Pedestrians' time-to-collision estimation and road crossing judgments differ between electric and conventional vehicles

Daniel Oberfeld-Twistel, Thirsa Huisman, Marlene Wessels

http://dx.doi.org/10.25673/103503

Abstract

Pedestrians can only safely cross the road before an approaching vehicle if the time remaining until the vehicle arrives at their position (time-to-collision, TTC) is longer than the time needed to cross. In real traffic, the sound of a vehicle provides important information about its motion. Using a virtual reality (VR) system that combines physically plausible acoustic simulations of approaching vehicles with visual VR simulations, we investigated if the sound differences between electric (EVs) and conventional vehicles (ICEVs) result in differences in the perception and behavior of pedestrians. In this paper, we present an overview of our results. 1) When vehicles approaching with a constant velocity are presented with the same TTC, participants estimate longer TTCs for softer compared to louder vehicles both in an auditory-only and an audiovisual condition. This indicates potential risks associated with quieter vehicles. 2) When the sound of an accelerating conventional vehicle is presented, this largely removes the inadequate consideration of acceleration (first-order estimation pattern, resulting in overestimated TTCs) observed in visual-only TTC estimation. 3) For accelerating EVs with and without AVAS, the benefit provided by the car sound is significantly reduced compared to ICEVs. 4) Compatible with these TTC estimation results, the collision probability in road-crossing decisions when interacting with accelerating vehicles increases significantly with the acceleration level for EVs with and without AVAS, but remains low for ICEVs. Taken together, auditory information is relevant for pedestrians, particularly so when the approaching vehicle accelerates. Our data indicate potential risks associated with EVs, and raise interesting questions concerning the design of acoustic vehicle alerting systems.

1. Introduction

Auditory perception and cognition are highly important for safe mobility, which requires the ability to avoid potentially dangerous collisions with objects in the environment. A pedestrian crossing a street must avoid being hit by an approaching vehicle. In such a situation, our sense of hearing provides important information. For example, we can auditorily detect a vehicle approaching us from outside our field of view. The acoustic detection of vehicles was investigated in a number of studies [e.g., 1,2,3] and some of these data are the basis for the current legal requirements for auditory vehicle alerting system (AVAS) technologies [4,5]. However, other aspects of auditory perception and cognition related to safe mobility are less well understood, because previous research in these areas focused on visual perception or presented auditory stimuli that were impoverished compared to the rich and dynamic sound field generated by an approaching vehicle in a real traffic scenario. To cross the road safely in front of an approaching vehicle, the time remaining until the vehicle arrives at the pedestrian's position (*time-to-collision, TTC*) must be longer than the time required to cross. Thus, pedestrians need to estimate the TTC as accurately as possible to adjust their crossing behavior. In real traffic, the sound of a vehicle provides important information about its motion. Here, we present key results from a recent series of experiments on *TTC estimation* and *street-crossing decisions* based on only auditory (A-only), only visual (V-only), or combined auditory and visual information (AV). The experimental conditions included constant-speed and accelerating approaches, and we studied pedestrians' perception and behavior in interaction with internal-combustion engine vehicles (ICEVs) and electric vehicles (EVs). The experiments were conducted using a novel simulation system that we describe next.

2. Interactive audio-visual virtual-reality simulation of approaching vehicles

When a vehicle approaches a pedestrian standing at the curb, a) the acoustic intensity increases dynamically as the car comes closer (due to spherical spreading and air absorption), b) the azimuthal position of the vehicle varies (because the vehicle is not on a direct collision course with the pedestrian), resulting in dynamic changes in interaural time and level differences (ITD and ILD), c) the auditory source width increases because from the pedestrian's perspective, the angle between, e.g., the left and right front tires is larger when the car is closer to their position, and d) the sound spectrum changes due to air absorption, dynamic comb-filter effects resulting from interference between direct and reflected sound, and Doppler frequency shifts (although the latter are generally small unless the vehicle is already rather close to the pedestrian). All of these dynamic acoustic changes potentially provide cues to the arrival time of the vehicle [6]. In addition to these motion-related effects, the vehicle sound varies depending on travel speed, engine type, rotational engine speed, engine load, etc. [7]. However, most previous studies on auditory or audiovisual TTC estimation and street-crossing decisions did not present realistic acoustic stimuli providing the full range of these auditory cues. To overcome these limitations, we designed and implemented a novel interactive audio-visual simulation system, described in detail in [8]. Because realistic simulation of tire, powertrain and aerodynamic noise in dynamic driving situations with changing speed, acceleration, and engine load is a formidable challenge, we opted for a source-based approach. The acoustic source signals are recordings made with microphones attached to the chassis of real vehicles (conventional and electric) while these were driving at defined constant speeds or at defined positive acceleration rates on a dry asphalt surface. The vehicles were two small passenger car models of the manufacturer Kia Motors. The ICEV was a gasoline-powered Kia Rio 1.0 T-GDI (2019, 1.0 l, 88 kW, 3 cylinders) with manual transmission. The tires on the ICEV were Continental summer tires (ContiSportContact 5, 205/45 R17). The EV was a Kia e-Niro (2019, 150 kW) with Michelin summer tires (Primacy 3, 215/55 R17). The EV was equipped with an AVAS that could be active at speeds between 0.5 km/h and 28 km/h, but could also be deactivated. The sound generated by the AVAS was compatible with the requirements described in UNECE R138 [4]. We made recordings of the EV with both active and inactive AVAS. During the acoustic recordings, the trajectory of the vehicle was measured with highly precise GPS position tracking, so that at each time point in the audio signals, the position, speed, and acceleration of the vehicle is known. In the experiments, the motion of the sound sources in space is simulated using the acoustic VR simulation software TASCAR [9], which provides a physically plausible interactive simulation of the dynamic spatial sound field, with dynamic processing of the geometry of the acoustic scene and acoustic modeling of the sound transmission from the sources to the receiver, and renders the scene using sound field synthesis. This simulation approach creates realistic vehicle sounds and provides all relevant monaural and binaural distance and motion cues such as such as dynamic changes in intensity, ITD and ILD, and frequency spectrum. In our current implementation, the simulated scenes are rendered on 40 Genelec 8020DPM loudspeakers plus Genelec 7360 APM subwoofer, arranged in an upper and a lower ring in a large acoustically treated space (see Figure 1, left). The direct sound of the vehicle is rendered on 32 loudspeakers positioned at ear-height and the subwoofer via 2D Higher-Order Ambisonics (15^{th} order) [10,11]. The reflected sound is rendered on the complete array, using 3D VBAP [12]. The auditory VR simulations can be combined with three-dimensional visual VR simulations presented stereoscopically on a head-mounted display with head-tracking (see Figure 1, right). The system provides interactive simulations because listeners can actively explore the simulated auditory and visual scene with head movements. Vehicles can be presented at arbitrary approach angles and distances, making it possible to present, e.g., exactly the same vehicle sound at different TTCs. Thus, the system can be used to conduct highly controlled VR experiments with a higher degree of realism compared to previous studies in this area, without challenging the participants' safety.

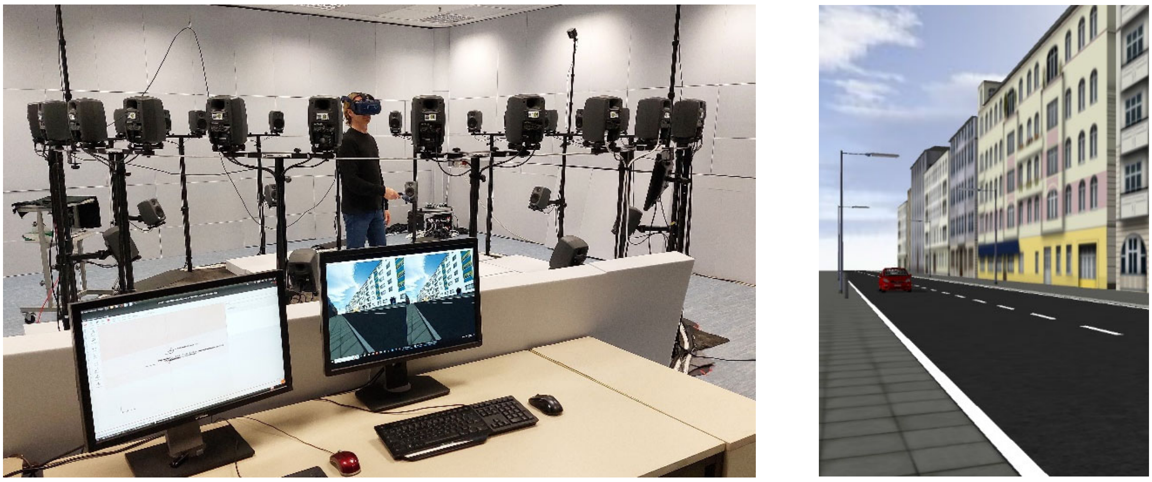

Figure 1: *First author wearing the head-mounted display in the loudspeaker array. The upper ring contains 32 loudspeakers positioned at approximately ear height. The lower ring contains 8 loudspeakers angled up towards the participant. The subwoofer is not visible in the picture. A screenshot of the visual scene is shown in the right panel.*

3. The effect of intensity on pedestrians' TTC estimation and road crossing decisions

Previous research from our lab showed an "*intensity-arrival effect*" [13,14]. At identical actual TTC, participants judged *softer* approaching sound sources to arrive *later* than louder sound sources. The intensity-arrival effect might indicate increased risks posed by quiet vehicles like electric cars: pedestrians might overestimate the TTC of a quiet electric car relative to a louder conventional vehicle with the same actual TTC, which in turn could result in risky road crossing decisions for an EV. However, the two previous experiments presented simple and somewhat artificial stimuli, which were impoverished compared to a real approaching vehicle that generates a dynamic spatial sound field, including sound from different sources (tire noise, powertrain noise, aerodynamic noise at higher speeds), as well as sound reflections from the ground surface, with variations in the vehicle sound depending on travel speed, rotational engine speed, or engine load. In **Exp. 1,** published in [8], we therefore investigated how vehicle loudness affects TTC estimation when highly realistic auditory or audiovisual simulations of the approaching vehicle are provided, using the simulation system described above. An internal-combustion engine vehicle (ICEV) and a loudness-matched electric vehicle (EV; AVAS not active) were approaching the participant standing at the curb of a simulated two-lane road in an urban setting, at different constant speeds (10, 30, 50 km/h). An auditory-only and an audio-visual condition were presented. The vehicle loudness levels were varied by 10 dB, independent of the other factors. In our experiments, we use a *prediction-motion task* [15], which is one of the best established procedures for research on TTC estimation. The simulated car approaches the participant for some seconds, and is then "occluded", i.e., it is no longer audible and disappears from the visual display. Participants press a response button to indicate the point in time at which the approaching vehicle would arrive at their position, had it continued to approach them with the same constant speed after it disappeared. The estimated TTC is defined as the time between the disappearance of the vehicle and the participant's button press. The TTC at "occlusion" was varied between 2.0 s and 5.0 s to present a range of TTCs relevant in daily street-crossing situations. Consistent with an intensity-arrival effect, participants estimated significantly *longer* TTCs when the cars were presented at the *lower* loudness level. This effect, while considerably stronger in the audio-only condition, persisted in the audio-visual condition, confirming that auditory information is used in TTC estimation even when full visual information is available [13,14]. There was no significant difference between the mean estimated TTCs for the ICEV and the loudness-matched EV, indicating that the sound quality differences between the vehicle types did not have a substantial effect on TTC estimation vehicles approaching at a constant speed.

In Exp. 1, the loudness level varied from trial to trial. Although this situation corresponds to an everyday street-crossing situation where the different approaching vehicles also vary in loudness, the trial-by-trial level variation might have directed the attention to the differences in loudness and might thus have amplified the effect of loudness on TTC estimation. To investigate if not only the loudness difference between trials, but also the “absolute” loudness of an approaching vehicle affects TTC estimation, we varied the vehicle loudness level in a blockwise fashion in **Exp. 2**. In an auditory-only condition, an ICEV approached at different constant velocities (10, 30, 50 km/h). The TTC at occlusion was varied between 1.25 and 5.0 s. Two different loudness levels were generated by presenting the car either at its original sound level as recorded on the test track, or at a loudness level increased by 10 dB. In experimental blocks 1 and 2, the same loudness level was presented on each trial of a given block, followed by a third block where the loudness level varied from trial to trial. Half of the participants started with the block presenting the higher loudness level and the other half started with the lower loudness level.

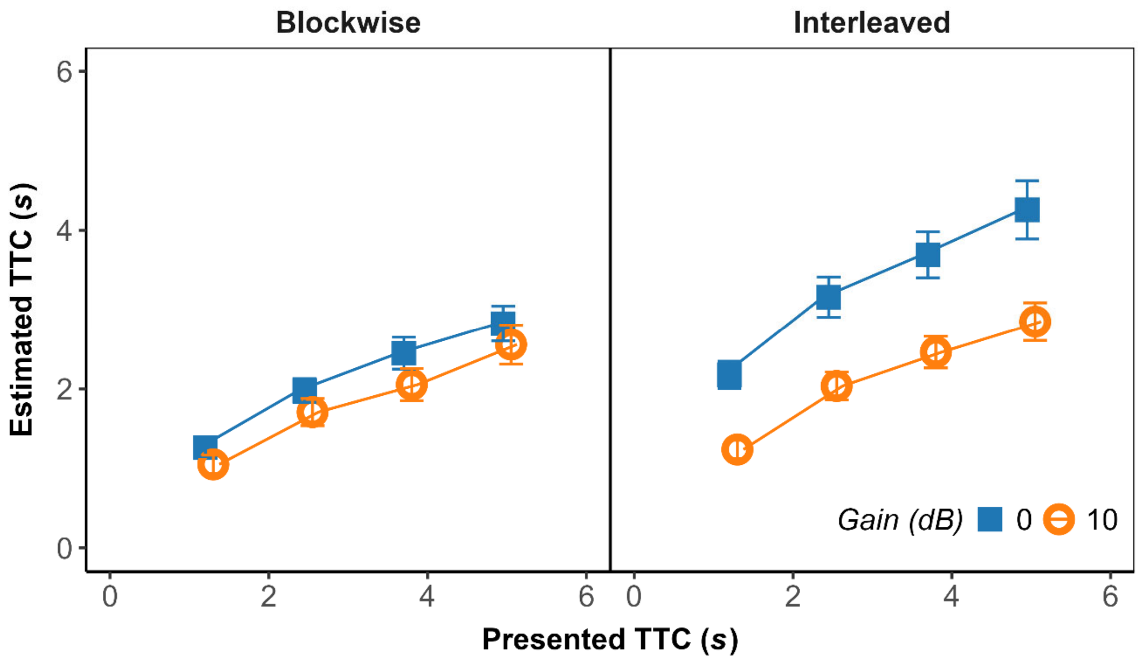

Figure 2: *Exp. 2. Mean estimated TTC as a function of the presented TTC. Blue squares: audio gain 0 dB (lower loudness level). Orange circles: gain 10 dB (higher loudness level). Left panel: only one loudness level presented per experimental block. Right panel: loudness levels randomly interleaved within the experimental block. Error bars show ± 1 standard error of the mean (SEM) across the 22 participants.*

As shown in the left panel of Fig. 2, in the blockwise condition, the estimated TTCs were significantly shorter in blocks presenting the higher loudness level than in blocks presenting the lower loudness level. Thus, the effect of vehicle loudness on TTC estimation is not limited to conditions where the sound level varies from trial to trial. However, the effect of loudness level was much stronger in the third block, where the two loudness levels were interleaved (right panel in Fig. 2).

Having established the effect of vehicle loudness on TTC estimation, **Exp. 3** investigated the effect of loudness on *road-crossing decisions*. An ICEV was presented, travelling at different constant velocities (30, 50 and 60 km/h). Two loudness levels differing by 10 dB were presented in an interleaved fashion. The experiment comprised both an auditory-only and an audio-visual condition to match Exp. 1. As in the TTC experiments, a vehicle approached for some seconds, and was then occluded. Participants were asked to indicate whether or not they would have crossed the road in front of the approaching vehicle at the moment of occlusion (positive or negative crossing decision, respectively). We measured the probability of a positive decision ("gap acceptance") across a range of TTCs at occlusion. If lower loudness results in longer estimated TTCs, as demonstrated by the above experiments, then at a given TTC at occlusion participants will think that they have more time available to cross the street in front of the vehicle when it's sound is softer and will thus make a positive crossing decision in a higher proportion of trials than for a louder vehicle. Compatible with this result, Fig. 3 shows that the probability (p_{coll}) that a positive crossing decision would have resulted in a collision' with the approaching vehicle because the TTC at occlusion was shorter than the time needed to cross the road was significantly higher for softer than for louder vehicles, both in the A-only and the AV condition. In the auditory-only condition, the effect of loudness was rather extreme, with p_{coll} close to zero in interaction with the louder vehicles but high collision probabilities for the quieter vehicles. The results of Exp. 3 confirm our hypothesis that quieter vehicles might cause riskier crossing decisions, even when full visual information is available.

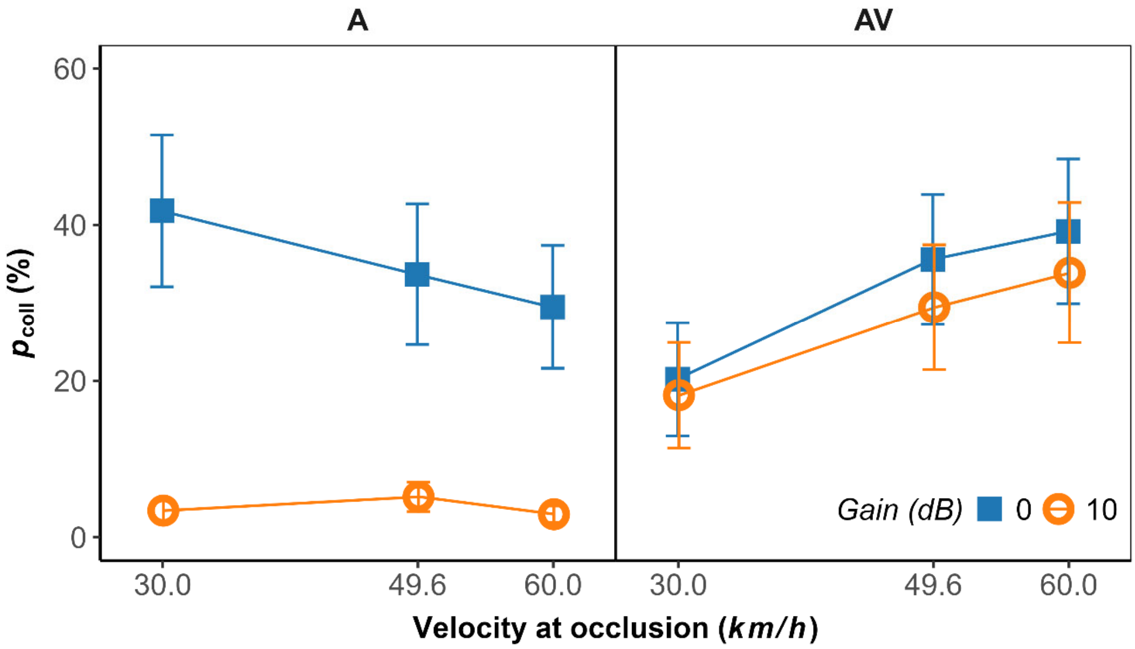

Figure 3: *Exp. 3. Mean collision probability (p_{coll}) as a function of the velocity at occlusion. Blue squares: audio gain 0 dB (lower loudness level). Orange circles: gain 10 dB (higher loudness level). Left panel: auditory-only condition. Right panel: audiovisual condition. Error bars show ± 1 SEM across the 13 participants.*

4. Auditory information improves TTC estimation and street-crossing decisions for accelerating vehicles

Vehicles often accelerate while they are approaching a pedestrian (e.g., when a vehicle pulls out of a parking space and heads toward the exit of the parking lot). The literature on *visual* TTC estimation consistently shows that humans have difficulty to account for the acceleration of an object [e.g., 16,17]. Instead, they estimate the TTC of an accelerating object as if it was moving at constant velocity. For positive acceleration rates, this so-called *first-order TTC estimation* results in an overestimated TTC, because the increase in velocity between the moment of estimation and the arrival of the object is ignored. However, when a vehicle accelerates, the resulting dynamic changes in the powertrain noise provide salient acoustic cues for acceleration. Can pedestrians use this auditory information to account for the acceleration?

In **Exp. 4**, published in [18], we compared TTC estimations for an ICEV approaching at either a constant speed (a = 0) or accelerating during the approach (a = 2 m/s^2) between a visual-only and an audio-visual condition. We used the same prediction-motion task as in the experiments described above. Participants pressed a response button to indicate the point in time at which the approaching vehicle would arrive at their position, had it continued to approach them with the same acceleration after it disappeared. In the visual-only condition, the TTC estimations showed a clear first-order pattern: with increasing presented TTC, participants increasingly overestimated the TTC, compatible with the literature on visual TTC estimation. However, if the sound of the accelerating ICEV was presented in addition to the visual information, this largely removed the first-order pattern, so that on average the estimated TTC was close to the veridical value. This result was compatible with our expectation that the salient acoustic signature of the ICEV sound during states of acceleration should help pedestrians to factor the acceleration into their TTC estimations.

Does this benefit provided by the vehicle sound also apply to electric vehicles? In **Exp. 5**, published in [19], we obtained TTC estimations for an accelerating ICEV and for an accelerating EV with or without activated AVAS, with acceleration rates between 0.4 and 2.6 m/s^2. At a given simulated TTC at occlusion, the mean estimated TTC increased significantly with the acceleration rate for the EV without AVAS, thus exhibiting a first-order pattern and indicating insufficient consideration of the acceleration. The increase in estimated TTC with the acceleration rate was still significant when the AVAS was activated on the EV, but was somewhat reduced compared to the condition without AVAS. In contrast, for the ICEV, the estimated TTC showed no significant effect of the acceleration rate, indicating that as in Exp. 4, participants were able to use the information about acceleration communicated by the vehicle sound.

In Exp. 5, the acceleration rates and speeds at occlusion of the accelerating electric and conventional vehicles were not identical because they exactly corresponded to the vehicles' motion during the recordings made for our simulation system on a test track. These manual drives showed deviations from the intended velocity profiles in conditions with acceleration, particularly so for the ICEV with manual transmission. In **Exp. 6**, we therefore presented the recorded sounds of the ICEV and of the EV with and without AVAS, but the motion of the sound source simulated in the virtual scene corresponded exactly to an initial phase of 2.0 s with a constant speed of 10 km/h, followed by an acceleration phase of 3.0 s with exactly a = 2.0 m/s^2. Thus, the motion was *identical* for all vehicle types. Also, we included a constant-speed approach matched to the speed at occlusion of the accelerating vehicle (v_{occ} = 31.6 km/h; note that at this speed the AVAS was not activated), and additionally presented a visual-only condition. Figure 4 plots the mean estimated TTC as a function of the distance between vehicle and participant at occlusion (D_{occ}). As a reference, the actual TTC is shown by the dotted lines. If participants use a first-order estimation strategy in the sense of $TTC_{est} = D_{occ}/v_{occ}$ for accelerating vehicles, then the function relating estimated TTC and D_{occ} would be identical in the constant speed (blue lines) and the acceleration condition (orange lines). The results show that in the visual-only condition (right panel) and for the EV without activated AVAS (muted loudspeaker symbols in the middle panel), this was indeed the case, compatible with a first-order TTC estimation strategy. In contrast, for the ICEV (left panel), the functions relating estimated TTC and distance at occlusion differed between a = 0 and a = 2.0 m/s^2, and on average the estimated TTC for the accelerating approaches was closer to the veridical value than for the EVs. For the accelerating EV with activated AVAS, the estimated TTCs lay again in between the pattern for the EV without AVAS and the ICEV, as in Exp. 5. The results thus confirm that the sound of an ICEV helps participants better to account for acceleration than the sound of an EV.

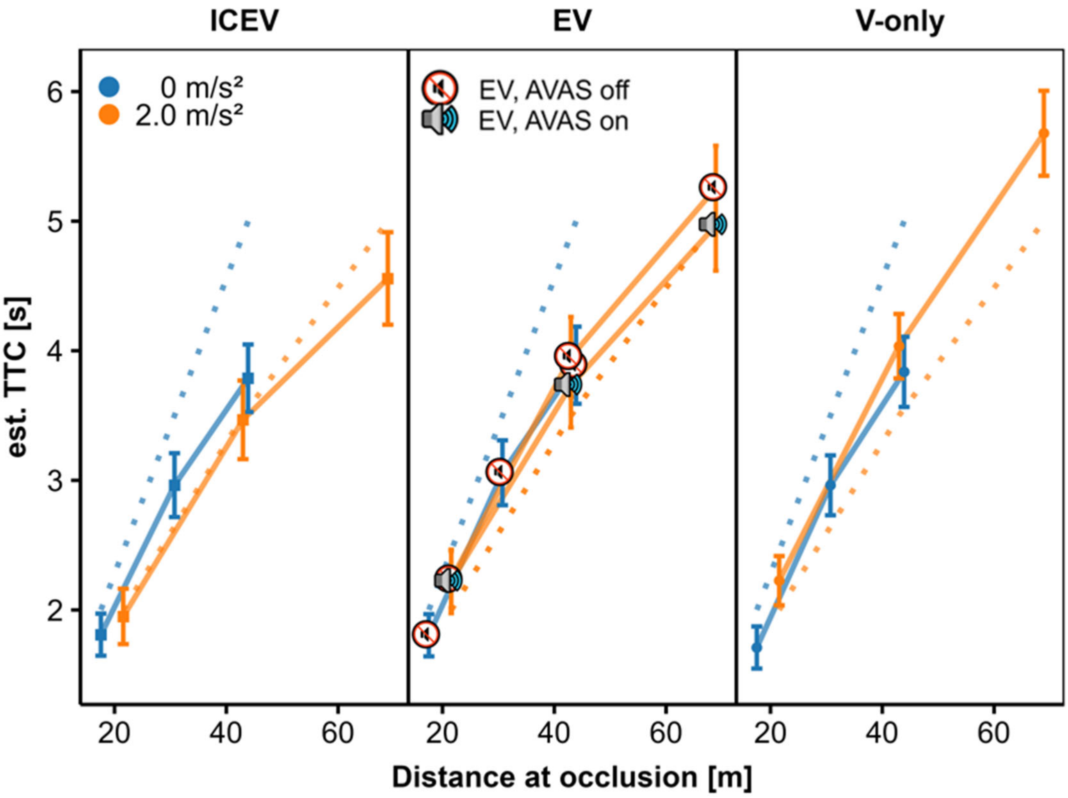

Figure 4: *Exp. 6. Mean estimated TTC (solid lines) as a function of the distance between vehicle and observer at occlusion. Blue symbols: constant-speed approaches. Orange symbols: accelerating approaches (a = 2 m/s^2). Dotted lines show the actual TTC in the constant-speed (blue) and acceleration condition (orange). Left panel: audiovisually presented ICEV. Middle panel: audiovisually presented EV. Loudspeaker symbols indicate whether the AVAS was activated or not. Right panel: visual-only presentation. Error bars show ± 1 SEM across the 15 participants.*

Although the simulated motion was identical for the three vehicle types in Exp. 6, the recorded sound of the ICEV contained an audible (manual) gear shift, while the sound of the EVs did not. Due to the selected presentation duration, the sound of the ICEV ended shortly after the gear shift in Exp. 6, so that the final 500 ms of the sound corresponded to a phase where the ICEV increased its acceleration rate again from less than 2 m/s^2 to over 2 m/s². In contrast, the acceleration rate of the EVs was nearly constant across the entire presented acceleration phase. Did the higher acceleration rate signaled during the final part of the ICEV sound contribute to the better consideration of acceleration during TTC estimation for this vehicle type? To answer this question, **Exp. 7** used the same approach as Exp. 6, i.e., presenting recordings of the three vehicle types but simulating identical motion (i.e., acceleration rate and speed at occlusion). However, in this experiment we extended the presentation duration for all vehicle types, so that after the gear shift, the ICEV was presented for one additional second during which the acceleration rate remained close to 2.0 m/s^2, just as for the EVs. Put differently, we made sure that the final 1 s of the presented sound corresponded to an acceleration rate of ~ 2.0 m/s^2 for all vehicle types. The left panel of Figure 5 shows that in the V-only condition (orange symbols), the mean estimated TTCs were close to the gray dashed line representing first-order TTC estimation. In contrast, the mean estimated TTCs for the ICEV (green symbols) were close to the black solid line representing veridical TTC estimation. The mean estimated TTCs for the EVs (blue lines) showed a first-order pattern because the amount of overestimation increased with the presented TTC, but less so than in the V-only condition. With activated AVAS (active loudspeaker symbols), the mean estimations were closer to the veridical value than without AVAS (muted loudspeaker symbols), but were still less accurate than for the ICEV. The results are similar to the patterns observed in our previous experiments and thus show that the differences in the estimated TTCs for an accelerating ICEV compared to accelerating EVs were not due to the fact that the ICEV sound signaled a higher acceleration rate at the end of the presentation than the EV sounds.

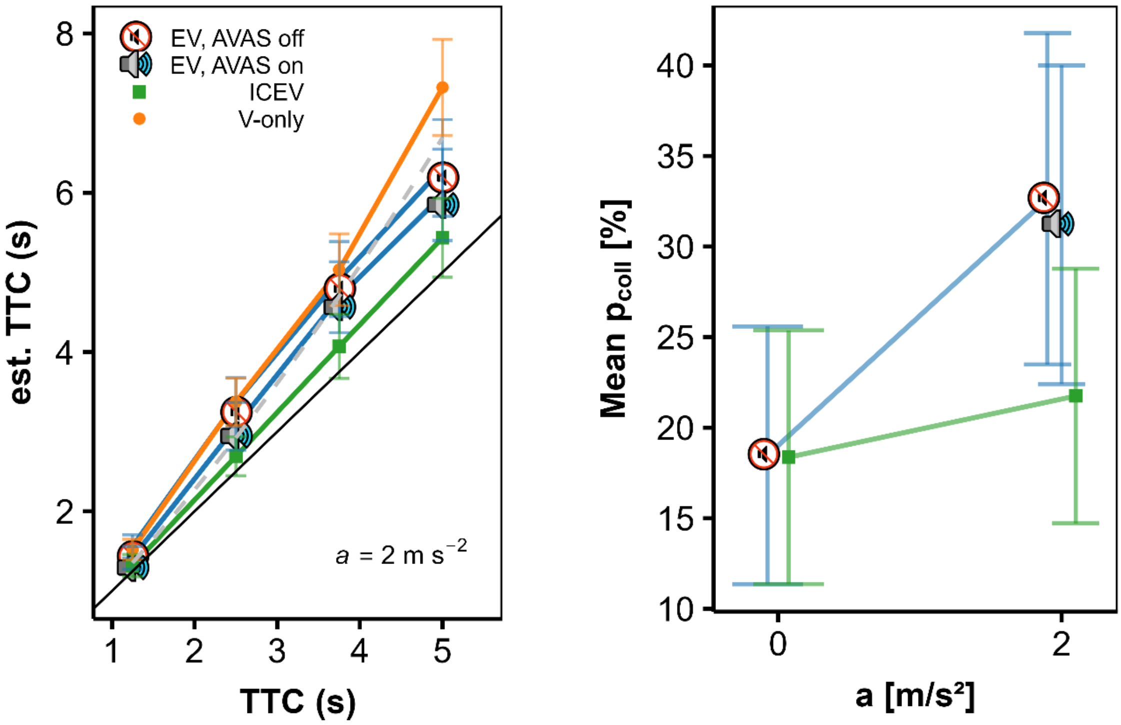

***Figure 5**: Left panel: Exp. 7. Mean estimated TTC as a function of presented TTC and vehicle type. The dashed gray line corresponds to first-order estimation. The solid block line represents the actual TTC. Orange symbols: visual-only condition. Green symbols: audiovisually presented ICEV. Blue symbols: audiovisually presented EV. . Loudspeaker symbols indicate whether the AVAS was activated or not. Right panel: Exp. 9s. Mean collision probability p_{coll} as a function of acceleration and vehicle type. Same color code as in left panel. Colors and symbols indicate vehicle type and error bars show ± 1 SEM across the 24 (Exp. 6) and 15 (Exp. 7) participants, respectively.*

In the final three experiments of this series, we measured street-crossing decisions in interaction with accelerating vehicles. We expected the decisions to reflect the pattern observed for TTC estimation. For instance, because Exp. 4-7 showed that participants *overestimate* the TTC for accelerating EVs, we expected *riskier* street-crossing decisions for EVs than when the sound of an accelerating ICEV is available.

In **Exp. 8**, published in [20], audiovisual simulations of an ICEV and an EV with or without activated AVAS were presented. We presented acceleration rates between 0.4 and 2.6 m/s^2 as well as constant-speed approaches. For the ICEV, the probability that a positive crossing decision would have resulted in a collision with the approaching vehicle because the TTC at occlusion was shorter than the time needed to cross the road did not increase with the acceleration rate but remained at a relatively low value, similar to the average p_{coll} for the constant-speed approaches. In interaction with the EV, however, p_{coll} was on average higher than in interaction with the ICEV and increased significantly with the acceleration rate. With activated AVAS, the mean p_{coll} was slightly lower than without AVAS, but the increase of p_{coll} with the acceleration rate was observed for both EV variants. This pattern is compatible with the TTC estimation results described above.

Because the simulations in Exp. 8 presented the actual vehicle trajectories driven on the test track, so that the acceleration rates differed somewhat between the ICEV and the EVs (as in Exp. 5), in **Exp. 9**, we used the same experimental design as in Exp. 6 and compared approaches with a = 2.0 m/s^2 to constant-speed approaches with matched v_{occ} = 31.6 km/h, simulating again exactly the same motion for the three vehicle types in an audiovisual condition. The right panel of Figure 5 shows the average collision probability p_{coll}, computed as in Exp. 3. As expected based on the TTC estimation data, p_{coll} was similar between the constant-speed and the acceleration condition for the ICEV (green line in the right panel of Fig 5). In contrast, for the EV with and without activated AVAS (loudspeaker symbols), p_{coll} was significantly higher when the vehicle accelerated, compared to the constant-speed condition.

In **Exp. 10**, we investigated street-crossing decisions in interaction with an ICEV that either approached at a constant speed or accelerated with a = 2 m/s^2. We compared an auditory-only and an audiovisual condition, and additionally investigated how vehicle loudness affects crossing decisions when the vehicle accelerates, with the same 10-dB gain variation as in Exp. 3. The presentation duration was relatively long, as in Exp. 7, to ensure that the final 1 s of the sound of the accelerating vehicle corresponded to a constant acceleration of a ~ 2 m/s^2.

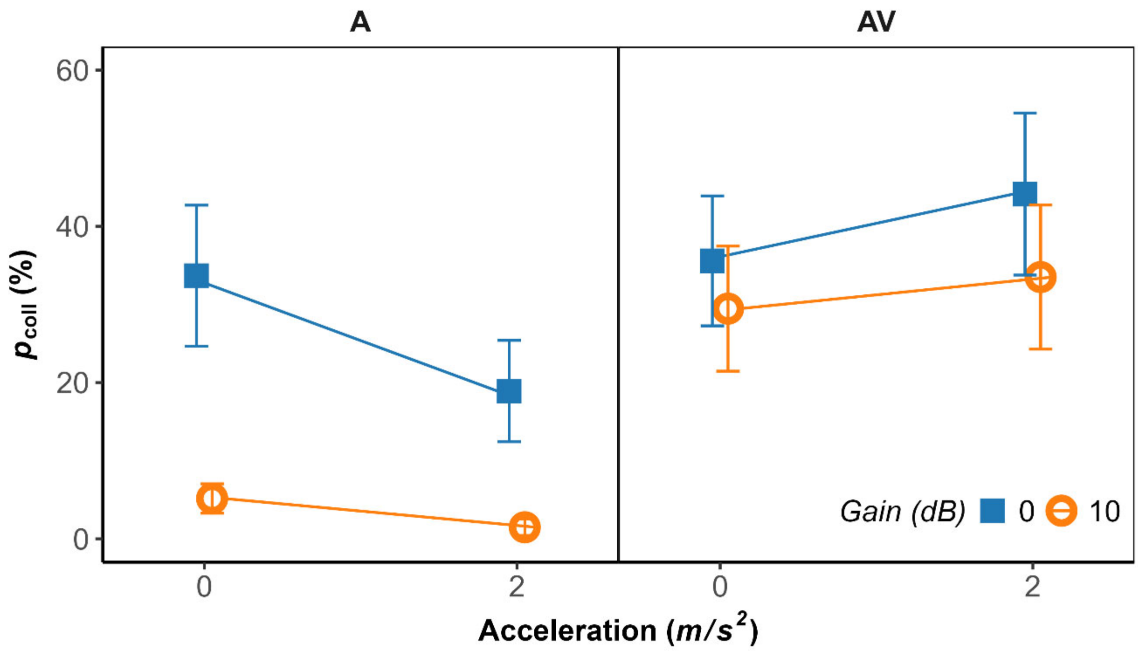

Figure 6: *Exp. 10. Mean collision probability (p_{coll}) as a function of the acceleration rate. Blue squares: audio gain 0 dB (lower loudness level). Orange circles: gain 10 dB (higher loudness level). Left panel: auditory-only condition. Right panel: audiovisual condition. Error bars show ± 1 SEM across the 13 participants.*

As shown in Fig 6, we again observed a strong effect of vehicle loudness, particularly so in the A-only condition (left panel). The average collision probability was significantly higher in the quieter (blue symbols) than in the louder condition (orange symbols). Somewhat surprising, in the A-only condition, p_{coll} was even lower in interaction with the accelerating vehicle than for the vehicle approaching at constant speed. This observation might indicate a particularly strong consideration of vehicle loudness when the gain is varied from trial to trial (the sound is louder when a vehicle accelerates), but additional data are necessary to confirm this hypothesis. In the AV condition, there was a small, non-significant increase in the average p_{coll} at a = 2 m/s^2 compared to a = 0. Taken together, the results confirm that the collision risk does not increase substantially due to acceleration when the sound of an ICEV is available. Also surprising, p_{coll} was on average higher in the audiovisual than in the A-only condition. This finding might be related to a more cautious decision strategy adopted when the cars were only audible, but not visible.

Taken together, the results of this series of experiments clearly show that the vehicle sound is important for pedestrians to account for the acceleration of an approaching vehicle, evident in TTC estimations and street-crossing decisions. However, this audiovisual benefit is reduced for EVs compared to ICEVs, even with activated AVAS.

5. Discussion and summary

The work program described briefly in this paper highlights the importance of acoustic information and auditory perception for the safe mobility of pedestrians and other road-users in interaction with motorized vehicles. Using a system providing more realistic auditory simulations of approaching vehicles than in previous studies, we studied time-to-collision estimation and road-crossing decisions. Our data clearly indicate that in traffic scenarios, auditory perception it not only relevant for the acoustic detection of vehicles, as it is already widely accepted, but that auditory information is also highly important for TTC estimation and road crossing decisions. The first series of experiments showed that at identical actual TTC, participants judged *quieter* vehicles to arrive *later* than louder vehicles, and that riskier road crossing decisions were made for quieter compared to louder vehicles, even when full visual information about the motion of the approaching car was available. The second series of experiments consistently indicated that the vehicle sound provides important information about acceleration that is not available in the visual domain. Only when the sound of an ICEV saliently signaled that the vehicle was positively accelerating as it approached them were participants able to make relatively accurate TTC estimations and safe road crossing decisions. However, for electric vehicles, this benefit provided by the car sound was significantly reduced compared to ICEVs, even when an AVAS compatible with UNECE R138 was active. Taken together, in street-crossing situations, pedestrians use auditory information for judging the motion of an approaching vehicle, particularly during acceleration, and even when the vehicle is in full view.

Our simulation system enables us to study auditory and audiovisual perception in street-crossing scenarios in controlled VR experiments, but with a considerably higher degree of realism than in previous studies in this area, and based on acoustic modeling of moving sound sources. By using recordings of real vehicles, we maximize the realism of the vehicle sounds. However, a limitation of this source-based approach is that in our experiments, we are can only use the vehicle recordings available in our database. We're hoping that realistic simulations of the exterior sound produced by ICEVs and EVs in dynamic driving situations (with changes in speed, acceleration, and engine load) will become available in the future, so that we can extend our studies to other car types, AVAS designs, acceleration levels and speeds, etc.

Our results raise a number of theoretical and practical questions, and our aim is to answer at least some of them during the continuation of our work program.

1) Even for constant-speed approaches, it is not yet clear which of the potential auditory cues discussed in the Introduction are used during TTC estimation and street-crossing decisions, how these cues are weighted relative to each other, and how they are combined with visual cues in an audiovisual condition. Measuring the importance of different potential cues requires to activate or deactivate some of the cues, or to shift cues against each other [e.g., 13]. Such experimental manipulations are only possible in virtual environments. Even in a simulated environment, it is challenging to, e.g., decorrelate the time-to-collision signaled by the dynamic increase in acoustic intensity from the TTC signaled by the dynamic change in source width, because both are by default linked to the simulated distance. We were already able to implement a subset of the required conditions in our simulation system and will conduct a series of experiments using this technology.

2) It remains to be investigated which auditory cues participants use for detecting the acceleration of a vehicle or for judging its acceleration rate. Which role do the potential different (psycho-) acoustic cues (e.g., intensity/loudness, frequency spectrum/pitch, modulation spectrum/roughness) play in this context? Identifying the relevant cues and their relative weights could be helpful for designing new AVAS concepts that are better suited for communicating an EVs state of acceleration to other road users.

3) Related to the preceding aspect, why did the sound of the EV (even with activated AVAS) presented in our experiments fail to enable our participants to make as accurate TTC estimations and safe crossing decisions as the ICEV sound did? One factor contributing to this finding could be that participants failed to detect that the EVs were accelerating, at least on a subset of trials. In fact, Exp. 3 in [20] showed that the probability of detecting that a vehicle accelerates based on only visual information was between 70 and 80% at an acceleration rate of 2 m/s^2, which is above the guessing rate but still far from perfect. If, however, the vehicle sound was presented in addition to the visual information, the detection probability increased to more than 90% for an ICEV, to slightly less than 90% for our EV with activated AVAS, but to only about 80% for the EV without AVAS. However, it is likely that not only a binary classification into constant speed versus accelerating is needed for accurate TTC judgments and crossing decisions, but that a more quantitative judgment of the acceleration rate provides an additional benefit. Additional experiments are required to test this hypothesis.

4) As a complement to the previous aspect, the exact mechanisms underlying the benefit provided by the sound of the accelerating ICEV also remain to be identified, as discussed in [18]. For instance, does the acoustic acceleration signal direct the attention to acceleration-related motion cues, or does it trigger a “correction” of an initial first-order TTC estimation?

5) How could AVAS designs be improved so that they communicate acceleration as effectively as an ICEV? A clear limitation of our experiments is that we studied only one AVAS variant so far, even though the increase in pitch caused by increases in travel speed was already rather salient in the car model we studied compared to the minimum requirements described in UNECE R138. Could the auditory perception of acceleration be improved by making the speed-related change in pitch – that indirectly signals acceleration - even more salient? Would it help to add more direct acceleration-related sound quality changes? Could it be beneficial to have the AVAS activated up

to higher speeds than it is currently required? We will conduct experiments to gain insight into these questions. In our simulation system, new AVAS sounds can be added to the recorded sound of the EV without AVAS simply by defining the AVAS loudspeakers as additional sound sources in the acoustic simulation. Ultimately, this design problem taps into the conflicting aims of reducing traffic noise while still maintaining pedestrian safety.

Acknowledgements

This work was supported by funding from Unfallforschung der Versicherer (udv.de) and Deutsche Forschungsgemeinschaft (www.dfg.de; OB 346/8-1; priority program AUDICTIVE - SPP2236: *Auditory Cognition in Interactive Virtual Environments*), granted to Daniel Oberfeld-Twistel.

References

[1] Altinsoy E, Landgraf J, Rosenkranz R, Lachmann M, Hagen K, et al. (2015) Wahrnehmung und Bewertung von Fahrzeugaußengeräuschen durch Fußgänger in verschiedenen Verkehrssituationen und unterschiedlichen Betriebszuständen. *FAT Schriftenreihe: Berichte der Bundesanstalt für Straßenwesen*.

[2] Poveda-Martinez P, Peral-Orts R, Campillo-Davo N, Nescolarde-Selva J, Lloret-Climent M, et al. (2017) Study of the effectiveness of electric vehicle warning sounds depending on the urban environment. *Applied Acoustics* 116: 317-328. doi:10.1016/j.apacoust.2016.10.003

[3] Emerson RW, Kim DS, Naghshineh K, Pliskow J, Myers K (2013) Detection of quiet vehicles by blind pedestrians. *Journal of Transportation Engineering* 139: 50-56. doi:10.1061/(asce)te.1943-5436.0000478

[4] UNECE R138 (2017) *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*.

[5] NHTSA 141 (2018) *Federal Motor Vehicle Safety Standard No. 141: Minimum Sound Requirements for Hybrid and Electric Vehicles*.

[6] Jenison RL (1997) On acoustic information for motion. *Ecological Psychology* 9: 131-151. doi:10.1207/s15326969eco0902_2

[7] Zeller P, editor (2018) Handbuch Fahrzeugakustik: Springer.

[8] Oberfeld D, Wessels M, Büttner D (2022) Overestimated time-to-collision for quiet vehicles: Evidence from a study using a novel audiovisual virtual-reality system for traffic scenarios. *Accident Analysis & Prevention* 175: 106778. doi:10.1016/j.aap.2022.106778

[9] Grimm G, Luberadzka J, Hohmann V (2019) A toolbox for rendering virtual acoustic environments in the context of audiology. *Acta Acustica United with Acustica* 105: 566-578. doi:10.3813/aaa.919337

[10] Daniel J (2000) *Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia*: Université Pierre et Marie Curie (Paris VI).

[11] Gerzon MA (1985) Ambisonics in multichannel broadcasting and video. *Journal of the Audio Engineering Society* 33: 859-871
[12] Pulkki V (1997) Virtual sound source positioning using vector base amplitude panning. *Journal of the Audio Engineering Society* 45: 456-466
[13] DeLucia PR, Preddy D, Oberfeld D (2016) Audiovisual integration of time-to-contact information for approaching objects. *Multisensory Research* 29: 365–395. doi:10.1163/22134808-00002520
[14] Keshavarz B, Campos JL, DeLucia PR, Oberfeld D (2017) Estimating the relative weights of visual and auditory tau versus heuristic-based cues for time-to-contact judgments in realistic, familiar scenes by older and younger adults. *Attention Perception & Psychophysics* 79: 929-944. doi:10.3758/s13414-016-1270-9
[15] Schiff W, Detwiler ML (1979) Information used in judging impending collision. *Perception* 8: 647-658. doi:10.1068/p080647
[16] Lee DN, Young DS, Reddish PE, Lough S, Clayton TMH (1983) Visual timing in hitting an accelerating ball. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology* 35: 333-346. doi:10.1080/14640748308402138
[17] Benguigui N, Ripoll H, Broderick MP (2003) Time-to-contact estimation of accelerated stimuli is based on first-order information. *Journal of Experimental Psychology: Human Perception and Performance* 29: 1083-1101. doi:10.1037/0096-1523.29.6.1083
[18] Wessels M, Zähme C, Oberfeld D (2022) Auditory information improves time-to-collision estimation for accelerating vehicles. *Current Psychology*. doi:10.1007/s12144-022-03375-6
[19] Wessels M, Kröling S, Oberfeld D (2022) Audiovisual time-to-collision estimation for accelerating vehicles: The acoustic signature of electric vehicles impairs pedestrians' judgments. *Transportation Research Part F: Traffic Psychology and Behaviour* 91: 191-212. doi:10.1016/j.trf.2022.09.023
[20] Oberfeld D, Wessels M, Kröling S (2021) Risiko hohe Beschleunigung? Straßenquerungsverhalten von Fußgänger:innen in Interaktion mit E-Fahrzeugen (mit und ohne AVAS) im Vergleich zu Verbrennern. Berlin: Gesamtverband der Deutschen Versicherungswirtschaft e.V. https://www.udv.de/resource/blob/84078/22741be085f0aa88579fe0c7362867c6/76-risiko-hohe-beschleunigung-d-data.pdf