Visual Feedback for In-car Voice Assistants

THOMAS MARINISSEN, Eindhoven University of Technology, The Netherlands
PAVLO BAZILINSKYY*, Eindhoven University of Technology, The Netherlands

This study presents ambient visual feedback for automotive voice assistants to enhance driver interaction and safety through peripheral visual cues. A user interface prototype incorporating ambient color feedback was evaluated through an online survey (N = 151 from 28 countries) and a lab-based user study (N = 24). Survey participants strongly preferred smartphone-integrated user interfaces, such as Android Auto and Apple CarPlay, over built-in manufacturer systems, indicating a desire for consistent digital ecosystems. In the user study, 18 participants favored the ambient feedback method over conventional or no visual feedback, citing improved visibility and assistance. Statistical analysis revealed that ambient feedback improved user visibility, position, and usefulness ratings. However, the need for auditory cues remained evident, confirming the importance of multimodal feedback in vehicles. These findings suggest that ambient visual feedback is a promising direction for improving the usability of voice assistants and driver satisfaction while supporting safer in-vehicle interaction.

CCS Concepts: • **Human-centered computing → Auditory feedback**; **Empirical studies in interaction design**.

Additional Key Words and Phrases: Automotive, User Interface, Voice Assistant, Visual Feedback, Speech Commands

ACM Reference Format:
Thomas Marinissen and Pavlo Bazilinskyy. 2018. Visual Feedback for In-car Voice Assistants. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation emai (Conference acronym 'XX)*. ACM, New York, NY, USA, 16 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

Modern automotive user interfaces (UIs) extend beyond traditional physical buttons and touchscreens by incorporating natural input modalities such as voice recognition [11, 12, 31, 32], driver state monitoring through eye-gaze tracking [16, 33, 51], and gesture controls [6, 7, 43], or combinations of these interaction modalities [4, 8]. Voice control has advanced from being a clunky gimmick to understanding natural language thanks to advancements in Large Language Models (LLMs) [34]. Modern voice control systems allow users to keep their eyes on the road while operating the in-vehicle infotainment system and have been shown to be safer to use than touch-screen controls [39]. However, a study conducted by the British transport consulting and research company TRL [46] found that voice-operated infotainment systems increase driver reaction times. Although the effect on driver reaction times is not as great as when using touch screens, it is worse than the impact of driving under the influence of alcohol or drugs [39]. Infotainment systems bring many desirable features to users, but they could take away attentional demand from the primary driving task if they are poorly designed [27]. According to Lentz et al. [24], the problems users might encounter with poorly designed UIs are difficulty locating the correct option they need, unintended invocation of actions, tedious sequences

Authors' Contact Information: Thomas Marinissen, thomas.j.marinissen@gmail.com, Eindhoven University of Technology, Eindhoven, The Netherlands; Pavlo Bazilinskyy, p.bazilinskyy@tue.nl, Eindhoven University of Technology, Eindhoven, The Netherlands.

of interactions, error-prone repetitive actions or being overwhelmed by too many choices. These problems distract the driver, as they increase the cognitive load required to operate the infotainment system. As driver distraction is estimated to be a factor in up to 30% of all road collisions in Europe [47], it is necessary to improve interactions with car infotainment systems and their voice assistants (VAs, often referred to as "virtual assistants"). EuroNCAP [13] recognized this need and made the 2026 safety regulations stricter on the design of touchscreen controls [14, 45].

1.1 Ambient Displays

One solution that has the potential to present visual information without requiring too much attention is the use of ambient displays and peripheral interactions [27]. Ambient displays can be defined as "*abstract and aesthetic peripheral displays portraying non-critical information on the periphery of a user's attention*" according to Mankoff et al. [28]. They designed a set of heuristics specifically for ambient displays, based on Nielsen's usability heuristics for UI design [35].

These design principles are in line with the design solutions offered by Lentz et al. [24], such as simplifying user interactions by limiting user choices or providing context-aware responses. The Web Content Accessibility Guidelines (WCAG) [50] provide principles for UI design that can also be applied to in-vehicle interfaces. The key principles are Perceivable, Operable, Understandable, and Robust. The Perceivable principle recommends a sufficient contrast between text and background and not solely relying on the use of color to convey important information, which is critical for color-blind users. However, in practical implementations in cars, red generally indicates critical alerts, yellow/amber indicates caution and green indicates normal operation [50]. As the visual feedback from VAs is non-critical, the color red should therefore be avoided. Using these design principles, it is possible to design visual feedback for VAs that does not demand as much cognitive load from the driver as compared to conventional VA feedback visualizations. This reduction in cognitive load should have a positive effect on road safety.

1.2 Visual Feedback for Voice Assistants

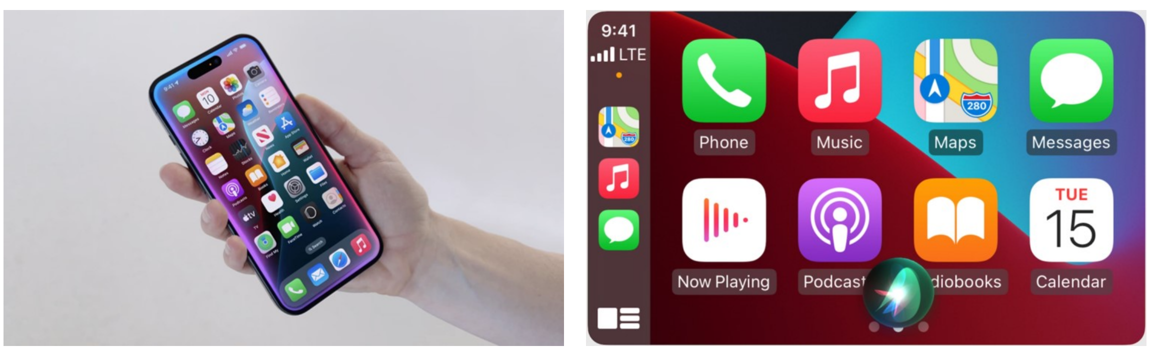


Fig. 1. Visual feedback for Apple's Siri in different environments. Left: Siri on iOS 18 [1]. Right: Siri on CarPlay [5].

Effective visual feedback for VAs is crucial to ensure that drivers can interact with these systems safely and efficiently. In the latest generation of Apple's mobile operating system, iOS 18, the visualization of their VA Siri will change from an interactive logo to an interactive ambient color-changing ring around the screen [3] (left in Figure 1). In Apple CarPlay [2], the "traditional" Siri logo remains as seen on the right side of Figure 1. This conventional, rather small, visualization is not ideal for drivers, as its size can be problematic in meeting the "Visibility of state" and "Peripherality of display" criteria set by Mankoff et al. successfully [28]. Google visualized their VA based on defined system states: Passive, Listening, Thinking, Replying, Incomprehension, and Confirmation, as seen in [25]. Each state has a distinct visualization designed to enhance user interactions, following design patterns from Google's Material 3 design system [18]. These six states employed by Google can also be effectively utilized in ambient design.

Fig. 2. Ambient lighting display in the Volkswagen ID3 [49]. Left: ambient navigation indicators. Right: lighting variations for different functions such as charging status or VA feedback.

Some cars, such as the 2024 Mercedes-Benz S-Class and the 2024 Volkswagen ID3 [15, 41, 49] (seen in Figure 2), already use ambient displays to give visual feedback to the VA or other functions. Although these advanced ambient lighting systems are very effective in bringing information to the periphery of a user's attention, few cars are equipped with them yet. As car manufacturers aim to reduce the number of parts, cost and complexity of their cars and strive to maximize the use of the hardware in the car, a more cost-effective approach could be to provide ambient information about the VA on the conventional infotainment display. Customizing the VA visualization for the user is necessary: research indicates that no single voice is universally suitable for all listeners and situations [21]. This could suggest that the same principle applies to its visualization characteristics, which may need to be tailored to individual preferences and contexts for optimal effectiveness.

1.3 The Importance of User Experience in Cars

Hassenzahl (2010) defined that interactions with products can be characterized by two distinct qualities [19]: (1) pragmatic qualities, which have practical and goal-oriented value and can be assessed through usability metrics, and (2) hedonic qualities, which gain value from the enjoyment or pleasure of using a product. These qualities are closely related to the usefulness (pragmatic) and satisfying (hedonistic) scores on the Acceptance scale [48]. As cars are both a utilitarian and emotional product for users [37], user interactions with cars should be reviewed on both aspects.

The user experience (UX) of cars has become increasingly important to users and is therefore considered by car manufacturers from the early stages of the development of new cars [22]. The infotainment system of a car and all interaction methods such as screens, buttons, and voice controls play an integral role in the car's UX. These systems are dependent on the hardware installed in cars, underscoring their importance as evidenced by the significant investments that car manufacturers make in their development and production. Electronics are projected to account for roughly 45% of the total cost of a new car in 2030, more than double the proportion in the year 2000 [10]. The investments made in improving the UX of cars go beyond hardware components: Mercedes-Benz, for example, is developing its operating system MB.OS and allocated 25% of its R&D budget on software by the mid-2020s [30]. Simultaneously, a study by

McKinsey & Co has found that (1) nearly half of car buyers would not purchase a car without Apple CarPlay [2] or Android Auto [17] and that (2) of those who have either of these systems in their car, 85% prefer it over the built-in system of the manufacturer. In other words, the user's need for phone connectivity in cars is dictating the battle in UI preferences in cars and is important to consider when improving the overall UX of cars.

1.4 Aim of Study

This study examines the effectiveness and satisfaction of a peripheral ambient feedback visualization for the VA of a prototype automotive UI. The novel visualization method is compared in a user study to a traditional method of visualizing VA in cars and not showing any visual feedback at all. Expert interviews with industry stakeholders and an online user questionnaire were conducted in preparation for the user study. New insights were included to improve the prototype over multiple iterations. The research questions are defined as follows:

RQ1: How does peripheral ambient visual feedback to automotive voice assistants impact driver satisfaction and effectiveness compared to traditional visual feedback methods?

RQ2: What effect do smartphone operating system, usage, and preferences have on user preferences for automotive user interfaces and voice assistants?

Overall, this study aims to design ambient visual VA feedback that users prefer over conventional feedback visualizations and can contribute to increased road safety.

2 Method

2.1 Online Questionnaire

The research was approved by the Human Research Ethics Committee of the Eindhoven University of Technology, and the participants gave their informed consent to use their data. To gain insights into people's preferences for infotainment systems and VA in cars, we created an online questionnaire. Participants were asked about their preferences for and usage of VAs in cars and smartphones and their preferences for visual feedback from VAs. The full list of questions can be found in the supplementary material. This questionnaire was shared with family and friends (snowball sampling) using Google Forms (https://docs.google.com/forms) and posted on the Appen crowdsourcing platform (http://appen.com). In Google Forms, no financial compensation was provided to the participants. In Appen, participants received 0.50 EUR.

A total of 151 people from 28 countries answered a questionnaire between July 5 and July 19, 2024. All participants (59 female and 92 male) were older than 18 with a mean age of 34.8 years (SD = 13.3, median = 29). 24 participants did not have a driver's license, 26 participants owned a driver's license for less than five years, and 100 participants owned a driver's license for more than five years.

2.2 User Study

A total of 24 people from eight different nationalities (Flemish Belgian, Brazilian, Polish, Dutch, Italian, Spanish, Greek, and Taiwanese) participated in a user study in July 2024. All participants (11 female and 13 male) were older than 18 with a mean age of 43.1 years (SD = 12.9, median = 46.5). 23 participants had a driver's license, one did not. None of the participants were native English speakers.

Figure 3 shows the apparatus for the user study in the two closed rooms used as test environments: location A in Leuven, Belgium and location B in Brussels, Belgium. The numbered items represent the following components: (1) primary screen with driving footage AOC 24B1H (23.6" monitor) in location A or Philips Brilliance 235PL (23" monitor) in location B and (2) secondary (laptop) screen with automotive UI prototype and integrated microphone Lenovo Legion 5 15ACH6H (15.6" monitor). This screen was positioned as close as possible to the position of the infotainment display in a left-hand drive car, offset to the right of the driver. The sound level in these rooms, measured shortly before each participant arrived, varied between 35 and 50 dB in both locations.

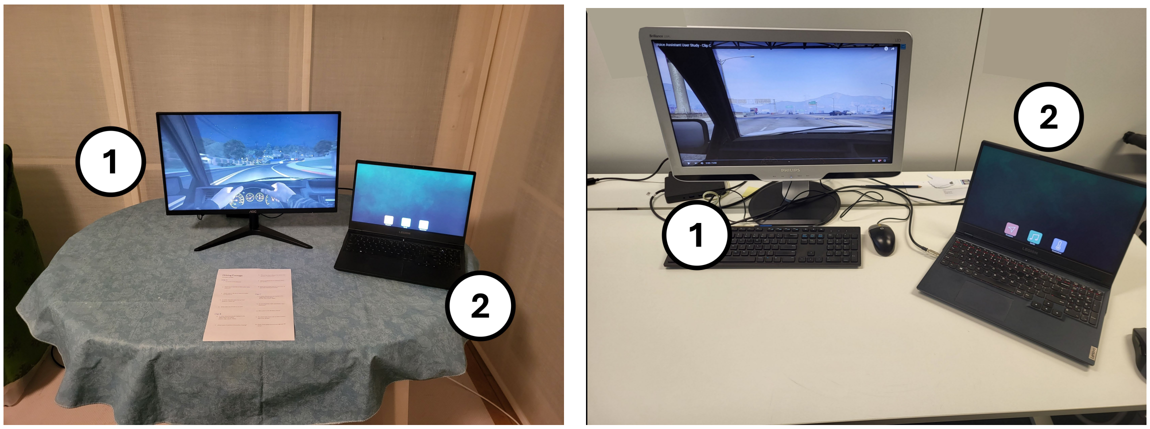


Fig. 3. User study setup. Left: location A. Right: location B.

The primary screen displayed first-person driving footage to simulate a real-life driving scenario in which the user must pay attention to the road. Three videos, recorded while driving according to the speed of traffic in first-person view in an unmodified version of the game Grand Theft Auto V [40], were shown to the participants. These videos, found in the supplementary material, were shown as videos playing on YouTube (https://www.youtube.com). The participants were aware that the examiner started the videos through a wireless keyboard. They saw the YouTube UI briefly before and after watching the videos. Each video lasted 5 minutes, consisted of mixed driving conditions (rural roads and highways in the southern half of the game map), and contained audible traffic and engine sound. A screenshot of each video can be seen in Figure 4.

(a) Video A: nighttime driving. (b) Video B: daytime driving. (c) Video C: driving during sunset.

Fig. 4. Screenshots of the driving footage shown on the primary screen.

The secondary screen displayed a simple automotive UI prototype, seen in Figure 5. This UI was specifically created for this study in the (paid) basic subscription version of ProtoPie [44]. The UI contained three usable applications: a navigation app, a music player, and temperature controls. These applications had limited functionality and could only be controlled by voice commands (see supplementary material for a complete list). By pressing a dedicated hotkey, the system was activated to start listening (hotkey mapping found in supplementary material). The standard ProtoPie speech recognition functionality was used for the voice commands: this meant that only preprogrammed voice commands were recognized by the system. The standard American English female voice of ProtoPie was used to give spoken feedback (confirmation when a task was understood or completed correctly) or to answer questions asked by the participants.

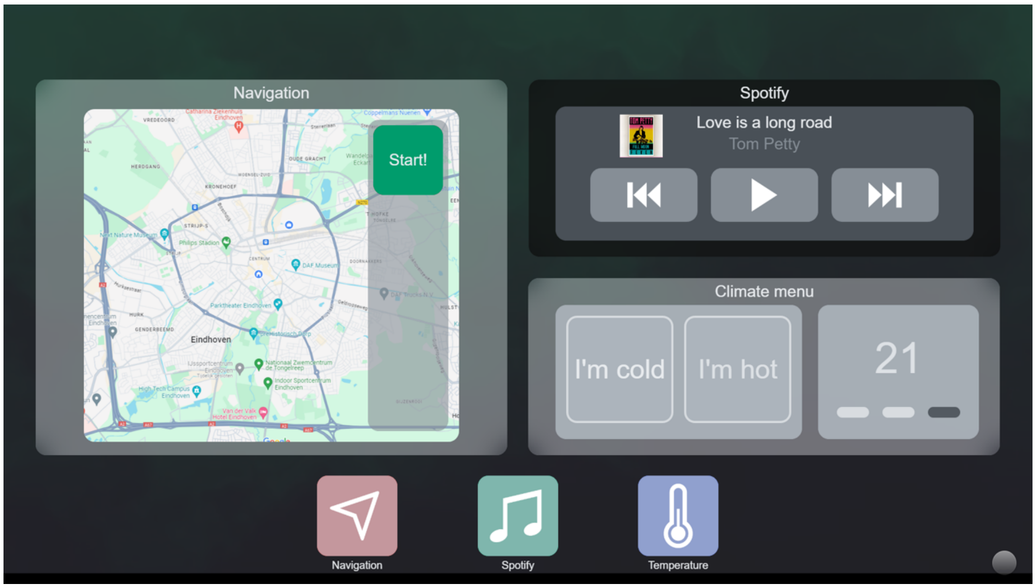


Fig. 5. UI prototype in the passive state with all apps opened simultaneously.

The UI was created with two variations for the visual feedback of the VA, called Concept 2 (C2) and Concept 3 (C3), which could be independently switched on or off through dedicated hotkeys. Turning both off resulted in no visual feedback: Concept 1 (C1). The selected visual feedback was automatically displayed when the space bar hotkey was pressed to start voice recognition. It could also be displayed by the click of another hotkey that did not start the voice recognition. The error and success states could also be triggered by separate hotkeys.

The participants' primary task was to give voice commands to the VA prototype. After every one-minute interval, the examiner told the participants which command to give, after which they had to repeat the command to trigger the system. Communication between the examiner and the participants was conducted in English, Dutch, or Portuguese depending on the preferences of the participants, but all voice commands were given exclusively in English. Table 1 lists the voice commands used. A "Hey car" command always preceded the voice commands, as participants were told that was the required trigger for the system to start listening to other voice commands; in reality, the examiner pressed a key on a separate keyboard to trigger the VA to start listening. This happened out of the participant's eyesight and they were not made aware of this so as not to break the immersion of the VA. After the trigger command "Hey car", the prototype accepted commands to control the UI. For Commands A3, B3, and C3 the examiner introduced one false error for the VA: instead of pressing the space bar after the trigger command "Hey car" was given, the hotkey was pressed, which did not start the voice recognition system and only the visual feedback of the listening state. This gave the participants the impression that the system was listening to their command but did not recognize it. This was done to ensure that each participant experienced an error scenario at least once for each visual feedback variation.

At the same time, participants were asked to complete a secondary task: a list of questions about the driving footage was provided on a printed sheet of paper (see supplementary material). Participants were given as much time as they needed to read the questions before the video started, could reread the questions during the video, and were allowed to write the answers on paper during the video or do it afterward. The purpose of the secondary task was to create the immersion of driving in a car and to give participants a reason to focus on the primary screen rather than the secondary screen. Participants were free to look at the secondary screen whenever they wanted, at the risk of missing an answer to the driving footage questions found on the primary screen; this mimicked a real-life driving scenario in which drivers can only keep their eyes off the road and on the infotainment screen for a short time. The order in which the visualization variants were shown was randomized according to the Latin square method [42]. The order of voice commands and videos was the same for all participants.

Table 1. Voice commands for giving instructions to the prototype during the user study.

Video	**Command**	**Actions**	**False error**
Video A	Command A1	Open navigation & Take me to Brussels	no
	Command A2	Give me a traffic update & Close navigation	no
	Command A3	What is the estimated time of arrival?	**yes**
	Command A4	Open navigation & Take me home	no
Video B	Command B1	Open Spotify & Play Michael Jackson	no
	Command B2	Close navigation	no
	Command B3	Play rock music & Close Spotify	**yes**
	Command B4	Open Spotify & Play Michael Jackson	no
Video C	Command C1	Open temperature & Temperature 20 degrees	no
	Command C2	Set fan speed to medium & Close temperature	no
	Command C3	Weather forecast & Close Spotify	**yes**
	Command C4	Open Spotify & I'm cold	no

The examiner kept track of whether the voice commands (tasks) given by the participant were understood correctly by the system, since the prototype did not log any type of data. If the voice command was understood on the first attempt, the task was marked as "pass". If the voice command was understood on the second or third attempt by the participant, the task was marked as an "eventual pass". If the voice recognition system could not recognize the voice command given by the participant correctly after three attempts, the task was marked as a "fail". The answer sheets for the questions about driving footage were used to check whether the participants found the correct answers. The answers to the questions about the driving footage were kept track of and marked either as "correct" or as "incorrect". All participants filled in a questionnaire (printouts of the forms used are available in supplementary material), which consisted of three parts:

(1) *Before experiment*: questions related to demographic data, namely age and gender, and the use of smartphones and cars.

(2) *After each condition of the user study*: participants were asked to answer the same set of four questions, give the method a grade, and fill in the Acceptance scale [48]. Participants were allowed to provide open feedback for each specific variant they had just experienced.

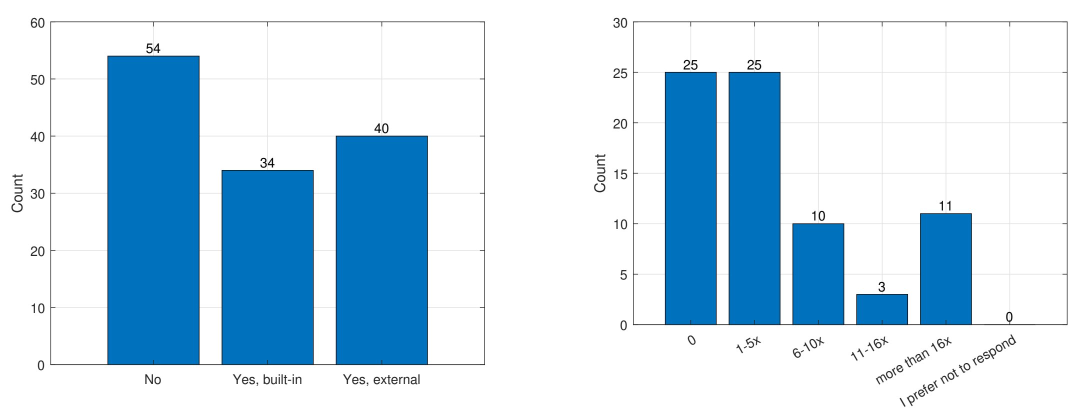
(3) *After experiment*: questions related to the likelihood of using VA for different functions in their car, under different circumstances, and the favorite variant of visual feedback was.

Data from both the online questionnaire and the user study was analyzed in MATLAB 2024A. The chi-square test or ANOVA test was performed to determine the significance of categorical and numerical data. An alpha level of 0.05 was used for all tests.

3 Results

3.1 Results from Online Questionnaire

The questionnaire resulted in a total of 286 responses, 79 from Google Forms and 207 from Appen. The results of Google Forms were all accepted as genuine participant data. On Appen, people from Venezuela were barred from joining the questionnaire: a study by Onkhar et al. (2022) shows an extreme case of Venezuelan people being overrepresented in the participant pool [36]. The ongoing economic crisis there may mean that the primary motivation for participants to participate would be financial compensation, which could lead to people answering questions randomly to reach the end of the questionnaire quickly. Not all Appen participants were trustworthy or suitable for the study. The questionnaire was filled out 85 times in two seconds with the same age and nationality. Although the responses came from different IP addresses, it was assumed that these results were not trustworthy and were therefore excluded from the final data set. Furthermore, 122 participants were excluded based on two additional filters: (1) not being able to respond to four test questions and (2) not meeting the age criteria allowed for 18 to 100 years. Among the responses from Appen, 72 met the requirements to be trusted as genuine participant data. After filtering, the online questionnaire resulted in a combined total of 151 participants. There were 23 participants who did not have access to a car, and the remaining 128 did; either their car or others. Figure 6a shows that out of the 128 participants who have access to a car, 54 participants did not have a VA in their car and the remaining 74 did. Figure 6b shows that out of the 74 participants who can use a VA in their car, 25 participants do not use it and 49 participants do.

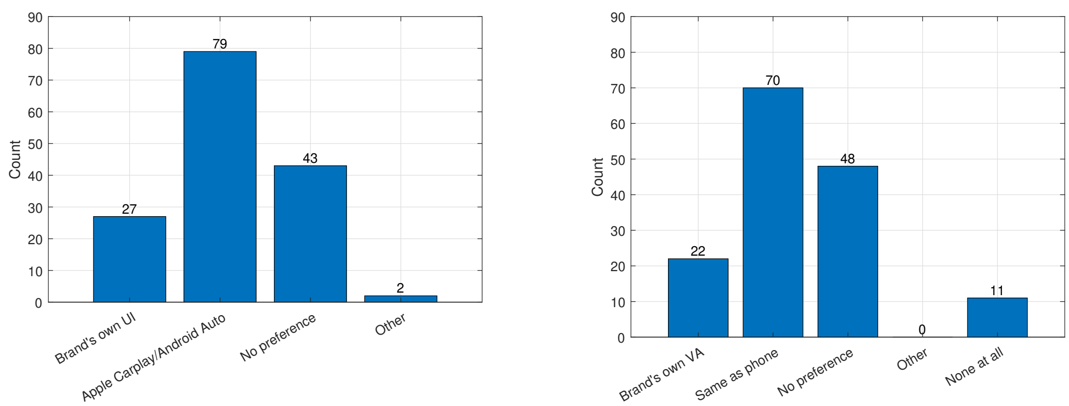


(a) "Does your car have a VA?" (N = 128).

(b) Monthly usage of VA in car (N = 74).

Fig. 6. Results for questions regarding car access and VA usage in cars. The results of Figure 6a and 6b are filtered: only respondents with access to a car and a VA in their car were included respectively.

Of 77 Android users, 34 prefer Android Auto and 13 prefer the manufacturer's interface. Of 72 iOS (iPhone) users, 42 preferred Apple CarPlay and 13 prefer the manufacturer's UI. 33 Android users preferred to use the same VA in the car as they used on their phone, and 13 prefer to use the manufacturer's own VA. 37 iOS (iPhone) users prefer to use the same VA in the car as they use on their phone, and eight prefer to use the manufacturer's own VA. Figures 7a and 7b show a strong overall preference for using UI and VA in the cars that participants use on their phones. These figures show the combined results for all smartphone OS categories. Table 3 shows a significant correlation between the participant's smartphone operating system (OS) and the preferred car UI and VA.



(a) "Which UI would you like to use in your ideal car?"

(b) "Which VA would you like to use in your ideal car?"

Fig. 7. Results for questions regarding preferred UI and VA in cars.

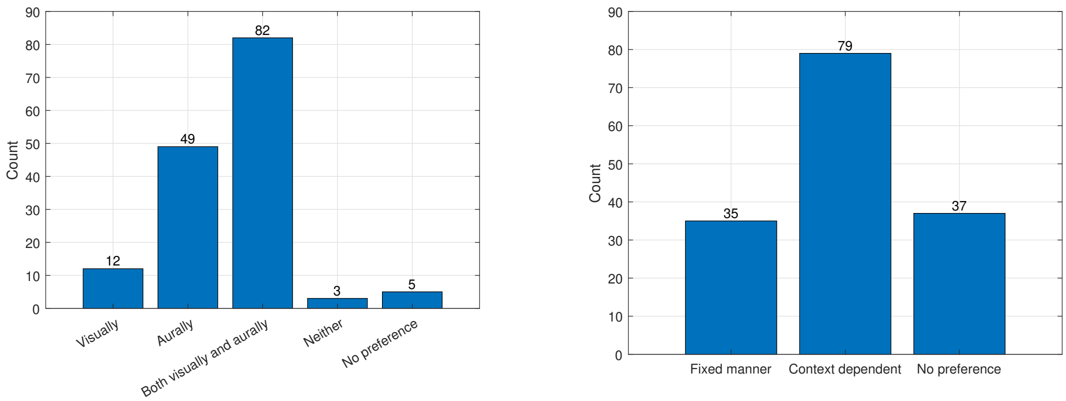
Figure 8a shows that most participants required aural (ie auditory) feedback: a total of 131 participants like auditory feedback, with or without visual feedback. Only 12 participants chose purely visual feedback. Figure 8b shows that 79 participants would be interested in context-dependent feedback.

Figure 9a shows that "context-dependent" and "feedback on top of the screen" score the highest (N = 49 and N = 36 respectively) for visual feedback placement. The background visual feedback scores are the lowest (N = 12). Figure 9b shows that the concept of personalizing the VA is received mostly positively by the participants.

3.2 Results from User Study

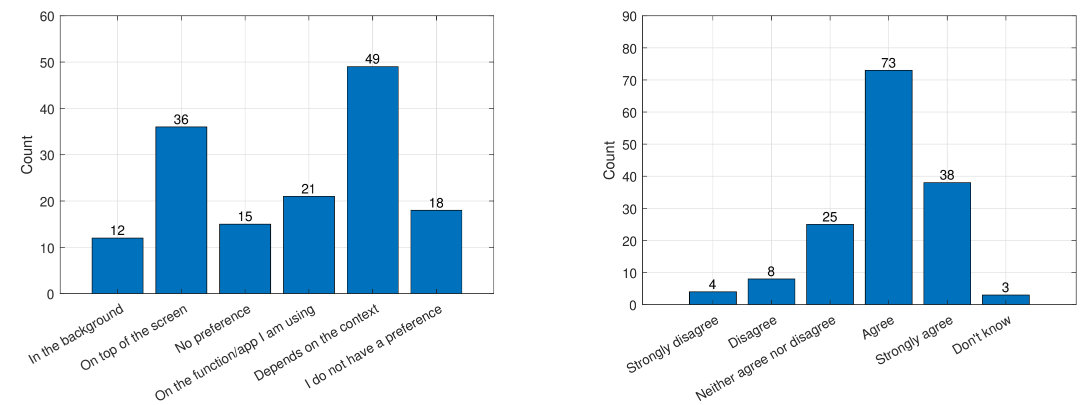
The data of all participants was retained. Figure 10 outlines the responses of participants for four questions. Table 3 shows that on average C3 scores higher than C1 and C2 for the visibility, suitableness of position, helpfulness, grade out of 10, and the Satisfying and Usefulness scores of the Acceptance scale. C3 scores lower than C1 and C2 in the need for auditory feedback. Table 3 shows that in the categories of visibility, helpfulness, grade, satisfaction and usefulness C3 scores significantly higher than C1. C3 scores significantly lower than C2 for needing auditory feedback.

Figure 10 shows the results for which participants in general gave each visual feedback method and the Usefulness and Satisfaction scores on the Acceptance scale [48]. It can be seen that C3 is rated higher than C1 and C2 in these three metrics. Table 3 shows that having some form of visual feedback from the VA (such as in C2 and C3) significantly increases the grade and the Satisfying and Usefulness scores of the Acceptance scale [48] over not having visual feedback (such as in C1). C3 does not improve significantly over C2 for these three metrics. The final question of the user study



(a) "How do you expect to receive feedback from a VA?" (b) "How do you expect to receive visual feedback from a VA?"

Fig. 8. Participants' preferences for VA feedback in cars.



(a) "Where do you expect to see visual feedback from a VA?" (b) "VAs in cars should be personalized to the user."

Fig. 9. Participants' preferences for VA placement and personalization in cars.

asked participants to mark their favorite visual feedback method. 18 participants marked C3 as their favorite visual feedback method, 4 responded C2, 1 responded C1, and 1 participant had no preference.

In open feedback, the visibility and position of the visual feedback were mentioned too: for C2, 4 participants mentioned that the position of the visual feedback should be closer to the driver to better appear in the peripheral vision. Not all changes in color for the different states of the system were perceived by the participants (mentioned 13 times for C2 and 7 times for C3): the error state (orange flashes) was generally perceived and understood, but the confirmation state was not seen by the participants making these remarks. The movement of the screen was better detected in the peripheral vision than the color changes: for C3, some participants did not remember how the system changed visually when switching from the passive state to the listening state, but they were aware that the system had responded to their voice command. One participant wrote: *"I did notice something was changing when the car was*

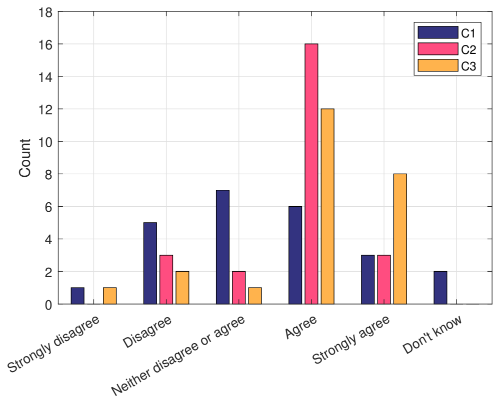
Table 2. Statistical analysis: chi-square test for categorical data. Significant scores marked in bold. (Note: FB stands for feedback)

	Gender	Driver's license	Transport method	Car access	Freq. car usage	Yearly km driven	VA in car	Freq. VA in car	Phone OS	Phone screen time	Freq. VA on phone	Pref. Car UI	Pref. car VA	Pref. VA FB	Pref. VA visual FB	Pref. place for VA FB
Gender	x															
Driver's license	0.322	x														
Transport method	**0.017**	**<0.001**	x													
Car access	**0.011**	**<0.001**	**<0.001**	x												
Freq. car usage	0.793	0.064	**<0.001**	**<0.001**	x											
Yearly km driven	0.214	**<0.001**	**0.001**	**<0.001**	**<0.001**	x										
VA in car	**0.001**	**0.024**	**0.048**	**0.001**	0.141	**0.010**	x									
Freq. VA in car	0.848	**<0.001**	0.288	0.163	0.219	**<0.001**	**<0.001**	x								
Phone OS	0.532	0.154	0.281	0.207	0.119	0.375	0.429	0.999	x							
Phone screen time	0.332	0.258	0.275	0.276	0.109	0.060	0.497	0.941	**<0.001**	x						
Freq. VA on phone	0.578	0.805	0.537	0.231	**0.037**	0.057	0.211	**<0.001**	0.862	0.170	x					
Pref. Car UI	0.416	0.599	0.272	0.395	0.183	0.133	**0.024**	**0.025**	**<0.001**	0.453	0.069	x				
Pref. car VA	0.145	0.510	0.336	0.804	0.080	0.573	0.452	0.396	**0.016**	**0.002**	**0.007**	**<0.001**	x			
Pref. VA FB	0.580	0.210	0.098	0.437	0.402	0.690	0.336	0.439	0.315	**0.015**	0.397	0.314	**<0.001**	x		
Pref. VA visual FB	0.798	0.681	0.782	0.686	0.252	0.586	0.072	0.097	0.283	0.621	0.191	0.227	**0.012**	**<0.001**	x	
Pref. place for VA FB	0.669	0.862	0.466	0.518	0.225	0.326	0.949	0.993	0.889	0.861	0.058	0.954	0.180	**0.045**	0.628	x

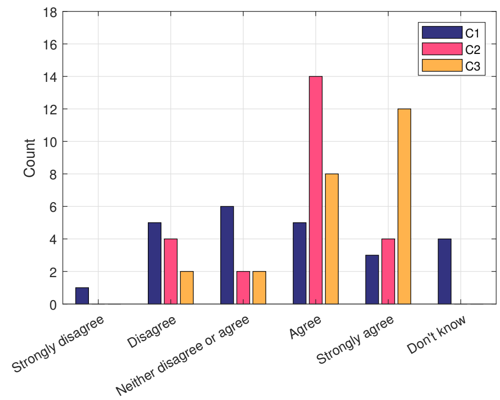
Table 3. Mean scores for user preferences and Acceptance scale [48], with standard deviation in parentheses, and post-hoc Anova test results. Significant scores are marked in bold.

	M (SD)			*p*-value of post-hoc ANOVA		
	C1	C2	C3	C1 vs. C2	C1 vs. C3	C2 vs. C3
Visibility	0.208 (1.062)	0.792 (0.833)	1.000 (1.063)	0.765	**0.025**	0.986
Position	0.167 (1.050)	0.750 (0.944)	1.250 (0.944)	**0.002**	0.689	0.736
Helpfulness	-0.167 (1.167)	0.958 (0.955)	1.125 (0.900)	0.476	**0.009**	0.068
Auditory	1.333 (1.050)	1.000 (0.933)	0.583 (1.100)	0.064	0.331	**0.014**
Grade	5.292 (2.236)	6.708 (1.429)	7.500 (1.720)	**0.024**	**<0.001**	0.296
Satisfaction	-0.063 (1.061)	0.750 (0.711)	0.938 (0.959)	**0.009**	**0.001**	0.762
Usefulness	-0.067 (0.965)	0.617 (0.736)	1.050 (0.893)	**0.022**	**<0.001**	0.203

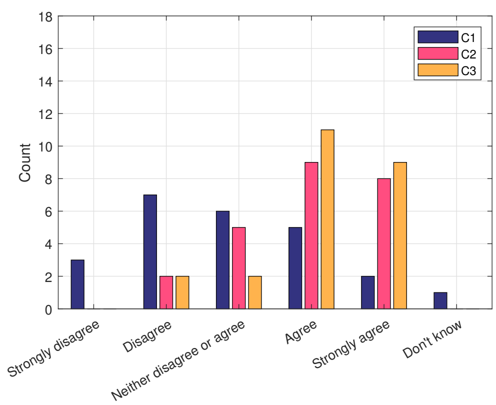
listening, but I was not fully aware of which color and how it changed on the screen exactly. When the car flashed yellow, I knew I had to repeat my command." The blinking effect of the listening state in C2 was specifically mentioned as a positive by 3 participants. One participant wrote for C2: *"The pulsing animation raised my attention but the dimension of the dot and the position made it not as visible as it was for other feedback."* For C1, 8 participants gave negative feedback on not having any visual feedback: there were mentions of needing auditory feedback as well as visual feedback and that this method might not be safe. One participant wrote: *"I need some kind of feedback from the system which does not stay quiet when it does not understand me. In this way, we can interact better when it did not understand me or if there was a delay."*



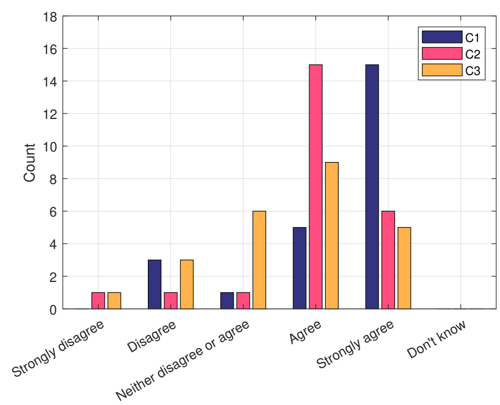
(a) "The visual feedback of the VA was clearly visible."



(b) "The visual feedback of the VA was located in a well-suited position on the screen."



(c) "The visual feedback of the VA helped me understand what the system was doing."



(d) "Auditory feedback is necessary with this specific type of visual feedback."

Fig. 10. Participant responses for questions regarding visual aspects of VA visualizations.

4 Discussion

This study aimed to evaluate the effectiveness of a novel ambient visual feedback system for VAs in cars, comparing it to conventional visual feedback and without visual feedback. In response to **RQ1**, the results indicate a strong preference for the ambient visualization method (C3, N = 18) over the conventional method (C2, N = 4) and no visual feedback (C1, N = 1), aligning with the literature that ambient displays can enhance user experience and safety in automotive environments [23, 38]. In particular, these results from the user study contrast with the insight of the online questionnaire, where the concept of background VA visualization was the least popular suggested method (as seen in Figure 9a).

The results of the survey demonstrated a strong preference of the respondents for Apple CarPlay and Android Auto over the proprietary UIs developed by the car manufacturers. These results address **RQ2** and are consistent with the literature, which highlights seamless integration and familiarity with these smartphone-based systems as key factors driving user satisfaction [26, 29]. The participants favored the continuity of using the same UI and VA in different environments, suggesting that automotive UIs should aim to achieve greater compatibility with popular smartphone platforms.

The ambient visualization method (C3) was preferred by 18 out of 24 participants, significantly outperforming both C1 and C2 on various metrics. C3 was rated higher in terms of visibility, positional suitability, and assistance value (Figures 10a-10c). This preference for C3 indicates that ambient visual feedback is more noticeable and helpful to drivers, possibly due to its less intrusive nature and better integration with peripheral vision [20]. Despite the strong preference for C3, the improvements over C2 were not statistically significant in some metrics, such as overall grade, satisfaction, and usefulness (Table 3). This suggests that while ambient visualization is favored, the conventional method still holds substantial value and may be effective in scenarios where less cognitive load is demanded from the driver. A critical insight from the study is the necessity of auditory feedback in conjunction with visual feedback. Most of the participants agreed or strongly agreed that auditory feedback is essential in all three conditions (Figure 10d). This finding underscores the importance of multimodal feedback systems in automotive environments, as purely visual feedback cannot replace the immediacy and clarity provided by auditory cues, especially when the driver's visual attention is primarily focused on the road [9].

4.1 Limitations and Future Work

The prototype's secondary screen did not align with typical infotainment display positions, affecting visualization visibility. This could be addressed through on-road testing or AR/VR simulations. Voice recognition was limited, often requiring repetition and misinterpreting commands. Improving voice technology is essential, as natural language interaction is key to effective use of virtual assistants. Visual feedback alone is insufficient if voice input remains unreliable. The prototype lacked features found in systems such as Apple CarPlay or Android Auto, which may have influenced participant impressions. Future studies should use more feature-complete or real-world UIs. The absence of gesture or gaze tracking and quantifiable interaction data also limited objective evaluation. Including these elements would strengthen the behavioral analysis. Some participants over-focused on the secondary task, unlike real-world drivers who would glance at displays more frequently. A more naturalistic setting would offer better insight. Future research should focus on improving voice recognition, testing in realistic contexts, and exploring multimodal feedback. The role of VA feedback in automated driving also deserves attention, as the shifting of cognitive load may change the way ambient displays are used.

5 Conclusion

This study explored the effectiveness of ambient visual feedback for voice assistants (VAs) in cars, comparing it with conventional visual feedback and no visual feedback. The results indicate a clear preference for the ambient visualization method (C3), which was favored by 18 out of 24 participants. C3 outperformed the other methods in terms of visibility, positional suitability, and assistance value. However, the need for auditory feedback alongside visual feedback (multimodal feedback) was also emphasized, as purely visual feedback was considered insufficient for effective interaction. Despite the promising results for ambient visual feedback, limitations such as restricted voice recognition capabilities and the prototype's limited functionality were noted. Future research should focus on improving voice recognition, developing feature-complete prototypes, and conducting studies in more naturalistic driving environments.

In general, the findings suggest that ambient visual feedback has significant potential to improve driver interaction and satisfaction with VAs in cars, pointing to a promising direction for future automotive UI development.

6 Supplementary Material

The supplementary material containing the questionnaire, videos and analysis code can be found at: https://www.dropbox.com/scl/fo/jmsn9g4y9g5zj8kc95xvc/ALltJ2SpPvNnskPCy7YSL2g?rlkey=toeyyqrdpjsfxxlr0a74x4s4f. The ProtoPie prototype can be accessed here: https://cloud.protopie.io/p/4dfbded4fb1d4793a626cf6c.

References

[1] 9to5mac. 2024. How iOS 18 will supercharge Siri with Apple Intelligence. https://9to5mac.com/2024/06/18/how-ios-18-will-supercharge-siri-with-apple-intelligence/. Accessed: 16 January 2025.

[2] Apple. 2023. Car keys and CarPlay. A smarter ride from start to finish. https://www.apple.com/ios/carplay/. Accessed: 16 January 2025.

[3] Apple. 2024. Siri. https://www.apple.com/siri. Accessed: 16 January 2025.

[4] Audi. 2022. Space travel in the heart of the megacity. https://www.audi-mediacenter.com/en/press-releases/space-travel-in-the-heart-of-the-megacity-14595. Accessed: 1 March 2025.

[5] AutoEvolution. 2021. Siri Suddenly Suffering From Amnesia on CarPlay, Forgetting Users' Names. https://www.autoevolution.com/news/siri-suddenly-suffering-from-amnesia-on-carplay-forgetting-users-names-159207.html. Accessed: 16 January 2025.

[6] BMW Group. 2015. The new BMW 7 Series. https://www.press.bmwgroup.com/global/article/detail/T0221224EN/the-new-bmw-7-series. Accessed: 4 April 2025.

[7] BMW Group. 2019. Get the most out of gesture control – BMW How-To. https://www.youtube.com/watch?v=_mGwJh4da5w. Accessed: 1 March 2025.

[8] Marie-Luce Bourguet. 2003. Designing and prototyping multimodal commands. In *Human-Computer Interaction – INTERACT'03*. IOS Press, Zurich, Switzerland, 717–720. https://doi.org/10.1007/978-3-540-45105-7_91

[9] Max Braun, Dominik Weber, and Klaus Bengler. 2020. Multimodal feedback in in-vehicle information systems: A review of current design strategies and future challenges. *Multimodal Technologies and Interaction* 4, 2 (2020), 12.

[10] Car and Driver. 2020. Electronics Account for 40 Percent of the Cost of a New Car. https://www.caranddriver.com/features/a32034437/computer-chips-in-cars/. Accessed: 16 January 2025.

[11] Cerence. 2021. How Mercedes is Creating Innovative Multi-Modal Experiences with Cerence Look. https://www.cerence.com/news-releases/news-release-details/how-mercedes-creating-innovative-multi-modal-experiences-cerence/. Accessed: 4 April 2025.

[12] Cerence. 2023. Gaze Detection. https://www.cerence.com/cerence-products/apps-multi-modality. Accessed: 4 April 2025.

[13] Euro NCAP. 2024. Euro NCAP - For Safer Cars. https://www.euroncap.com/en Accessed: 21 Januar 2025.

[14] European Transport Safety Council. 2024. Cars will need buttons, not just touchscreens, to get a 5-star Euro NCAP safety rating. https://etsc.eu/cars-will-need-buttons-not-just-touchscreens-to-get-a-5-star-euro-ncap-safety-rating/ Accessed: 21 Januar 2025.

[15] Everything Electric Show. 2021. Robert, Kryten The ID.3 Voice Control | SUBSCRIBE to Fully Charged PLUS. https://www.youtube.com/watch?v=FWOmWmrz1zQ. Accessed: 16 January 2025.

[16] Rafael C. Gonçalves, Courtney M. Goodridge, Natasha Merat, Jonny Kuo, and Mike G. Lenné. 2024. Using driver monitoring to estimate readiness in automation: a conceptual model based on simulator experimental data. *Cognition, Technology Work* (2024). https://doi.org/10.1007/s10111-024-00777-3

[17] Google. 2024. Android Auto. Maak kennis met Android Auto. https://www.android.com/intl/nl_nl/auto/. Accessed: 16 January 2025.

[18] Google. 2024. Material Design. https://m3.material.io/ Accessed: 5 March 2025.

[19] Marc Hassenzahl. 2010. *Experience Design: Technology for All the Right Reasons*. Synthesis Lectures on Human-Centered Informatics, Vol. 3. Morgan Claypool Publishers. https://doi.org/10.2200/S00261ED1V01Y201003HCI008

[20] Qing Huang, Shuo Zhao, and Wei Wang. 2022. Peripheral interaction for in-car systems: Ambient displays and visual distraction in simulated driving. *IEEE Transactions on Human-Machine Systems* 52, 1 (2022), 88–97.

[21] Marie Jonsson and Nils Dahlbäck. 2009. Impact of voice variation in speech-based in-vehicle systems on attitude and driving behaviour. In *Human Factors and Ergonomics Society Europe Chapter (HFES)*. HFES.

[22] Alexander Kreis, Daniel Fragner, and Mario Hirz. 2023. User Experience in Modern Cars – Definition, Relevance and Challenges of Digital Automotive Applications. *Usability and User Experience* 110 (2023), 1–8. https://doi.org/10.54941/ahfe1003172

[23] Jisoo Lee and Woohun Ju. 2023. Visual feedback for in-vehicle voice assistants: Exploring ambient and foveal modalities in multimodal interactions. *International Journal of Human–Computer Interaction* 39, 2 (2023), 103–117.

[24] Lentz, Alison and Schlesinger, Benny and DiMartile III, John Thomas and Taubman, Gabriel and O'Dell Regina. 2018. A logical layer to interpret user interactions. https://www.tdcommons.org/dpubs_series/1223/. Accessed: 16 January 2025.

[25] Royen Lock. 2015. Google dots. https://www.youtube.com/watch?v=IYyRpZglZP4 Accessed: 5 March 2025.

[26] Maolin Lyu, Jiali Yang, Yifan Zhang, and Yifan Liu. 2020. CarUX: A comparative study of smartphone-based and built-in automotive user interfaces. *AutomotiveUI '20: Proceedings of the 12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (2020), 50–60.

[27] Andreas Löcken, Shadan Sadeghian, Heiko Müller, Thomas Gable, Stefano Triberti, Cyriel Diels, Christiane Glatz, Ignacio Alvarez, Lewis Chuang, and Susanne Boll. 2017. *Towards Adaptive Ambient In-Vehicle Displays and Interactions: Insights and Design Guidelines from the 2015 AutomotiveUI Dedicated Workshop*. Springer, 325–348. https://doi.org/10.1007/978-3-319-49448-7_12

[28] Jennifer Mankoff, Anind K. Dey, Gary Hsieh, Julie Kientz, Scott Lederer, and Morgan Ames. 2003. Heuristic Evaluation of Ambient Displays. In *Proceedings of the Conference on Human Factors in Computing Systems – CHI'03*, Gilbert Cockton and Panu Korhonen (Eds.). ACM Press, Ft. Lauderdale, Florida, USA, 169–176. https://doi.org/10.1145/642611.642652

[29] McKinsey Company. 2023. How do consumers perceive in-car connectivity and digital services? https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/how-do-consumers-perceive-in-car-connectivity-and-digital-services#/. Accessed: 16 January 2025.

[30] Mercedes-Benz AG. 2023. Mercedes-Benz previews its operating system MB.OS - Software architects. https://group.mercedes-benz.com/investors/events/2023-02-mercedes-benz-group-strategy-update.html. Accessed: 16 January 2025.

[31] Mercedes-Benz AG. 2023. Mercedes-Benz takes in-car voice control to a new level with ChatGPT. https://group.mercedes-benz.com/innovation/digitalisation/connectivity/car-voice-control-with-chatgpt.html. Accessed: 24 January 2025.

[32] Mercedes-Benz USA. 2020. Meet the S-Class DIGITAL: My MBUX (Mercedes-Benz User Experience). https://media.mbusa.com/releases/release-9e110a76b364c518148b9c1ade19bc23-meet-the-s-class-digital-my-mbux-mercedes-benz-user-experience. Accessed: 4 April 2025.

[33] MotorTrend. 2023. The 2024 BMW 5 Series Lets You Steer With Just Your Eyes. https://www.motortrend.com/news/2024-bmw-5-series-eye-lane-change-tech/. Accessed: 4 April 2025.

[34] Humza Naveed, Asad Ullah Khan, Shi Qiu, Muhammad Saqib, Saeed Anwar, Muhammad Usman, Naveed Akhtar, Nick Barnes, and Ajmal Mian. 2023. A Comprehensive Overview of Large Language Models. *arXiv preprint arXiv:2303.07381* (2023). https://arxiv.org/abs/2303.07381

[35] Jakob Nielsen. 1994. Enhancing the Explanatory Power of Usability Heuristics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Boston, Massachusetts, USA, 152–158. https://doi.org/10.1145/191666.191729

[36] V. Onkhar, P. Bazilinskyy, D. Dodou, and J.C.F. de Winter. 2022. The effect of drivers' eye contact on pedestrians' perceived safety. *Transportation Research Part F: Psychology and Behaviour* 84 (2022), 194–210. https://doi.org/10.1016/j.trf.2021.10.017

[37] Egon Ostrosi, Jean-Bernard Bluntzer, Zaifang Zhang, Josip Stjepandić, Bernard Mignot, and Hugues Baume. 2020. Emotional Design: Discovering Emotions Across Cars' Morphologies. In *Emotional Engineering*, Shunji Fukuda (Ed.). Vol. 8. Springer Nature Switzerland AG, 157–175. https://doi.org/10.1007/978-3-030-38360-2_10

[38] Bastian Pfleging, Stefan Schneegass, and Niels Broy. 2016. Design and evaluation of automotive ambient light displays. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (2016), 4066–4070.

[39] R. Ramnath, N. Kinnear, S. Chowdhury, and T. Hyatt. 2020. *Interacting with Android Auto and Apple CarPlay when driving: The effect on driver performance*. PUBLISHED PROJECT REPORT PPR948. TRL Limited, Wokingham, UK. https://trl.co.uk/uploads/trl/documents/PPR948-_IAM-RoadSmart---infotainment-sim-study.pdf

[40] Rockstar Games. 2024. Grand Theft Auto V. https://www.rockstargames.com/gta-v. Accessed: 5 March 2025.

[41] Roland Togonon. 2021. HOW TO USE THE INCREDIBLE VOICE CONTROL ON NEW MERCEDES-BENZ S-CLASS W223 | Incredible Voice Command. https://www.youtube.com/watch?v=Qbw7DJbzA00. Accessed: 16 January 2025.

[42] Valentin Schwind, Stefan Resch, and Jessica Sehrt. 2023. The HCI User Studies Toolkit: Supporting Study Designing and Planning for Undergraduates and Novice Researchers in Human-Computer Interaction. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI EA '23)* (Hamburg, Germany). ACM, New York, NY, USA, 7. https://doi.org/10.1145/3544549.3585890 as found on https://hci-studies.org/balanced-latin-square/.

[43] Screens. 2023. Li Auto L9 | Gesture Control. https://www.youtube.com/watch?v=q09VOemNji8. Accessed: 1 March 2025.

[44] Studio XID. 2024. Protopie website. https://www.protopie.io/. Accessed: 2024-01-13.

[45] The Times. 2024. Stop making dangerous touchscreens, car firms told. https://www.thetimes.com/business-money/technology/article/stop-making-dangerous-touchscreens-car-firms-told-xv3gmpdc6 Accessed: 21 Januar 2025.

[46] TRL. 2024. TRL: Who We Are. https://www.trl.co.uk/about-us/who-we-are. Accessed: 16 January 2025.

[47] TRL, TNO, RAPP-TRANS. 2015. *Study on Good Practices for Reducing Road Safety Risks Caused by Road User Distractions*. Final report. European Commission, Directorate-General for Mobility and Transport, Brussels, Belgium. https://doi.org/10.2832/88265

[48] Jinke D. Van Der Laan, Adriaan Heino, and Dick De Waard. 1997. A simple procedure for the assessment of acceptance of advanced transport telematics. *Transportation Research Part C: Emerging Technologies* 5, 1 (1997), 1–10. https://doi.org/10.1016/S0968-090X(96)00025-3

[49] Volkswagen.nl. [n. d.]. ID. Light - LED's talk. https://www.volkswagen.nl/features/id-light. Accessed: 16 January 2025.

[50] World Wide Web Consortium (W3C). 2018. Web Content Accessibility Guidelines (WCAG) 2.1. https://www.w3.org/WAI/standards-guidelines/wcag/ Accessed: 5 March 2025.

[51] Xunfei Zhou, Tobias Wingert, Maximilian Sauer, and Subrata Kundu. 2020. Development of a Camera-Based Driver State Monitoring System for Cost-Effective Embedded Solution. In *SAE Technical Paper 2020-01-1210*. SAE International, WCX SAE World Congress Experience. https://doi.org/10.4271/2020-01-1210

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009