Int. J. Human–Computer Studies 186 (2024) 103247

Contents lists available at ScienceDirect

International Journal of Human - Computer Studies

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

Effects of context-sensitive distraction warnings on drivers' smartphone use and acceptance: A long-term naturalistic field study

Tuomo Kujala[a,*], Hilkka Grahn[a], Jakke Mäkelä[b], Johanna Silvennoinen[a], Timo Tokkonen[c]

[a] Faculty of Information Technology, University of Jyväskylä, Jyväskylä, Finland
[b] Social Insurance Institution of Finland (Kela), Turku, Finland
[c] University of Oulu, Oulu, Finland

A R T I C L E I N F O

Keywords:
Driver distraction
Smartphone
Attention monitoring
Warning system
Acceptance

A B S T R A C T

Driver distraction by smartphone use is a major contributor in traffic accidents. Context-sensitive driver (in) attention monitoring and warning systems might mitigate the associated risks. However, few naturalistic studies are yet available on the effects of such systems. In this paper, utility of context-sensitivity in inattention monitoring was studied by a smartphone-based context-sensitive distraction warning system. The warnings were based on driver's phone use and the attentional demands of the upcoming traffic environment. The system's effects on 26 heavy smartphone users' phone usage and acceptance were analyzed after a within-subject naturalistic study with 12 weeks of control (warnings off) and 12 weeks of interventions (warnings on). The system decreased odds that the drivers would touch their smartphones in reminder areas that were defined a priori as high demanding for attention. Against expectations, the system had no effect in urban road environments. The drivers reported that they had paid more attention to traffic because of the system and that the warnings were acceptable and useful, even if annoying. Similar systems' safety effects should be further studied. No eye-tracking or driving performance measures were collected and thereby it is questionable if there was a true positive effect on participants' attention. However, the findings suggest that (in)attention warning systems might benefit from adaptation of the warnings to the upcoming driving demands. The findings can be utilized for the development of proactive and context-sensitive (in)attention monitoring and distraction mitigation systems.

1. Introduction

Driver monitoring systems (DMS) are developed and utilized to detect a driver's state in terms of drowsiness, emotions, inattention, and distraction, which is influenced by different factors, such as phone usage. The purpose of DMS is to observe the driver's status in real time and provide necessary assistance and alerts (Khan and Lee, 2019). Driver (in)attention monitoring systems, in particular, are meant to detect driver inattention and warn the driver to pay attention to the forward roadway. However, current commercial attention monitoring systems do not utilize contextual information to an extent that could be useful for the valid and reliable classification of the driver as attentive or inattentive based on the variable attentional demands of the driving task (Anon., AAA, 2022; Ahlström et al, 2021).

In general, three main types of data input are used to recognize a driver's state (Ahlström et al., 2021; Tran et al., 2018; McDonald et al., 2020): physiological data, vehicle control data, and visual data. Physiological data are obtained from the driver's body using electrodes to measure the driver's attentiveness (Khan and Lee, 2019). Vehicle control data are used to observe driving patterns and detect declines in driving performance (Ramzan et al., 2019; Wang et al., 2022). Systems based on vehicle control data are trained to recognize, for instance, potential forward collisions (Iranmanesh et al., 2018), distracted, drowsy, or drunk driving (Shahverdy et al., 2020), or hands removed from the steering wheel (AAA, 2022). Visual data are obtained using images or videos to recognize the driver's facial expressions and eye and body movements (Tran et al., 2018). Typically, computer vision methods are utilized for detecting these behaviors (Shahverdy et al., 2020).

Besides drowsiness, visual data are used to detect driver distraction based on gaze direction estimates and derived gaze/head direction (Ahlstrom et al., 2013; Herbers et al., 2023). *Eyes off the road* is perhaps the most used metric to determine whether a driver is distracted (Halin et al., 2021). However, only foveal vision can be measured with these

* Corresponding author at: Mattilanniemi 2, 40014 University of Jyväskylä, Finland.
E-mail address: tuomo.kujala@jyu.fi (T. Kujala).

https://doi.org/10.1016/j.ijhcs.2024.103247
Received 16 October 2023; Received in revised form 13 February 2024; Accepted 19 February 2024
Available online 20 February 2024

techniques, not peripheral vision (Ahlström et al., 2021). Furthermore, the first gaze-based DMS labeled drivers inattentive or distracted as soon as the driver's gaze was directed away from the forward roadway (Ahlström et al., 2021). As this is quite a strict approach because of spare visual capacity in driving (Kujala et al., 2023), later attention monitoring systems have adopted an approach that allows drivers to gaze shortly away from the forward roadway, with the common threshold being 2 s (Ahlström et al., 2021; 2013; AAM, 2006; Klauer et al., 2006; NHTSA, 2013).

However, current commercial driver attention monitoring techniques or systems (Anon., AAA, 2022; Fredriksson et al., 2021) do not utilize contextual information, such as traffic density, weather, and type of situation ahead (e.g., straight road or a junction), from the dynamic context in which driving happens. The attentional demands of driving vary depending on context, and therefore, the spare attentional capacity varies too (Kujala et al., 2023). For instance, 2 s of glance time off forward can be too strict, too allowing, or even necessary (e.g., to glance at mirrors), depending on the situational demand (e.g., Ahlström et al., 2013; Han et al., 2023). A recent study by Han et al. (2023) suggested that different distraction alert timer settings are required for different roads based on their geometric characteristics. Accordingly, Ahlström et al. (2021) suggested that driver attention monitoring systems should also consider the situational context. As a solution, they introduced an algorithm called AttenD2.0, which incorporates a level of context dependency. The algorithm uses both eye-tracking data and a model of the surrounding environment to classify the driver as attentive or inattentive. At present, the AttenD2.0 algorithm has been purely theoretical and has been studied so far only in well-controlled experimental conditions (Ahlström et al., 2021). The difficulties in implementing a real-world version of AttenD2.0 are related to the need for a computationally expensive world model that is also challenging to create.

The lack of context dependency can affect how driver attention monitoring systems are accepted and, therefore, their effectiveness. According to Lubkowski et al. (2021), drivers may find the driver state monitoring system confusing and may not necessarily understand why the system gives alerts, or they may feel that the system is not measuring *what matters*. Strand et al. (2011) found that drivers have problems understanding what activates the driver state monitoring system and that it is even turned off if the alerts are not understood. Additionally, the false alarms of monitoring systems diminish trust, compliance, and acceptance (Lees and Lee, 2007; Roberts et al., 2012).

In response to these challenges, Kujala et al. (2016) designed and tested a context-sensitive distraction warning application (VisGuard: “Vision Guard”) that ran in the background on a driver's smartphone and tracked the driver's gaze and location. The results of a track study suggested that while warnings based on in-car glance duration were found to be ineffective in decreasing glance durations, location-based warnings based on upcoming demanding driving situations seemed to increase drivers' glance time on the road. The warnings were also found to be acceptable to the participants. However, whether the findings of this track study can be generalized to longer-term daily driving in real traffic is questionable.

In this research, the effects of a smartphone-based context-sensitive distraction warning system (VisGuard 2.0) are studied in a controlled naturalistic study involving 26 drivers for the duration of six months (three months: control, three months: intervention). The main purpose is to study the utility of context sensitivity in driver (in)attention monitoring systems to mitigate driver distraction caused by smartphone use. The system was context sensitive in that its display of visual reminders and auditory warnings on smartphone use was based on where the driver was located on map data. The driver was visually reminded about phone usage in areas defined as attentionally high demanding (e.g., when approaching an intersection) and they were auditorily warned if there was continued use within these areas. Even if not as sophisticated, this level of context dependency is computationally much simpler than that of the world model that would be required for an efficient AttenD2.0 implementation (Ahlström et al., 2021).

As a general effect, it was expected that the system would decrease the number of touches on the smartphone in areas with high attentional demands, where reminders and warnings were provided (cf. Kujala et al., 2016). Studies by Caird et al. (2008) and Wang et al. (2024) indicated that in-vehicle visual warning signs can facilitate drivers' attentiveness when approaching intersections. The density of intersections and crosswalks and, therefore, the density of areas with high attentional demands are much more intense in urban than in other road environments (especially on highways or rural roads). Therefore, it was expected that the drivers would reduce their phone use the most in these areas on urban roads. Finally, a possible effect could be an overall decrease in smartphone usage while driving. Based on these expected effects, the following hypotheses were proposed:

- H1. The system reduces smartphone touches in areas of high attentional demands (i.e., reminder areas).
- H2. The system reduces smartphone touches the most in reminder areas on urban roads compared to other road environments.
- H3. The system decreases smartphone usage while driving.

A negative outcome would be an increase in phone usage while driving because of a possibly false sense of security provided by the system. Further, it was expected that the drivers would understand the reminders and warnings, and that they would experience these as acceptable and useful.

2. VisGuard 2.0

The broader design principles behind the context-sensitive VisGuard 2.0 distraction warning system were based on naturalistic studies indicating an association between drivers' phone use and safety-critical events in traffic (Bálint et al., 2020; Klauer et al., 2014), occlusion studies indicating that the attentional demands of driving vary by driving context (Kujala et al., 2023), drivers' general requirements and abilities to be able to stop their cars in time before crashing into other road users (American Association of State Highway and Transportation Officials [AASHTO], 2011), proactive instead of reactive warnings (Kujala et al., 2016), and the utilization of multimodal warnings instead of warnings relying on the use of focal vision (e.g., Wang et al., 2022). The general aim of the system design was to promote safer driving by monitoring the driver's phone use and reminding a multitasking driver to be attentive when necessary, based on the detection of attentionally high-demanding areas ahead. These are areas where there is a high probability of another road user intersecting the driver's trajectory, and where the driver should visually observe whether crossing the area is safe before entering it. In the current implementation, intersections, junctions, roundabouts, crosswalks, and railroad crossings were chosen as such areas based on the ease of recognition of these locations from map data. The idea was to remind the driver to be attentive, not based on the detection of objects but based on the increased probability of an object crossing the driver's trajectory in these locations ahead.

2.1. Technical implementation

The VisGuard 2.0 prototype consisted of three parts:

1. A smartphone (Samsung Galaxy XCover3) was attached rigidly to the dashboard of the participant's car in such a position that it would be in the line of sight whenever the driver was looking forward without blocking the driver's view (Fig. 1). Whenever motion of the car was detected, this dashboard phone measured the location of the car at 1 s intervals based on the global positioning system (GPS) signal. At every GPS fix, the phone's software calculated its location and direction on the road network, and the distance to the next intersection, junction, roundabout, crosswalk, or railroad crossing (i.e., area

Fig. 1. Typical arrangement of the dashboard phone and the visual reminder. The exact position depended on the car and the height of the driver; the target was to find a central location that did not interfere with the driver's line of sight but was still visible to the driver, and where nothing blocked the view of the rear camera.

of high attentional demand) based on Open Street Map (https://www.openstreetmap.org/) data that was stored on the phone. Based on velocity information, the dashboard phone then estimated whether the car was approaching an area of high attentional demand and, if needed, showed a visual reminder on its display or gave an auditory warning (Figs. 1 and 2). The dashboard phone collected and stored the GPS, acceleration and the participant's phone usage data, for later upload to the backend.

2. A tailored background application was installed on each participant's personal smartphone, and it monitored touches on the phone's touchscreen in the background. The phone was automatically connected by Wi-Fi to the dashboard phone. Whenever the keylock state changed (open or locked) or a touch was registered, the phone sent a message to the dashboard phone with information about the application that was used. The dashboard phone also recorded a photo of the forward roadway on each touch on a participant's smartphone.
3. Data from the dashboard phone were uploaded to a backend system. As the positioning software required high real-time performance, data were not uploaded when the car was in motion but only after it had been stationary for a predefined period.

2.2. Reminder area definition

The function of the VisGuard 2.0 software was simple in principle—to determine, at any given moment, whether the driver could safely come to a full stop, if needed, before the next area of high attentional demand in the map data (i.e., intersection, junction, roundabout, crosswalk, or railroad crossing). These are areas where there is a high demand to observe whether there are other road users who could come into the driver's trajectory and, in the worst-case scenario, before which the driver should be able to stop the car. The intervention to the driver's attentiveness by the system was intended to happen before there was an urgent need to react to an object ahead. There were two parameters in the associated calculation:

1. Brake reaction time (BRT): This refers to how long it takes for the driver to initiate braking. This was assumed to be 1 s after Summala (2000).
2. Braking distance (BD): This refers to how long it takes for the car to come to a full stop with a deceleration rate of a. From basic physics, we know that the stopping distance of an object from an initial velocity of v and a deceleration rate a equals $0.5*v^2/a$ (where v is in m/s and a is in m/s^2); that is, $BD = 0.5*v^2/a$. According to Anon. AASHTO

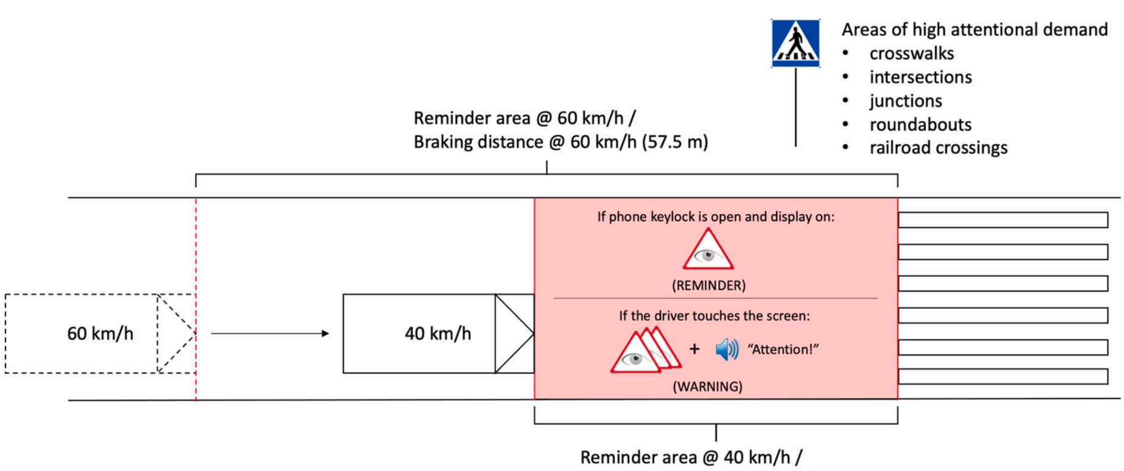

Fig. 2. Illustration of the reminder area's definition, a reminder, and a warning in the case of a crosswalk.

(2011), approximately 90 % of all drivers decelerate at rates greater than 3.4 m/s^2 when confronted with the need to stop because of an unexpected object in the roadway. Drivers can still maintain the maneuverability of the car under this acceleration. Therefore, 3.4 m/s^2 was used for *a*.

The distance the car then travels before halting is

$$D = BRT^{*}v + BD \quad (1)$$

or, in our case, $D = v + 0.5^{*}v^2/3.4$. Distance D was used to define the length of a reminder area until the area of high attentional demand in the map data. If the distance to the next area of high attentional demand was smaller than this, the driver was considered to be inside an area in which any distraction was considered risky. Based on the BD, the length of this reminder area was dependent on the speed of the car. Some technical corrections to this basic algorithm had to be made to account for the 1 s sampling rate of the GPS, but fundamentally, the equation above was used. The accuracy of the timing of the reminders on the dashboard phone was confirmed with various speeds (20–100 km/h), an in-car video camera, and sticks on the roadside at an intersection on an unused road.

2.3. Functionality – reminder vs. warning

What happened in a reminder area depended on the driver's phone use. There were two types of possible actions by the dashboard phone (Fig. 2): reminder and warning.

1. A reminder (static visual alert only, Fig. 1) was given if the keylock in the driver's phone was open while driving within a reminder area, but the driver did not touch their smartphone – which still indicated possible intention of use. The visual reminder was meant to be perceivable by peripheral vision in the normal visual field of the driver (Fig. 1) and to serve as a reminder that the driver is in an area where smartphone use might distract them from observing critical road events ahead.
2. A warning (the reminder icon blinking twice, accompanied by a loud female voice saying "attention!" in Finnish ["huomio!"]) was given for each touch on the driver's smartphone within a reminder area. The animated icon and voice warning were meant to stop the driver from interacting with the phone if the driver continued using the phone, regardless of the visual reminder.

The visual reminder icon was designed and located in such a way that it could be detected by peripheral vision but that it would not draw the driver's attention when the purpose was to direct the driver's gaze toward the forward roadway. In the previous version of VisGuard, the icon was displayed on the driver's phone (Kujala et al., 2016). Again, this could have drawn the driver's attention to the phone instead of directing it toward the road. In addition, an icon blocking the use of the phone would have been an unwanted confounding variable. The blinking warning icon, together with the auditory warning, was meant to support multimodal correspondence, that is, the integration of spatially and temporally equivalent warning information that does not require the use of focal vision for detection (Spence, 2011). A schematic of the working principles of the reminders and warnings is provided in Fig. 2.

3. Methods

3.1. Participants

The participants were recruited via social media and newspaper advertisements. Over 200 applications were received. Based on the pool of applicants, the best attempt was made for a balanced sample of males and females, as well as a wide range of ages. The aim was to recruit a representative random sample from the target population, but drivers who reported driving in both urban and rural areas were favored in order to have a representative sample for testing H2. A sample of 32 volunteers was initially selected based on multiple criteria, including how much they reported driving per year. All of the recruited participants were Finnish drivers who reported driving almost daily and using their smartphones frequently and regularly while driving. The participants had to have insured cars and Android-based smartphones in use for the study period.

There were technical issues with the data collection with some participants' smartphones (e.g., application freezing, security updates, or disabled Wi-Fi connectivity), resulting in fewer days of data collection than expected, but the number of recorded locations and the number of touches on the smartphone were considered sufficient for 26 participants to enable reliable analyses (see Descriptive statistics). This final sample of 26 participants was composed of 18 male and 8 female participants with a mean age of 39 years ($SD = 11.8$, range: 18–64). Their self-reported driving experiences varied from 10,000 to 65,000 km ($M = 28{,}522$, $SD = 14{,}491$ km).

This research complied with the tenets of the Declaration of Helsinki and was approved by the Institutional Review Board at the University of Jyväskylä. Informed consent was obtained from each participant. The participants were allowed to withdraw from the study at any point. They were rewarded with the dashboard smartphone, its holder, and the car charger used in the experiment (approximately worth of 180 EUR in total as new).

3.2. Experimental design

The experimental design was a within-subject experiment with the experimental condition as an independent variable (IV) with two levels (control vs. experiment, i.e., intervention by the distraction warning system). The control phase preceded the experiment phase, and both lasted for 12 calendar weeks. The order of the conditions was not counterbalanced, as the purpose was to first measure the drivers' normal behavior with the smartphone in traffic and then study the effects of the intervention by the system on these behaviors. Their behaviors would probably have been affected by the system use if the control condition followed the experiment condition.

1. Furthermore, the road environment was used as an additional IV for testing H2 (a 2 × 4 design: experimental condition × road environment). The driven road environments were identified from the map data based on the Finnish Digiroad classification system (https://vayla.fi/vaylista/aineistot/digiroad) as follows:
2. Highway (default 120 km/h, sometimes 100 km/h). No opposing traffic. As an example of the temporal density of reminder areas, at 120 km/h, the time between junctions that were 400 m apart was 12 s.
3. Main rural road (default 80 km/h, sometimes 100 km/h, or 60 km/h at some intersections). By default, the driver has a right-of-way. For example, at 80 km/h, the time between intersections or junctions that were 200 m apart was 9 s.
4. Local rural road (speed limit varies, default either 80 km/h or 50 km/h). The driver may or may not have a right-of-way. The distances and times between reminder area centers were often smaller than those on main rural roads but larger than those on urban roads.
5. Urban road (default 40 or 50 km/h, sometimes 30 km/h). The minimum distances between two intersections in a city center could be around 40 m, corresponding to 3.6 s with a speed of 40 km/h.

Naturally, the densities of the reminder areas per road environment varied based on the specific geometry of the environment and the driving speed.

Three dependent variables (DVs) were analyzed to test the hypotheses: 1) touch on the smartphone in a reminder area (binary variable: a touch was done outside the reminder area or in the reminder area, 0/1, H1 and H2); 2) number of touches on the smartphone while driving; and 3) number of touches on the smartphone per hour driven as two measurements of the total usage of smartphone on the road (H3).

3.3. Procedure

The data were collected between June 2016 and December 2016 in 24-week periods for each driver, with the target of 12 calendar weeks for both conditions. The researchers installed the dashboard phones on the participants' cars, as well as the background application on the participants' personal smartphones. During the installations, the participants received information about the two phases of the study and instructions for the control phase. They were given a description of the system, its purpose, and its functionalities, both verbally and in text. They were also provided with instructions on what to try if they noticed a malfunction in the system and to contact the researchers, if needed. Instructions on how to shut down the application were provided for situations in which the participant was a passenger in the car. The same type of data was collected in the control phase as in the experiment phase, but no reminders or warnings were provided to the driver when they were located in a reminder area or when they touched their phone. The reminder areas were calculated in the exactly same manner in both phases.

After 12 calendar weeks, the experiment phase was started automatically with a timer in the dashboard phone. Before the beginning of the experiment phase, each participant was contacted by email and provided with information on the purpose of the system, the meanings of the reminders and warnings, and the reminder areas. At this point, the participants were also reminded that the warning system was not perfect, that it had known limitations, and that in any case, the participants should pay attention to the road as is required for safe driving. The participants were instructed to drive and behave as they normally would and to deal with the reminders and warnings as they saw appropriate. After the study, the participants were provided with a survey link via email to report their experiences with the system.

3.4. Data processing and analysis

The collected data were tabulated in two data matrices and filtered to include only the four road environments of interest and for a GPS speed > 2 m/s because of the inaccuracy of the GPS signal. For an aggregated data matrix, the observation unit was participant ($N = 26$). For the other more detailed data matrix, the observation unit was touch on the smartphone ($N = 175{,}593$).

Hypothesis 1 was tested with the detailed data matrix and mixed effects logistic regression (GLMM in SPSS v28) by modeling a binary outcome variable of whether a touch was made on the smartphone in a reminder area (0/1), that is, the odds of touching the phone in a reminder area based on a fixed effect of the experimental condition (control vs. experiment). The participant was added as a random effect into the model to control for inter-individual differences, and the road environment was added as a fixed effect to control for its effect in testing H1 and for comparison of the odds of touches on reminder areas between the road environments in general. This produced a multilevel binary logistic regression model with participant as a Level 2 grouping variable into which the touches were nested. The intercepts of the participants were allowed to vary.

For testing H2, the interaction effects of the experimental condition and the road environment were studied similarly with a multilevel logistic regression by modeling a binary outcome variable of a touch on the smartphone in a reminder area and with participant as a random effect. Hypothesis 3 was tested with the aggregated data matrix and a nonparametric Wilcoxon signed rank test, as the data in the control and experiment conditions were non-normally distributed at the aggregated (i.e., participant) level. Accordingly, differences in the descriptive statistics of the recorded data per condition were tested with the Wilcoxon signed rank test to ensure fairness of the comparison between the conditions.

4. Results

4.1. Descriptive statistics

Based on the descriptive statistics, the comparability of the touch data between the experimental conditions and the road environments was at a fair level for further analyses. No significant differences were found with the Wilcoxon signed rank test between the control and experiment conditions in the recorded number of days of driving ($Z = 0.727$, $p = .467$) or the number of recorded GPS locations (both stopped and driving, $Z = 0.767$, $p = .443$, Table 1). The mean number of days with data recording was 42 days for the control ($SD = 16$, range: 7–69 days) and 46 days for the experiment ($SD = 15$, range: 12–80 days). These descriptives represent the number of days with recorded driving data during the 12 calendar weeks of data collection. There were touches in general and touches in reminder areas (control: $M = 1957$, $SD = 3207$, experiment: $M = 1614$, $SD = 2615$, min $= 34$ on the main rural road) on all road environments from all 26 drivers included in the analyses to enable reliable multilevel modeling of the data (Richter, 2006).

4.2. H1: system's effects on touches in reminder areas

Hypothesis 1 was supported by the data (Table 2). A multilevel binary logistic regression model with touch on the smartphone in a reminder area (outside the reminder area/in the reminder area, 0/1) as the DV, the experimental condition and road environment as fixed effects, and participant as a random effect, indicated the significant effects of all the predictors (Table 2).

A total of 175,593 touches were included in the analysis. In the intercept-only model, there was significant variance between the intercepts of the participants ($\sigma^2 = 0.249, p < .001$, ICC $= 0.07$), indicating that a multilevel model with observations nested in the participants is appropriate for the modeling. According to the intercept-only model, the expected odds across all the data that a touch on the smartphone is done in a reminder area were 0.41, and the expected probability of this type of touch is 29 % (Crowson, 2020).

The odds for a smartphone touch to be located in a reminder area decreased significantly from the control to the experiment (odds ratio [OR] $= 0.84$, Table 2), with the other predictors being held constant. The corresponding expected probability decreased from 29 % in the control to 26 % in the experiment (i.e., a 10.3 % decrease in probability).

The odds for a smartphone touch to be done in a reminder area decreased slightly from urban to the main rural road (OR $= 0.90$) and more strongly from urban to highway (OR $= 0.67$), with the other predictors being held constant (Table 2, Fig. 3). There was no significant difference in the odds between urban and local rural roads.

4.3. H2: interaction effects of the road environment and the experimental condition on touches in reminder areas

Against expectations, the urban environment was the only road environment in which the system did not affect the probability of smartphone touches in reminder areas (H2 not supported, Table 3 and Fig. 4, $N = 175{,}593$ touches). There was a significant decrease in the odds of these touches from the control to the experiment in all the other road environments.

4.4. H3: effects of the system on smartphone usage while driving

Hypothesis 3 was rejected. The number of touches on the smartphone while driving did not decrease significantly from the control to

Table 1
Number of days, locations, and touches on the smartphone while driving per condition ($N = 26$), Mean (SD).

Condition	Days	Locations	Touches*	Touches* – Highway	Touches* – Main Rural Road	Touches* – Local Rural Road	Touches* – Urban
Control	42 (16)	218,010 (143,006)	4212 (5931)	623 (1141)	815 (1353)	832 (1071)	1074 (2198)
Experiment	46 (15)	245,553 (134,094)	4103 (6253)	1472 (4329)	750 (320)	495 (619)	770 (670)

* Recorded touches on the smartphone while on the move (GPS speed >= 2 m/s).

Table 2
Multilevel binary logistic regression model predicting the odds for a smartphone touch in a reminder area (DV: 0/1).

Fixed effects	Coefficient	Standard Error	p	Odds ratio (OR)	95 % CI Lower Bound (odds)	95 % CI Upper Bound (Odds)
Intercept	−0.76	0.10	< 0.001	0.47	0.38	0.57
Experiment	−1.75	0.01	< 0.001	0.84	0.82	0.86
Control*	0*			1.00		
Highway	−0.40	0.02	< 0.001	0.67	0.65	0.69
Main rural road	−0.10	0.02	< 0.001	0.90	0.88	0.93
Local rural road	−0.01	0.01	.384	0.99	0.96	1.02
Urban*	0*			1.00		
Random effects		σ^2			**ICC**	
Intercept (participant)		0.25	.07	< 0.001	.07	

Note. *The factors above are compared to the factor that obtains a value of zero.

the experiment, $Z = 0.114$, $p = .909$ (Wilcoxon signed rank test, $N = 26$, Table 1). The same result was obtained for a normalized measurement of touches per hour driven (control: $M = 79$, $SD = 99$, experiment: $M = 62$, $SD = 61$, $Z = 0.982$, $p = .326$, $N = 26$), which considered a possible difference in the driving time between the control and the experiment phases.

4.5. *Survey – key findings*

The survey contained items targeted at measuring the acceptance and usefulness of the warning system from various perspectives on a Likert scale (1–5). Most of the survey items (see the Technical Appendix) were based on the survey in Kujala et al. (2016) study. On average, the descriptives indicated highly positive attitudes toward the warning system (means > 3.5 with significant differences to 3), except for two items focused on technical reliability ("There are no technical errors in the application": $M = 3.46$, $SD = 1.14$, and "The application did not always remind me of a situation requiring attention": $M = 3.12$, $SD = 1.31$) and an item on the usefulness of the reminders during calls ("The reminders during calls were useful for me": $M = 3.38$, $SD = 1.14$). Textual data on the open-ended questions were analyzed using a qualitative content analysis (Mayring, 2000) to examine the arguments behind the numerical survey data. The written opinions were mostly in line with the quantitative metrics. Twenty out of the 26 participants stated that they would be willing to use the system and found it important for road safety, but they would like to have the ability to modify the auditory warnings. These warnings were experienced as annoying but still acceptable. The following statements are examples of the participants' feedback: "It is the whole point of warnings! They need to be annoying so that they are not listened to only for fun" and "The app was annoying but in the right way so that it made me decrease my phone usage while driving." By contrast, the visual reminders were not considered annoying ($n = 21$). The majority of the participants ($n = 24/26$) felt that the system supported them in driving more safely. The following lines are examples of the participants' feedback: "I noticed a notable decrease in my phone usage while driving," "The auditory warnings did not encourage me to use my phone; on the contrary, they encouraged me to stop at least for the time being," and "My attention was directed more to traffic than to the phone."

5. Discussion

The smartphone-based context-sensitive distraction warning system had a significant decreasing effect on the odds of the participants touching their phones while approaching areas of high attentional demand (e.g., intersections, junctions, roundabouts, crosswalks, or railroad crossings; H1 supported). The drivers also reported paying more attention to traffic because of the system and found the reminders and

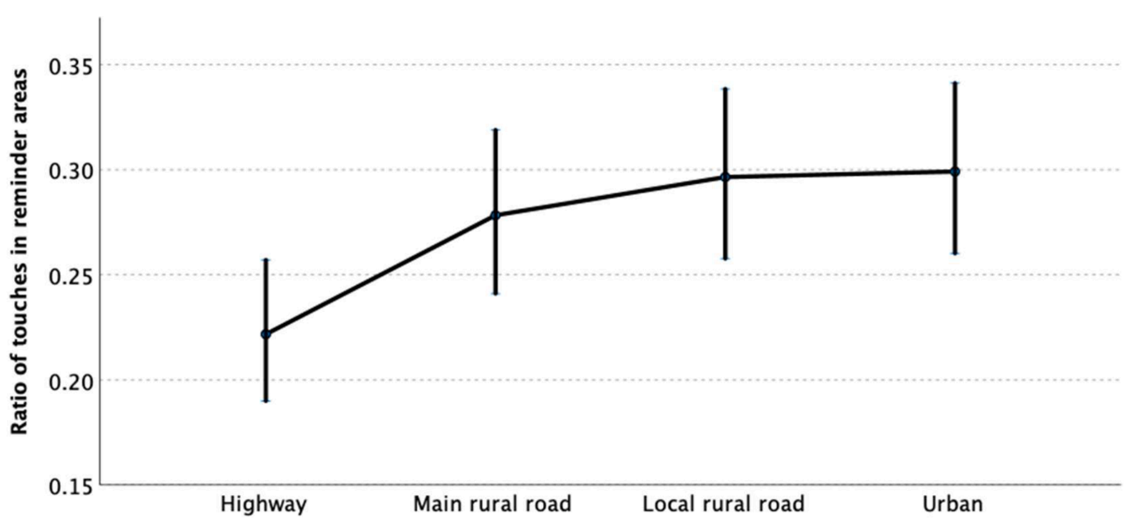

Fig. 3. Estimated marginal means (0–1) for the effect of the road environment on the ratio of smartphone touches in reminder areas to all touches.

Table 3
Multilevel binary logistic regression model of the interactions between the experimental condition and the road environment predicting the odds for a smartphone touch in a reminder area (DV: 0/1).

Fixed Effects (Interaction)	Coefficient	Standard Error	*p*	Odds Ratio (OR)	95 % CI Lower Bound (Odds)	95 % CI Upper Bound (Odds)
Intercept	−0.82	0.10	< 0.001	0.44	0.36	0.53
Highway / Experiment	−0.53	0.03	< 0.001	0.59	0.56	0.62
Main rural road / Experiment	−0.38	0.02	< 0.001	0.68	0.65	0.72
Local rural road / Experiment	−0.18	0.03	< 0.001	0.83	0.79	0.88
Urban / Experiment	0.30	0.02	.162	1.03	0.99	1.07
Highway / Control	−0.36	0.02	< 0.001	0.70	0.67	0.72
Main rural road / Control	0.01	0.02	.671	1.01	0.97	1.04
Local rural road / Control	0.06	0.02	< 0.001	1.06	1.03	1.09
Urban / Control*	0*			1.00		

Random effects	**σ^2**			**ICC**
Intercept (participant)	0.25	.07	< 0.001	.07

Note. *The factors above are compared to the factor that obtains a value of zero.

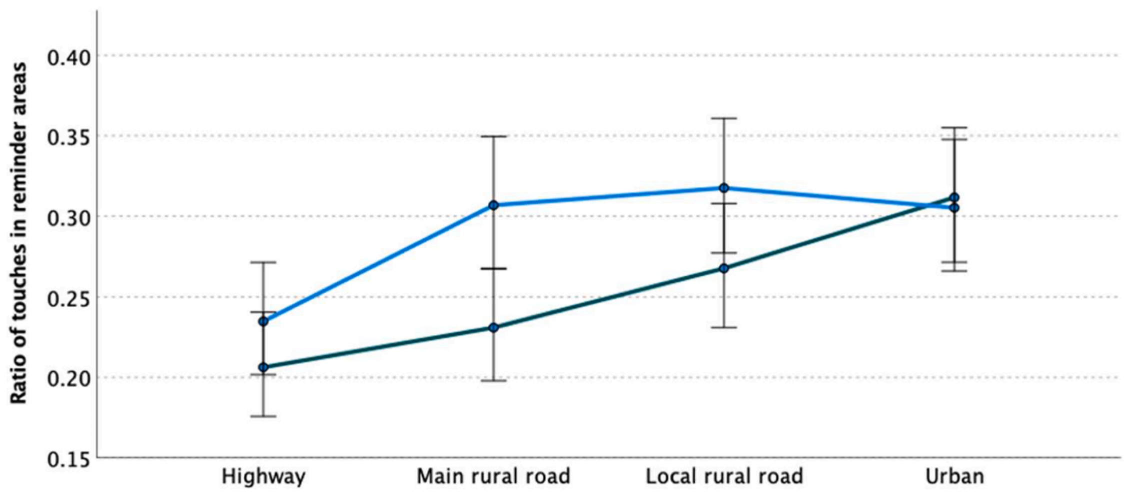

Fig. 4. Estimated marginal means (0–1) for the effect of the experimental condition by road environment on the ratio of smartphone touches in reminder areas to all touches. Blue: control, Green: experiment.

warnings to be acceptable and useful.

Against expectations, the system did not decrease the participants' touches on their smartphones in areas of high attentional demand in urban road environments (H2 rejected). The assumption of the high densities of these areas on urban and local roads is supported by the higher probability of touches on the smartphone in these areas in general than on main and highway roads (Fig. 3). In the experiment, the high density resulted in a number of warnings when driving through an urban environment and touching the phone. However, the warnings did not seem to decrease the participants' interactions with their phones on city roads. This finding might be a result of the high density itself; there is much less *spare time* between the reminder areas in urban environments than in other areas and thus fewer opportunities to postpone phone use. There are also more frequent stops in city environments (e.g., by traffic lights), which may encourage phone use that may not be ceased when moving again. As the volunteer drivers were highly keen to use their phones while driving, they were unlikely to fully stop their phone use in urban road environments. However, increased attention on the road was also possible while using a smartphone in reminder areas in cities if the participants, for instance, distributed their visual attention more extensively, appropriately, and carefully, or if they decreased speed accordingly (Ahlström et al., 2021; Kircher and Ahlström, 2017).

The participant sample was composed of heavy smartphone users, so the findings may not be generalizable beyond these types of drivers. On average, the drivers made 79 touches per hour on their phones while driving in the control condition. In general, the system did not seem to affect their phone usage while driving (H3 rejected). This finding is positive in the sense that the system did not either seem to encourage phone use based on a probably false sense of security. The study was conducted in 2016, but because of the small effects, it seemed to be worth publishing only at the present time when driver attention monitoring systems are becoming compulsory in new cars (Fredriksson et al., 2021). PokémonGo was launched during the time the study was conducted, and it seemed to have also drawn the attention of the participants in this study (Kujala and Mäkelä, 2018). This highlights the attitudes toward phone usage in traffic of some of the participants who played the game while driving. Many of them reported in the open-ended questions that they would like to change their behaviors and that this type of system could help them. Some of the participants wrote that the auditory warnings were annoying but necessary. Naturally, it cannot be ruled out that the participants wanted to please the researchers as the designers of the system, even if they were asked to behave normally and as they see appropriate.

There were technical issues in play that might have decreased the reliability and effects of the system. The accuracy of the reminders and warnings was based on the accuracy of Open Street Map data. For instance, rarely used farm and forest road junctions on main roads in Finland were regarded as reminder areas because these were mapped as junctions in the map data. Some crosswalks were missing from the map data, but fortunately, these often coincided with intersections. The definition of reminder areas in the map data could also consider other types of high attentional demands (e.g., schools, road construction, or other local information). The context sensitiveness of the system was still limited in these aspects. It would be useful to augment this deficiency using vehicle sensors, machine vision, or other means (Kashevnik et al., 2021) for better access to relevant contextual data (e.g., headway distance, weather and road conditions, and visibility) in future implementations of similar systems.

There is a need for an improved understanding of the variability of drivers' spare attentional capacity to enhance the effectiveness of driver inattention monitoring systems (Fredriksson et al., 2021). The concept of spare visual capacity (Ahlström et al., 2021; Kircher and Ahlström, 2017; Kujala et al., 2023) challenges the common assumption that drivers are inattentive whenever they look away from the forward roadway. Instead of relying on gross probabilities and static thresholds, inattention detection could better consider the specific situational factors in each driving scenario. For instance, the popular 2 s off-forward glance duration (Klauer et al, 2006) as a threshold for inattentiveness was derived based on ORs on safety-critical event statistics as a gross measure over a variety of different situations. Even if these odds are valuable for risk assessment, they may fail to account for the variability in driving situations required for effective driver attention monitoring and intervention. The current findings emphasize the need for a more nuanced and context-dependent approach to driver inattention detection.

The significant decreasing effect of the system on phone touches in areas of high attentional demand and the fact that the drivers reported paying more attention to traffic suggest that these kinds of warning systems may have the potential to mitigate the risks associated with driver distraction. At least, the participants seemed to accept the context-sensitive reminders and warnings as part of their daily driving, perhaps because they understood the need to pay attention to the safety-critical environments observed ahead. These types of alerts seem to be rarely regarded as false positives. This contrasts with distraction warnings based on off-forward glance durations only, without utilizing any contextual information about concurrent driving demands (e.g., AttenD; Ahlstrom et al., 2013). However, further research is needed to determine the objective safety effects of such context-sensitive systems. The small sample size, the focus on heavy smartphone users, and the lack of eye-tracking or driving performance measures limit the generalizability of the findings. Future studies could consider incorporating these measures to provide a more complete picture of the effects of context-sensitive (in)attention monitoring systems on driver attention. Furthermore, a between-subject experiment with a control group driving on the same roads at the same time than an experiment group could better control for possible time-of-year effects.

It has been observed that cars equipped with advanced driver assistant systems (ADAS), such as adaptive cruise control and lane assist, may increase driver inattention (Reagan et al., 2021). Thus, it has been proposed that all cars with ADAS should be equipped with driver-facing cameras to detect inattention and alert the drivers (AAA, 2022). The main takeaway message here is that besides driver-facing cameras and other in-car sensors, inattention warning systems might benefit from the adaptation of the warnings to the upcoming demands of the driving environment, especially in terms of drivers' acceptance and thus the system's potential effectiveness.

CRediT authorship contribution statement

Tuomo Kujala: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Hilkka Grahn:** Writing – original draft, Writing – review & editing. **Jakke Mäkelä:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft. **Johanna Silvennoinen:** Data curation, Formal analysis, Writing – review & editing. **Timo Tokkonen:** Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing.

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.

Data availability

Data will be made available on request.

Acknowledgments

This research was supported by VisGuard, an innovation and research project funded by TEKES, the Finnish Funding Agency for Technology and Innovation (grant number 813/31/2015], and the Academy of Finland (Appropriate Uncertainty in Manual and Automated Driving, grant number 343259). The VisGuard 2.0 application was programmed by Ficonic Solutions Ltd.

Declaration of Generative AI and AI-assisted technologies in the writing process

None.

Technical Appendix.

Survey items

(Each Likert item with scale 1–5 (items 3–31, 33–49) were followed by an open-ended question for providing justification for the given scale value.)

Please describe briefly what purpose you think the VisGuard application is intended for. (open-ended question)

Have you noticed regularities or patterns in the operation of the application in certain specific situations? (open-ended question)

The definition of the application's purpose by the developers is as follows: The purpose of the application is to promote safe driving by monitoring the driver's phone use while driving and reminding the driver to be attentive when necessary. Please choose your level of agreement with the following statement: The application works well in its task.

There are no technical errors in the application.

The operation of the application is consistent and logical.

The application is reliable.

The intentions of the application's designers are good.

The application was generally useful for me.

The application supported my safe driving.

The application had a detrimental effect on my driving.

The reminders provided by the application (warning sign) were useful for me.

The reminders supported my safe driving.

The reminders had a detrimental effect on my driving.

The reminders during calls were useful for me.

The reminders came when there was a reason for them.

The application did not always remind me of a situation requiring attention.

The timing of the reminders was effective.

The reminders came in a timely manner.

The reminders came too early.

The reminders came too late.

The reminders stayed on the screen for too long.

The audio warnings presented by the application ('Attention!') were useful for me.

The audio warnings supported my safe driving.

The audio warnings had a detrimental effect on my driving.

The audio warnings came when there was a reason for them.

I am satisfied with the application.

It was pleasant to use the application.

I could recommend the application to my acquaintances.

The application was annoying.

The reminders were annoying.

The audio warnings presented by the application ("Attention!") were annoying.

If the annoyance from the application was caused by some other reason, what was it? (open-ended question)

I accepted the application as part of my driving during the study.

The application supported my driving activities.

I could use the application in everyday driving in the future now that the study is over.

The application enabled safe driving and simultaneous use of the phone.

The application increased my alertness in traffic.

Any ideas on how to develop the system further? (open-ended question)

References

Ahlström, C., Georgoulas, G., Kircher, K., 2021. Towards a context-dependent multi-buffer driver distraction detection algorithm. IEEE Trans. Intel. Transport. Systems 23 (5), 4778–4790.

Ahlström, C., Kircher, K., Kircher, A., 2013. A gaze-based driver distraction warning system and its effect on visual behavior. IEEE Trans. Intell. Transport. Systems 14 (2), 965–973. https://doi.org/10.1109/TITS.2013.2247759.

Anon. Alliance of Automobile Manufactures. (2006). Statement of Principles, Criteria and Verification Procedures On Driver Interactions With Advanced In-Vehicle Information and Communication Systems. Washington, D.C.: Alliance of Automobile Manufactures.

Anon. American Association of State Highway and Transportation Officials. (2011). A Policy on Geometric Design of Highways and Streets. Washington, D.C.: American Association of State Highway and Transportation Officials.

Anon. American Automobile Association. (2022). *Effectiveness of Driver Monitoring Systems*. https://newsroom.aaa.com/wp-content/uploads/2022/01/Driver-Monitoring-Full-Report-February-2022-Final.pdf.

Bálint, A., Flannagan, C.A., Leslie, A., Klauer, S., Guo, F., Dozza, M., 2020. Multitasking additional-to-driving: prevalence, structure, and associated risk in SHRP2 naturalistic driving data. Accident Anal. Prevent. 137, 105455.

Caird, J.K., Chisholm, S.L., Lockhart, J., 2008. Do in-vehicle advanced signs enhance older and younger drivers' intersection performance? Driving simulation and eye movement results. Int. J. Hum. Comput. Stud. 66 (3), 132–144.

Crowson, M. (2020). *Multilevel Binary Logistic Regression Using IBM SPSS*. https://drive.google.com/file/d/16UJsWJodaVFdxJesu7OTQFgGWtrsITzv/view.

Fredriksson, R., Lenné, M.G., van Montfort, S., Grover, C., 2021. European NCAP program developments to address driver distraction, drowsiness and sudden sickness. Front. Neuroergon. 2, 786674.

Halin, A., Verly, J.G., Van Droogenbroeck, M., 2021. Survey and synthesis of state of the art in driver monitoring. Sensors 21 (16), 5558. https://doi.org/10.3390/s21165558.

Han, S., Guo, F., Klauer, C., 2023. In-Depth Analysis of Crash Risk Associated with Eyes-off-Road Duration by Road Control type and Intersection type. National Surface Transportation Safety Center For Excellence (NSTSCE). Report # 23-UM-122.

Iranmanesh, S.M., Nourkhiz Mahjoub, H., Kazemi, H., Fallah, Y.P., 2018. An adaptive forward collision warning framework design based on driver distraction. IEEE Trans. Intell. Transport. Syst. 19 (12), 3925–3934. https://doi.org/10.1109/TITS.2018.2791437.

Kashevnik, A., Shchedrin, R., Kaiser, C., Stocker, A., 2021. Driver distraction detection methods: a literature review and framework. IEEE Access. 9, 60063–60076. https://doi.org/10.1109/ACCESS.2021.3073599.

Khan, M.Q., Lee, S., 2019. A comprehensive survey of driving monitoring and assistance systems. Sensors 19 (11), 2574. https://doi.org/10.3390/s19112574.

Kircher, K., Ahlstrom, C., 2017. Minimum required attention: a human-centered approach to driver inattention. Hum. Fact. 59 (3), 471–484.

Klauer, S.G., Dingus, T.A., Neale, V.L., Sudweeks, J.D., Ramsey, D.J., 2006. The Impact of Driver Inattention on Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data (No. DOT HS 810 594). Department of Transportation – National Highway Traffic Safety Administration, Washington, D.C.

Klauer, S.G., Guo, F., Simons-Morton, B.G., Ouimet, M.C., Lee, S.E., Dingus, T.A., 2014. Distracted driving and risk of road crashes among novice and experienced drivers. N. Engl. J. Med. 370 (1), 54–59.

Kujala, T., Karvonen, H., Mäkelä, J., 2016. Context-sensitive distraction warnings–effects on drivers' visual behavior and acceptance. Int. J. Hum. Comput. Stud. 90, 39–52.

Kujala, T., Kircher, K., Ahlström, C., 2023. A review of occlusion as a tool to assess attentional demand in driving. Hum. Factors 65 (5), 792–808.

Kujala, T., Mäkelä, J., 2018. Naturalistic study on the usage of smartphone applications among Finnish drivers. Accident Anal. Prevent. 115, 53–61.

Lees, M.N., Lee, J.D., 2007. The influence of distraction and driving context on driver response to imperfect collision warning systems. Ergonomics 50 (8), 1264–1286. https://doi.org/10.1080/00140130701318749.

Lubkowski, S.D., Lewis, B.A., Gawron, V.J., Gaydos, T.L., Campbell, K.C., Kirkpatrick, S. A., Reagan, I.J., Cicchino, J.B., 2021. Driver trust in and training for advanced driver assistance systems in real-world driving. Transport. Res. Part F: Traffic Psychol. Behav. 81, 540–556. https://doi.org/10.1016/j.trf.2021.07.003.

Mayring, P., 2004. Qualitative content analysisFlick, U., von Kardoff, E., Steinke, I. (Eds.), In: A Companion to Qualitative Research. SAGE Publications, pp. 159–176.

McDonald, A.D., Ferris, T.K., Wiener, T.A., 2020. Classification of driver distraction: a comprehensive analysis of feature generation, machine learning, and input measures. Hum. Factors 62 (6), 1019–1035. https://doi.org/10.1177/0018720819856454.

Anon, National Highway Traffic Safety Administration, 2013. Visual-Manual NHTSA Driver Distraction Guidelines for In-Vehicle Electronic Devices. Department of Transportation – National Highway Traffic Safety Administration, Washington, D.C.. *NHTSA-2010-0053*.

Ramzan, M., Khan, H.U., Awan, S.M., Ismail, A., Ilyas, M., Mahmood, A., 2019. A Survey on state-of-the-art drowsiness detection techniques. IEEE Access. 7, 61904–61919. https://doi.org/10.1109/ACCESS.2019.2914373.

Reagan, I.J., Teoh, E.R., Cicchino, J.B., Gershon, P., Reimer, B., Mehler, B., Seppelt, B., 2021. Disengagement from driving when using automation during a 4-week field trial. Transport. Res. Part F: Traffic Psychol. Behav. 82, 400–411.

Richter, T., 2006. What is wrong with ANOVA and multiple regression? Analyzing sentence reading times with hierarchical linear models. Discourse Process. 41 (3), 221–250.

Roberts, S.C., Ghazizadeh, M., Lee, J.D., 2012. Warn me now or inform me later: drivers' acceptance of real-time and post-drive distraction mitigation systems. Int. J. Hum. Comput. Stud. 70 (12), 967–979.

Shahverdy, M., Fathy, M., Berangi, R., & Sabokrou, M. (2020). Driver behavior detection and classification using deep convolutional neural networks. Expert Syst. Appl., 149, 113240. https://doi.org/10.1016/j.eswa.2020.113240.

Spence, C., 2011. Crossmodal correspondences: a tutorial review. Attent., Percept. Psychophys. 73, 971–995.

Strand, N., Nilsson, J., Karlsson, I.C.M., Nilsson, L., 2011. Interaction with and use of driver assistance systems: a study of end-user experiences. In: 18th World Congress on Intelligent Transport Systems.

Summala, H., 2000. Brake reaction times and driver behavior analysis. Transport. Hum. Fact. 2 (3), 217–226.

Tran, D., Do, H.M., Sheng, W., Bai, H., Chowdhary, G., 2018. Real-time detection of distracted driving based on deep learning. IET Intell. Transport Syst. 12 (10), 1210–1219. https://doi.org/10.1049/iet-its.2018.5172.

Wang, J., Yang, J., Fu, Q., Zhang, J., Zhang, J., 2024. A new dynamic spatial information design framework for AR-HUD to evoke drivers' instinctive responses and improve accident prevention. Int. J. Hum. Comput. Stud. 183, 103194.

Wang, X., Xu, R., Zhang, S., Zhuang, Y., Wang, Y., 2022a. Driver distraction detection based on vehicle dynamics using naturalistic driving data. Transport. Res. Part C: Emerg. Technol. 136, 103561 https://doi.org/10.1016/J.TRC.2022.103561.

Wang, Y., Wu, B., Ma, S., Wang, D., Gan, T., Liu, H., Yang, Z., 2022b. Effect of mapping characteristic on audiovisual warning: evidence from a simulated driving study. Appl. Ergon. 99, 103638.