

Exploring Speech and Non-Speech Audio Feedback in Autonomous Vehicle Handoff Situations

Yan Tian, Adithya Muralikrishna, Christopher Hovey, Jaeri Suh
School of Psychology
Georgia Institute of Technology, Atlanta, Georgia, USA

ABSTRACT

The current study investigated the effectiveness of speech and nonspeech audio feedback on takeover performance and user experience in proactive and reactive scenarios within Level 3 autonomous vehicles. Employing a 2x2 mixed factorial design, the study aimed to determine preferences for audio feedback types and their impact on drivers' takeover time during autonomous vehicle handoff situations. Participants ($N=8$) were subjected to simulated driving scenarios that lacked the criticality to evoke distinct responses between proactive and reactive situations. Contrary to initial hypotheses, results indicated that neither audio feedback type nor situation type significantly affected takeover times or user experience scores. The findings suggest that the type of audio feedback does not influence driver performance or preference in the context presented. These results challenge existing assumptions about the superiority of speech feedback in facilitating quicker driver responses and underscore the need for further research into more personalized feedback systems that consider individual user characteristics and real-world driving complexities. Future studies should expand on scenario realism, diversify participant demographics, and explore the long-term effects of audio feedback on driver habituation to enhance the ecological validity and applicability of findings in autonomous vehicle technology.

Keywords: Autonomous Vehicle, Level-3 AV, Auditory Modality, Human Factors

INTRODUCTION

Autonomous Vehicle Levels

With autonomous vehicles evolving at a rapid rate, ensuring safety and interoperability is paramount. The development of autonomous vehicles relies on comprehensive standards established by organizations like the Society of Automotive Engineers (SAE). SAE describes 6 levels, ranging from zero to five. We will focus on levels 1-3 as they are the most common in cars today. Cars at Levels 1 and 2 necessitate full driver attention for monitoring. These vehicles may feature safety functions like "lane keep" and "adaptive cruise control" to assist drivers. Level 3 vehicles use sensor arrays for environment monitoring, as shown by Tesla's base self-driving package. Here, drivers are expected to intervene as necessary in response to situations the car cannot handle on its own.

Potential Situations

In considering various situations and potential scenarios that may arise during L3 autonomous vehicle (AV) driving, our team conducted explorations to categorize these possibilities. Schneider et al. (2021) introduced terminology encompassing proactive, reactive, critical, and non-critical scenarios to generalize driving conditions. Proactive situations allow ample reaction time, while reactive scenarios demand swift responses. Critical situations entail risks to human lives, whereas non-critical contexts ensure safety. These classifications yield four overarching situations. For instance, a proactive/non-critical scenario might involve an autonomous vehicle encountering a construction site,

requiring anticipation but posing no danger to human life. Conversely, a proactive/critical instance might entail the vehicle preemptively braking to avert a collision in oncoming traffic. Reacting to a sudden obstruction like a dog on the road represents a reactive/non-critical situation, while an emergency brake to avoid endangering lives in oncoming traffic exemplifies a reactive/critical event. Additionally, the transition between autonomous and manual driving, known as the "SAD" model (Stationary, Autonomous, Driving), has been studied by Van der Heiden et al. (2018), who utilized audio cues to signal hand-offs. Yun and Yang (2020) expanded on this work, incorporating visual, auditory, and haptic feedback in experiments focused on AV takeover requests, primarily in reactive/non-critical and proactive/non-critical scenarios.

User Group

Building on the understanding of Automated Vehicles' different levels and four types of scenarios, our research team also considered the appropriate user groups for our investigation. Our team actively researches AV adoption among the younger generation, leveraging insights from various studies. Ruggeri et al. (2018) identified five adopter groups, with the largest proportion of early adopters aged 18-25, while those over 65 tend to be avoiders. Understanding the behaviors and perceptions of the younger generation is crucial for the future of the AV industry. Establishing trust in automated systems is paramount, as misjudgments can hinder adoption rates. Age moderates

technology adoption, necessitating targeted studies across demographics. Factors like driving experience, safety perceptions, privacy concerns, and social influence shape attitudes toward AV adoption. Notably, young drivers often use handheld devices while driving (Kaye & Demmel, 2021), prompting our investigation into how various AV feedback mechanisms influence their takeover time in conditional scenarios. Through this analysis, we aim to shed light on AV adoption dynamics among the younger demographic (ages 18-26), informing future technological implementations and safety guidelines.

In light of our focus on the young generation aged 18-26 as our user group, our research team delves into various feedback types tailored to this demographic's preferences and needs.

Feedback Modalities

Research into auditory feedback within autonomous vehicles reveals its significant role in driver perception and interaction. Rukonic et al. (2021) showcase that speech warnings effectively reduce cognitive load and enhance situational awareness and trust, particularly among elderly drivers, by offering clear, actionable advice. This is corroborated by Cohen-Lazry et al. (2020), who find that clear verbal instructions significantly aid navigation and decision-making. Furthermore, Zang et al. (2022) emphasize the critical role of transparent audio feedback in fostering drivers' trust and improving takeover performance and situational awareness, highlighting the preference for detailed information on the vehicle's automation status, including specifics of current actions, reasons for those actions, and how the system predicts or reacts to situations.

Ko et al. (2022) and Bazilinskyy & de Winter (2015) explore non-speech auditory cues, showing they effectively convey crucial information without distraction, reducing cognitive workload and improving drivers' physiological responses, thereby enhancing user experience and safety. These cues are appreciated for their non-intrusive delivery of urgent information, leading to high user acceptance. Cohen-Lazry et al. (2020) and van der Heiden et al. (2018) further emphasize the cognitive benefits of well-designed non-speech cues, noting their ability to boost driver engagement without taxing cognitive processing.

As discussed above, many past research studies on auditory alert systems in automated vehicles have investigated the impacts of speech and non-speech auditory feedback and mixed modal feedback on driver awareness, preference, and performance. However, the differentiated testing conditions and findings proposed a new challenge – which is how we can adopt a nuanced approach to designing auditory cues based on the

context of their application and the complexity of the information they need to convey. Some literature tried to solve the challenge of balancing the effectiveness of auditory feedback with user preferences and safety considerations by underscoring the need for both informative and adaptable (while non-intrusive) feedback systems to accommodate different driving contexts and user needs (Bazilinskyy & de Winter, 2015; Ko et al., 2022). Overall, current literature still reveals a nuanced preference among drivers for auditory feedback types. Speech warnings are preferred for their clarity and informativeness, especially in complex driving scenarios, while non-speech cues are valued for their subtlety and ability to convey urgency without overwhelming the driver.

Due to the nature of the multifaceted role of auditory feedback in autonomous vehicles, it is suggested that vehicle feedback systems could potentially be personalizable so that they can adjust based on individual driver characteristics and preferences, thus enhancing the overall user experience (Rukonic et al., 2021; Di Campli San Vito et al., 2020). This implies the need for further research into personalizing auditory feedback to accommodate diverse user preferences and conditions. To better understand the interaction between humans and automation, future studies can look at how nonspeech cues compare to speech feedback when varying the characteristics and content of cues given such as the frequency of sound and richness/transparency of information. This research avenue may further our understanding of how automobile feedback mechanisms may be refined and thus provide preliminary evidence for how we could set up the system feedback for default and personalizations in future automated vehicles for better user experience.

Experimental Setup

Simulation Approach Choice

Our team deemed it important to explore the variety of options for the experimental setup. We analyzed three studies, each focusing on one of the three main types of experimental simulation used in autonomous vehicle research: simulators, Wizard of Oz techniques, and video-based setups.

The research performed by Yeo et al. (2020) examined several simulation methods and primarily looked at their ability to establish a sense of presence and reduce motion sickness. The paper examines six different combinations of self-driving simulation, including CGI vs real-world images, VR vs a 3-screen display, a chair with kinesthetic features vs a standard chair, and a motionless setup vs participants being driven in a car while wearing a headset. Presence was gauged with established questionnaires. The scenario

presenting a virtual reality car interior alongside real-world exterior video yielded the highest level of presence.

Another study employed Wizard of Oz techniques to simulate autonomous driving (Baltodano et al., 2015). Participants sat in a car with a fake steering wheel and a partition blocking the passenger seat's view, while a study member controlled the vehicle from the passenger seat. Using a car with the steering wheel on the opposite side can offer a cost-effective alternative to high-tech simulations.

A study conducted by Gerber et al. (2020), offers guidance on how to set up a video-based simulator study without the need for user control. Several cameras were used when capturing the video data, three attached to the front of the car, cameras for the rearview mirrors, and another capturing video behind the car. Audio was recorded from both sides of the car to aid participants in spatial awareness. The researchers recommend that while driving, the car should maintain speeds at or below the speed limit and switch lanes well in advance of turns. Researchers suggest capturing video along familiar road routes to enhance participant engagement with the simulation.

After considering these options utilized in previous studies, we plan to primarily adapt the video-based experimental setup while potentially incorporating features found in the other methods. While drivers may interact differently with these techniques compared to a self-driving vehicle, safety and technical constraints make them the best available options.

Auditory Feedback Choice

In order to explore best practices for speech and non-speech audio feedback, we decided to conduct a literature review to aid our decisions for choosing in-vehicle audio feedback.

Speech Audio. The study by (Dong et al., 2020) examined perceptions of voice agents in fully autonomous vehicles, revealing that female agents were rated higher in likability and comfort compared to males. However, male agents scored higher in perceived safety, echoing findings from (Bazilinskyy et al., 2018). While participants generally preferred female voices, males were deemed more urgent in emergency situations. These results align with traditional stereotypes favoring male voices for safety-critical contexts, as noted in prior research (Nass & Moon, 2000). In light of our research experiment's focus on scenarios involving urgent situations such as construction signs or similar prompts requiring immediate action, the choice of voice feedback becomes critical. Given the association of the male voice with perceptions of safety in high-stakes contexts, our team

opted to utilize a male voice for providing feedback during our experiments. The specific phrase we chose was "takeover requested", due to its brevity and generality.

Non-speech Tone. We opted to investigate optimal approaches concerning the frequency and duration of tones in non-verbal audio contexts. A literature review conducted by the U.S. Department of Transportation recommends a base frequency higher than 1000 Hz and a pulse duration greater than 200 ms. The paper also had recommendations for elements we had not yet considered, such as the number of beeps at ≥ 3 and the time between beeps at ≤ 125 ms (Jeon et al. 2022). Continuing our research, we found another paper indicating that two frequencies resulted in better outcomes than one. The study utilized frequencies of 800 Hz and 1750 Hz, a tone duration of 50 ms, an interval of 30 ms, and a volume of about 20 dB louder than the surrounding environment (Yun & Yang 2022). We decided to utilize the same metrics for our study.

The Present Study

This study aims to explore the preferences and effects of audio feedback types on drivers in autonomous vehicle (AV) handoff situations. Specifically, it addresses two research questions:

1. R1: What types of audio feedback do drivers prefer under combinations of proactive and reactive situations?
2. R2: Do different types of audio feedback affect the driver's takeover time during AV handoff?

Given the research questions, this study will explore the effect of the type of feedback (speech audio feedback vs. non-speech tone feedback) and type of situation (proactive situations vs. reactive situations) on takeover performance (measured by the takeover time) and user experience (assessed through subjective questionnaires focusing on satisfaction, perceived safety, and comfort).

Based on the research questions and background literature, we proposed the following two hypotheses:

1. H1: Drivers will have faster takeover times with speech audio feedback compared to nonspeech tones in proactive situations due to the intuitive perception of speech.
2. H2: In reactive scenarios, drivers will exhibit a preference for nonspeech tone feedback over speech audio feedback, as nonspeech tones are expected to minimize cognitive load and facilitate immediate attention allocation.

METHOD

Experimental Design

This study aims to investigate the impact of different audio feedback types on driver takeover time and overall user experience in AV handoff scenarios. The motivation stems from the need to understand how to best keep drivers informed and ready to take over control from an AV under necessary driving situations while optimizing safety and comfort.

A 2x2 mixed factorial design will be employed, with the type of feedback (speech vs. non-speech) as a between-subjects variable and the type of situation (proactive vs. reactive) as a within-subjects variable. Participants will be randomly assigned to one of the feedback conditions and exposed to both situation types.

To simulate proactive situations where the AV has a sufficient reaction to a situation, auditory warnings will be provided to the driver a few seconds before a critical event appears on the screen (~6-7 seconds). This timing allows the driver to prepare for the takeover in a non-urgent manner, reflecting scenarios where the AV can predict and inform about a required handover well in advance. For reactive situations where the AV has insufficient reaction time for a situation, auditory warnings will be given to the driver much closer to the occurrence of a critical event on the screen (2-3 seconds). This simulates more urgent or sudden scenarios requiring the driver to quickly take over control from the AV.

The study uses video-based simulation to measure drivers' takeover reactions. Participants will be presented with pre-edited video clips of various driving situations. We drove in our local city, Atlanta, Georgia, and procured the video situations ourselves. All situations were heuristically assessed on having a similar level of risk. We then edited and combined clips together to include our audio warnings.

Experimental Task

We expect the total session to take 30 minutes, with 10 minutes allocated to the experimental video simulation, 10 minutes for the post survey, and 10 minutes for training and questions. In a series of simulation videos presented with a set order, participants will engage in simulated driving tasks, experiencing a series of controlled driving scenarios that necessitate an AV handoff. Each scenario will be designed to represent real-world driving conditions, with variations introduced to simulate proactive and reactive situations. More specifically, participants are expected to undergo the following situations: scenarios requiring the autonomous vehicle to navigate complex urban environments and scenarios involving interactions with other road users. More specifically, we included

eight scenarios in the videos: merging, construction zones, lane changing, road work ahead, bike interruptions, ambulance yielding, unexpected road detours, and truck interruptions.

While watching the videos, participants will be distracted by a secondary task. According to (Kuhlen et al, 2017), a game-based task is most engaging and would likely distract drivers from the videos the most. As the study explicitly mentioned Tetris, we will have participants play a mobile version of Tetris while they wait to receive the audio warning (either nonspeech tone or speech audio) they are randomly assigned to. Once they have received the audio feedback, participants will be asked to react to the sound by responding to the handoff request and taking over from the AV. After going through all video clips and takeover reaction tasks, participants will be asked to fill out the UEQ-S.

The experimental task closely mirrors real-world contexts by simulating driving scenarios within a Level-3 AV environment, incorporating urban complexities, traffic conditions, and interactions with other road users. Participants encounter a variety of realistic situations crafted to necessitate AV handoff. This design ensures a realistic approximation of driving conditions, enhancing the ecological validity of the study. The subsequent completion of the UEQ-S by participants allows for the assessment of user experience in response to these simulated AV handoff scenarios, bridging the gap between experimental conditions and real-world applications of AV technology.

Measurements

The study includes both behavioral measures and a subjective scale measurement.

The first construct, Takeover Performance, will be measured by the takeover time, which is defined as the duration between the issuance of a takeover request (TOR) (i.e., when the driver receives an audio warning from the AV) and the driver's initiation of the control action. Takeover time will be recorded to assess reaction speeds in different feedback and situation conditions. To measure participants' takeover time accurately, this study uses a custom Python keylogging software to record the timestamp of the driver's response (via "shift" key press) and compare it to the known timestamp of the warning.

The second construct, User Experience, is evaluated using the User Experience Questionnaire - Short (UEQ-S), as suggested by Schneider et al. (2021). The UEQ-S will be employed to evaluate participants' satisfaction, perceived safety, and comfort, providing insights into the effectiveness of each feedback type from the user's perspective. It provides a standardized

method to assess the subjective experience of users, covering aspects such as attractiveness, perspicuity (ease of understanding), efficiency, dependability, stimulation, and novelty.

Experimental Procedure

Participants will be briefed on the study's purpose and will give their consent before the study begins. Participants will also complete a demographic questionnaire that includes age, gender, and driving experience. Then, they will undergo a brief training session to familiarize themselves with the video simulations and task requirements. This will consist of multiple situations similar to the ones presented in the experiment, however, participants will be asked to press the takeover key (the “shift” key in our case) when the warning flashes on the screen. Note, no audio feedback will be presented in this case.

Participants will be randomly assigned to two conditions: they will receive either speech audio or non-speech tone feedback while watching the simulation videos, while all participants will experience driving scenarios that include both proactive and reactive situations. Video clips demonstrating various driving scenarios will be presented in the same order for all participants.

Participants will be asked to fill out the survey instrument throughout the study - they will do UEQ-S halfway through the study, and another UEQ-S plus some demographic questions upon completion of the driving task. A debriefing session will follow, where the study's purpose and the expected contributions will be explained. Participants will receive food compensation for their time and contribution to the study.

Participants

The study targets licensed drivers with varying levels of experience in using AV technologies. Participants will be recruited using convenient sampling using the online social platform like Slack. We aimed to recruit participants with a broad representation of age, gender, and driving experience, aiming for a sample size that provides adequate power to detect significant effects of the independent variables on the dependent measures.

The study will be conducted in accordance with ethical guidelines, ensuring participants' safety and confidentiality. All data will be anonymized, and participants will have the right to withdraw from the study at any point without penalty.

RESULTS

Descriptive Statistics

Descriptive statistics were calculated for each combination of experimental conditions involving two

independent variables (IVs): type of audio feedback (nonspeech vs. speech) and type of situation (reactive vs. proactive). These were measured against two dependent variables (DVs): takeover time and user experience (UX) scores. These are summarized in Table 1, providing an overview of the central tendency and variability for each experimental group.

Table 1

Descriptive Statistics for Takeover Time and UX Scores by Experimental Group

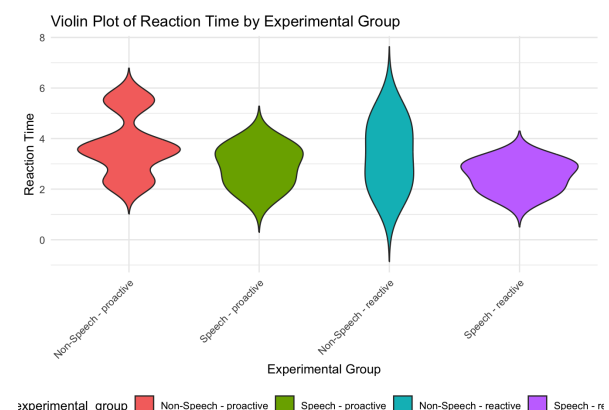
Experimental Group	Response Type	M	SD
Nonspeech - Proactive	Takeover Time	3.74	1.33
	UX Score	35.8	4.99
Speech - Proactive	Takeover Time	2.90	0.80
	UX Score	41.2	5.06
Nonspeech - Reactive	Takeover Time	3.26	1.37
	UX Score	34.8	1.50
Speech - Reactive	Takeover Time	2.53	0.62
	UX Score	33.0	12.20

Data Visualizations

Violin plots were created to visually represent the distribution of takeover times and UX scores across the different experimental conditions. Figure 1 depicts the distribution of takeover times, and Figure 2 illustrates the distribution of UX scores.

Figure 1

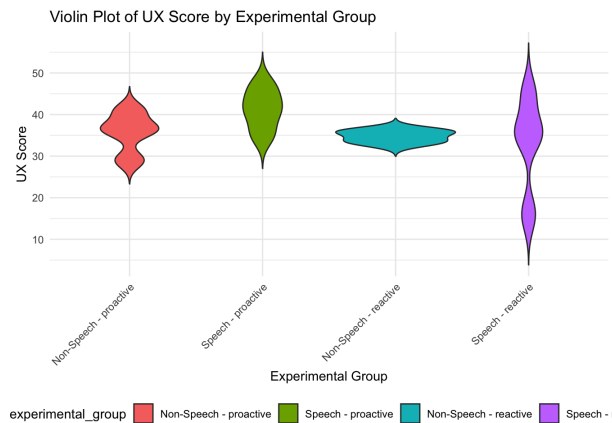
Violin Plot of Takeover Time by Experimental Group



This plot displays the distribution of takeover times for each condition, emphasizing the spread and density of the data, which reveals variability in response times across the different audio feedback and situation types.

Figure 2

Violin Plot of UX Scores by Experimental Group



This plot illustrates the distribution of UX scores within each experimental condition, highlighting differences in the spread and central tendency, particularly the broader range of scores observed in the speech-reactive condition.

Inferential Statistics

Assumption Testing

Normality was checked using Shapiro-Wilk tests, indicating non-normality only for UX scores in the reactive condition ($p = .02$). Levene's test for equality of variances confirmed homogeneity of variances across groups, $F(3, 28) = 0.57, p = .64$.

Analysis of Variance

A two-way ANOVA was conducted for each DV, with no significant effects found for takeover time or UX scores. The main effects analysis showed that the effect of feedback type on UX scores was not significant, $F(1, 11) = .26, p = .62, \eta^2 = .02$. The main effect of situation type on UX scores was also not significant, $F(1, 11) = 1.57, p = .24, \eta^2 = .11$. Specifically, the interaction of audio feedback type and situation type on UX scores was not significant, $F(1, 11) = .96, p = .35, \eta^2 = .07$, indicating that these variables do not interact to influence user experience.

Similarly, the effect of feedback type on takeover time was not significant, $F(1, 11) = 2.40, p = .15, \eta^2 = .16$. The main effect of situation type on takeover was also not significant, $F(1, 11) = .50, p = .50, \eta^2 = .03$. Also, no significant interaction effects were found for takeover time, $F(1, 11) = .00, p = 1.00, \eta^2 = .00$, suggesting that the type of audio feedback and the

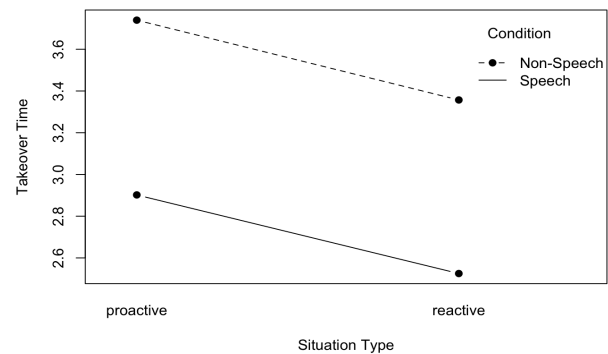
situation type do not significantly impact the speed of response in takeover scenarios.

Data Visualizations

Interaction plots were generated to explore the effects of audio feedback type and situation type on the takeover time and UX scores. These plots serve as a visual confirmation of the statistical findings from the ANOVAs, which did not show significant interaction effects (see Figures 3 and 4).

Figure 3

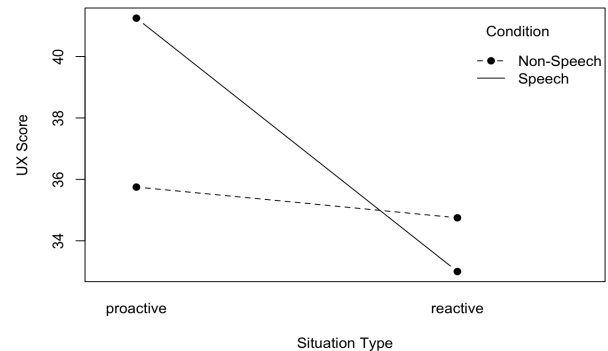
Interaction Plot for Takeover Time by Experimental Group



This plot illustrates the interaction between the type of audio feedback (nonspeech vs. speech) and the type of situation (proactive vs. reactive) on takeover time. Each line represents one of the audio feedback conditions, showing how takeover time varies with the situation type within each audio feedback category. The plot suggests minimal interaction between these variables, as indicated by the relatively parallel lines, suggesting that the influence of situation type on takeover time does not differ significantly between audio feedback types.

Figure 4

Interaction Plot for UX Scores by Experimental Group



Similarly, this plot displays the interaction effects for UX scores. It depicts how UX scores vary with the type of situation within each audio feedback condition.

The greater divergence between the lines in this plot compared to the takeover time plot suggests a more pronounced interaction effect, although this was not statistically significant in the ANOVA results. This visual discrepancy indicates potential areas for further exploration, particularly in understanding how different types of audio feedback might influence user experience in varying situational contexts.

Independent Samples T-Tests

The differences in mean scores for takeover time and UX scores between less experienced drivers (≤ 5 years) and more experienced drivers (> 5 years) were analyzed using independent samples T-tests. These tests showed no significant differences in UX scores, $t(8.48) = -.51$, $p = .62$, with a Cohen's d of .29, nor in takeover times, $t(11.02) = 1.81$, $p = .10$, with a Cohen's d of .82, indicating that driving experience does not significantly influence these outcomes.

DISCUSSION

The results of the present study provide supporting evidence that the type of audio feedback—either speech or nonspeech—does not significantly affect takeover time or user experience in proactive versus reactive situations in autonomous vehicle handoff scenarios. These findings are consistent with prior research that suggests the effectiveness of auditory cues may not significantly differ across feedback types under varying situational contexts (Cohen-Lazry et al., 2020).

Key Findings

First, the hypothesis that speech audio feedback would facilitate faster takeover times in proactive situations was not supported. This outcome contrasts with expectations derived from literature suggesting that speech cues could enhance response times due to their clarity and the intuitive nature of verbal instructions (Ko et al., 2022).

Second, the anticipation that nonspeech tones might be preferred in reactive scenarios, purportedly reducing cognitive load, was also not substantiated. This indicates a complex interaction between auditory feedback type and the situational context that might not directly influence user preferences or performance as previously assumed (Zang et al., 2022).

Third, despite varying experimental conditions, no significant differences were observed in user experience scores, challenging the notion that different audio feedback types might lead to differential perceptions of safety, comfort, or satisfaction (Seppelt & Lee, 2019).

Inferential Statistics

These results suggest a reevaluation of the assumptions regarding audio feedback's role in

autonomous vehicle handoff situations. Theoretically, the findings challenge models that predict distinct advantages of speech over nonspeech audio cues in improving takeover readiness and performance. Practically, the results imply that designers of autonomous vehicle interfaces may need to consider broader aspects of human factors and user context when choosing between speech and nonspeech audio cues. This could involve more personalized feedback systems that adjust to individual user characteristics and preferences, as suggested by Rukonic et al. (2021).

Theoretical and Practical Implications

These findings necessitate a reevaluation of theoretical models that distinctly advantage speech over nonspeech cues in facilitating driver readiness and performance in autonomous vehicles. From a practical standpoint, the results suggest that autonomous vehicle interface designers might need to consider a more holistic approach that incorporates individual differences and situational variables when selecting auditory feedback types.

Study Limitations

Experimental Realism and Scenario Design

Participants reported difficulties in distinguishing between the proactive and reactive scenarios presented in the simulation videos. This lack of differentiation could have significantly impacted their response times and subjective experiences, as the experimental conditions may not have been distinct enough to elicit different reactions. The scenarios were also perceived as non-critical, leading many participants to view the audio alerts as unnecessary. This suggests that the experimental stimuli might not have adequately represented the urgency or importance typically associated with Level 3 autonomous vehicle interventions, which could dilute the perceived utility and impact of different types of audio feedback.

Relevance and Criticality of Scenarios

Perceived Importance: The feedback that the scenarios were not critical enough implies that the simulated driving situations failed to convey the real-world stakes that might necessitate quick and decisive action from the driver. This perception could undermine the effectiveness of any audio feedback, as participants might not feel compelled to respond with the urgency that true reactive situations would demand.

Participant Familiarity with Autonomous Vehicles

Of the eight participants, only one had prior experience with Level 3 autonomous vehicles. This lack of familiarity with autonomous vehicle technologies among the majority of participants could skew their perceptions and responses to the audio feedback. Participants who are inexperienced with how

autonomous vehicles operate might not have the contextual knowledge or expectations that seasoned users might have, potentially affecting their ability to assess the utility and appropriateness of the audio cues provided.

Sample Size and Diversity

With only eight participants, the study's sample size is quite small, limiting the generalizability of the results. This would also enable the examination of demographic variables such as age, driving experience, and familiarity with technology, which could influence responses to autonomous vehicle technology and feedback mechanisms.

Future Directions

Future research should further explore the conditions under which different types of auditory feedback may be more effective, potentially examining the long-term effects of continuous exposure to auditory cues and their impact on driver habituation. Additionally, incorporating a broader range of demographic characteristics and driving contexts could enhance the generalizability of the findings and provide deeper insights into the preferences and performances of diverse user groups.

Based on these limitations, future research should consider the following improvements:

First, develop more distinct and critically engaging proactive and reactive driving scenarios that clearly convey the stakes and required responses. This could involve using real-life critical incident footage or more detailed scenario scripting to heighten realism.

Second, include participants with varying levels of experience with autonomous vehicles to better understand how familiarity with such technologies might influence the perception and effectiveness of audio feedback.

Third, conduct the study with a larger, more diverse sample to improve the statistical power and generalizability of the findings. This would also help in conducting subgroup analyses to explore how different groups may uniquely respond to the experimental manipulations.

Fourth, employ an iterative design process where initial feedback from participants about the experimental setup can be used to refine the scenarios and simulation environment before conducting the full-scale study.

CONCLUSION

In conclusion, this study did not find significant differences in takeover performance or user experience based on audio feedback type across proactive and reactive driving scenarios. This suggests that the impact of audio feedback may be more nuanced and dependent on factors not fully captured in this study. The findings

contribute to the ongoing discourse on how best to design auditory feedback in autonomous vehicles, emphasizing the need for further research into personalized and context-sensitive feedback systems. The theoretical and practical insights gained underscore the complexity of human-automation interaction and the necessity of a nuanced approach to the design of feedback modalities in enhancing the safety and user experience of autonomous vehicle technologies.

REFERENCE

- Baltodano, S., Sibi, S., Martelaro, N., Gowda, N., & Ju, W. (2015). The RRADS platform: A real road autonomous driving simulator. *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 281–288. <https://doi.org/10.1145/2799250.2799288>
- Basantis, A., Miller, M., Doerzaph, Z., & Neurauter, M. L. (2021). Assessing Alternative Approaches for Conveying Automated Vehicle “Intentions.” *IEEE Transactions on Human-Machine Systems*, 51(6), 622–631. <https://doi.org/10.1109/THMS.2021.3106892>
- Bazilinskyy, P., & de Winter, J. (2015). Auditory interfaces in automated driving: An international survey. <https://doi.org/10.7287/peerj.preprints.1069v1>
- Cohen-Lazry, G., Borowsky, A., & Oron-Gilad, T. (2020). The impact of auditory continual feedback on take-overs in Level 3 automated vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour*, 75, 145–159. <https://doi.org/10.1016/j.trf.2020.10.003>
- Di Campli San Vito, P., Brown, E., Brewster, S., Pollick, F., Thompson, S., Skrypchuk, L., & Mouzakitis, A. (2020). Haptic Feedback for the Transfer of Control in Autonomous Vehicles. *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 34–37. <https://doi.org/10.1145/3409251.3411717>
- Dong, J., Lawson, E., Olsen, J., & Jeon, M. (2020). Female voice agents in fully autonomous vehicles are not only more likeable and comfortable, but also more competent. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 64(1), 1033–1037. <https://doi.org/10.1177/1071181320641248>
- Gerber, M. A., Schroeter, R., & Vehns, J. (2019). A Video-Based Automated Driving Simulator for Automotive UI Prototyping, UX and Behaviour Research. *Proceedings of the 11th International Conference on Automotive User Interfaces and*

- Interactive Vehicular Applications*, 14–23.
<https://doi.org/10.1145/3342197.3344533>
- Jeon, M., Lautala, P., Nadri, C., & Nelson, D. N. (2022). In-vehicle auditory alerts literature review.
- Kaye, S.-A., Demmel, S., Oviedo-Trespalacios, O., Griffin, W., & Lewis, I. (2021). Young drivers' takeover time in a conditional automated vehicle: The effects of hand-held mobile phone use and future intentions to use automated vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour*, 78, 16–29.
<https://doi.org/10.1016/j.trf.2021.01.012>
- Ko, S., Kutchek, K., Zhang, Y., & Jeon, M. (2022). Effects of Non-Speech Auditory Cues on Control Transition Behaviors in Semi-Automated Vehicles: Empirical Study, Modeling, and Validation. *International Journal of Human-Computer Interaction*, 38(2), 185–200.
<https://doi.org/10.1080/10447318.2021.1937876>
- Kuehn, M., Vogelpohl, T., & Vollrath, M. (2017, May). Takeover times in highly automated driving (level 3). In 25th International technical conference on the enhanced safety of vehicles (ESV) national highway traffic safety administration (pp. 1-11).
- Nass, C., & Moon, Y. (2000). Machines and Mindlessness: Social responses to computers. *Journal of Social Issues*, 56(1), 81–103.
<https://doi.org/10.1111/0022-4537.00153>
- Ruggeri, K., Kácha, O., Menezes, I., Kos, M., Franklin, M., Parma, L., Langdon, P., & Miles, J. (2018). In with the new? Generational differences in technology adoption differences in population technology adoption in the age of self-driving vehicles. *Journal of Engineering and Technology Management*, 50.
<https://doi.org/10.1016/j.jengtecman.2018.09.001>
- Rukonic, L., Mwange, M.-A., & Kieffer, S. (2021). UX Design and Evaluation of Warning Alerts for Semi-autonomous Cars with Elderly Drivers: Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, 25–36.
<https://doi.org/10.5220/0010237000250036>
- Schneider, T., Ghellal, S., Love, S., & Gerlicher, A. R. S. (2021). Increasing the User Experience in Autonomous Driving through different Feedback Modalities. *26th International Conference on Intelligent User Interfaces*, 7–10.
<https://doi.org/10.1145/3397481.3450687>
- Seppelt, B. D., & Lee, J. D. (2019). Keeping the driver in the loop: Dynamic feedback to support appropriate use of imperfect vehicle control automation. *International Journal of Human-Computer Studies*, 125, 66–80.
<https://doi.org/10.1016/j.ijhcs.2018.12.009>
- van der Heiden, R. M. A., Janssen, C. P., Donker, S. F., Hardeman, L. E. S., Mans, K., & Kenemans, J. L. (2018). Susceptibility to audio signals during autonomous driving. *PLoS ONE*, 13(8), e0201963.
<https://doi.org/10.1371/journal.pone.0201963>
- Yeo, D., Kim, G., & Kim, S. (2020). Toward Immersive Self-Driving Simulations: Reports from a User Study across Six Platforms. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–12.
<https://doi.org/10.1145/3313831.3376787>
- Yun, H., & Yang, J. H. (2020). Multimodal warning design for take-over request in conditionally automated driving. *European Transport Research Review*, 12(1), 34.
<https://doi.org/10.1186/s12544-020-00427-5>
- Zang, J., Jeon, M., & Link to external site, this link will open in a new tab. (2022). The Effects of Transparency and Reliability of In-Vehicle Intelligent Agents on Driver Perception, Takeover Performance, Workload and Situation Awareness in Conditionally Automated Vehicles. *Multimodal Technologies and Interaction*, 6(9), 82. <https://doi.org/10.3390/mti6090082>