Human Response to Electric Vehicle Noise: Laboratory Findings on Localization, Annoyance, and Non-Auditory Effects

Leon Müller
Chalmers University of Technology, Division of Applied Acoustics, 41296 Göteborg, Sweden, leon.mueller@chalmers.se

Electric vehicles (EVs) are increasingly present in urban environments and typically emit little noise at low driving speeds, requiring artificial warning sounds via acoustic vehicle alerting systems (AVAS). While these sounds are intended to improve pedestrian traffic safety, they also become part of the urban acoustic environment. This contribution summarizes a series of laboratory studies on human responses to electric vehicle noise, focusing on auditory localization, subjective appraisal, and physiological and cognitive responses. The first study addresses the localization of electric vehicles in a static parking lot scenario. Using a concealed array of 24 loudspeakers and auralized stimuli derived from vehicle recordings, localization accuracy and response time were examined while simulating EVs with different AVAS signal types and configurations. Results from 52 participants show that AVAS design can strongly affect localization performance. In particular, highly tonal signals that may be effective in single-vehicle detection tasks showed poor localizability, especially when multiple vehicles with similar AVAS are present. The second experiment focused on residents exposed to indoor traffic noise from EVs driving slowly on a nearby street. In a living-room-like laboratory environment, 60 participants performed a combined Eriksen Flanker and spatial Stroop attention test while being exposed to low-level EV noise at realistic closed-window levels ($L_{A,eq} \leq 21.5$ dB). In addition to subjective workload and annoyance ratings, electrodermal activity was recorded as an indicator of physiological arousal. While attention performance was not measurably affected, highly tonal warning sounds elicited higher annoyance ratings, larger perceived workload, and greater physiological arousal than other signals. Taken together, these findings demonstrate that human responses to electric vehicle noise extend beyond detectability and depend on signal characteristics and context. The results suggest that AVAS design choices can influence multiple aspects of human response, underlining the usefulness of quantitative human-response measurements in AVAS development and evaluation.

1 Introduction

With the ongoing transition to electromobility [1], the character of urban road traffic noise is changing. At low driving speeds, electric vehicles (EVs) typically emit less noise than internal combustion engine vehicles (ICEVs) [2,3], which may be beneficial from an environmental noise perspective. However, this reduced sound emission also diminishes acoustic cues for pedestrians, cyclists, and other vulnerable road users. Accident statistics from several countries indicate an increased collision risk involving slow-moving EVs in urban environments [4–8], and human subject studies confirm that, without visual cues or acoustic countermeasures, EVs are often detected later than ICEVs [2,9,10].

To address this, regulations in many countries now require EVs to be equipped with an acoustic vehicle alerting system (AVAS). In the US, FMVSS No. 141 [11] specifies minimum third-octave band levels for defined speed ranges and maneuvers. The EU and other countries follow UNECE Regulation No. 138 [12], which similarly defines spectral and

level requirements and additionally mandates a speed-dependent pitch shift while limiting the maximum overall level to 75 dBA at 2 m distance. Both frameworks leave considerable design freedom, resulting in diverse AVAS implementations ranging from engine-like sounds to highly tonal or broadband synthetic signals.

Much of the existing AVAS research has focused on the detectability of single vehicles, typically quantified as detection time or time-to-vehicle arrival [13–17]. These studies have shown that adding warning sounds improves detection probability and that synthetic sounds can outperform recorded engine sounds. Some findings suggest that tonal signals are particularly detectable [14]. However, the detectability of a single vehicle in a controlled pass-by scenario does not capture the full range of perceptual demands that arise in real urban traffic, where multiple vehicles may be present simultaneously and where localization accuracy may be equally important, for example, in parking lot environments.

Furthermore, the environmental impact of AVAS on non-involved listeners has received comparatively little attention. While overall level changes due to AVAS have been estimated to be small [18], sounds with identical levels can differ substantially in perceived annoyance [19], and no studies have examined AVAS effects on cognitive or physiological responses. This is particularly relevant for indoor environments, where most people spend the majority of their time [20], and where sound transmission through building facades can alter the spectral character of incoming traffic noise.

This contribution summarizes two laboratory studies that address these gaps by investigating human responses to EV noise beyond detectability. The first study [21] examines the auditory localization of EVs with different AVAS types in single- and multi-vehicle scenarios, using combustion engine noise as a reference condition. The second study [22] examines the effects of low-level indoor EV traffic noise on attention, electrodermal activity, perceived workload, and noise annoyance in a living-room laboratory environment, comparing three AVAS conditions against silence. Both studies use the same three representative AVAS signal types – a narrowband noise AVAS, a multi-tone AVAS, and a two-tone AVAS – enabling a combined evaluation across traffic safety and environmental noise dimensions.

2 Localization of Electric Vehicles

The first study [21] investigated how accurately and quickly pedestrians can localize electric vehicles solely based on their warning sounds. Using auralized stimuli and a concealed loudspeaker array, localization performance was compared across AVAS types and combustion engine noise for scenarios with one, two, or three simultaneously present vehicles.

2.1 Methods

To evaluate how well different EV warning sounds can be localized, a listening experiment was conducted in an anechoic chamber using a circular array of 24 Genelec 8020 loudspeakers at a 3 m radius, concealed behind acoustically transparent curtains, as shown in Figure 1. Participants stood in the center of the array and were handed a custom-built motion controller whose horizontal orientation was mapped to a red dot on an LED strip mounted around the array. They were instructed to aim the controller toward each perceived vehicle direction and lock in the position by pulling a trigger. Each trial lasted 10 seconds, and participants were told that up to three vehicles could be present simultaneously.

The stimuli comprised three AVAS signal types and a combustion engine recording. The AVAS signals were generated using the *Electric Vehicle Auralization Toolbox* [23], simulating a static vehicle at 7.5 m distance in free-field conditions. The noise AVAS, based on recordings of a *Tesla Model Y 2019*, consists of two narrow-band noise components. The two-tone AVAS, based on a reversing *Tesla Model Y 2019*, consists of two amplitude-modulated pure tones. The multi-tone AVAS, based on a *Volkswagen ID.3 Pro Performance 2021*, consists of around 30 different amplitude-modulated tones. The combustion engine reference was recorded from an idling Volkswagen Golf VII. All stimuli were normalized to equal loudness and comply with both UNECE and US regulatory requirements, as illustrated in Figure 2. Four slightly different versions of each stimulus were generated by varying the vehicle velocity to avoid identical signals in multi-vehicle trials. In addition, a recording of parking lot ambient noise was played back continuously throughout the experiment.

Four scenario types were investigated: a single vehicle, two vehicles with the same sound, three vehicles with the same sound, and two vehicles with different sounds. The stimuli were distributed across 72 trials with controlled but partially randomized spatial positions. After excluding three outliers, 52 participants (24 female, 26 male, 2 other; median age 25.5 years; self-reported normal hearing) were included in the analysis. Localization accuracy was quantified as the absolute angular error between the playback position and the marked response, and localization time was measured as the duration from trial onset (or previous response) to each answer. Responses with errors exceeding 90° were classified as failed localizations.

Figure 1: Participant performing the localization experiment, placed in a circular array of 24 hidden loudspeakers, indicating the perceived direction with a motion controller.

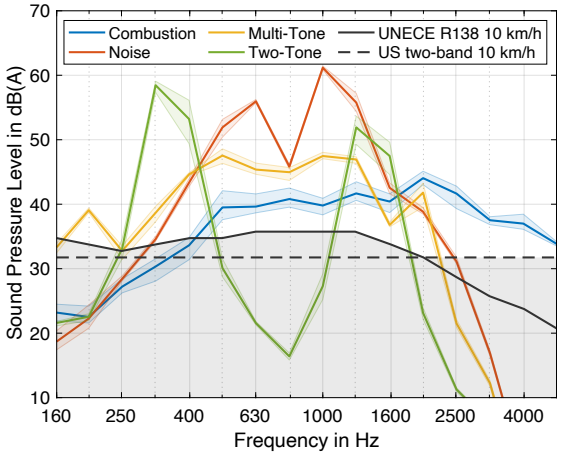


Figure 2: Third-octave spectra of combustion engine, noise AVAS, two-tone AVAS, and multi-tone AVAS stimuli, as well as UN and US regulation minimum levels.

2.2 Results

The effect of vehicle sound type on localization error and time was analyzed using repeated measures analyses of variance with Bonferroni-corrected post-hoc paired comparisons. Figure 3 provides an overview of the most relevant localization results. For single vehicles, the sound type had a significant effect on both localization error and localization time ($p < .001$). Post-hoc tests revealed that the two-tone AVAS was localized significantly less accurately than all other signals, with a mean error approximately 2° larger than the combustion engine noise. All three AVAS signals were localized more slowly than the combustion noise, with the two-tone AVAS being slowest (0.66 s slower on average). However, the absolute differences in the single-vehicle case were relatively small, with mean errors below 6° for all signal types.

The differences became substantially larger when multiple vehicles with the same sound type were presented simultaneously. For two same-sound vehicles, the effect of sound type on localization error was significant ($p < .001$), and the two-tone AVAS showed a mean error approximately 16° larger than the combustion noise. For three same-sound vehicles, the two-tone AVAS reached a mean localization error of 27° and a mean localization time of 4.1 s. The number of failed localizations – defined as responses with an angular error exceeding 90° or no response within the 10-second trial – also increased drastically in multi-vehicle scenarios: when three two-tone AVAS signals were presented, more than half of the participants failed to localize at least 40% of the vehicles, and not a single participant successfully localized all of them. In contrast, 45 out of 52 participants achieved zero failed localizations for three combustion engine vehicles.

When two vehicles with different sound types were combined, the localization of each signal was primarily affected by the presence of a second vehicle with the same sound type, while cross-type interference was minimal. This suggests that the degradation in multi-vehicle scenarios is driven by acoustic similarity rather than general masking.

3 Non-auditory effects of indoor EV noise

While the first study addressed traffic safety by evaluating how well AVAS signals can be localized, the second study [22] focused on the perspective of non-involved listeners exposed to EV noise in a residential environment. In a laboratory setting designed to replicate a typical living room, participants performed a cognitive attention task while being exposed to low-level EV traffic noise at realistic closed-window levels. In addition to attention performance, electrodermal activity, perceived workload, and noise annoyance were assessed across three AVAS conditions and a silent reference.

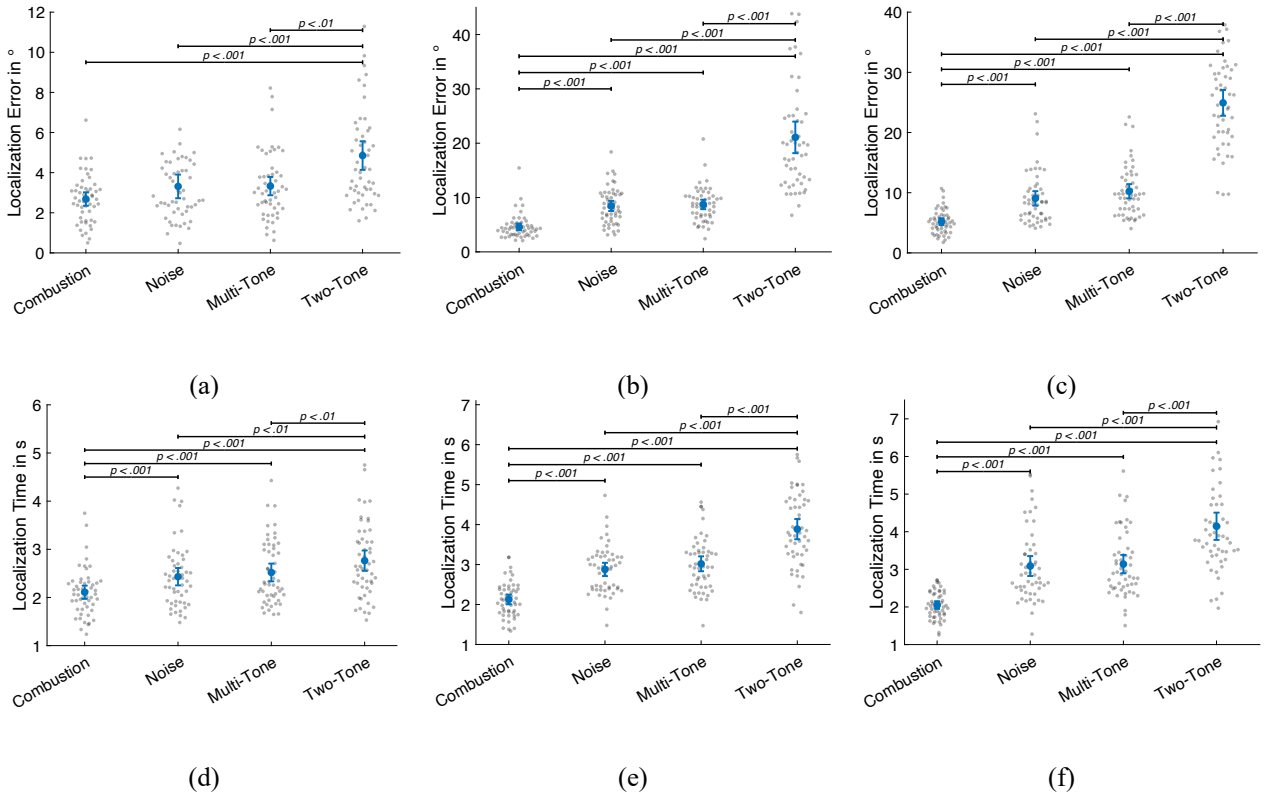


Figure 3: Localization error for single vehicle condition (a), for two vehicles (b) and for three vehicles with the same sound type (c), as well as localization time for single vehicle condition (d), for two vehicles (e) and for three vehicles with the same sound type (f). The horizontal bars indicate Bonferroni corrected *p*-values for significant paired comparisons. The gray dots show individual subject results. The blue error bars show arithmetic means with 95% confidence intervals.

3.1 Methods

The experiment was conducted in the Chalmers *Living Room Lab* (Figure 4), a facility specifically designed for indoor traffic noise studies [24]. The lab consists of two acoustically isolated rooms coupled via a double wall with an insulating window. The receiving room is furnished to resemble a typical living room, while the sending room houses a linear array of 24 Neumann KH80 DSP loudspeakers and heavy acoustic curtains to reduce reflections. Using wave field synthesis (WFS), the loudspeaker array projects the sound field of a simulated vehicle passage onto the outside of the window, which then propagates through the glass into the receiving room, creating a realistic spatial distribution of indoor traffic noise, including angle-of-incidence-dependent transmission effects. The window's sound reduction was adjusted via FIR filtering of the WFS signals to match a standard double-pane unit (weighted sound reduction index $R_W = 31$ dB). Artificial background noise resembling a quiet ventilation system was continuously present at $L_{A,eq} = 19$ dB, to avoid an unnaturally silent environment.

The acoustic stimuli were generated using the same electric vehicle auralization toolbox [23] and the same three AVAS signal types as in the localization study, combined with synthesized tire-road noise. Each stimulus simulated a continuous flow of EV traffic on a nearby street. The outdoor level in front of the window was normalized to $L_{A,eq} = 60$ dB, which is realistic based on reference measurements for the evaluated vehicles. After transmission through the window, indoor A-weighted equivalent continuous sound pressure levels ranged from 19.1 dB to 21.5 dB, depending on the AVAS type. A silent condition with only the artificial background noise served as a reference. Participants were seated at a table facing the curtain-covered window and could not see any loudspeakers.

The main task was a combined Eriksen Flanker and spatial Stroop attention test [25,26]. Participants indicated the direction of a central arrow, using left and right buttons on a keyboard, while ignoring surrounding flanker symbols and the arrow's on-screen position, which could be congruent, incongruent, or neutral with respect to the target direction, as shown in Figure 5. This paradigm probes selective attention, conflict monitoring, and spatial attention. Each block

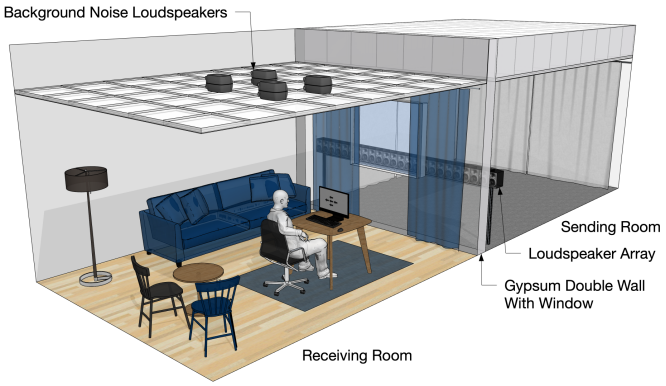


Figure 4: Visualization of *Living Room Lab*, including loudspeaker array in sending room, subject at participant position, and hidden ceiling background noise loudspeakers.

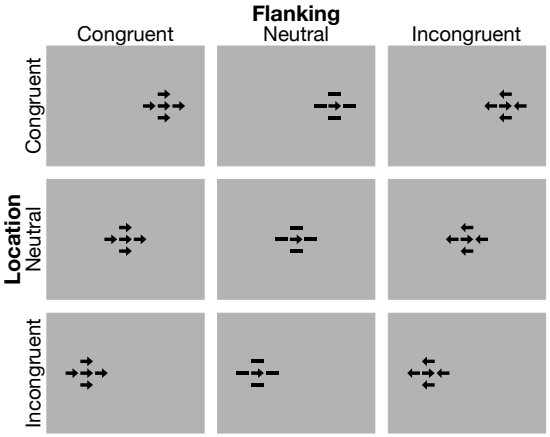


Figure 5: Visual stimuli of combined Eriksen Flanker and spatial Stroop attention test.

consisted of 180 randomized trials and lasted approximately 5 minutes. Response accuracy and response time were recorded for each trial.

Following each attention test block, participants completed a questionnaire assessing perceived workload using the NASA Task Load Index [27] and noise annoyance on a continuous scale ranging from *Not at all* to *Very*. The full procedure – attention test followed by questionnaire – was repeated for each of the four sound conditions in randomized order to control for potential order effects, resulting in a within-subject design.

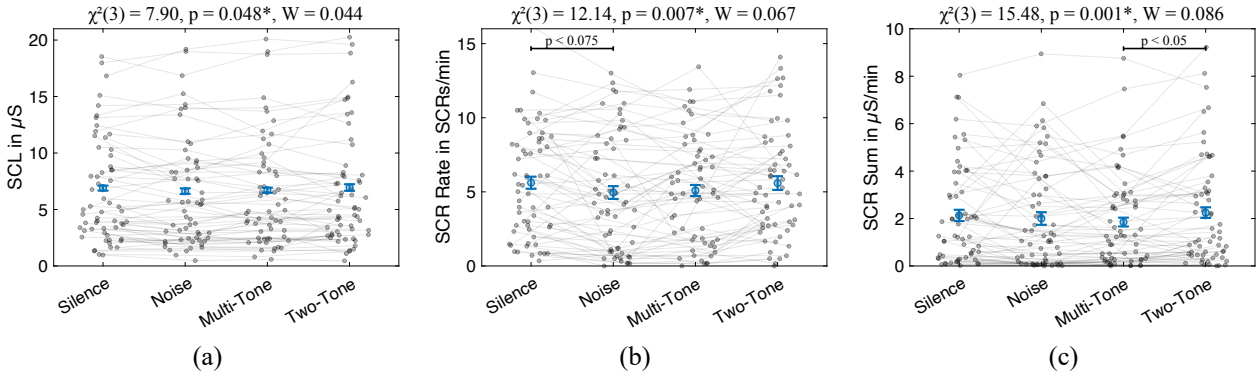
Throughout the experiment, electrodermal activity (EDA) was continuously recorded using Ag/AgCl electrodes placed on the middle and ring fingers of the left hand. The EDA signal was decomposed into tonic and phasic components using Butterworth low-pass and high-pass filters at 0.05 Hz. For the tonic component, the mean skin conductance level (SCL) was computed for each condition. From the phasic component, the number of skin conductance responses (SCRs) per minute and their summed amplitude per minute were extracted. Sixty participants completed the experiment (30 male, 30 female), with a median age of 25 years and self-reported normal hearing.

3.2 Results

The effects of sound condition on all outcome measures were analyzed using Friedman tests, followed by Bonferroni-corrected Wilcoxon signed-rank post-hoc comparisons. No significant effect of sound condition was found on any of the attention test metrics, including mean response time, response time variability, and error count. This suggests that the low-level EV traffic noise did not measurably impair attention performance as assessed by the combined Flanker and spatial Stroop paradigm.

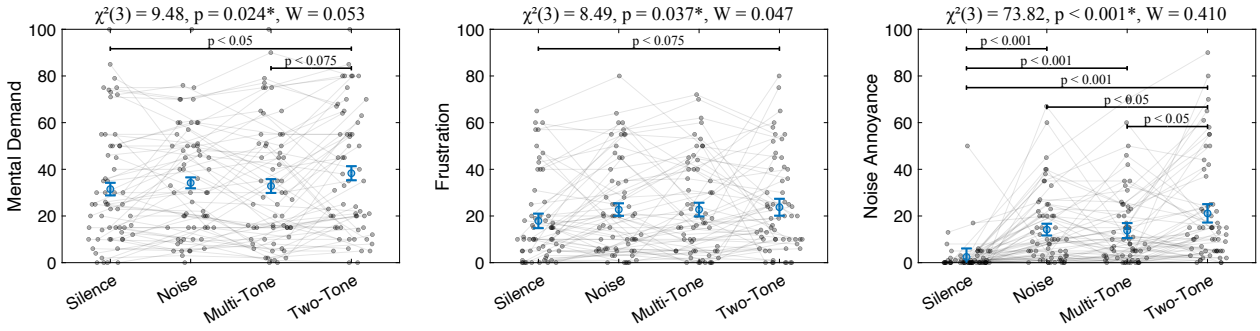
In contrast, the sound condition had a significant effect on all three EDA measures: skin conductance level, SCR rate, and summed SCR amplitude per minute ($p < .05$ for all), as shown in Figure 6. While effect sizes were small and most post-hoc pairwise comparisons did not reach significance after correction, the two-tone AVAS consistently produced the highest values across all three metrics. The only significant post-hoc difference was found between the two-tone and multi-tone conditions for the summed SCR amplitude. Overall, the EDA results indicate that even at low indoor levels that barely exceed the background noise, EV traffic noise can elicit subtle but measurable differences in physiological arousal, with the two-tone AVAS evoking the strongest response.

The subjective results, partly shown in Figure 7, revealed a significant main effect of sound condition on *Noise Annoyance*, *Mental Demand*, *Temporal Demand*, and *Frustration* ($p < .05$ for all). All three EV noise conditions were rated as significantly more annoying than silence, with large effect sizes. Crucially, the two-tone AVAS was rated as significantly more annoying than both the noise AVAS and the multi-tone AVAS. A similar pattern emerged for perceived mental demand, where the two-tone condition was rated significantly higher than silence and near-significantly higher than the multi-tone condition. The two-tone AVAS also showed a near-significant increase in frustration compared to silence. Perceived performance in the attention task did not differ across conditions, consistent with the objective attention results.



(a) (b) (c)

Figure 6: EDA results, showing individual subject data and arithmetic mean with within-subject 95% confidence intervals for SCL (a), SCR rate (b), and sum of SCR amplitudes per minute (c). The titles report Friedman test results for the main effect of sound condition. The horizontal bars represent significant and near-significant Bonferroni corrected Wilcoxon signed-rank post hoc tests.



(a) (b) (c)

Figure 7: Questionnaire results for *Mental Demand* (a), *Frustration* (b) and *Noise Annoyance* (c). The titles report Friedman test results for the main effect of sound condition. The horizontal bars represent significant and near-significant Bonferroni corrected Wilcoxon signed-rank post hoc tests.

4 Discussion

The two studies presented in this paper examine human responses to electric vehicle warning sounds from complementary perspectives – one focused on traffic safety through localization performance, the other on the impact on non-involved listeners through cognitive, physiological, and subjective measures. Considered together, a consistent picture emerges: the two-tone AVAS, composed of two amplitude-modulated pure tones, performed worst across virtually all evaluated dimensions. It was the least localizable signal, particularly in multi-vehicle scenarios where most participants failed to localize a substantial proportion of the presented vehicles. At the same time, it produced the highest annoyance ratings, the greatest perceived mental demand, and the strongest physiological arousal among the tested AVAS types, even at indoor levels barely above background noise levels. The noise AVAS and multi-tone AVAS, in contrast, performed comparably well on both localization and subjective measures.

This convergence is noteworthy because the two studies address fundamentally different listening situations – an active localization task in an anechoic environment versus passive exposure to low-level traffic noise in an indoor living environment – yet both point to the same signal as problematic. The finding that a highly tonal signal performs poorly in both contexts challenges the assumption, based on some earlier research [14], that tonal AVAS signals are generally advantageous due to their high detectability. While tonal signals may indeed be easier to detect in a single-vehicle pass-by scenario, the presented results suggest that this benefit does not extend to localization in multi-vehicle situations and comes at the cost of increased annoyance and physiological burden on bystanders.

These results have practical implications for AVAS design and regulation. Current regulatory frameworks [11,12] primarily define spectral and level requirements aimed at ensuring detectability but leave broad freedom in signal design. The findings of the presented studies suggest that this design freedom should be informed by a more comprehensive set of human-response criteria. Signals optimized solely for detectability or brand identity may perform poorly when evaluated from a broader perceptual perspective. In particular, the US regulation's sameness requirement, which demands that vehicles of the same model use the same AVAS sound [11], may be counterproductive in light of the multi-vehicle localization results, where acoustic similarity between vehicles was identified as the primary driver of degraded localization performance. Introducing a degree of randomness or variation between individual vehicles could mitigate this effect, although the required degree of dissimilarity remains to be investigated.

More generally, the present findings illustrate the value of employing multiple types of human-response measurements in AVAS evaluation. Detectability alone does not capture localizability, and sound pressure level alone does not predict annoyance or physiological response. The dissociation observed in the second study – where subjective ratings and physiological arousal differed across conditions while attention performance did not – further demonstrates that different outcome measures can lead to different conclusions about the same stimulus. A holistic evaluation framework that integrates perceptual, cognitive, physiological, and subjective measures across both outdoor safety and indoor environmental noise contexts would therefore provide a more complete basis for AVAS design decisions.

Several limitations should be acknowledged. Each study evaluated only one representative signal per AVAS type, and alternative implementations may yield different results. The localization study was limited to static free-field scenarios, and the indoor study used a single tire-road noise recording and relatively short exposure durations. Both studies were conducted with young, normal-hearing participants, and generalizability to other populations, particularly older adults or individuals with hearing impairments, remains to be established. Future work should explore a broader range of AVAS designs, more diverse acoustic scenarios, and longer exposure durations to further inform the balance between traffic safety and environmental acceptability.

5 Conclusion

This paper summarizes two laboratory studies on human responses to electric vehicle noise, covering auditory localization, noise annoyance, perceived workload, and physiological arousal. Across both studies, the two-tone AVAS – a signal composed of two amplitude-modulated pure tones – consistently produced the worst outcomes, being the least localizable and the most annoying, demanding, and physiologically arousing of the tested signal types. These findings demonstrate that human responses to AVAS extend well beyond detectability and that signal design choices can have consequences across multiple perceptual and physiological dimensions. The results underscore the importance of incorporating quantitative human-response measurements into AVAS development and evaluation, complementing current regulatory criteria that primarily focus on sound pressure levels and spectral content.

Acknowledgements

This research was funded by FORMAS – a Swedish research council for sustainable development – under grant agreement FR-2020-01931 as well as the HEAD-Genuit-Foundation under grant P-22101-W.

References

[1] International Energy Agency (IEA). Global EV Outlook 2025.

[2] Garay-Vega L, Hastings A, Pollard JK, Zuschlag M, Stearns M. Quieter Cars and the Safety of Blind Pedestrians: Phase 1. Washington, DC: National Highway Traffic Safety Administration; 2010.

[3] Pallas M-A, Bérengier M, Chatagnon R, Czuka M, Conter M, Muirhead M. Towards a model for electric vehicle noise emission in the European prediction method CNOSSOS-EU. Appl Acoust 2016;113:89–101.

[4] Hanna R. Incidence of Pedestrian and Bicyclist Crashes by Hybrid Electric Passenger Vehicles. Washington, DC: National Highway Traffic Safety Administration; 2009.

[5] Wu J, Austin R, Chen C-L. Incidence rates of pedestrian and bicyclist crashes by hybrid electric passenger vehicles: an update. Washington, DC: National Highway Traffic Safety Administration; 2011.

[6] Morgan PA, Morris L, Muirhead M, Walter LK, Martin J. Assessing the perceived safety risk from quiet electric and hybrid vehicles to vision-impaired pedestrians. UK Department for Transport; 2011.

[7] Hou X, Su M, Liu C, Li Y, Ma Q. Examination of the Factors Influencing the Electric Vehicle Accident Size in Norway (2020–2021). World Electr Veh J 2023;15:3. https://doi.org/10.3390/wevj15010003.

[8] Edwards PJ, Moore S, Higgins C. Pedestrian safety on the road to net zero: cross-sectional study of collisions with electric and hybrid-electric cars in Great Britain. Journal of Epidemiology & Community Health 2024;78:487–92.

[9] Goodes P, Bai YB, Meyer E. Investigation into the Detection of a Quiet Vehicle by the Blind Community and the Application of an External Noise Emitting System. SAE Tech Pap Ser 2009;1:63–9.

[10] Kim DS, Emerson RW, Naghshineh K, Pliskow J, Myers K. Impact of adding artificially generated alert sound to hybrid electric vehicles on their detectability by pedestrians who are blind. J Rehabil Res Dev 2012;49:381–93.

[11] National Highway Traffic Safety Administration (NHTSA). FMVSS No. 141 - Minimum Sound Requirements for Hybrid and Electric Vehicles; 2016.

[12] UNECE. Regulation No 138 of the Economic Commission for Europe of the United Nations (UNECE) — Uniform provisions concerning the approval of Quiet Road Transport Vehicles with regard to their reduced audibility [2017/71]; 2017.

[13] Hastings A, Pollard JK, Garay-Vega L, Stearns M, Guthy C. Quieter Cars and the Safety of Blind Pedestrians, Phase 2 : Development of Potential Specifications for Vehicle Countermeasure Sounds, NHTSA; 2011.

[14] Hastings A, McInnis C. Detectability of Alert Signals for Hybrid and Electric Vehicles: Acoustic Modeling and Human Subjects Experiment. Washington, DC: National Highway Traffic Safety Administration; 2015.

[15] Mendonça C, Freitas E, Ferreira JP, Raimundo ID, Santos JA. Noise abatement and traffic safety: The trade-off of quieter engines and pavements on vehicle detection. Accid Anal Prev 2013;51:11–7.

[16] Parizet E, Ellermeier W, Robart R. Auditory warnings for electric vehicles: Detectability in normal-vision and visually-impaired listeners. Appl Acoust 2014;86:50–8. https://doi.org/10.1016/j.apacoust.2014.05.006.

[17] Roan MJ, Neurauter L, Song M, Miller M. Probability of detection of electric vehicles with and without added warning sounds. J Acoust Soc Am 2021;149:599–611. https://doi.org/10.1121/10.0003386.

[18] National Highway Traffic Safety Administration (NHTSA). Minimum Sound Requirements for Hybrid And Electric Vehicles - Final Environmental Assessment; 2016.

[19] Altinsoy ME. The Evaluation of Conventional, Electric and Hybrid Electric Passenger Car Pass-By Noise Annoyance Using Psychoacoustical Properties. Appl Sci 2022;12:5146. https://doi.org/10.3390/app12105146.

[20] World Health Organization (WHO). Combined or multiple exposure to health stressors in indoor built environments; 2014.

[21] Müller L, Forssén J, Kropp W. Auditory localization of multiple stationary electric vehicles. J Acoust Soc Am 2025;157:2029–41. https://doi.org/10.1121/10.0036248.

[22] Müller L, Forssén J, Kropp W. Effects of low-level electric vehicle noise on attention, electrodermal activity, workload, and annoyance. J Acoust Soc Am 2026;159:285–99. https://doi.org/10.1121/10.0042190.

[23] Müller L, Kropp W. Auralization of electric vehicles for the perceptual evaluation of acoustic vehicle alerting systems. Acta Acust 2024;8:27. https://doi.org/10.1051/aacus/2024025.

[24] Müller L, Ahrens J, Kropp W. Loudspeaker Array-Based Auralization of Electric Vehicle Noise in Living Environments. Proceedings of the 11th Convention of the European Acoustics Association Forum Acusticum 2025, Málaga, Spain: 2025. https://doi.org/10.61782/fa.2025.0531.

[25] Eriksen BA, Eriksen CW. Effects of noise letters upon the identification of a target letter in a nonsearch task. Percept Psychophys 1974;16:143–9. https://doi.org/10.3758/bf03203267.

[26] Viviani G, Visalli A, Montefinese M, Vallesi A, Ambrosini E. The Stroop legacy: A cautionary tale on methodological issues and a proposed spatial solution. Behav Res Methods 2024;56:4758–85.

[27] Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. Adv Psychol 1988;52:139–83. https://doi.org/10.1016/s0166-4115(08)62386-9.