

Research article

Can you hear the collision risk? A VR study on beamforming warnings for cyclists at urban intersections

Rodney Leitner^{1,a}, Christoph Ende², Timo Jahns², Michaela Rehm³, Marek Junghans³, Christian Weissig², Thomas Jürgensohn¹

¹HFC Human-Factors-Consult GmbH, Germany

²Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute, Germany 

³German Aerospace Center (DLR), Germany 

Keywords: beamforming, human factors in traffic, noise pollution, Virtual Reality (VR), Western Europe

<https://doi.org/10.55329/vtcg9685>

Right turns at urban intersections pose a high collision risk for cyclists, but existing infrastructure-based warning systems often suffer from high false alarm rates and a lack of specificity. Following a PICO-based research design, this study addresses the research question of whether targeted acoustic warnings using beamforming technology can effectively warn cyclists (target) while simultaneously minimizing the disturbance for pedestrians and residents (non-targets). Using a fully immersive VR environment with spatial audio simulations and realistic traffic noise recordings, the study examined how cyclists, pedestrians, and residents perceive and accept these targeted signals. The experiment compared different warning sounds at varying distances from the intersection to determine their perceptibility and effectiveness. The results show that while all acoustic warnings were reliably perceived as relevant, verbal warnings proved to be the most effective, as speech is immediately understood regardless of the cyclist's distance from the intersection. In terms of environmental impact, the study found that beamforming can significantly minimize disturbance to pedestrians and residents, with narrow-beam signals perceived as the least disturbing. This work contributes to this field of research by demonstrating that spatially controlled acoustic systems offer a more specific and socially acceptable solution for protecting vulnerable cyclists than conventional omnidirectional or purely visual methods. By combining technical precision with intuitive communication, this approach effectively counteracts the “cry wolf” effect and improves overall road safety.

1. Introduction

Infrastructural measures have significantly improved cyclist safety on medium-sized road sections in recent years, such as physical separations, flexible bollards, and coloured cycle paths. However, intersections in urban environments remain critical accident hotspots. In these areas, interaction between cyclists and motorised traffic is unavoidable. Of particular concern are conflicts involving cyclists travelling straight ahead and right-

turning motor vehicles. Accident analyses by the German Insurers Accident Research (UDV) indicate that in turning accidents involving cyclists, motor vehicle drivers are assigned primary fault in over 90% of cases. This finding underscores the prevalence of perceptual and attentional deficits on the part of turning drivers (Kolrep-Rometsch et al., 2013).

Concurrently, the significance of cycling within the urban transport system is growing markedly. In major German cities, the

^a Corresponding author: rodney.leitner@human-factors-consult.de

proportion of cycle traffic in the modal split has increased significantly over the last decade. Data from the “Mobility in Cities” (SrV) project for comparable major cities show that the average proportion of distances covered by bicycle increased from 12.9% in 2013 (Ahrens *et al.*, 2015) to 18.0% in 2023 (Hubrich *et al.*, 2025). This development increases not only the exposure of cyclists but also the frequency of potential conflicts, particularly given the limited traffic space available in many locations.

Technical countermeasures have primarily focused on turning vehicles. In recent years, new safety regulations have changed for heavy commercial vehicles. Since 2022, electronic turn-assist systems have become a legal requirement for new truck types across Europe (Regulation (EU) 2019/2144). This regulation addresses a major safety issue, namely collisions between cyclists and heavy vehicles turning right. As highlighted by the UDV (2019), these accidents are often fatal, and heavy trucks are involved far more often than other vehicles, making automated assistance a critical safeguard. Nevertheless, two structural gaps remain: First, the existing fleet of older vehicles will only be covered over long transition periods; second, conflicts with passenger cars, which account for a large proportion of interactions in terms of numbers, remain a significant problem. Consequently, a substantial level of risk persists, even with the increasing proliferation of vehicle-based assistance systems. From a Safe System perspective, this residual risk requires additional protective layers for vulnerable road users, rather than shifting legal responsibility away from turning drivers.

Since vehicle-based systems cannot provide ubiquitous protection in the near future, infrastructure-based approaches must fill this gap. In practice, such warnings are frequently implemented using broadly emitting signals (e.g., flashing amber lights). However, these systems are often nonspecific, addressing not only the individuals at risk but also uninvolved road users. This presents a fundamental problem: If warnings are triggered repeatedly in the absence of an acute threat from the

user’s perspective, subjective false alarm rates increase. This leads to habituation, reduced compliance, and a loss of trust in the warning system. This phenomenon is described in research as the ‘cry-wolf’ effect or alarm fatigue (Breznitz, 2013). Consequently, while nonspecific signalling ensures constant visibility, such a system diminishes its warning effect at critical moments while simultaneously causing unnecessary disturbance to bystanders. While visual infrastructure warnings have been explored, they often suffer from inattention blindness in visually cluttered urban environments. Acoustic signals, conversely, possess a higher potential to attract attention but have traditionally been avoided due to their contribution to noise pollution in cities. This requires a solution that combines the urgency of the sound with spatial precision.

For this reason, it is useful to understand such strategies not merely as “warnings” but as specific support for the situational awareness of vulnerable road users. The basic idea is not to shift responsibility from turning drivers to cyclists, but to add an additional, proactive layer that can mitigate the consequences of human error, similar to helmets, ABS or airbags. Even though the cyclist does not typically cause the conflict, an early hazard cue of a potentially dangerous situation can help de-escalate the situation through small behavioural changes (e.g., reducing speed, redirecting visual scanning, increasing braking readiness), rather than relying exclusively on last-second reactions by the driver. One approach that can overcome the constraints of widely distributed infrastructure warnings is targeted acoustic hazard information using beamforming. Beamforming uses an array of loudspeakers to shape sound waves spatially and focus them in a specific direction. Conceptually, such systems are designed as event-triggered cues that are only activated when a high-risk conflict is anticipated based on sensor-derived prediction, rather than as continuous long-duration warnings to every passing cyclist. As a result, it is possible to target

vulnerable cyclists while reducing the impact on non-target groups, such as pedestrians and residents. Beyond spatial focusing, the semantic quality of the signal plays a crucial role in overcoming alarm fatigue. The effectiveness of a targeted warning may depend significantly on whether the information is presented as an abstract tone or as clear verbal instructions, which minimize the cognitive load during the interpretation process.

In the present study, the central objective is to evaluate the perceptual-psychological effectiveness and the directional specificity of beamforming-based hazard information in a realistic VR environment. The study does not aim to evaluate a deployable warning system or to identify the optimal warning signal. Instead, it should be understood as a proof-of-concept investigation of whether beamforming technology can generate perceptually distinguishable acoustic fields that selectively address the intended target group while minimising disturbance to non-target listeners. To ensure a rigorous scientific approach, the study design follows the PICO framework: the Population consists of urban cyclists as well as uninvolved pedestrians and residents; the Intervention involves targeted acoustic warnings via infrastructure-based beamforming; the Comparison is drawn between different signal types (verbal vs. tonal) and spatial configurations (narrow-beam vs. wide-beam); and the Outcomes are measured through warning interpretation, behavioural intention, and perceived disturbance. Within this framework, the following research questions are addressed:

- **RQ1:** To what extent are beamforming signals perceived as warnings by cyclists and interpreted correctly?
- **RQ2:** How does the type of signal influence the immediacy of the warning effect and the resulting behavioural intention?
- **RQ3:** Is the directionality of beamforming sufficiently specific to significantly

minimize interference for non-target groups (pedestrians and residents)?

To investigate these questions empirically, a user study was conducted in a fully immersive virtual reality (VR) environment. Virtual reality was selected as the research environment because it enables the precise auralisation of complex beamforming patterns and the systematic manipulation of traffic variables under strictly controlled, reproducible and safe conditions. This methodology is particularly suitable for evaluating safety-critical acoustic interfaces where real-world testing would pose an unacceptable risk to participants and where perceptual effects need to be isolated from behavioural influences. Participants experienced the scenes from different perspectives, both as warned cyclists and as pedestrians and residents, in order to reliably measure both the safety-related effectiveness and social acceptance of the concept. The study was conducted in accordance with the ethical guidelines for research involving human subjects. Since the virtual reality environment posed no physical risk to participants and no sensitive personal data beyond standard demographic information were collected, the study was deemed minimal risk.

2. Technical approach

To auralise the sound immissions of the acoustic scenes for various listening positions, a hybrid approach was chosen, combining simulations of the beamformed warning signals with measurements of traffic noise immissions at the desired locations. For the simulations, a realisable array configuration was defined, and beamforming filters were derived to control the sound radiation in the horizontal plane. The transfer function from each emitter to each immission location was estimated using the highly realistic geometrical solver of the acoustic simulation tool by Treble Technologies. Both the simulations and measurements were conducted in the Ambisonics domain,

enabling spatially accurate binaural sound reproduction in the VR test environment.

2.1 Simulation of loudspeaker array

First, a suitable loudspeaker arrangement must be defined that is both optimised for focused sound beams in the audible frequency range and installable at traffic intersections. As steering the beam in the vertical direction is not required, a linear array with two horizontal rows of 16 loudspeakers each was selected (see Figure 1). Using two rows allows for smaller horizontal membrane spacing and, therefore, a higher boundary frequency with respect to spatial aliasing. The total number of 32 loudspeakers represents a reasonable trade-off between material efficiency and sufficient beamforming performance.

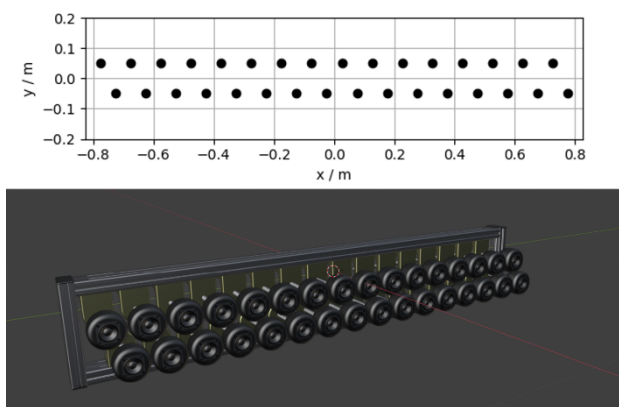


Figure 1. Defined loudspeaker array: positioning (top) and derived 3D model with housings and mounting (bottom)

The beamforming filters were calculated using the pressure-matching method (Olivieri *et al.*, 2015). In this study, a constant beam directed towards the cycleway was used; however, in practice, the beam could also track a cyclist with a variable radiation angle. For the employed method, so-called “dark points” and “bright points” must be defined, where minimal sound radiation is desired or where the signal should be clearly audible, respectively. Two sets of filters were calculated using an iterative optimisation with a variable regularisation parameter and different configurations of dark and bright points: one set targeting an optimal radiation pattern for broadband

signals and one for narrow-band signals containing only frequencies below the spatial aliasing frequency. The resulting radiation patterns for three different warning signals are depicted in Figure 2.

2.2 Auralisation of warning sounds

In addition to beamforming, auralisation must also account for sound propagation effects caused by surrounding objects. Therefore, a simplified 3D model of the intersection was created to perform sound propagation calculations for each loudspeaker–listening position pair. These calculations were carried out using the geometrical solver of the sound propagation model developed by Treble Technologies. By assigning realistic material parameters to all surfaces, physical effects such as reflection, air absorption, distance attenuation, and scattering can be modelled. The resulting impulse responses are available in second-order spatial Ambisonics format and are therefore suitable for 3D auralisations in VR applications using a binaural decoder with head tracking. By convolving a warning sound with the beamforming FIR filter of each loudspeaker and the corresponding spatial impulse response for a given listening position, and subsequently superposing the resulting Ambisonics signals, the warning sound can be rendered for the selected position. As two different loudspeaker array positions were used due to the unusually curved route of the cycleway at the test intersection, and 14 listening positions were considered, a total of 896 impulse responses had to be calculated.

The sound emitters in the simulations were adapted to represent the actual frequency-dependent directivity patterns of the loudspeakers with 3D-modelled enclosures. For this purpose, a boundary element method (BEM) simulation of the loudspeaker model was conducted using mesh2HRTF (Ziegelwanger *et al.*, 2015a,b; Brinkmann *et al.*, 2023) and validated with AKABAK (Panzer, 2026). This step is necessary to avoid overemphasising reflections from the rear side of the loudspeakers, where

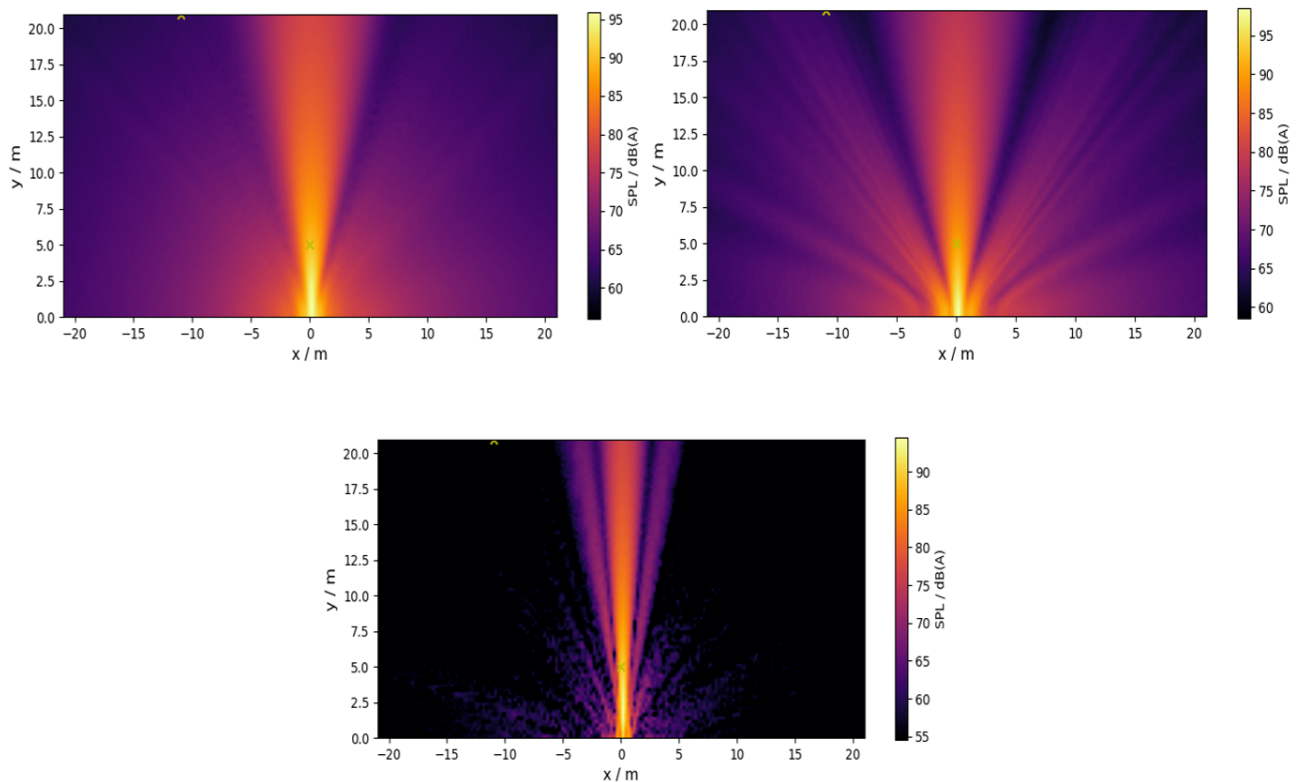


Figure 2. Radiation patterns in front of the speaker array for three different warning sounds: verbal (left), broad-band tone (right) and narrow-band tone (bottom)

sound emissions are significantly attenuated compared to frontal radiation.

2.3 Auralisation of traffic noise

To obtain realistic acoustic scenes for all auralised scenarios, authentic traffic noise immissions were added. As the most realistic approach is to use recordings from the actual listening positions in the test environment in the city of Braunschweig, an acoustic measurement campaign was conducted (see Figure 3). Ideally, the recordings should also be available in second-order Ambisonics. Therefore, Core Sound OctoMic™ microphones were used, which provide second-order Ambisonics in the azimuth plane and first-order Ambisonics in elevation. As not all 14 listening positions could be equipped with a spatial microphone, each immission location was assigned one measurement microphone, and five groups of nearby positions shared one OctoMic™ each. The measurement microphones were calibrated, allowing for accurate sound pressure level

reproduction. Spectral differences between the microphones were evaluated during post-processing and used to apply appropriate spectral corrections to the Ambisonics recordings for each position. In this way, spectral and localisation errors resulting from the simplified spatial recording approach could be minimised.

2.4 Finalisation of auditory stimuli

The simulated warning signals for each listening position were calibrated in accordance with ISO 7731 (ISO, 2003), which recommends that danger signals be at least 15 dB(A) louder than the background noise within the danger area. With a maximum sound pressure level of 74.8 dB(A), the warning signals were adjusted to yield 90 dB(A) at a distance of 5 m in front of the loudspeaker arrays. To ensure correct binaural sound pressure levels during the VR study, a comparative calibration measurement was conducted in an acoustically optimised environment using



Figure 3. Measurement site with microphones and safety measures

a measurement microphone and a Neumann KU 100 dummy head. The finalised signals were imported into a live binaural decoding application developed using the software Max by [Cycling '74 \(2025\)](#). This audio application is controlled by the visual VR application via MQTT and receives head-tracking data corresponding to the current head rotation as well as a sound file selection for playback.

3. Methodology

3.1 Study design

A controlled, within-subject VR study was conducted to assess (1) the perception of the beamforming warning signals from a cyclist's perspective and (2) the perceived disturbance and perceivability of these signals when heard from pedestrians and residents. The study deliberately focused on perceptual and attitudinal responses rather than on actual cycling behaviour and should be understood as a proof-of-concept investigation of beamforming in a realistic VR environment. All participants experienced all positions and warning stimuli.

3.2 Participants

A total of $N = 25$ subjects participated in the study (12 male, 13 female). The participants

ranged in age from 22 to 71, with an average age of $M = 46.2$ years ($SD = 15.6$). All participants provided written consent and received compensation for their time. In terms of cycling habits, the participants represented various cycling experiences. The self-reported frequency of cycling was distributed as follows: 4% cycled daily, 36% cycled several times a week, 28% cycled one to three times a month, 16% rarely cycled, and 16% never cycled. Due to the stationary nature of the experiment, no cases of simulator sickness or physical discomfort were reported. This setup successfully avoided the sensory conflicts between perceived and actual motion that often occur in dynamic VR environments. One participant did not complete the final preference ranking from the cyclist's perspective; analyses of this ranking are therefore based on $N = 24$, whereas all other cyclist-related analyses use the full sample of $N = 25$.

3.3 Virtual environment and listening positions

The study was carried out in a realistic 1:1 virtual simulation that replicates a complex urban intersection with several main roads in Braunschweig, Germany. The high complexity of the junction, combined with its heavy use by a wide range of road users, makes it inherently risky. The junction would directly benefit from further safety measures and is therefore a suitable test environment for the evaluation of a directional warning system to protect vulnerable road users. Within this environment, eight specific positions were defined to represent the perspectives of different groups (see Figure 4).

Four positions were located on the cycle path approaching the intersection. Because the cycleway features a slight bend, two loudspeaker array locations were implemented.

- **Cyclist positions.** Four positions were located on the cycle path approaching the intersection. Due to the slight curve in the cycle path, two loudspeaker



Figure 4. Positions of all measurement points (C, P, and R) within the user study

arrays were implemented. Position C1 represented the farthest position, located approximately 35–40 m upstream of the intersection. Positions C2 and C3 were placed 15 and 25 m closer to the intersection on the cycle path. Position C4 was located directly in front of the intersection, right next to one of the loudspeaker systems.

- **Pedestrian positions.** Position P1 was located on the pavement beside the cycle path at a similar distance from the array as C2. Position P2 was located directly at the pedestrian crossing. Each participant experienced both pedestrian positions, resulting in two ratings per warning stimulus from a pedestrian perspective (P1 and P2).
- **Resident positions.** The positions R1 and R2 were located on a ground-floor balcony of an adjacent building, with R2 positioned closer to the intersection than R1. Each participant experienced both resident positions, resulting in two

ratings per warning stimulus from a balcony perspective (R1 and R2).

The virtual scene was presented using a Pimax 8K head-mounted display (HMD) featuring a dual-4K resolution (3840 x 2160 pixels per eye), a 90 Hz refresh rate, and an ultra-wide field of view of up to 200° (diagonal) to ensure a high level of immersion and peripheral perception. The VR system was equipped with head tracking, and the audio signal was dynamically adapted to the participant's head orientation. During the study, participants remained seated in the physical environment and were teleported between predefined positions in the virtual world. This design intentionally separated perceptual evaluation from active cycling behaviour: participants could explore the directional characteristics of the signals by turning their heads, while other motion-related influences were excluded. As a result, observed differences can be attributed to the acoustic characteristics of the beamforming signals.

3.4 Procedure

At each position, participants were presented with three warning stimuli representing different directivity characteristics: a narrow-beam warning, a wide-beam warning, and a verbal warning signal. The verbal signal consisted of the spoken phrase “Achtung, Gefahr!” (Attention, Danger!). At each cyclist position, participants first experienced a baseline scene with authentic traffic noise and no warning signal. These baseline scenes served as an acoustic reference for the intersection but were not rated on the warning-related scales, because no warning stimulus was present; the analyses focus on differences between the three warning signals rather than on “with vs. without” comparisons. Each stimulus was looped for a maximum duration of 30 seconds to allow participants sufficient time to explore its perceptual and directional characteristics, including head movements in the VR environment. The warning sound itself was short and simply repeated within this interval; participants could request early termination once they felt able to rate the stimulus. This experimental playback duration does not represent the intended design of real-world warnings.

Participants were informed in advance that the acoustic signals may represent a warning related to a dangerous situation for a specific cyclist. Beamforming-based infrastructure warnings are currently unknown to most road users; without a short conceptual explanation, many participants would have struggled to understand the purpose of the signals. The experiment was structured into three consecutive blocks reflecting different perspectives. The block sequence (cyclist - pedestrian - resident) was chosen intentionally, so that participants first understood the cyclist's warning concept before assessing the potential disturbance as pedestrians and residents. Importantly, pedestrian and resident ratings did not address warning interpretation or behavioural intention, but only perceptibility and disturbance, which limits the impact of the block order. Within each block, the

order of positions was randomised (C1–C4 randomised; P1/P2 randomised; R1/R2 randomised). Within each position, the order of the three stimuli was randomised to reduce order and habituation effects.

The experimental flow was controlled by a custom-built state machine, communicating via MQTT with a low-latency audio engine in Max 8 (Cycling '74, 2025). This ensured that the spatial audio cues remained synchronized with the visual head-tracking data at all times, maintaining a high level of immersion and directional stability.

3.5 Measures

After each stimulus (excluding baseline traffic-noise scenes), participants rated role-specific statements using a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree, with 4 = neither agree nor disagree). In the role of a cyclist, participants rated three dimensions: (1) perceptibility (“The acoustic signal is very clearly perceivable to me.”), (2) interpretation of the warning (“The acoustic signal means a warning to me.”), and (3) behavioural intention (“I would change my riding behaviour due to the acoustic signal.”). In the role of a pedestrian or resident, two dimensions were assessed: (1) perceptibility (“The acoustic signal is very clearly perceivable to me.”), and (2) disturbance (“I perceived the acoustic signal as disturbing.”). After the participants completed their evaluations for all positions, they were asked to rate the three warning signals from different perspectives. From the perspective of a cyclist, they rated the signals from most preferred (rank 1) to least preferred (rank 3). From the perspective of a resident, they rated the signals from least disruptive (rank 1) to most disruptive (rank 3).

4. Results

The following sections present the statistical analysis of the user study. The results are structured by role perspective, starting with cyclists followed by pedestrians and residents, and concluding with direct comparison rankings.

4.1 Cyclist perspective

4.1.1 Perceptibility

The first dimension assessed was the acoustic perceptibility of the signals. Descriptive statistics reveal consistently high ratings across all positions and stimuli. The grand mean across all conditions was $M=6.6$ ($SD=0.7$), indicating a distinct ceiling effect. Friedman tests confirmed that neither the position nor the stimulus type had a significant impact on perceptibility.

4.1.2 Interpretation of the warning

The second dimension assessed the extent to which participants interpreted the acoustic signal specifically as a warning. In contrast to the assessments of perceptibility, the results on the effect of the warning showed significant differences between the signal types at all four positions. Friedman tests revealed a significant main effect of the signal type at all locations: C1 ($\chi^2(2)=12.74$, $p=.0017$), C2 ($\chi^2(2)=16.62$, $p=.0002$), C3 ($\chi^2(2)=8.82$, $p=.0122$), and C4 ($\chi^2(2)=11.54$, $p=.0031$). Post-hoc pairwise comparisons (Wilcoxon rank sum tests with Bonferroni correction) were performed to identify differences between the stimuli for each position. The descriptive statistics for warning interpretation, including means (M) and standard deviations (SD), are summarized in Table 1.

Table 1. Means (M) and Standard Deviations (SD) for warning interpretation across signal types and positions (N=25)

Position	M/SD	Narrow-beam signal	Wide-beam signal	Verbal signal
C1	M	4.92	5.48	6.52
	SD	1.61	1.26	0.65
C2	M	5.16	5.76	6.76
	SD	1.62	1.33	0.66
C3	M	5.88	6.08	6.88
	SD	1.24	1.12	0.60
C4	M	5.40	5.84	6.92
	SD	1.71	1.31	0.40

At every position (C1–C4), the verbal signal was rated significantly higher than both beam signals (see Figure 5). The comparison of verbal signals vs. narrow-beam signals

showed significance at all points (C1: $p=.0008$, C2: $p=.0008$, C3: $p=.0212$, and C4: $p=.0023$). Pairwise comparisons with the wide-beam signal showed similar significant differences at all positions (C1: $p=.0078$, C2: $p=.0029$, C3: $p=.0260$, C4: $p=.0044$). In contrast, the differences between the two beam signals were mostly non-significant. A significant difference was only found at position C2 ($p=.0405$), where the wide-beam signal received higher ratings.

4.1.3 Behavioural intention

The third dimension evaluated the participants' self-reported behavioural intention based on the warning interpretation. Similar to the results for warning interpretation, the behavioural intention showed significant differences between the signal types at all four positions. Friedman tests confirmed a significant main effect of the signal type at all positions: C1 ($\chi^2(2)=8.78$, $p=.0124$), C2 ($\chi^2(2)=11.06$, $p=.0040$), C3 ($\chi^2(2)=8.54$, $p=.0140$), and C4 ($\chi^2(2)=6.48$, $p=.0392$). Post-hoc pairwise comparisons were conducted. The descriptive statistics for behavioural intention, including means (M) and standard deviations (SD), are summarized in Table 2.

Table 2. Means (M) and Standard Deviations (SD) for behavioural intention across signal types and positions (N=25)

Position	M/SD	Narrow-beam signal	Wide-beam signal	Verbal signal
C1	M	4.80	5.24	6.12
	SD	1.80	1.45	0.83
C2	M	5.00	5.56	6.36
	SD	1.63	1.58	0.76
C3	M	5.56	5.8	6.44
	SD	1.29	1.22	0.77
C4	M	5.32	5.84	6.64
	SD	1.77	1.49	0.70

At nearly all positions, the verbal signal showed a significantly higher behavioural intention than both beam signals. The comparison of verbal vs. narrow-beam showed significance at all points (C1: $p=.0032$, C2: $p=.0021$, C3: $p=.0140$, and C4: $p=.0091$). Pairwise comparisons with the wide-beam signal were also significant at three positions

(C1: $p=.0226$, C2: $p=.0343$, and C3: $p=.0490$). Consistent with the findings on warning interpretation, the behavioural intention did not significantly differ between the two beam signals. A significant difference was only observed at position C2 ($p=.0213$), where the wide-beam signal received significantly higher ratings for behavioural intention compared to the narrow-beam signal (see Figure 5).

4.2 Pedestrian perspective

This section examines the impact of warning signals from the perspective of pedestrians. Each of the 25 participants experienced both pedestrian positions (P1 and P2), resulting in 50 observations per warning stimulus. Initial analyses showed no significant differences between positions P1 and P2 for any of the stimuli. Therefore, the data from both locations were pooled ($N=50$) for the analyses, in order to focus on differences between the warning signals rather than on positional effects.

4.2.1 Perceptibility

The first dimension assessed from the pedestrian's perspective was the perceptibility of the signals. Friedman tests confirmed a significant main effect of the signal type ($\chi^2(2)=59.43$, $p<.0001$). The descriptive results show a statistically significant reduction in perceptibility for the narrow-beam signal ($M=4.78$, $SD=1.37$) compared to the wide-beam signal ($M=6.66$, $SD=0.56$) and the verbal signal ($M=6.72$, $SD=0.54$). Post-hoc pairwise comparisons revealed a clear distinction in how the signals were perceived from a pedestrian perspective. The narrow-beam signal was rated significantly less perceptible than both the wide-beam signal ($p<.0001$) and the verbal signal ($p<.0001$). The results indicate a statistically significant reduction in perceptibility for the narrow-beam signal compared to the other stimuli. In contrast, no significant difference was found between the wide-beam signal and the verbal signal.

4.2.2 Disturbance

The second dimension assessed the perceived disturbance of the signals. Consistent with the perceptibility results, the level of perceived disturbance was significantly influenced by the warning signal (Friedman test: $\chi^2(2)=24.33$, $p<.0001$). Similar to the perceptibility, the narrow-beam signal was rated as significantly less disturbing ($M=3.1$, $SD=1.82$) than the wide-beam signal ($M=4.52$, $SD=1.76$; $p<.0001$) and the verbal signal ($M=4.74$, $SD=1.98$; $p<.0001$). No significant difference in disturbance was found between the wide-beam signal and the verbal signal (see Figure 6).

4.3 Residents' perspective

This section examines the impact of warning signals from the perspective of residents. Each of the 25 participants rated both balcony positions (R1 and R2), resulting in 50 observations per warning stimulus. Initial analyses showed only a minor statistical difference in perceived disturbance for the wide-beam signal R1 and R2. Because our primary focus was on differences between the warning stimuli rather than on a detailed positional analysis of the balconies, ratings from R1 and R2 were pooled ($N=50$) to maintain consistency across all stimuli.

4.3.1 Perceptibility

The first dimension assessed the perceptibility of the signals from the residents' perspective. A Friedman test revealed a significant main effect of the signal type ($\chi^2(2)=27.21$, $p<.0001$). The narrow-beam signal was rated as significantly less perceptible ($M=5.24$, $SD=1.41$) than the wide-beam signal ($M=6.37$, $SD=0.83$; $p<.0001$) and the verbal signal ($M=6.20$, $SD=0.93$; $p<.0001$). No significant difference in perceptibility was found between the wide-beam and the verbal signal.

4.3.2 Disturbance

The second dimension assessed the perceived disturbance of the signals. Consistent with the perceptibility results, the level

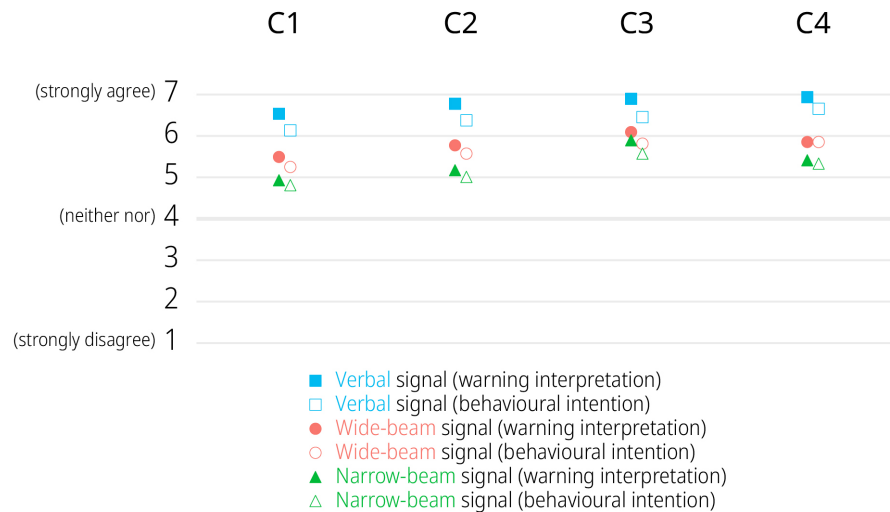


Figure 5. Warning interpretation and behavioural intention from the cyclists' perspective (N=25)

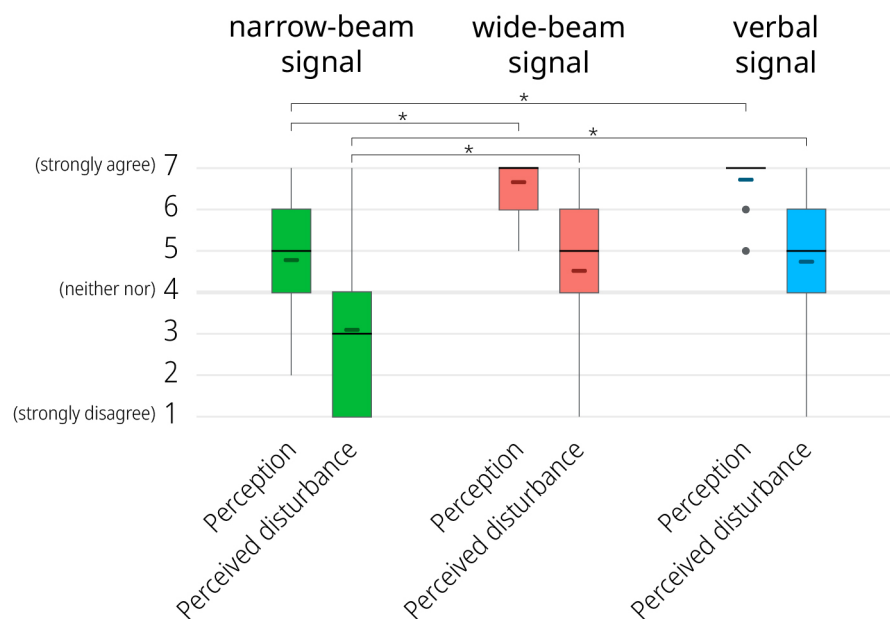


Figure 6. Perceptibility and disturbance of warning signals from the pedestrians' perspective (pooled data, N=50)

of perceived disturbance was significantly influenced by the signal type (Friedman test: $\chi^2(2)=15.97$, $p=.0003$). Post-hoc pairwise comparisons showed that the narrow-beam signal ($M=4.7$, $SD=1.72$) was rated significantly less disturbing than the wide-beam signal ($M=5.9$, $SD=1.15$; $p<.0001$). However, the difference between the narrow-beam and the verbal signal did not reach statistical significance ($p=.0663$), a tendency toward lower perceived disturbance was observed for the narrow-beam signal. In addition,

a significant difference was found between the wide-beam and verbal signal ($p=.0296$), with the verbal signal being perceived as less disturbing from the residents' perspective (see Figure 7).

4.4 Subjective warning signal preference

In the final stage of the evaluation, participants were asked to rank the three warning signals based on their personal preference (1 = most preferred, 3 = least

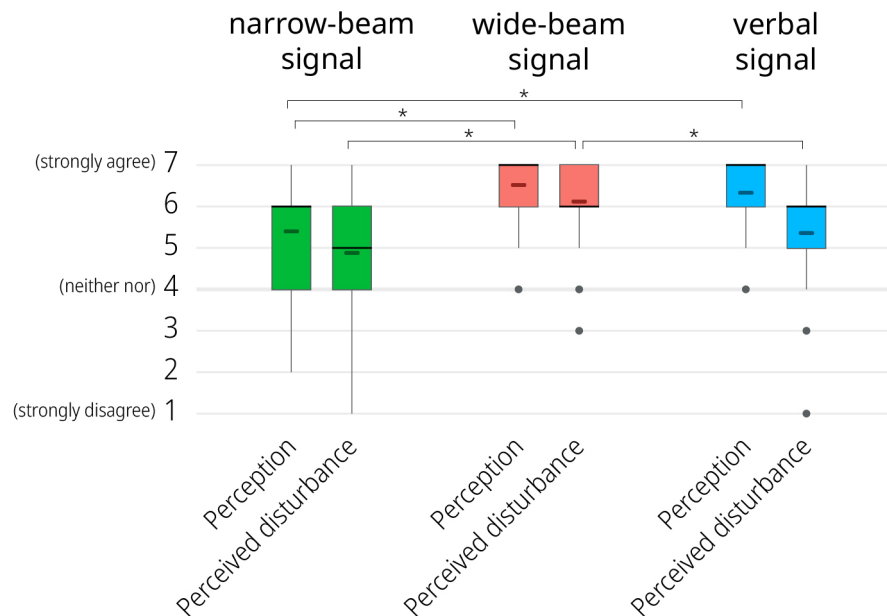


Figure 7. Perceptibility and disturbance of warning signals from the residents' perspective (pooled data, N=50)

preferred). The ranking from the cyclists' perspective shows that the verbal signal was most frequently rated as the preferred warning signal (14 out of 24 participants). The wide-beam signal was rated best by 8 participants, whereas the narrow-beam configuration was only ranked first by 2 participants. Conversely, the narrow-beam signal was also most frequently chosen as the least preferred. From the residents' perspective, the ranking shows that the narrow-beam signal was most frequently ranked first and is therefore the most preferred (11 out of 25 participants). Both the wide-beam signal and the verbal signal were rated as preferred by 7 participants each, with the verbal signal most often receiving the least preferred rank 3. Figure 8 shows the comparative distribution of these preferences.

5. Discussion

5.1 Methodological reflection

A key objective of this study was to investigate whether the directional characteristics of warning signals can be accurately reproduced and evaluated in a virtual environment as a proof-of-concept, focusing on perception

and attitudes rather than on full behavioural simulation. The results of the study can be considered a success in this regard: the clear differences in perception between the narrow-beam and wide-beam stimuli by non-target groups (pedestrians and residents) prove that the acoustic simulation model successfully converted the physical differences into subjective perceptible impressions. The consistency of the evaluations across different spatial positions confirms that the VR environment could reliably transmit the directionality of the warning signals. The high correlation between the physical simulation and the subjective ratings indicates a strong auditory immersion. This confirms that the complex auralisation was perceived by the participants as a valid and realistic representation of urban acoustic environments in safety-critical research.

This also demonstrates that VR-based testing is a valid and efficient tool for prototyping safety-critical acoustic systems. Beyond the field of traffic safety, this opens up a wide range of testing possibilities for all applications in which directional sounds play a role. VR offers the advantage of being able to manipulate environmental variables under

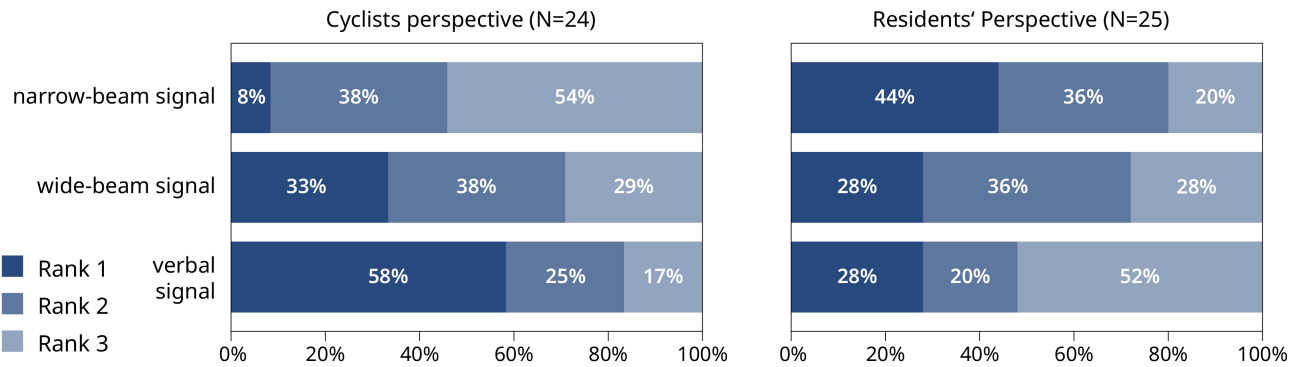


Figure 8. Subjective preference ranking of the three warning signals from the cyclists' (N=24, left) and residents' perspective (N=25, right)

controlled conditions, which would be nearly impossible in real-world field tests. Despite these advantages, certain limitations must be considered. Although statistical significance was determined with a sample size of N=25, the study does not cover the entire diversity of road users (e.g. older people, athletic speeders or children). Furthermore, although the VR environment offered a high degree of acoustic control, it lacked the multisensory complexity (e.g. vibrations, wind, changing weather) of actual city traffic. Future research should therefore focus on validating these results in field trials to confirm the validity of beamforming. In addition, dynamic VR scenarios and field studies will be needed to investigate trigger timing, real-world warning durations and actual behavioural adaptation (e.g. speed reduction, braking, swerving, gaze changes), which were intentionally outside the scope of this first proof-of-concept study.

5.2 Signal effectiveness of the three stimuli

Regarding RQ1 and RQ2, the results confirm that all stimuli were perceived very well from the perspective of the cyclists. This demonstrates that the observed differences in warning interpretation (RQ1) and the resulting behavioural intention (RQ2) are not due to a lack of perceptibility, but rather to the clarity and intuitiveness of the signal itself. The results show a clear advantage of verbal signals over beam warnings. From a cognitive point of view, this is very plausible, as verbal information has unique semantics. In this

proof-of-concept setting, it represents an upper bound of warning clarity instead of a general solution for all user groups and situations. Regardless of the distance or complexity of the traffic situation, a verbal message such as "Attention, danger!" is immediately understood as a warning. In contrast, narrow and wide-beam signals are initially abstract to the user and require a moment of cognitive processing to be interpreted as a warning about a potentially dangerous situation. Although the data show that the interpretation of the two beam signals improved with decreasing distance to the intersection, it never reached the effectiveness of the verbal signal. Consequently, the intention to change cycling behaviour was higher with a verbal warning than with either of the beam signals. Verbal warnings offer a practical solution to the perception and attention deficits identified by [Kolrep-Rometsch et al. \(2013\)](#). While turning drivers often fail to notice cyclists, verbal warnings can effectively compensate this deficit. The urgent clarity of the verbal warning shortens the cyclist's cognitive reaction time, enabling a proactive response to the potentially dangerous situation.

The descriptive data also shows that the interpretation of narrow and wide-beam signals as warnings peaked at position C3, but showed a slight decline directly at the intersection position C4. Qualitative feedback from some participants revealed that the beam signals sounded similar to familiar urban traffic noises, such as the reverse

warning signals of trucks or the acoustic signals for visually impaired pedestrians at crosswalks. This suggests that the chosen beam signals could lead to confusion or even misinterpretations and would need to be investigated in more detail in follow-up studies. Alternatively, the slight decrease in warning interpretation at position C4 could be due to acoustic masking effects. The closer the cyclist gets to the intersection, the closer they get to other sound sources, such as turning motor vehicles. This could have led to subtle masking of the beam stimuli compared to the more robust frequency range of the verbal signal. From the cyclists' perspective, verbal communication offers the greatest clarity and is the most robust trigger for behavioural adjustments. This can be underpinned by the fact that they preferred the verbal signal as the primary warning option.

5.3 Safety versus noise pollution

With respect to RQ3, which investigated whether beamforming can effectively minimise disturbance to non-target groups, the results show that the narrow-beam signal was highly effective. Within a Safe System perspective, this supports the idea of adding an extra protective layer for vulnerable road users without shifting legal responsibility away from turning drivers. Compared to the wide-beam and verbal signals, the narrow-beam signal significantly reduced both the perceptibility and subjective disturbance for pedestrians and residents. Although the residents' ratings did not yet reach the expected ideal low level, the results provide proof of concept: beamforming works in principle, even if there is still room for optimisation in the specific design of the acoustic beam patterns. In real-world applications, such systems should be understood as event-triggered hazard information that activates only in predicted high-risk situations, rather than as continuous long-duration warnings for every approaching cyclist or pedestrian. Questions of trigger criteria, lead time and warning duration were not addressed in this study and will need to be investigated in dedicated follow-up work.

Interestingly, the perceived disturbance was significantly lower among pedestrians than among residents. Qualitative feedback from participants suggests that some of them did not find the signal disturbing, but rather perceived it as potentially helpful information. This observation is crucial when considering the 'cry wolf' effect (Breznitz, 2013). The fact that cyclists in this study showed a high behavioural intention to respond to the signals suggests that the targeted warnings were actively perceived due to a potentially dangerous situation. In contrast to conventional, non-specific infrastructure warnings, which are often ignored over time due to high false alarm rates, the spatial specificity of beamforming appears to promote clearer situational awareness. The high intention to adjust cycling behaviour suggests that the signals were not dismissed as irrelevant noise, but were considered in the participants' decision-making process.

However, the subjective preference rankings reveal a clear conflict of interest. While cyclists prioritise the high information content of the verbal signal, residents and pedestrians generally prefer the narrow-beam signal to minimise disturbance. This discrepancy demonstrates that there is no simple solution for complex urban environments. The results indicate a necessary technical compromise: the ideal solution would probably be a highly directional (narrow-beam) acoustic signal, if technically feasible. Such a system would meet cyclists' need for clear verbal information while respecting local residents' requirements for noise reduction. It should be emphasised that this concept is intended as an additional safety layer for vulnerable road users at documented high-risk intersections, not as a mechanism to transfer responsibility for yielding from motor vehicle drivers to cyclists or pedestrians.

6. Conclusion

This proof-of-concept study demonstrates that acoustic beamforming can address key safety-related and social acceptance

challenges of infrastructure-based warnings. The core findings of this study can be summarized as follows:

- **Advantages of verbal warnings (RQ1 & RQ2):** Verbal signals are significantly more effective than acoustic signals because they provide immediate semantic clarity and trigger stronger behavioural intentions without requiring cognitive decoding. In this setting, spoken warnings thus represent an upper bound of warning clarity, not a single universal solution for all user groups and situations.
- **Social acceptance (RQ3):** Narrow-beam warnings provide proof of concept for a significant reduction in disturbance for pedestrians and residents without compromising the safety of the target group.
- **Methodological reliability:** The VR-based methodology proved to be a robust tool for the auralisation and evaluation of directional acoustic systems at a perceptual level prior to physical prototype development.

While the method has clear advantages, some research questions need to be addressed in the future. The participants noted that some beam signals sounded like existing urban noises, like the acoustic signals at crosswalks for visually impaired pedestrians. Therefore, future work must focus on developing unique signal characteristics that prevent confusion with existing warning or notification tones. In order to avoid habituation and to reduce noise further, development should move towards context-dependent and event-based warnings, where signals are only given when a high-risk conflict is detected. This should also include adaptive volume control depending on the ambient noise level. In addition, further optimisation of the beam width and frequency response could sharpen the focus of the signal and ensure that the verbal information is intended exclusively for the target cyclist. Ultimately, the ideal “time to collision” at

which a warning should be triggered remains an open question that requires more detailed investigation. This will require dynamic VR scenarios and field trials that explicitly model trigger timing, lead time and actual behavioural responses (e.g. speed reduction, braking, swerving, gaze changes), which were intentionally outside the scope of the present study. By combining the clarity of verbal warnings with the precision of beamforming technology, urban environments can become safer for vulnerable traffic participants without impacting the quality of life for residents. Beyond the experimental setting, these findings offer important insights for the implementation of smart city infrastructures. This research enables local authorities to work towards the goals of ‘Vision Zero’. It offers a promising solution for protecting vulnerable road users at accident hotspots without the need for vehicle-based equipment or increasing noise pollution in cities, and should be understood as an additional safety layer within a Safe System approach rather than as a shift of responsibility from drivers to cyclists or pedestrians.

CRediT contribution

Rodney Leitner: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. **Christoph Ende:** Formal analysis, Investigation, Methodology, Supervision, Writing – review & editing. **Timo Jahns:** Data curation, Formal analysis, Investigation, Visualization. **Michaela Rehm:** Conceptualization, Writing – review & editing. **Marek Junghans:** Funding acquisition, Supervision, Writing – review & editing. **Christian Weissig:** Conceptualization, Funding acquisition. **Thomas Jürgensohn:** Conceptualization, Funding acquisition, Supervision.

Acknowledgements

The authors would like to thank Treble Technologies for providing free access to their

acoustic simulation software and for their technical support.

Declaration of competing interests

The authors report no competing interests.

Declaration of generative AI use

During the preparation of this work, the authors used Grammarly and Trink AI in order to improve the linguistic quality and ensure grammatical accuracy. These tools were used exclusively for language editing, stylistic polishing, and scientific writing enhancement to ensure the highest quality of the final draft. The output was reviewed and revised by the authors, who take full responsibility for the content of the publication.

Prior dissemination declaration

An earlier version of this work was presented at the 37th ICTCT conference, held in Berlin, Germany, on 23–24 October 2025.

Ethics statement

The present study did not require formal approval by an ethical review authority, as it involved non-invasive behavioural observations in a virtual reality environment and did not pose any physical or psychological risk to the participants. All subjects participated voluntarily and were fully

informed about the nature of the study, the data collection process, and their right to withdraw at any time without prejudice. Prior to the experiments, written informed consent was obtained from all participants. Furthermore, all data were anonymized to ensure the privacy and confidentiality.

Funding statement

This research was conducted within the project "SOUNDWARN" and was funded by the Federal Ministry for Digital and Transport (BMDV) as part of the mFUND research initiative under grant number 19F1204.

Data availability statement

The data are available on request to the authors.

Code availability statement

No custom code was developed for this study. Statistical analyses were performed using standard functions in R.

Editorial information

Handling editor: **Lai Zheng**, Harbin Institute of Technology, China.

Reviewers: **Ceri Woolsgrove**, European Cyclists' Federation, Belgium; **Xinyu Liang**, Hebei University of Technology, China.

Submitted: 30 January 2026; Accepted: 9 July 2026; Published: 17 July 2026.



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-4.0). View this license's legal deed at <https://creativecommons.org/licenses/by/4.0> and legal code at <https://creativecommons.org/licenses/by/4.0/legalcode> for more information.

References

- Ahrens, G. A., Wittwer, R., Hubrich, S., Wittig, S., Ließke, F. (2015). 'Sonderauswertung zum Forschungsprojekt „Mobilität in Städten – SrV 2013“: Städtevergleich'. Technical report, Technische Universität Dresden. https://tu-dresden.de/bu/verkehr/ivs/srv/ressourcen/dateien/SrV2013_Staedtevergleich.pdf
- Breznitz, S. (2013). 'Cry wolf: The psychology of false alarms'. in . Psychology Press.
- Brinkmann, F., Kreuzer, W., Thomsen, J., Dombrovskis, S., Pollack, K., Weinzierl, S., Majdak, P. (2023). 'Recent advances in an open software for numerical HRTF calculation'. *Journal of the Audio Engineering Society*, 71, 502–514. <https://doi.org/10.17743/jaes.2022.0078>
- Cycling '74 (2025). 'Max (Version 9.1.2) [Computer software]'. <https://cycling74.com>
- European Parliament and Council (2019). 'Regulation (EU) 2019/2144 on type-approval requirements for motor vehicles and their trailers, and systems, components and separate technical units intended for such vehicles, as regards their general safety and the protection of vehicle occupants and vulnerable road users'. *Official Journal of the European Union*, L 325(1). <http://data.europa.eu/eli/reg/2019/2144/oj>
- Hubrich, S., Wittig, S., Ließke, F., Wittwer, R., Gerike, R. (2025). 'Sonderauswertung zum Forschungsprojekt „Mobilität in Städten – SrV 2023“ für Große SrV-Vergleichsstädte'. Technical report, Technische Universität Dresden. https://tu-dresden.de/bu/verkehr/ivs/srv/ressourcen/dateien/SrV2023_Staedtevergleich_v2.pdf
- International Organization for Standardization (2003). 'ISO 7731:2003 Ergonomics—Danger signals for public and work areas—Auditory danger signals'. <https://www.iso.org/standard/33590.html>
- Kolrep-Rometsch, H., Leitner, R., Platho, C., Richter, T., Schreiber, A., Schreiber, M., Butterwegge, P. (2013). 'Abbiegeunfälle Pkw/Lkw und Fahrrad (UDV Forschungsbericht Nr. 21)'. Technical report, Gesamtverband der Deutschen Versicherungswirtschaft e. V. / Unfallforschung der Versicherer. <https://www.udv.de/resource/blob/78322/b3dd00fc1e86e9cd7164fa8872a45932/21-abbiegeunfaelle-pkw-lkw-und-fahrrad-data.pdf>
- Olivieri, F., Fazi, F. M., Shin, M., Nelson, P. A. (2015, May 6–9). 'Pressure-matching beamforming method for loudspeaker arrays with frequency dependent selection of control points'. *Audio Engineering Society Convention 138*. Warsaw, Poland: .
- Panzer, J. (2026). 'AKABAK (Version v324b126) [Computer software]'. <https://randteam.de/AKABAK3/AKABAK-Examples.html>
- Unfallforschung der Versicherer (UDV) (2019). 'Unfälle schwerer Lkw mit Fußgängern und Radfahrern (Unfallforschung kompakt Nr. 94)'. Technical report, Gesamtverband der Deutschen Versicherungswirtschaft e. V. <https://www.udv.de/resource/blob/79174/3ebf94ad78105c7c032c19684d27858d/94-unfaelle-lkw-fg-rf-data.pdf>
- Ziegelwanger, H., Kreuzer, W., Majdak, P. (2015a, July 12–16). 'Mesh2HRTF: Open-source software package for the numerical calculation of head-related transfer functions'. *22nd International Congress on Sound and Vibration*. Florence, Italy: <https://doi.org/10.13140/RG.2.1.1707.1128>
- Ziegelwanger, H., Majdak, P., Kreuzer, W. (2015b). 'Numerical calculation of listener-specific head-related transfer functions and sound localization: Microphone model and mesh discretization'. *The Journal of the Acoustical Society of America*, 138(1), 208–222. <https://doi.org/10.1121/1.4922518>