Investigating the impacts of auditory and visual feedback in advanced driver assistance systems: a pilot study on driver behavior and emotional response

Zhao Zou[1], Aila Khan[2]*, Michael Lwin[2], Fady Alnajjar[3] and
Omar Mubin[1]

[1]School of Computer, Data and Mathematical Sciences, Western Sydney University, Sydney, NSW,
Australia, [2]School of Business, Western Sydney University, Sydney, NSW, Australia, [3]College of
Information Technology, United Arab Emirates University, Al Ain, United Arab Emirates

In the autonomous vehicle industry, Advanced Driver Assistance Systems (ADAS) are recognized for their capacity to enhance service quality, improve on-road safety, and increase driver comfort. Driver Assistance Systems are able to provide multi-modal feedback including auditory cues, visual cues, vibrotactile cues and so on. The study will concentrate on assessing the impacts of auditory and visual feedback from assistive driving systems on drivers. A group consisting of five participants ($N = 5$) was recruited to take part in two sets of driving experiments. During the experimental sessions, they were exposed to several reminders designed for drivers in audio-only format and audio-visual format, respectively. Their driving behaviors and performances were under researcher's observation, while their emotions were evaluated by YOLO v5 detecting model. The results reveal that the participants higher compliance rate and strong emotional reactions (especially the feelings of anger, sadness and surprise) towards the unimodal feedback of audio-only driving reminders. There is no strong evidence showing that the bimodal ADAS feedback of audio-visual cues effectively improve drivers' performance during driving period. However, both the emotion data and user satisfaction results indicate that participants experienced an increase in feelings of happiness when they were able to visualize the AI assistant while hearing the audio reminders from the assistant. The study serves as one of the pioneering studies aimed at enhancing the theoretical foundation in the field of automotive user interface design, particularly concerning the design of auditory functions.

KEYWORDS

robot voice, emotion detecting, advanced driver assistance systems, human-robot interaction, humanoid agents

1 Introduction

In recent decades, the significance of AI applications and automated technologies, such as chatbots, physical robots in the public sector, and virtual automated systems in vehicles, has become increasingly apparent (Fernandes and Oliveira, 2021). AI-based products and social robots are commonly used to communicate and interact with users through visual and verbal cues, which drives to the fact that the AI-based Digital Voice Assistants are becoming more prevalent (Kaplan and Haenlein, 2019). Advanced humanoid-assisted products are introduced into the market at the contemporary fast-developing society. In the realm of autonomous

vehicle industry Advanced Driver Assistance Systems (ADAS) are notable for their ability to improve service quality, increase on-road safety, and enhance driver comfort, making them a prominent feature in modern vehicle designs (Jamson et al., 2013). It is reported that the adoption of assistance system and vehicle automation technology contribute greatly to the reduction of road accidents (Biondi et al., 2018). As vehicle automation levels increase, the dynamics of driver-vehicle interactions undergo significant transformation. Consequently, drivers are increasingly becoming aware of the nuances involved in effectively operating modern semi-autonomous vehicles (Mulder et al., 2008). Vehicle designers incorporating AI-based supportive systems into semi-autonomous vehicles also contemplate strategies to enable the most naturalistic communication between vehicles and drivers (Chattaraman et al., 2019).

ADAS has gained popularity because it integrates both automation systems and supportive systems, enabling drivers to maintain control while benefiting from the assistance provided by automated functionalities (Bayly et al., 2007). According to the recommendations and guidelines regarding safe and efficient in-vehicle information and communication system released by European Commission in 2006 (Commission of the European Communities, 2008), modern autonomous vehicles are increasingly being designed with interactive devices and features such as touchscreens, voice recognition systems, and information visual display panels. This facilitates interaction between humans, particularly drivers in this context, and the machine-powered system. Consequently, the transition between human drivers and semi-autonomous vehicles is a critical scenario (Shi et al., 2019). A growing number of drivers are transitioning from conventional operators to passive monitors, indicating a shift in their role from actively driving to monitoring the ADAS while driving (Banks and Stanton, 2019). This shift is primarily due to the driving feedback issued by ADAS, which enhance drivers' awareness of safety issues. Driving feedback from ADAS involves auditory signals, visual information presented or other types of signals and reminders. In other words, a semi-autonomous driving system is designed to communicate with drivers through vision, audio, and haptics (Mulder et al., 2008).

Wickens' multiple resource model proposes that human cognitive processing involves multiple, distinct resources that operate simultaneously but independently, including stages of processing (perception, cognition, response), input modalities (visual, auditory), and types of codes (verbal, spatial). The model suggests that tasks drawing on the same resource create higher interference and cognitive load, while tasks using different resources can be managed more effectively in parallel (Wickens, 2008). In the context of driving, Wickens' model implies that auditory feedback is more suitable than visual feedback for delivering reminders or messages, as driving is predominantly a visual task. Adding visual feedback may increase cognitive load due to interference with the visual resources needed for monitoring the driving scene.

Besides, current ADAS designs predominantly utilize advanced sensors, cameras, radar, and software to deliver real-time assistance (Jumaa et al., 2019; Mishra and Kumar, 2021). However, the complexity of these algorithms can reduce system interpretability, making it challenging for drivers to fully understand the system, especially in dynamic driving contexts. Enhancing the usability, user-friendliness, and overall user experience of ADAS is crucial in supporting the public's transition 2 to higher automation levels.

Effective engagement with robotic agents often relies on a combination of visual cues, such as facial appearance, and verbal cues, including auditory features (Kaplan and Haenlein, 2019; Park, 2009). This underscores the importance of understanding which method of message delivery is most effective for drivers, as optimizing these cues could enhance driver comprehension and interaction with ADAS. It is readily apparent that appropriate and timely driving reminders can positively support drivers particularly in enhancing safety (Lu et al., 2004), improving situational awareness (Kridalukmana et al., 2020), reducing cognitive load (Cades et al., 2017), and promoting better decision-making while driving. Conversely, inadequately designed signals from Advanced Driver Assistance Systems (ADAS) can adversely affect drivers (Biondi et al., 2018). For instance, ambiguous or unclear messages may lead to confusion, while sudden, loud, or high-pitched sounds can induce panic. Additionally, excessively frequent and repetitive alerts or reminders can cause disruptions in the driving experience. Delivering correct alarm reminders and capture the attention when necessary is quite important for improving drivers' performance (Wiese and Lee, 2004). To be specific, the abrupt and unexpected auditory signals may startle drivers which will reduce their wiliness of compliance (Adell et al., 2008). It has been observed that the impact of verbal tasks on driver performance can vary depending on the driving conditions, such as speed and traffic complexity. According to Recarte and Nunes, the impact of verbal tasks may differ from that observed in standard driving scenarios, particularly in situations involving high speeds or complex traffic patterns (Recarte and Nunes, 2000). Specifically, visually demanding tasks, such as imagery tasks, tend to reduce drivers' ability to visually monitor important cues, such as continuously checking rearview mirrors or responding to critical traffic events. In contrast, verbal tasks may cause less interference, especially in complex or high-speed driving situations. Thus, while verbal tasks are likely to have a less pronounced impact compared to spatial-imagery tasks, they can still negatively affect driving performance, particularly when drivers are already managing high cognitive demands in challenging traffic conditions (Recarte and Nunes, 2000). Therefore, it is worthwhile to examine the contribution of auditory feedback and visual feedback in ADAS designs, focusing on minimizing cognitive interference and enhancing driver performance. Scott and Gray have compared the effectiveness of unimodal warnings (vibrotactile or auditory) and bimodal warning (vibrotactile and auditory), and the result shows that bimodal warnings work better in the emergency situations (Scott and Gray, 2008). The results indicate that it is worthwhile to explore and design AI-based digital voice assistance tailored to individual drivers (McLean and Osei-Frimpong, 2019), especially when ADAS is involved. Given that voice-based technology may affect user preferences (Kaplan and Haenlein, 2020), a voice-based support system in autonomous vehicles may influence drivers' emotions and behaviors.

However, there is limited research investigating the differential impacts of task instructions presented in audio-only format versus those accompanied by visual elements on drivers' performance and acceptance. In order to fill this gap, the current study aims to provide insights for designing in-vehicle driver assistance systems and to contribute to the theoretical foundation in this area by exploring the effectiveness of driving task instructions in visual and verbal formats. This study will focus on the impact of varying ADAS message delivery methods on drivers' performance and behavior within a driving

context. This study examines the effects of multi-modal feedback in ADAS, with particular emphasis on auditory and visual feedback, and their influences on drivers. The research questions guiding this study are as follows:

1 Which advanced driving assistant system reminder is more effective in improving drivers' recognition and compliance: auditory-only feedback or audio-visual feedback?
2 Which advanced driving assistant system reminder is more impactful to drivers' emotions: auditory-only feedback or audio-visual feedback?
3 Which advanced driving assistant system reminder wins higher acceptance from drivers: auditory-only feedback or audio-visual feedback?

To investigate these questions, a comprehensive car simulator was utilized, offering participants an immersive driving experience (Ultralytics Team, 2023). During the driving sessions, the participants received various driving-related reminders in auditory and visual formats. Subsequently, their performance was closely observed, and their emotions were detected. Through thorough evaluation of these interconnected elements, the study aimed to gain essential insights for the development of human-machine interfaces. Furthermore, it is important to note that this study adopts a largely exploratory approach, aiming to investigate foundational questions about the effectiveness of auditory and visual feedback in ADAS. As such, the use of a small sample size is appropriate to establish initial insights and identify key areas for future, more extensive investigations.

2 Experiment

The project received approval from the relevant Human Research Ethics Committee (Approval Number: H15278). All experimental procedures were conducted in accordance with applicable guidelines and regulations. Informed consent was obtained from all participants or their legal guardians for the use and publication of information and images in an online open-access format.

2.1 Hardware and software

Utilization of driving simulation systems is common in experiments targeted at assessing drivers' behaviors. This type of high-fidelity driving simulator is featured with safety, reliability, and user-friendliness, render car simulators viable in diverse experimental contexts (Shi et al., 2022). Notably, car simulators reduce the likelihood of encountering potential risks inherent in on-road driving scenarios (De Winter et al., 2012). By adopting driving simulators, researchers can readily adjust observational variables and replicate experiments, thereby facilitating efficient investigation, particularly in high-risk settings (Robbins et al., 2019). While car simulators offer valuable insights, they inherently diverge from authentic on-road driving scenarios, thus presenting certain potential disadvantages. One of the potential problems is the occurrence of simulator sickness among users navigating through 3D animated environments, with older drivers being particularly vulnerable to driving discomfort (Brooks et al., 2010). Furthermore, drivers utilizing car simulators may encounter constraints related to both physical and behavioral fidelity (De Winter et al., 2012). Moreover, drivers' awareness that simulator-based hazards pose no actual harm may constrain their reactions and sense of responsibility (De Winter et al., 2012). Through attentive efforts to enhance both physical and behavioral fidelity, the validity of car simulators can be bolstered, rendering them valuable tools for researchers (Godley et al., 2002).

A custom-built high-fidelity driving simulator was designed to serve as the cockpit for the experiment. It is suggested that a high-fidelity driving simulator should allow participants to see, hear, and feel the impact of automated vehicle handling (Jamson et al., 2013). The driving simulator designed for the study includes a monitor to display driving routes, a custom-built driving system and a driving chair, as displayed in Figure 1A. The driving system consists of a Logitech G29 PlayStation version racing wheel with pedals, a Logitech gear shifter designed for the G29 driving system, and a driving chair, as indicated in Figure 1B. The racing wheel and pedals were affixed to an adjustable support stand and adjusted to the standard height of human drivers. Furthermore, the selected steering wheels could provide participants with different vibration responses, recreating the feeling of driving on various road surfaces and uplifting the overall realism of the driving experience. The large monitor is placed in front
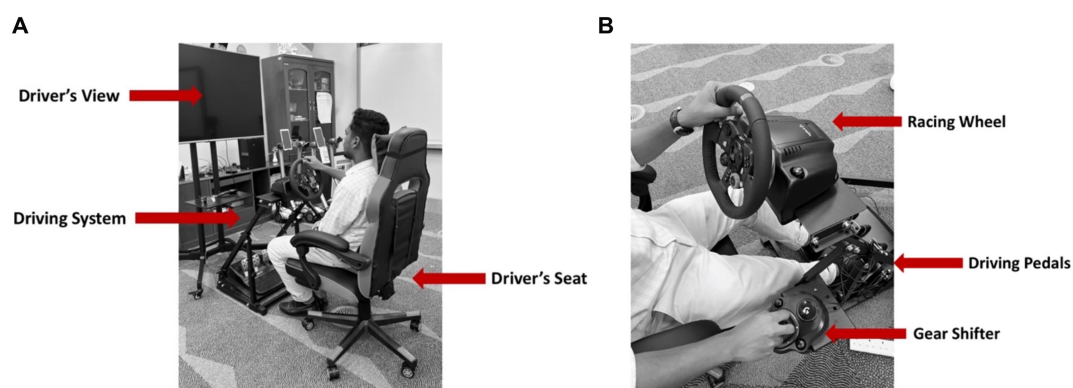

FIGURE 1
(A) Experimental simulator setup. **(B)** Display of driving panel on simulator.

of the participants to provide a clear view of the driving environment. Additionally, an external tablet is integrated into the driving simulator to effectively deliver designed reminders through both audio and visual cues.

The software component consists of a driving game, an emotion detecting model and an AI-based assisted system programed with driving reminders. The driving game selected for the study is DiRT Rally 2.0. This game is available on the STEAM gaming platform and is compatible with the Logitech G29 PlayStation. The game offers various maps and driving routes with realistic environments, allowing participants to immerse themselves in the simulated driving experience by utilizing the hand wheel, pedals and gear shifter. The participants will stay informed with reminders from the humanoid agent driving assisted system displayed on a tablet-sized screen in front of them. Additionally, a pre-programed emotion detection model is integrated in the driving system to monitor the emotional state throughout the drive session.

2.2 Experiment design

The experiment was utilising a within-subjects design, incorporating two conditions: (1) auditory feedback and (2) multimodal feedback. The condition of "auditory feedback" consisted of the audio-only ADAS for delivering driving reminders. Conversely, the "multimodal feedback" condition utilized an ADAS equipped with both audio and visual systems, allowing participants to hear the audio reminders and simultaneously view an avatar providing the reminders. The participants were required to complete two experimental driving sessions following a 5-min testing trial. During each experimental session, participants were exposed to six different reminders from the driving assisted system. Their behaviors were observed by the researcher and noted on a structured observation sheet, while their emotional responses were detected and recorded by an emotion detection model. Upon completion of the two driving sessions, the participants were invited to complete a brief online questionnaire.

2.2.1 Content of driving reminders

The participants received six reminders pertaining to their driving behavior, which included prompts to fasten seatbelts, check fuel levels, inspect rear-view mirrors, reduce speed when exceeding limits, exercise caution regarding obstacles ahead, and take a rest when experiencing drowsiness. The content of the reminders was generated by a Learning Language Model (LLM) which is listed below. Research suggests that content generated or modified by LLMs can positively influence users' emotions, particularly when the LLM is instructed to produce content in a friendly manner (Zou et al., 2024). For this experiment, ChatGPT (GPT-3.5) provided multiple options for each prompt. The first option on the LLM response list was selected as the reminder content for the experiment.

> Reminder 1
>
> Human Prompt: Please write a reminder for the drivers to check the seatbelt.
>
> LLM Response: Before starting the car, make sure your seatbelt is securely fastened

> Reminder 2
>
> Human Prompt: Please write a reminder for the drivers to check the fuel level
>
> LLM Response: It is always good to check if you have enough fuel for your journey

> Reminder 3
>
> Human Prompt: Please write a reminder for the drivers to check the rear mirrors
>
> LLM Response: Please adjust your rear mirrors and make sure they provide a clear view

> Reminder 4
>
> Human Prompt: Please write a reminder for the drivers when they are over-speeding
>
> LLM Response: You are over speeding now. Please slow down

> Reminder 5
>
> Human Prompt: Please write a reminder for the drivers when there are obstacles ahead
>
> LLM Response: Kindly remind you that there are obstacles ahead

> Reminder 6
>
> Human Prompt: Please write a reminder for the drivers when they have been driving for a long time
>
> LLM Response: You must feel tired after the long journey. Let us have a rest

2.2.2 Driving reminders auditory feature

The auditory cues from humanoid agents do influence the emotions of participants. Both voice pitches (alto and bass) and speech styles (casual and frozen) play a crucial role. Our prior work [hidden for review] indicates that an alto voice pitch, resembling that of an adult female, evokes a more cheerful response from participants. Additionally, compared with frozen speech, casual speech is more effectively bring the participants with sense of happiness. In order to create a comfortable environment for the participants in a driving setting, we developed a humanoid-agent-assisted system with an alto voice and a casual speaking style.

2.2.3 Driving reminders visual feature

The assisted ADAS was designed with an alto voice pitch, so a female avatar was chosen to match the voice. The avatar was created using Soul Machines, a platform capable of autonomously generating animated digital humans using AI. The avatar we created for the system is shown in Figure 2. The visual and audio systems were integrated to ensure that the lip movements of the AI-generated figure synchronized with the speech content.

FIGURE 2
AI-generated avatar figure designed for ADAS.

In the experiment, participants attended two separate driving sessions with the order of the two conditions (audio-only and audio-visual) randomized. In the audio-only session, they were exposed solely to auditory driving reminders, with the AI-generated face remaining invisible to them. The participants received the reminders exclusively in audio format. In the audio-visual session, they were exposed to both auditory and visual driving reminders. To be specific, the animated figure will be displayed on a tablet-sized screen positioned in front of the participants.

2.2.4 Emotion detecting model

The YOLO (You Only Look Once) object detecting model was trained and tested as emotion detector in the present experiment. YOLO series are currently the most popular object detection algorithms in academic studies (Nepal and Eslamiat, 2022). The YOLO series is superior to the earlier models in terms of running speed, accuracy and its capacity to detect small objects, even facial expressions and emotions (Ali et al., 2022). YOLOv5, released in 2020, offers a range of pre-trained object detection architectures based on the MS COCO dataset (Jocher, 2024). We conducted data augmentation on a subset of the publicly accessible AffectNet dataset, which is an expansive dataset comprising approximately 400,000 images, categorized into eight distinct emotion categories: neutral, angry, sad, fearful, happy, surprised, disgust, and contempt (Mollahosseini et al., 2017). For our endeavors, we employed the YOLOv5 model with a prototype precisely curated 10,459 images for training purposes, and the results obtained were indeed promising. Out of the eight categories, five prevalent emotions, including sad, angry, happy, neutral, and surprised were selected for training the model.

The model was trained and assessed by using standard metrics including precision-recall curve and F1-Confidence Curve. From the outcomes in precision-recall curve, the precision-recall values for the emotions "Happy" ($V_h = 0.864$) and "Surprise" ($V_s = 0.775$) exceed the baseline mean Average Precision (mAP) value ($V_b = 0.718$), suggesting that the model exhibits superior performance in predicting these two emotions. The values for "Angry" ($V_a = 0.717$) and "Sad" ($V_d = 0.701$) being close to the baseline mAP@50 value indicates that the model performs adequately in detecting the angry emotion. These outcomes are further validated by the F1-Confidence results. Although the YOLO series has been upgraded to YOLOv10, YOLOv5 still demonstrates superior processing speed, which is one of the reasons we selected YOLOv5 for our study (Ali et al., 2022). And the model has

robust documentation, community support, and numerous resources available for troubleshooting and development (Jocher, 2024).

2.2.5 Task performance observation

Three researchers (observing researcher, coding researcher and adjudicating researcher) participated in the observation and coding process, ensuring thorough documentation of participants' performance throughout the entire study. The observing researcher systematically observed participants' performance and behavior during the experimental sessions, documenting their response behavior using an observation sheet, as illustrated in Figure 3. During the experiment session, to minimize participant anxiety caused by being observed by a crowd, only one researcher conducted on-site observation. Table 1 illustrates the detailed criteria used for categorizing participants' behaviors into three different groups. The observing researcher adhered to predefined criteria and provided detailed textual descriptions on the observation sheet.

Consequently, the observing researcher and the coding researcher both follow the below coding rules and transferring the textual documents into measurable codes. The documented observation results were categorized into three distinct codes: did not notice the reminder at all (coded as 1), noticed the reminder but did not comply (coded as 2), and noticed the reminder and complied (coded as 3). The two blind observation sheets were compared to create the final observation sheet. In cases of category conflicts, a third researcher which is the adjudicating researcher was invited to review the documented observation sheet and determine the final classification. Given the small sample size for this pilot study, this observational method was both feasible and effective in reducing bias and errors.

After completion of each experimental session, participants' performance were encoded following specific predefined rules, allowing for the efficient translation of textual descriptions into a numeric, visualized format. For example, to calculate the "recognition rate" for Reminder 1, we included observations coded as 2 and 3, representing cases where participants either "noticed the reminder but did not comply" or "noticed the reminder and complied." Conversely, to calculate the "compliance rate" for Reminder 1, we considered only instances coded in 3, representing cases where participants "noticed the reminder and complied." This approach enables a quantitative analysis of recognition and compliance, facilitating a clear and objective assessment of participant response patterns.

2.2.6 Questionnaire content

In the follow-up experiment questionnaire, participants were asked to rate their preference for different formats of driving reminders using a five-point rating scale. The two questions are listed below:

- How much do you appreciate reminders from avatars featuring voice only?
- How much do you appreciate reminders from avatars featuring voice and face?

2.3 Participants

A group of five participants ($N = 5$) were recruited to join the experimental sessions. The participants with demographic features of

Group A – Audio Only Reminder

Start Time: 15:59:45   **End Time:** 16:14:30

Task Performance:

Reminder 1: Seatbelt	**Reminder 2: Fuel Tank**	**Reminder 3: Rear Mirror**
Not noticed at all	Yes, noticed and followed	Yes, noticed and followed
Reminder 4: Overspeed	**Reminder 5: Obstacle**	**Reminder 6: Drowsy**
Yes, noticed and followed	Yes, noticed but not followed	Yes, noticed but not followed

Group B – Audio and Visual Reminder

Start Time: 16:15:45   **End Time:** 16:28:30

Task Performance:

Reminder 1: Seatbelt	**Reminder 2: Fuel Tank**	**Reminder 3: Rear Mirror**
Yes, noticed and followed	Yes, noticed and followed	Yes, noticed and followed
Reminder 4: Overspeed	**Reminder 5: Obstacle**	**Reminder 6: Drowsy**
Yes, noticed and followed	Yes, noticed but not followed	Yes, noticed but not followed

FIGURE 3
Samples of researcher observation sheet.

university students from a local university in one of Middle East countries, age ranging from 17 to 35 years old and with multiple culture background. All participants attended two rounds of driving sessions, resulting in the collection of 10 datasets for analysis. In the first driving session, they were exposed to driving reminders of audio format only. While in the second driving session, the participants received the same driving reminders in both audio and visual format. The sequence of the driving sessions was randomly balanced. Specifically, some participants began with the audio-only group, while others started with the audio-visual group.

3 Results

3.1 Participants' performance

Participants' performance observation sheet provides results of their driving performance and responses towards the driving reminders. The frequency counts for each note category in observation sheet— "did not notice the reminder at all," "noticed the reminder but did not comply," and "noticed the reminder and complied" —are presented in Figure 4.

3.1.1 Recognition rate

In this study, the recognition rate was calculated based on the observation results corresponding to code 2 (notice but not follow)

and code 3 (notice and follow). The results indicate that all participants recognized both types of reminders presented in the audio-only format. Specifically, the recognition rate for the two types of reminders was 100%, suggesting that participants noticed all reminders delivered in either audio-only or audio-visual formats. This finding underscores that designing audio-only reminders could be a more cost-effective approach while still ensuring successful communication with the driver when designing in-vehicle driving assistant systems.

3.1.2 Compliance rate

The compliance rate was calculated based on the observation results regarding to code 3 only. In other words, only the participants' behaviors were categorized under "notice and follow" were included in the calculation. Participants' responses to the two types of reminders exhibit variation in compliance rates. Specifically, during the pre-driving stage involving reminders concerning seatbelt check, fuel level check, and rear mirror check, participants complied with audio-only reminders a total of eight times (compliance frequency = 8) compared to six instances for audio-visual reminders (compliance frequency = 6).

In the driving stage, which included reminders for over-speeding, obstacles detection and drowsy driving, the compliance frequency for audio-only reminders (compliance frequency = 11) than for audio-visual reminders (compliance frequency = 10). These results suggest that participants were more likely to comply with audio-only reminders in both the pre-driving and during-driving stages.

TABLE 1 Criteria of categorizing participants' behaviors.

	Code 1 Not noticed at all	Code 2 Noticed but not followed	Code 3 Noticed and followed
Reminder 1	No action at all	Verbal response such as "okay," "got it," or may even ask questions related to the ADAS. A distinct pause or hesitation in the ongoing actions or tasks. A visible physical reaction, such as nodding the head or other body movements. No observable physical action or movement made by the participant within 5 s after hearing the instruction, indicating a failure to follow through with the reminder.	Demonstrated the behavior of looking for and fastening the seatbelt within 5 s after hearing the instruction.
Reminder 2			Directed their head and eye movement toward the fuel level indicator, which is located on a specific position of the large screen, within 5 s after hearing the instruction.
Reminder 3			Switched the simulator equipment button that controls rear mirror checking within 5 s after hearing the instruction.
Reminder 4			Reduced speed by 10 km/h or more within 5 s after hearing the instruction.
Reminder 5			Begun to slow down the vehicle by 10 km/h and more, or sits straighter for better vision within 5 s after hearing the instruction.
Reminder 6			Started slowing down within 5 s after hearing the instruction and eventually stops the car within 15 s.

3.2 Participants' emotion responses

A total of number of $n = 17,240$ emotion data points were collected from the five participants, comprising n (ad) = 8,430 data points for the audio-only reminder group and n (ad-vi) = 8,810 data points for the audio-visual reminder group. This distribution indicates that the data points are nearly equally divided between the two groups. Table 2 presents the mean values for each emotion reported by participants, along with their corresponding 95% confidence levels and confidence intervals (CIs). Figure 5 displays the mean values for each emotion, alongside with standard deviations (SD) to convey the reliability and variability within the dataset.

For the emotions Angry, Sad, and Surprised, the data for both types of reminders exhibit relatively small standard deviation (SD) values. An analysis of the 95% confidence intervals for the three emotions under the two conditions (audio-only and audio-visual) reveals relatively narrow confidence intervals. This suggests that we can expect the true mean levels of these emotions to lie within these ranges, indicating a degree of precision in the estimated mean values and a high level of confidence in the stability of the findings within this context. The mean values of emotional responses for Angry [Ma (ad) = 0.080], Sad [Mc (ad) = 0.123], and Surprise [Md (ad) = 0.114] indicate that participants exhibited stronger reactions to audio-only reminders, with a difference of approximately 0.02 compared to the mean values of these emotions in response to audio-visual reminders.

Regarding the emotional responses for happiness, the mean value [Mb (ad) = 0.076] to audio-only reminders is 0.1 lower than the mean value [Mb (ad-vi) = 0.171] provoked by audio-visual reminders. This suggests that the participants showed a stronger emotional response when they were exposed to audio-visual reminders compared to audio-only reminders. The standard deviation for the happy emotion in response to audio-visual reminders is 0.204, indicating substantial variability in participants' responses. Additionally, the confidence interval is relatively wide, suggesting limited precision in estimating the true mean happiness level. This wide interval reflects the high variability within the data, indicating that participants' emotional responses to the audio-visual reminders were not consistent.

3.3 Participants' preference

The participants were given a five-point likert rating scale (strongly dissatisfied, dissatisfied, neutral, satisfied, very satisfied) to provide their preferences for both audio-only reminder group and audio-visual reminder group. The Likert coding scheme is shown in Figure 6 where smaller numbers represent lower levels of satisfaction. Figure 6 presents a radar chart illustrating the preferences of five participants across two conditions. Each axis represents a participant, with the innermost circle corresponding to the lowest preference level and the outermost circle representing the highest preference level. The blue line denotes participants' preference for audio-only feedback, while the grey line indicates their preference for audio-visual feedback. The use of a radar chart is particularly appropriate for small sample sizes, as it allows for a clear and concise presentation of all data.

Through the calculation of the mean values of the Likert rating points, the median value (value =4.0) of satisfaction with audio-only reminders is slightly lower than that with audio-visual reminders (median value =4.2), which suggests the stronger preference on the ADAS with both audio and visual features.

4 Discussion and limitation

The results of this pilot study were assessed from three primary perspectives: participants' performance and responses to ADAS reminders, changes in their emotional states, and their preference ratings.

The findings on participants' recognition rates indicate that participants were able to perceive the instructions effectively when they were voiced, regardless of the presence or absence of visual feedback. However, their willingness to comply with the instructions is influenced by the presence of visual cues. In the present context of driving tasks, there is a consistent tendency of greater compliance rates from the participants to audio-only reminders across different types of driving reminders in both pre-driving and during-driving stages. The auditory cues in the autonomous assistive system place stronger influences on the drivers' behaviors than the visual cues. In other words, in highly
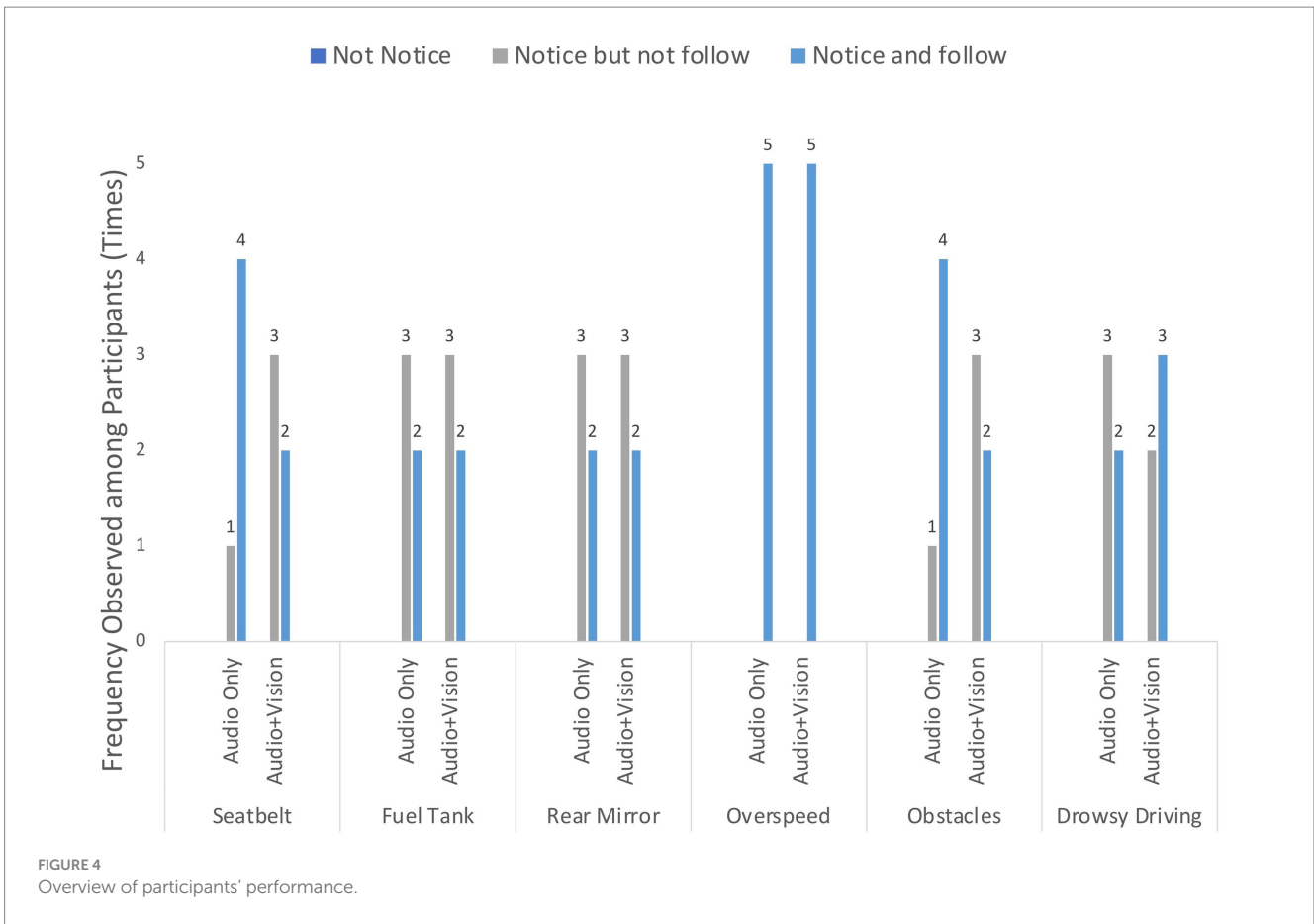
FIGURE 4
Overview of participants' performance.

TABLE 2 Mean values of participant' emotions with CI.

		Angry	Happy	Sad	Surprise
Audio only	Mean value	0.0802	0.0756	0.1233	0.1145
	Confidence level (95%)	0.0410	0.0393	0.0394	0.1136
	Upper CI (95%)	0.1225	0.1114	0.1651	0.2282
	Lower CI (95%)	0.0405	0.0327	0.0862	0.0011
Audio_vision	Mean value	0.0687	0.1706	0.1064	0.0889
	Confidence level (95%)	0.0288	0.2537	0.0416	0.0551
	Upper CI (95%)	0.0987	0.4028	0.1494	0.1405
	Lower CI (95%)	0.0412	−0.1047	0.0662	0.0304

dynamic and demanding contexts, participants may be better equipped to process unimodal instructions, such as audio-only reminders. While multimodal instructions, such as audio-visual reminders, can occasionally enhance user engagement and satisfaction, they may also introduce challenges during the dynamic driving environment. The added visual component has the potential to cause cognitive overload or distract participants, especially in scenarios requiring constant attention, such as driving. This suggests that Advanced Driver Assistance Systems should prioritize well-designed auditory feedback mechanisms to ensure driver compliance and responsiveness. Audio-only reminders can be highly effective, especially in situations where visual attention is already occupied by driving tasks.

The findings from emotion response of the participants also offer valuable insights. It is not hard to find that participants exhibited

stronger emotional reactions to auditory cues in advanced driver assistance systems, especially the feelings of anger, sadness and surprise. Auditory cues are proven to be effective in provoking emotional responses. The emotion of "Anger" may be attributed to the frustration elicited by specific reminders which interrupted their excitement during driving, such as reminders to reduce speed when participants are engaged in high-speed driving. "Sadness" is linked to the perception of reminders with negative feedback, which may make them feel being criticized. And the unexpected audio alert or reminder may elicit participants' "Surprise." However, a contrasting pattern was detected in the emotional responses to happiness, suggesting that involving visual cues may bring positive responses in certain circumstances. The findings suggest that a balanced integration of visual cues alongside auditory reminders has the potential to enhance
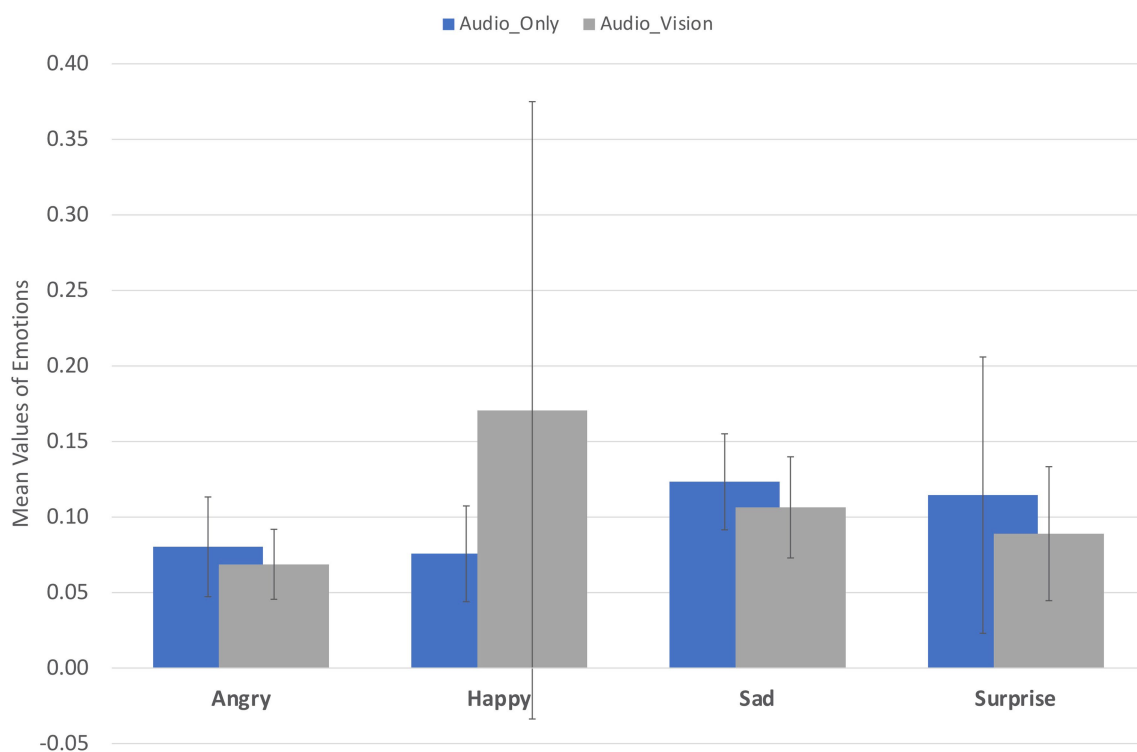
FIGURE 5
Mean values of participants' emotions with SD.

the overall emotional experience for users. Selective incorporation of visual cues could foster positive emotional responses and alleviate the stress associated with relying solely on auditory feedback. However, an excessive use of visual cues may lead to distractions, making such an approach unsuitable for contexts that require sustained attention, such as driving.

The data on participants' satisfaction ratings indicate that, within the driving context, audio-only cues are slightly more preferred compared to audio-visual cues. According to Wickens (2008), adding visual feedback to a predominantly visual task competes for the same resource pool, increasing cognitive load and the likelihood of interference. In contrast, auditory signals utilize a separate resource channel, allowing drivers to process instructions without overburdening their visual system. This theoretical perspective explains why participants demonstrated higher compliance rates and satisfaction with audio-only reminders, as these allowed them to maintain better focus on driving tasks while still receiving effective guidance. These findings prompt further investigation into the specific attributes of auditory cues that contribute to their higher acceptance and effectiveness. Understanding these characteristics could provide valuable insights for designing more user centric ADAS that enhance driver satisfaction and compliance. Future research should focus on identifying which auditory features, such as tone, frequency, or timing, are most effective and well-received in order to optimize the design of auditory feedback in ADAS.

One of the limitations for this study is the relatively small sample size. A larger sample size is always better for generalization of a study. However, pilot study sample size greatly depends on the purpose of the study (Johanson and Brooks, 2010) and the context of the study

(Lenth, 2001). As a pilot study, this paper serves as a pioneering research effort in exploring the auditory features of advanced driver assistance systems and their impact on drivers' responses. With participants ($N = 5$) in each of the two groups, contributing to a total dataset of 10 copies, this study closely aligns with the scholarly recommendation of a sample size of 12 for such research. This sample size is deemed sufficient for the purposes of this pilot study (Johanson and Brooks, 2010). Besides, it is recognized that the experiment was conducted in a Middle East country, therefore, the cultural implications may also introduce certain limitations to the present study. To enhance the generalizability, the study could be incorporated with more participants from diverse cultural backgrounds and age groups in the future studies. What's more, our current study, the impact of reminders on participants' emotional states was not directly assessed. We acknowledge this as a limitation, as understanding the emotional responses elicited by auditory or multimodal reminders could provide deeper insights into drivers' reactions and overall performance. In future research, we plan to incorporate time-stamped emotion tracking to evaluate participants' emotional states before and after the reminders are issued. This approach will allow us to more precisely measure the emotional influence of auditory and visual feedback in driving contexts.

A potential direction for future research could involve expanding the detection of various driving behaviors, with a particular focus on specific demographic groups, such as older drivers or apprentice drivers. This approach could yield insights into how different user groups respond to ADAS cues and inform tailored system designs that cater to diverse needs. Additionally, it would be valuable to explore which auditory characteristics of humanoid agents are most

FIGURE 6
Participants' preference feedback.
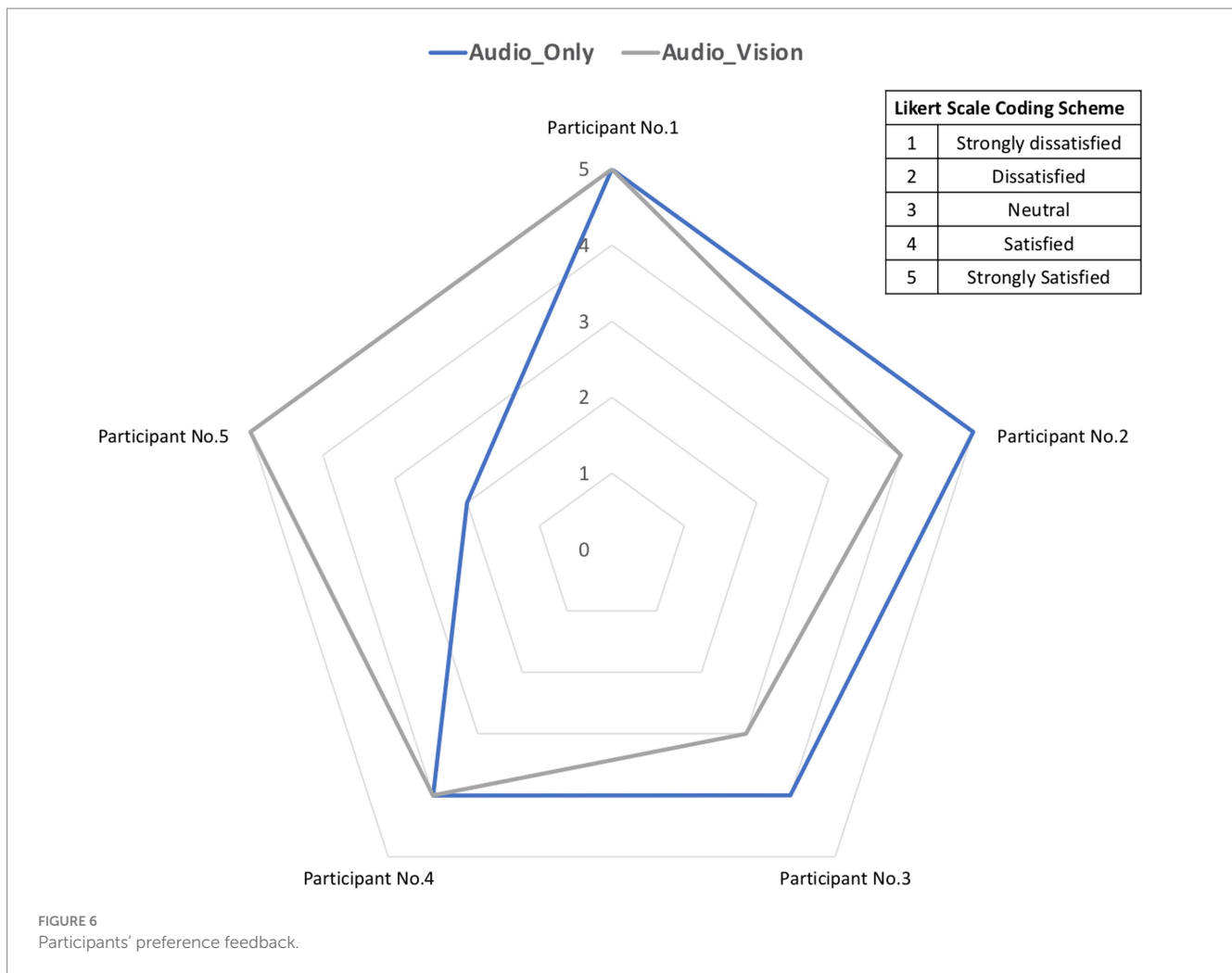
effective for ADAS in semi-autonomous and fully autonomous vehicles. Investigating these auditory features, such as voice tone, pitch, and emotional expressiveness, could enhance the design of ADAS systems, making them more intuitive and effective for a broader range of users. Moreover, a review of existing literature indicates that essential executive functions, such as working memory and inhibitory control (Engström et al., 2017), are strongly linked to risky driving behaviors and errors that can result in accidents. This insight suggests a potential future research direction focused on examining the impact of auditory feedback from ADAS on drivers' cognitive load and their levels of attention or distraction. Given that drivers' cognitive load is task-dependent, this finding underscores the importance of designing task-specific experiments for future investigations (Bamicha et al., 2024).

This pilot study investigated the impacts of auditory and visual feedback in Advanced Driver Assistance Systems (ADAS) on driver behavior and emotional response. After in total of 10 driving sessions from five participants on a well-designed high-fidelity driving simulator, results indicated a higher compliance rate and stronger emotional reactions, particularly anger, sadness, and surprise, towards audio-only reminders. While bimodal feedback did not significantly enhance performance, participants reported increased happiness when visualizing the AI assistant alongside audio reminders. This is

possibly due to the novelty of avatar technology, with drivers being unaccustomed to interacting with or observing these virtual assistants.

Overall, this study contributes to the extended research in the field of human-machine interaction, specifically in the study of Advanced Driving Assistance Systems. Future research will aim to address the limitations identified in this pilot study by refining the study protocol to enhance robustness and generalizability. In future research, we plan to investigate specific auditory cue types, such as voice pitch, speech style, and tone, that may optimize the driver experience and contribute to developing more intuitive, human-centered interfaces to enhance driver-assistant interaction. Additionally, expanding the sample size in future studies will improve data accuracy and enhance the generalizability of the quantitative findings. Through these advancements, this research aims to strengthen the theoretical foundation of automotive user interface design, highlighting the critical role of auditory and multimodal feedback in improving driver engagement and satisfaction.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Author contributions

Funding

Conflict of interest

Publisher's note

References

Adell, E., Várhelyi, A., and Hjälmdahl, M. (2008). Auditory and haptic systems for in-car speed management–a comparative real life study. *Transport. Res. F: Traffic Psychol. Behav.* 11, 445–458. doi: 10.1016/j.trf.2008.04.003

Ali, L., Alnajjar, F., Parambil, M. M. A., Younes, M. I., Abdelhalim, Z. I., and Aljassmi, H. (2022). Development of YOLOv5-based real-time smart monitoring system for increasing lab safety awareness in educational institutions. *Sensors* 22:8820. doi: 10.3390/s22228820

Bamicha, P. V., Chaidi, I., and Drigas, A. (2024). Driving under cognitive control: the impact of executive functions in driving. *World Electric Vehicle J.* 15:474. doi: 10.3390/wevj15100474

Banks, V. A., and Stanton, N. A. (2019). Analysis of driver roles: modelling the changing role of the driver in automated driving systems using EAST. *Theor. Issues Ergon. Sci.* 20, 284–300. doi: 10.1080/1463922X.2017.1305465

Bayly, M., Fildes, B., Regan, M., and Young, K. (2007). Review of crash effectiveness of intelligent transport systems. *Emergency* 3:14.

Biondi, F. N., Getty, D., McCarty, M. M., Goethe, R. M., Cooper, J. M., and Strayer, D. L. (2018). The challenge of advanced driver assistance systems assessment: a scale for the assessment of the human–machine interface of advanced driver assistance technology. *Transp. Res. Rec.* 2672, 113–122. doi: 10.1177/0361198118773569

Brooks, J. O., Goodenough, R. R., Crisler, M. C., Klein, N. D., Alley, R. L., Koon, B. L., et al. (2010). Simulator sickness during driving simulation studies. *Accid. Anal. Prev.* 42, 788–796. doi: 10.1016/j.aap.2009.04.013

Cades, D.M., Crump, C., Lester, B.D., and Young, D., (2017). Driver distraction and advanced vehicle assistive systems (ADAS): investigating effects on driver behavior. In Advances in Human Aspects of Transportation: Proceedings of the AHFE 2016 International Conference on Human Factors in Transportation, July 27–31, 2016, Walt Disney World®, Florida, USA (pp. 1015–1022). Springer International Publishing.

Chattaraman, V., Kwon, W. S., Gilbert, J. E., and Ross, K. (2019). Should AI-based, conversational digital assistants employ social-or task-oriented interaction style? A task-competency and reciprocity perspective for older adults. *Comput. Hum. Behav.* 90, 315–330. doi: 10.1016/j.chb.2018.08.048

Commission of the European Communities (2008). ESoP - European statement of principles on human-machine interface. *Off. J. Eur. Union* 1, 1–30.

De Winter, J., Van Leeuwen, P. M., and Happee, R. (2012). Advantages and disadvantages of driving simulators: a discussion. *Proceed. Measuring Behav.* 2012, 28–31.

Engström, J., Markkula, G., Victor, T., and Merat, N. (2017). Effects of cognitive load on driving performance: the cognitive control hypothesis. *Hum. Factors* 59, 734–764. doi: 10.1177/0018720817690639

Fernandes, T., and Oliveira, E. (2021). Understanding consumers' acceptance of automated technologies in service encounters: drivers of digital voice assistants adoption. *J. Bus. Res.* 122, 180–191. doi: 10.1016/j.jbusres.2020.08.058

Godley, S. T., Triggs, T. J., and Fildes, B. N. (2002). Driving simulator validation for speed research. *Accid. Anal. Prev.* 34, 589–600. doi: 10.1016/S0001-4575(01)00056-2

Jamson, A. H., Merat, N., Carsten, O. M., and Lai, F. C. (2013). Behavioural changes in drivers experiencing highly-automated vehicle control in varying traffic conditions. *Transport. Res. Part C* 30, 116–125. doi: 10.1016/j.trc.2013.02.008

Jocher, G. (2024). Yolov5. Code Repository. Available at: https://github.com/ultralytics/yolov5. (Accessed March 4, 2024).

Johanson, G. A., and Brooks, G. P. (2010). Initial scale development: sample size for pilot studies. *Educ. Psychol. Meas.* 70, 394–400. doi: 10.1177/0013164409355692

Jumaa, B. A., Abdulhassan, A. M., and Abdulhassan, A. M. (2019). Advanced driver assistance system (ADAS): a review of systems and technologies. *IJARCET* 8, 231–234.

Kaplan, A., and Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Bus. Horiz.* 62, 15–25. doi: 10.1016/j.bushor.2018.08.004

Kaplan, A., and Haenlein, M. (2020). Rulers of the world, unite! The challenges and opportunities of artificial intelligence. *Bus. Horiz.* 63, 37–50. doi: 10.1016/j.bushor.2019.09.003

Kridalukmana, R., Lu, H. Y., and Naderpour, M. (2020). A supportive situation awareness model for human-autonomy teaming in collaborative driving. *Theor. Issues Ergon. Sci.* 21, 658–683. doi: 10.1080/1463922X.2020.1729443

Lenth, R. V. (2001). Some practical guidelines for effective sample size determination. *Am. Stat.* 55, 187–193. doi: 10.1198/000313001317098149

Lu, M., Wevers, K., van der Heijden, R., and Heijer, T. (2004). "ADAS applications for improving traffic safety" in 2004 IEEE international conference on systems, man and cybernetics IEEE cat, (IEEE Cat. No. 04CH37583). The Hague, Netherlands: IEEE. 4, 3995–4002.

McLean, G., and Osei-Frimpong, K. (2019). Hey Alexa… examine the variables influencing the use of artificial intelligent in-home voice assistants. *Comput. Hum. Behav.* 99, 28–37. doi: 10.1016/j.chb.2019.05.009

Mishra, M., and Kumar, A. (2021). ADAS technology: a review on challenges, legal risk mitigation and solutions. *ADAS*, 15, 401–408. doi: 10.1201/9781003048381-21

Mollahosseini, A., Hasani, B., and Mahoor, M. H. (2017). Affectnet: a database for facial expression, valence, and arousal computing in the wild. *Transact. Affective Comput. IEEE.* 10, 18–31. doi: 10.1109/TAFFC.2017.2740923

Mulder, M, Abbink, D.A, and Boer, E.R. (2008). The effect of haptic guidance on curve negotiation behavior of young, experienced drivers. In 2008 IEEE international conference on systems, man and cybernetics. IEEE. 4, 804–809. doi: 10.1109/ICSMC.2008.4811377

Nepal, U., and Eslamiat, H. (2022). Comparing YOLOv3, YOLOv4 and YOLOv5 for autonomous landing spot detection in faulty UAVs. *Sensors* 22:464. doi: 10.3390/s22020464

Park, S. (2009). Social responses to virtual humans: The effect of human-like characteristics. United States: Georgia Institute of Technology.

Recarte, M. A., and Nunes, L. M. (2000). Effects of verbal and spatial-imagery tasks on eye fixations while driving. *J. Exp. Psychol. Appl.* 6, 31–43. doi: 10.1037/1076-898X.6.1.31

Robbins, C. J., Allen, H. A., and Chapman, P. (2019). Comparing drivers' visual attention at junctions in real and simulated environments. *Appl. Ergon.* 80, 89–101. doi: 10.1016/j.apergo.2019.05.005

Scott, J. J., and Gray, R. (2008). A comparison of tactile, visual, and auditory warnings for rear-end collision prevention in simulated driving. *Hum. Factors* 50, 264–275. doi: 10.1518/001872008X250674

Shi, Y., Boffi, M., Piga, B. E., Mussone, L., and Caruso, G. (2022). Perception of driving simulations: can the level of detail of virtual scenarios affect the Driver's behavior and emotions? *Trans. Vehicul. Technol. IEEE.* 71, 3429–3442. doi: 10.1109/TVT.2022.3152980

Shi, Y, Maskani, J, Caruso, G, and Bordegoni, M. (2019). Explore user behaviour in semi-autonomous driving. In Proceedings of the Design Society: International Conference on Engineering Design. Cambridge University Press. 1. 3871–3880.

Ultralytics Team. (2023). Deploy YOLOv5 With Neural Magic's DeepSparse for GPU-Class Performance on CPUs. Available at: https://www.ultralytics.com/blog/deploy-yolov5-with-neural-magics-deepsparse-for-gpu-class-performance-on-cpus. (Accessed on 01 June 2024).

Wickens, C. D. (2008). Multiple resources and mental workload. *Hum. Factors* 50, 449–455. doi: 10.1518/001872008X288394

Wiese, E. E., and Lee, J. D. (2004). Auditory alerts for in-vehicle information systems: the effects of temporal conflict and sound parameters on driver attitudes and performance. *Ergonomics* 47, 965–986. doi: 10.1080/00140130410001686294

Zou, Z., Mubin, O., Alnajjar, F., and Ali, L. (2024). A pilot study of measuring emotional response and perception of LLM-generated questionnaire and human-generated questionnaires. *Sci. Rep.* 14:2781. doi: 10.1038/s41598-024-53255-1