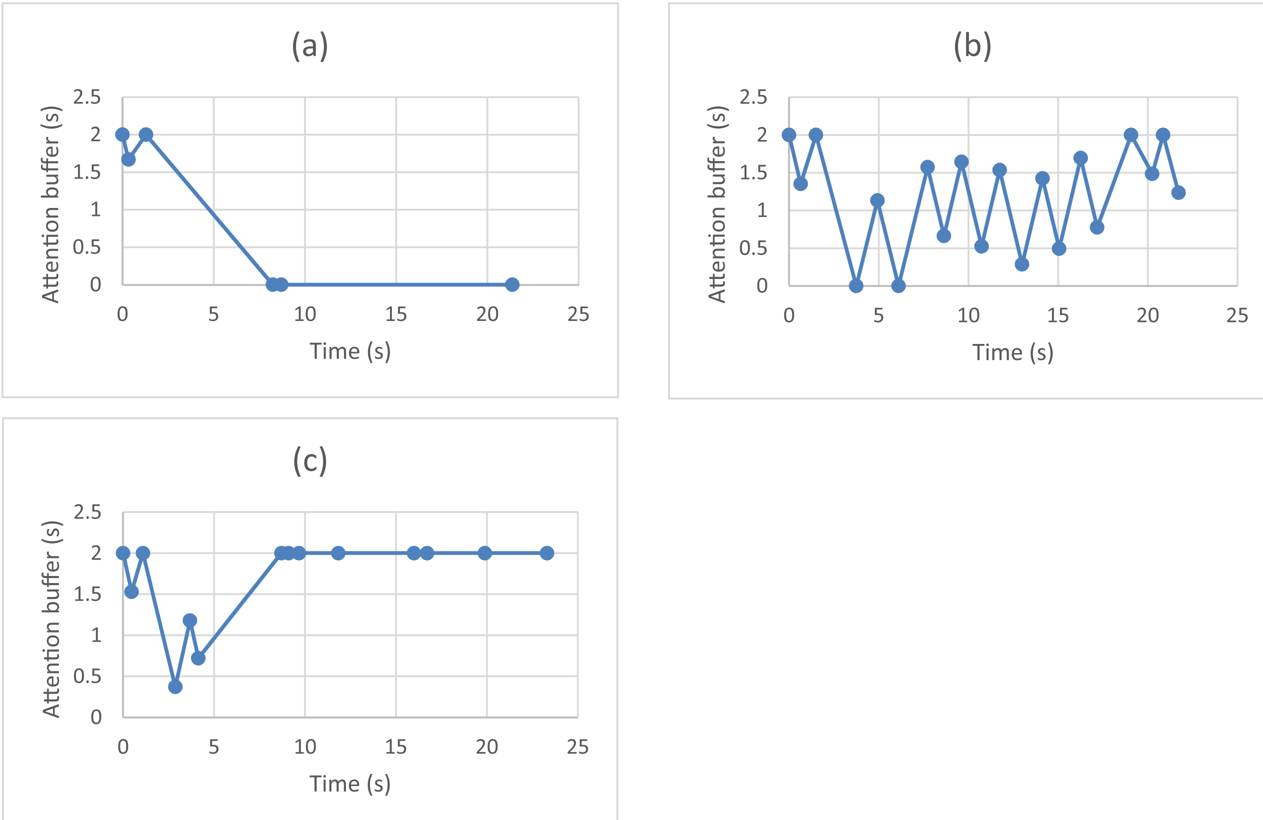


Fig. A1. Distration visualiation by attenD algorithm (a) Risky driver. (b) Safer driver. (c) Extra safer driver.

References

Ahlström, C., Georgoulas, G., Kircher, K., 2021. Towards a context-dependent multi-buffer driver distraction detection algorithm. IEEE Trans. Intell. Transp. Syst.

Ahlstrom, C., Kircher, K., Kircher, A., 2013. A gaze-based driver distraction warning system and its effect on visual behavior. IEEE Trans. Intell. Transp. Syst. 14 (2), 965–973.

Borowsky, A., Horrey, W.J., Liang, Y., Garabet, A., Simmons, L., Fisher, D.L., 2016. The effects of brief visual interruption tasks on drivers' ability to resume their visual search for a pre-cued hazard. Accid. Anal. Prev. 93, 207–216. https://doi.org/10.1016/j.aap.2016.04.028.

Burge, R., Chaparro, A., 2018. An investigation of the effect of texting on hazard perception using fuzzy signal detection theory (fSDT). Transp. Res. Part F: Traffic Psychol. Behav. 58, 123–132.

Carsten, O., Lai, F. C. H., Barnard, Y., Jamson, A.H., 2012. Control Task Substitution in Semiautomated Driving : Does It Matter What Aspects Are Automated ? 1985. doi: 10.1177/0018720812460246.

Choi, J.K., Ji, Y.G., 2015. Investigating the importance of trust on adopting an autonomous vehicle. Int. J. Human-Comput. Interaction 31 (10), 692–702. https://doi.org/10.1080/10447318.2015.1070549.

Cunningham, M.L., Regan, M.A., 2018. Driver distraction and inattention in the realm of automated driving. IET Intel. Transport Syst. 12 (6), 407–413.

De Winter, J.C.F., Stanton, N.A., Price, J.S., Mistry, H., 2016. The effects of driving with different levels of unreliable automation on self-reported workload and secondary task performance. Int. J. Veh. Des. 70 (4), 297–324. https://doi.org/10.1504/ijvd.2016.076736.

Fisher, D.L., Lohrenz, M., Moore, D., Nadler, E.D., Pollard, J.K., 2016. Humans and intelligent vehicles: the hope, the help, and the harm. IEEE Trans. Intell. Veh. 1 (1), 56–67. https://doi.org/10.1109/TIV.2016.2555626.

Ginsburg, A.P., 2003. Contrast sensitivity and functional vision. Int. Ophthalmol. Clin. 43 (2), 5–15.

Hart, S.G., 2006. NASA-task load index (NASA-TLX); 20 years later. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 50 (9), 904–908.

Hart, S.G., Staveland, L.E., 1988. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In: Advances in Psychology, Vol. 52. Elsevier, pp. 139–183.

Hoover, A., Singh, A., Fishel-Brown, S., Muth, E., 2012. Real-time detection of workload changes using heart rate variability. Biomed. Signal Process. Control 7 (4), 333–341.

Horswill, M.S., Hill, A., Wetton, M., 2015. Can a video-based hazard perception test used for driver licensing predict crash involvement? Accid. Anal. Prev. 82, 213–219.

Krishnan, A., Samuel, S., Yamani, Y., Romoser, M.R.E., Fisher, D.L., 2019. Effectiveness of a strategic hazard anticipation training intervention in high risk scenarios. Transp. Res. Part F: Traffic Psychol. Behav. 67, 43–56. https://doi.org/10.1016/j.trf.2019.10.004.

Lin, R., Liu, N., Ma, L., Zhang, T., Zhang, W., 2019. Exploring the self-regulation of secondary task engagement in the context of partially automated driving: A pilot study. Transp. Res. Part F: Traffic Psychol. Behav. 64, 147–160. https://doi.org/10.1016/j.trf.2019.05.005.

Lohani, M., Payne, B.R., Strayer, D.L., 2019. A review of psychophysiological measures to assess cognitive states in real-world driving. Front. Hum. Neurosci. 13, 57.

Martens, M. H., Van Winsum, W., 2000. Measuring distraction: the peripheral detection task. TNO Human Factors, Soesterberg, Netherlands.

Naujoks, F., Purucker, C., Neukum, A., 2016. Secondary task engagement and vehicle automation – comparing the effects of different automation levels in an on-road experiment. Transp. Res. Part F: Psychol. Behav. 38, 67–82. https://doi.org/10.1016/j.trf.2016.01.011.

Nilsson, E.J., Victor, T., Ljung Aust, M., Svanberg, B., Lindén, P., Gustavsson, P., 2020. On-to-off-path gaze shift cancellations lead to gaze concentration in cognitively loaded car drivers: a simulator study exploring gaze patterns in relation to a cognitive task and the traffic environment. Transp. Res. Part F: Traffic Psychol. Behav. 75, 1–15. https://doi.org/10.1016/j.trf.2020.09.013.

Noble, A.M., Miles, M., Perez, M.A., Guo, F., Klauer, S.G., 2021. Evaluating driver eye glance behavior and secondary task engagement while using driving automation systems. Accid. Anal. Prev. 151, 105959.

Olsson, S., Burns, P.C., 2000. Measuring driver visual distraction with a peripheral detection task. Obtained from August.

Ortiz-Peregrina, S., Oviedo-Trespalacios, O., Ortiz, C., Casares-López, M., Salas, C., Anera, R.G., 2020. Factors determining speed management during distracted driving (WhatsApp messaging). Sci. Rep. 10 (1), 1–10.

Oviedo-Trespalacios, O., Haque, M.M., King, M., Washington, S., 2017. Effects of road infrastructure and traffic complexity in speed adaptation behaviour of distracted drivers. Accid. Anal. Prev. 101, 67–77.

Paxion, J., Galy, E., Berthelon, C., 2014. Mental workload and driving. Front. Psychol. 5, 1344.

Peters, R., Stavrinos, D., 2017. Technology Transfer: Distracted Driving, Overview Summary of Ways to Alleviate.

Pohlmeyer, B. M. B. R. A. E., Coughlin, J. F., 2008. The association between heart rate reactivity and driving performance under dual task demands in late middle age drivers. Advances in Transportation Studies, Int. J., Special Issue,(2008), 53–70.

Recarte, M.A., Nunes, L.M., 2003. Mental workload while driving: effects on visual search, discrimination, and decision making. J. Experimental Psychol.: Appl. 9 (2), 119.

Rosenzweig, J., Bartl, M., 2015. A review and analysis of literature on autonomous driving. E-Journal Making-of Innovation 1–57.

Rubio, S., Díaz, E., Martín, J., Puente, J.M., 2004. Evaluation of subjective mental workload: a comparison of SWAT, NASA-TLX, and workload profile methods. Appl. Psychol. 53 (1), 61–86. https://doi.org/10.1111/j.1464-0597.2004.00161.x.

SAE. (2018). SAE International Releases Updated Visual Chart for Its "Levels of Driving Automation" Standard for Self-Driving Vehicles.

Salvucci, D.D., Taatgen, N.A., 2009. Toward a unified theory of the multitasking continuum : from concurrent performance to task switching. Interruption, and Resumption. 1819–1828.

Savage, S.W., Potter, D.D., Tatler, B.W., 2013. Does preoccupation impair hazard perception? A simultaneous EEG and Eye Tracking study. Transp. Res. Part F: Traffic Psychol. Behav. 17, 52–62. https://doi.org/10.1016/j.trf.2012.10.002.

Schömig, N., Metz, B., 2013. Three levels of situation awareness in driving with secondary tasks. Saf. Sci. 56, 44–51. https://doi.org/10.1016/j.ssci.2012.05.029.

Solís-Marcos, I., Ahlström, C., Kircher, K., 2018. Performance of an additional task during level 2 automated driving: an on-road study comparing drivers with and without experience with partial automation. Hum. Factors 60 (6), 778–792. https://doi.org/10.1177/0018720818773636.

Stapel, J., Mullakkal-Babu, F.A., Happee, R., 2019. Automated driving reduces perceived workload, but monitoring causes higher cognitive load than manual driving. Transp. Res. Part F: Traffic Psychol. Behav. 60, 590–605. https://doi.org/10.1016/j.trf.2018.11.006.

Taubman-Ben-Ari, O., Mikulincer, M., Gillath, O., 2004. The multidimensional driving style inventory – scale construct and validation. Accid. Anal. Prev. 36 (3), 323–332. https://doi.org/10.1016/S0001-4575(03)00010-1.

Tobii, A.B., 2018. Tobii pro glasses 2.

Victor, T., 2005. Keeping eye and mind on the road. Acta Universitatis Upsaliensis.

Vlakveld, W., van Nes, N., de Bruin, J., Vissers, L., van der Kroft, M., 2018. Situation awareness increases when drivers have more time to take over the wheel in a Level 3 automated car: a simulator study. Transp. Res. Part F: Traffic Psychol. Behav. 58, 917–929. https://doi.org/10.1016/j.trf.2018.07.025.

Wickens, C.D., Gutzwiller, R.S., Santamaria, A., 2015. Discrete task switching in overload: a meta-analyses and a model. Int. J. Hum Comput Stud. 79, 79–84.

Wilson, K.M., Yang, S., Roady, T., Kuo, J., Lenné, M.G., 2020. Driver trust & mode confusion in an on-road study of level-2 automated vehicle technology. Saf. Sci. 130 (June), 104845 https://doi.org/10.1016/j.ssci.2020.104845.

Winter, J.C.F. De, Happee, R., Martens, M.H., Stanton, N.A., 2014. Effects of adaptive cruise control and highly automated driving on workload and situation awareness: a review of the empirical evidence. Transp. Res. Part F: Psychol. Behav. 27, 196–217. https://doi.org/10.1016/j.trf.2014.06.016.

Yamani, Y., Samuel, S., Knodler, M.A., Fisher, D.L., 2016. Evaluation of the effectiveness of a multi-skill program for training younger drivers on higher cognitive skills. Appl. Ergon. 52, 135–141. https://doi.org/10.1016/j.apergo.2015.07.005.

Young, K., Regan, M., Hammer, M., 2007. Driver distraction: a review of the literature. Distracted Driving 2007, 379–405.