



Auditory Information Improves Time-to-collision Estimation for Accelerating Vehicles

Marlene Wessels¹ · Carolina Zähme¹ · Daniel Oberfeld¹

Accepted: 18 June 2022
© The Author(s) 2022

Abstract

To cross a road safely, pedestrians estimate the time remaining until an approaching vehicle arrives at their location (time-to-collision, TTC). For visually presented accelerated objects, however, TTC estimates are known to show a first-order pattern indicating that acceleration is not adequately considered. We investigated whether added vehicle sound can reduce these estimation errors. Twenty-five participants estimated the TTC of vehicles approaching with constant velocity or accelerating, from a pedestrian's perspective at the curb in a traffic simulation. For visually-only presented accelerating vehicles, the TTC estimates showed the expected first-order pattern and thus large estimation errors. With added vehicle sound, the first-order pattern was largely removed, and TTC estimates were significantly more accurate compared to the visual-only presentation. For constant velocities, TTC estimates in both presentation conditions were predominantly accurate. Taken together, the sound of an accelerating vehicle can compensate for erroneous visual TTC estimates presumably by promoting the consideration of acceleration.

Keywords Acceleration · Audiovisual perception · Time-to-collision · Pedestrian safety

Introduction

Successful mobility requires, among others, to avoid collisions with objects. Pedestrians estimate the time approaching cars need to arrive at their position (time-to-collision, TTC)¹ to adjust their behavior and avoid collisions (Butler et al., 2016; Dommès et al., 2013; Petzoldt, 2014). What role does a vehicle's sound play in addition to visual information in this context? To answer this question, it is necessary to investigate the use and weighting of visual and auditory information in TTC estimation (DeLucia et al., 2016; Keshavarz et al., 2017; Prime & Harris, 2010; Zhou et al., 2007), which has received little attention so far in research and theorizing. At the same time, this question has important practical implications for traffic safety, for instance in the context of wearing

noise-cancelling headphones as a pedestrian or when confronted with quieter electric vehicles. In this study, we shed light on TTC estimation for *accelerated* objects and demonstrate that auditory information plays an important role.

The Impact of Auditory Information on TTC Estimation

The impact of auditory information on TTC estimation has been investigated only in a small number of studies, which reported largely consistent results (DeLucia et al., 2016; Hofbauer et al., 2004; Keshavarz et al., 2017; Prime & Harris, 2010; Zhou et al., 2007). For instance, DeLucia et al. (2016) presented objects approaching the observers on a collision course and at constant velocities, in three different conditions – auditory-only, visual-only, and audiovisual. Whereas TTC estimates in the auditory-only were significantly shorter than in the visual condition, TTC estimates for visual and audiovisual objects did not differ substantially. In the audiovisual condition, participants based their estimates on both auditory and visual cues, but relied more strongly on visual information (e.g., optical size) compared to auditory

✉ Marlene Wessels
mwessels@uni-mainz.de

Carolina Zähme
czaehme@students.uni-mainz.de

Daniel Oberfeld
oberfeld@uni-mainz.de

¹ Institute of Psychology, Section Experimental Psychology, Johannes Gutenberg-Universität Mainz, Wallstrasse 3, 55122 Mainz, Germany

¹ Although the vehicle is sometimes not on a direct collision course with the virtual observer, we use the term time-to-collision for simplicity.

cues (e.g., sound pressure level). The stronger reliance on visual than auditory cues in audiovisual TTC estimation was also found for a frontally approaching vehicle in a virtual traffic scenario (Keshavarz et al., 2017). In sum, additional auditory information was not found to play a major role in TTC estimation for constant velocity object approaches.

However, these studies do not reflect the complex nature of common everyday scenarios. In urban traffic, road users are often confronted with accelerating vehicles, for example, when drivers try to catch a green traffic light. The sound of an accelerating vehicle with internal combustion engine with its dynamically increasing engine speed and load is particularly salient and clearly communicates the vehicle's accelerating state. Thus, it is conceivable that the auditory information plays a stronger role when the approaching vehicle accelerates, compared to a constant velocity approach. It is noteworthy that previous studies have consistently shown that the visual processing of acceleration is poor, (e.g., Calderone & Kaiser, 1989; Snowden & Braddick, 1991; Werkhoven et al., 1992), and that TTC estimates for accelerating objects show pronounced errors, (e.g., Benguigui et al., 2003; Benguigui & Bennett, 2010; Bennett & Benguigui, 2016; Kaiser & Hecht, 1995; Lee et al., 1983; Rosenbaum, 1975; Senot et al., 2003), as we will discuss in greater detail below. For this reason, it is conceivable that the characteristic sound profile of an accelerating vehicle provides information that is not readily available in vision, thus significantly improving TTC estimation. However, the role of auditory information during TTC estimation for accelerating vehicles has not yet been studied, but the present study aims to fill this gap.

Visual TTC Estimation for Accelerated Approaches

For objects approaching with constant velocity, at each moment in time t , the TTC between the object and the position of arrival is determined by the ratio of instantaneous distance $D(t)$ to the first derivative of distance with respect to time, i.e., its instantaneous velocity $v(t)$, $TTC(t) = \frac{D(t)}{v(t)}$. Lee and Young (1985) termed this the “tau-margin”. For accelerating objects, however, the TTC also depends on the acceleration. If the object accelerates at a constant rate a , then $TTC(t) = \frac{-v(t) + \sqrt{2a \cdot D(t) + (v(t))^2}}{a}$. Thus, both velocity and acceleration have to be considered to accurately estimate the TTC of an accelerating object. Previous studies on TTC estimation for accelerating objects have quite consistently shown that humans do not adequately consider the acceleration in their visual TTC estimates, which may be related to their relatively low visual sensitivity for detecting acceleration, (e.g., Gottsdanker et al., 1961; Snowden & Braddick, 1991; Werkhoven et al., 1992). If acceleration is not considered (or

not perceived), the TTC estimate equals the tau-margin and represents a so-called first-order estimation (Tresilian, 1995). For a positive acceleration, the first-order estimation results in overestimated TTCs (Benguigui et al., 2003; Benguigui & Bennett, 2010; Bennett & Benguigui, 2016; López-Moliner et al., 2003), because the vehicle would take longer to reach the arrival position if it were to move constantly at the instantaneous velocity obtained at the moment of estimation than if it were continuously increasing its speed as it accelerates. The first-order TTC at time t is related to the actual TTC as $TTC_1(t) = TTC(t) + \frac{a \cdot (TTC(t))^2}{2v(t)}$. Thus, the deviation of the first-order estimate from the actual TTC increases linearly with acceleration, increases as a quadratic function of actual TTC, and decreases with the instantaneous velocity at the moment of estimation t .

In traffic, if the remaining time to an approaching vehicle is estimated to be longer than it actually is, the probability of a collision increases. In fact, observers in a virtual traffic environment overestimated the TTC of an oncoming accelerating vehicle (Tharanathan & DeLucia, 2006), and initiated an evasive maneuver to avoid a collision later when the approaching vehicle accelerated than when it was traveling at a constant speed (Tharanathan, 2009). In sum, temporal estimation of accelerating objects based on visual information has proven to be quite erroneous. However, there are no results whether auditory information can improve the TTC estimates for accelerating objects and, thus, enhance traffic safety.

Theoretical models of interceptive actions, e.g., catching a ball, assume that the prediction of future object motion involves both the processing of current object information as well as an expectation of object dynamics represented in a prediction model (for review, see Zago et al., 2009). Such a predictive model is based on prior knowledge, for instance, about gravitational forces that lead to acceleration of falling objects (Zago et al., 2004), or the memory of previously seen motion (Rust & Palmer, 2021). Even though these models focus on the visual domain and interceptive tasks, it is conceivable that the familiar sound of an accelerating vehicle could trigger expectations about the future vehicle motion. Against the background of low visual acceleration sensitivity (e.g., Gottsdanker et al., 1961; Snowden & Braddick, 1991; Werkhoven et al., 1992), such expectations may not be based on purely visual information.

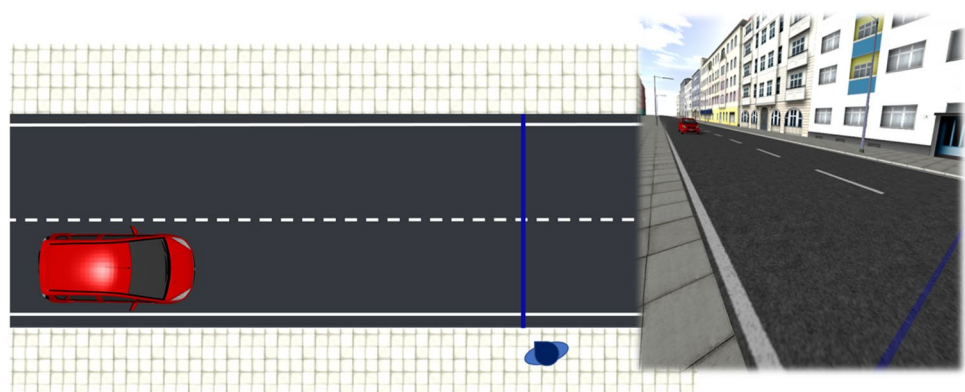
In the present experiment, we measured TTC estimations for vehicles presented only visually or audiovisually. The vehicles either accelerated or traveled at a constant speed during the approach. In line with previous research, we hypothesized no significant benefit of audiovisual over visual-only presentation for constant-velocity vehicle approaches. In contrast, we expected a significant improvement in TTC estimates for *accelerating* vehicles when the vehicle sound is presented.

Methods

Auditory Stimuli and Simulation

Participants observed an approaching car in a virtual urban scenario (Fig. 1), which was presented audiovisually or only visually. The auditory stimuli were based on acoustic recordings of a three-cylinder gasoline-powered Kia Rio 2019 with manual transmission driving on a test track. During the test drives, we recorded the sound emission using four free-field microphones mounted to the vehicle's chassis to capture the dynamic tire and engine noise. The position and velocity of the vehicles was measured with high-precision GPS tracking. In the VR simulations, we followed a source-based approach and modeled the propagation of the sound from the sources (microphone signals) to the receiver (participant) in the simulated scene, using the software TASCAR (Grimm et al., 2019). The software dynamically updates the scene geometry (i.e., the position of the vehicle sound sources relative to the receiver) according to the trajectory of the real vehicle on the test track (GPS data). Doing so, the acoustic stimuli accounted for the distance-dependency of the sound level and propagation time. They also provided the acoustic reflection and absorption characteristics of the simulated scene. The simulated dynamic spatial sound field was presented using sound field synthesis (Ahrens et al., 2014). On a 16-channel circular array of speakers with a radius of 2 m, the sound field was rendered with 7th order 2D higher-order Ambisonics (Zotter & Frank, 2019). As background noise, we presented a first-order Ambisonics recording of a quiet residential neighborhood ($L_{Aeq} = 37.5$ dB). The acoustic simulations contained all auditory cues (intensity and spectral changes, interaural time and level differences, Doppler frequency shifts, etc.) as in a corresponding real-world scenario. Further details about the recording procedure and acoustic simulation can be found in a previous article (Oberfeld et al., 2022).

Fig. 1 Bird's-eye view of the simulated traffic scene (left) and the participant's view of it (right). The vehicle approached the participants along the road. Participants observed the vehicle and judged its TTC. The blue line 50 cm to the left of the participant's position served as reference location in the virtual scene



Visual Stimuli and Simulation

The interactive auditory simulation was synchronized in time with a visual simulation of an approaching vehicle in an urban traffic scene. The approaching car was a red Mitsubishi Colt and the two-lane street was modeled after the Eislebener Straße, Berlin (https://www.stadtentwicklung.berlin.de/planen/stadtmodelle/de/digitale_innenstadt/3d/index.shtml), 3D model provided by the Senate Department for Urban Development, Building and Housing of Berlin). The car approached the participants along the street from the left-hand side (Fig. 1) with different driving profiles. Their position in the virtual scene was 0.5 m away from the curb. A blue line on the road marked the participants' position and served as orientation in the virtual environment. Participants stood in the middle of the loudspeaker array and wore a head-mounted display (HTC Vive Pro). Head tracking allowed them to explore the traffic scene and to track the vehicle's motion. The visual simulations were provided by the VR software WorldViz Vizard 5. The auditory and the visual simulations were synchronized via the OpenSoundControl network protocol (opensoundcontrol.org).

Task

During each trial, the participants observed the approaching car for 3 s. Then, the car was occluded from the scene and was no longer visible or audible. They were instructed to press the controller button to indicate the moment, they thought the vehicle's front would have arrived at their position (prediction-motion task) (Schiff & Detwiler, 1979). In making their estimates, participants were instructed to assume that the vehicles traveling at constant velocity before occlusion would continue to travel at this constant velocity after occlusion, while vehicles that accelerated would continue to accelerate thereafter. The time between occlusion and the button press served as a measure for the estimated TTC.

Experimental Design

In a four-factorial within-subjects design, all participants completed all combinations of the following experimental conditions. We presented the vehicle either audiovisually (AV) or visually-only (V-only). The vehicle approached with different velocity profiles, that is, with or without acceleration. As the presented driving profiles were based on real drives on a test track, they deviated from the intended velocity/and or acceleration to a certain extent. Therefore, we report the mean presented initial and constant velocities and the mean presented acceleration within the last second before occlusion in brackets (see also Supplement). Without acceleration, the vehicle traveled for 3 s either at a constant velocity of 10 km/h ($M = 10.45$ km/h, $SD = 0.22$ km/h) or 40 km/h ($M = 40.51$ km/h, $SD = 0.59$ km/h). With acceleration, the car initially approached at a constant velocity of 10 km/h ($M = 11.22$ km/h, $SD = 0.40$ km/h) or 40 km/h ($M = 40.22$ km/h) for 1.5 s and then accelerated for another 1.5 s with 2 m/s^2 ($M = 2.16 \text{ m/s}^2$, $SD = 0.45 \text{ m/s}^2$). The TTC, i.e., the actual time the vehicle would have needed to arrive at the participant's position after it had been occluded if it had continued its motion with the same constant velocity or the same acceleration as before the occlusion, was varied (1.25, 2.5, 3.75 or 5.0 s). We set the vehicle distance at occlusion so that the desired TTC at occlusion resulted. For constant velocity approaches, the calculation was based on the mean presented velocity. For accelerated approaches, it was based on the presented velocity at occlusion and the mean acceleration within the last second before occlusion. Each of the 32 combinations of the factors presentation condition, acceleration, velocity and actual TTC was presented 8 times in the experiment and twice within each of the 4 blocks. Each participant completed the 256 experimental trials.

Procedure

Prior to testing, all participants passed the required vision and hearing tests. We measured the participant's ear height and inter-pupillary distance to align the simulated sound field and adjust the displays of the head-mounted display, respectively. Participants were familiarized with the virtual scene and the experimental task in 16 training trials before they started with the experimental trials. Throughout the experiment, the experimenter regularly checked on the participants' well-being using the Fast Motion Sickness Scale (Keshavarz & Hecht, 2011). Participants could take a break at any time, as they controlled the start of each trial themselves. The experiment consisted of four experimental blocks and lasted for approximately 50 minutes.

Participants

We expected a relatively large effect size of at least $d_z = 0.8$ (Cohen, 1988) because the added sound of an accelerating vehicle was expected to strongly improve the TTC estimates compared to a visual-only presentation. A power analysis showed that a sample size of $n = 23$ was sufficient to detect this effect size with a α -level of 5% and a power of 95%. We collected data of 25 volunteers (12 female, 12 male, 1 diverse) with a mean age of 26.21 years ($SD = 7.51$ years). All had (corrected-to-)normal visual acuity (Freiburg Visual Acuity test (Bach, 1996)), normal stereoscopic vision (Titmus test (A. G. Bennett & Rabbetts, 1998)) and normal hearing (thresholds better than 20 dB HL in each ear at frequencies between 125 and 4000 Hz). The participants volunteered for course credit and gave written informed consent prior to testing. The experimental procedure was in accordance with principles outlined in the Declaration of Helsinki and the local ethics committee of the Johannes Gutenberg-University Mainz approved the study (approval number: 2019-JGU-psychEK-S011). Prior to testing, all participants received experimental information and gave written informed consent to participate and to publish their data.

Statistical Analyses

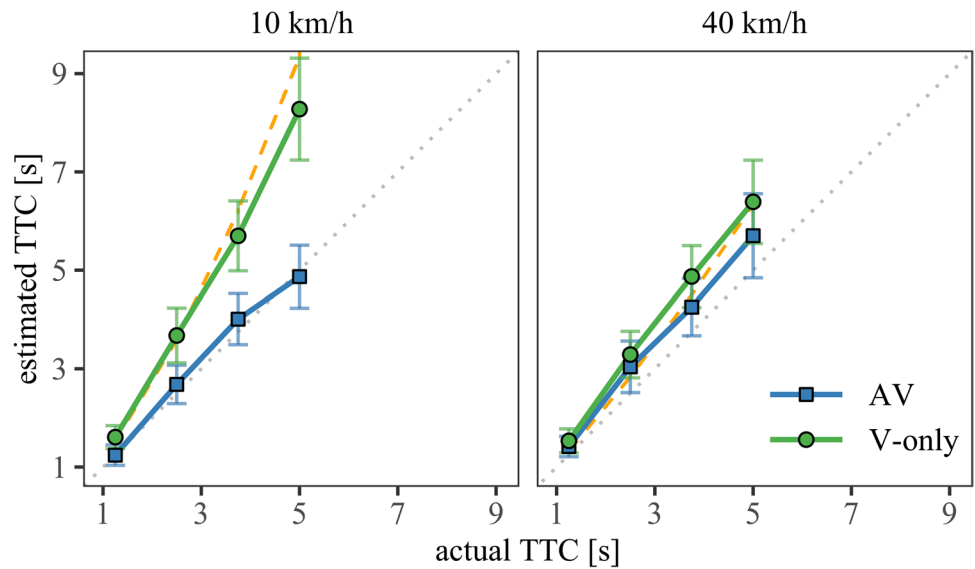
To exclude extreme data points, we applied a Tukey criterion to the data collected per combination of participant and experimental condition, excluding 0.89% of the trials (57 of a total of 6400 trials), which were 3 interquartile ranges below the first or above the third quartile. We aggregated the TTC estimates ($n = 6343$) for each combination of participant and experimental condition. We conducted two separate repeated-measures ANOVAs (rmANOVAs) for conditions with constant velocity and acceleration, using a univariate approach and Huynh-Feldt (Huynh & Feldt, 1976) correction for the degrees of freedom. Significant results ($p < .050$) were followed up with two-sided paired-samples Bonferroni-corrected t -tests where necessary.

Results

TTC Estimation for Accelerating Vehicles

Figure 2 illustrates the mean estimated TTC as a function of the actual TTC, presentation condition, and initial velocity. It also visualizes the relationship between the actual TTC or a perfectly accurate TTC estimation (grey dashed diagonal) and a first-order estimation. In contrast to the actual TTCs, as explained in the Introduction, the first-order TTC estimates (orange dashed line) increase as a quadratic function of actual TTC and decrease as a function of the velocity at occlusion, which was higher at the faster compared to the

Fig. 2 Mean estimated TTC for accelerating vehicles as a function of actual TTC and initial velocity (left: 10 km/h; right: 40 km/h). Blue squares: audio-visual (AV) condition. Green circles: visual-only (V-only) condition. Orange dashed line: first-order estimation. Dotted grey line: perfectly accurate TTC estimation. Error bars indicate 95% confidence intervals



slower initial speed in our experiment. Hence, a first-order estimation results in TTC overestimation increasing with the actual TTC, and a stronger TTC overestimation for vehicles accelerating from a slower than from a faster initial velocity.

For the V-only condition (green circles), participants increasingly overestimated the TTC for accelerating vehicles with increasing actual TTC and decreasing initial velocity – similar to a first-order estimation pattern. In contrast, the added vehicle sound (AV, blue squares) eliminated this overestimation and resulted in largely accurate TTC estimates close to the veridical values, as expected.

Table 1 displays the results of the rmANOVA for accelerating vehicles, in which all main and interaction effects were statistically significant, $p < .007$. As hypothesized, we observed a strong effect of the presentation condition ($d_z = 2.04$). Participants estimated the TTC to be significantly longer in the V-only ($M = 4.42$ s, 95% CI = [3.88 s, 4.95 s]) than in the AV presentation ($M = 3.40$ s, 95% CI = [2.94 s, 3.86 s]). The TTC estimates for the V-only presentation increased with increasing actual TTC, which is compatible with a first-order estimation for visually

presented objects (Benguigui et al., 2003; Benguigui & Bennett, 2010; Bennett & Benguigui, 2016; Kaiser & Hecht, 1995; Senot et al., 2003). However, in the AV condition, TTC estimates remained largely accurate even at longer TTCs. The difference between the V-only and AV condition thus increased with increasing actual TTC (presentation \times actual TTC interaction). Additionally, it decreased with increasing initial velocity (presentation \times velocity interaction). Consistent with a smaller deviation of the first-order estimate from the actual TTC at a higher compared to a lower initial velocity (see orange lines in Fig. 2), the TTC estimates of the V-only condition at the higher initial velocity were closer to the more accurate TTC estimates of the AV condition. Also, the actual acceleration at the higher initial velocity was lower than at the lower initial velocity (see Supplement), which further reduced the difference between the first-order estimates and actual TTCs at the higher initial velocity. The TTC estimates in the V-only and AV condition differed most substantially at an initial speed of 10 km/h and an actual TTC of 5 s (presentation \times velocity \times actual TTC interaction). Here, the participants estimated the TTC

Table 1 Results of the rmANOVA on the mean TTC estimates for accelerating vehicles

| | df_{Num} | df_{Den} | $\tilde{\epsilon}$ | F | p | η^2_p |
|--|------------|------------|--------------------|--------|--------|------------|
| Presentation | 1 | 24 | | 104.09 | < .001 | .81 |
| Velocity | 1 | 24 | | 8.73 | .007 | .27 |
| Presentation \times velocity | 1 | 24 | | 64.59 | < .001 | .73 |
| Actual TTC | 3 | 72 | 0.44 | 229.84 | < .001 | .91 |
| Presentation \times actual TTC | 3 | 72 | 0.45 | 53.59 | < .001 | .69 |
| Actual TTC \times velocity | 3 | 72 | 0.55 | 6.17 | .007 | .20 |
| Presentation \times actual TTC \times velocity | 3 | 72 | 0.45 | 16.63 | < .001 | .41 |

Displayed are uncorrected numerator degrees of freedom (df_{Num}), denominator degrees of freedom (df_{Den}), the Huynh-Feldt multiplier for sphericity correction ($\tilde{\epsilon}$), F -values, p -values and partial η^2 (η^2_p)

to be, on average, 3.28 s longer than the veridical value. With added vehicle sound, the mean estimate deviated by 0.13 s from the actual TTC. In sum, with increasing actual TTC and slower initial velocity, the amount of overestimation of the TTC of accelerating vehicles increased in the V-only condition, showing a pattern similar to first-order TTC estimation, while the additional vehicle sound largely removed this pattern and resulted in substantially more accurate TTC estimates.

Apart from the above effects involving the presentation condition, which are most important for the present study, the rmANOVA also showed a significant effect of TTC, indicating that the TTC estimates increased with the actual TTC. Participants also on average estimated the TTC to be significantly longer for vehicles accelerating from 10 km/h ($M=4.01$ s, 95% $CI=[3.54$ s, 4.48 s]) compared to 40 km/h ($M=3.81$ s, 95% $CI=[3.30$ s, 4.32 s]). The comparably longer TTC estimates for slower velocities occurred mainly at longer actual TTCs (velocity \times actual TTC interaction). These statistical effects can again be attributed to the estimates in the V-only condition, where the amount of overestimation increased with the actual TTC and was smaller at the faster initial speed, showing a pattern compatible with first-order TTC estimation (Fig. 2).

TTC Estimation for Vehicles with Constant Velocity

When the vehicles approached with constant velocity (Fig. 3), TTC estimates for audiovisually and visually presented vehicles were rather similar but differed descriptively between the two velocities. On average, participants estimated the TTC of slower vehicles to be shorter than for faster ones.

Table 2 shows the results of the rmANOVA for constant-velocity approaches. Contrary to our expectation, we observed a significant effect of presentation condition for vehicles at constant velocity. However, it followed quite a different pattern than for accelerating vehicles. The mean TTC estimates were slightly shorter in the V-only ($M=2.80$ s, 95% $CI=[2.44$ s, 3.15 s]) than in the AV presentation ($M=2.89$ s, 95% $CI=[2.50$ s, 3.28 s]). Thus, additional auditory information slightly influenced the TTC estimates for constant velocity approaches. However, this effect was relatively small both in terms of the difference in seconds and in terms of effect size ($d_z=0.76$). The effect of added vehicle sound was similar for all actual TTCs and velocities (non-significant presentation \times velocity, presentation \times actual TTC, and presentation \times velocity \times actual TTC interactions).

Besides, participants estimated the TTC to be longer for faster ($M=3.14$ s, 95% $CI=[2.71$ s, 3.57 s]) than for slower vehicles ($M=2.55$ s, 95% $CI=[2.21$ s, 2.88 s]). The main effect of velocity is consistent with a size-arrival effect (DeLucia, 1991, 2013; DeLucia et al., 2020; DeLucia & Warren, 1994), or a distance bias (Law et al., 1993), that predicts longer TTC estimates for objects of a smaller optical size or at a greater distance than for larger or close objects. In a prediction-motion paradigm, a faster vehicle is farther away at occlusion than a slower one at the same actual TTC. Thus, at identical TTC, the optical size at occlusion is also smaller for the faster compared to the slower velocity. The significant velocity \times actual TTC interaction is also in line with this explanation. The difference in vehicle distance at occlusion, and thus also in optical size, increased more strongly with increasing actual TTC at a higher than at a lower velocity. Consistent with the size-arrival effect, the mean TTC estimates consequently increased more strongly

Fig. 3 Mean estimated TTC for vehicles approaching at a constant velocity as a function of actual TTC and velocity (left: 10 km/h; right: 40 km/h). Blue squares: audiovisual (AV) condition. Green circles: visual-only (V-only) condition. Dotted grey line: perfectly accurate TTC estimation. Error bars indicate 95% confidence intervals

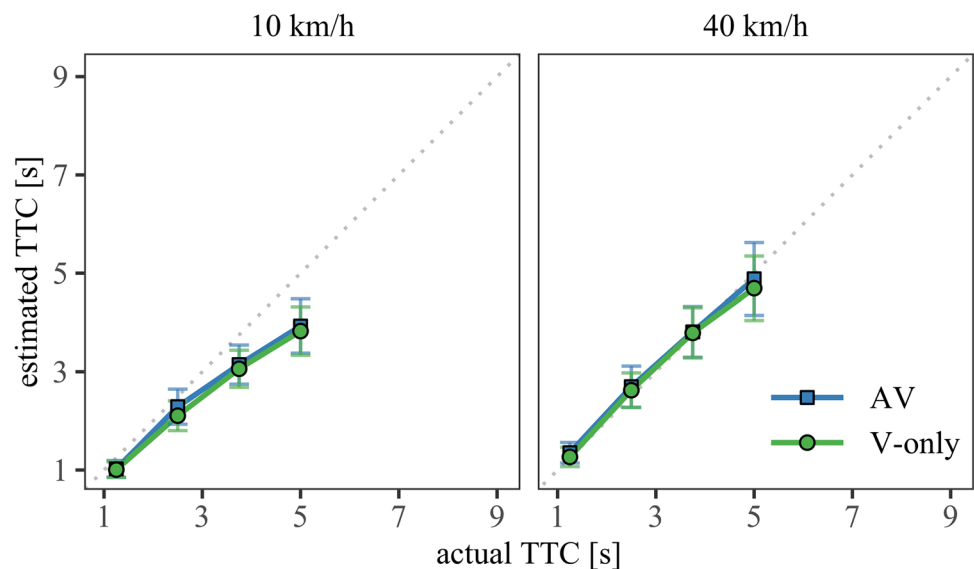


Table 2 Results of the rmANOVA on the mean TTC estimates for vehicles with constant velocity

| | df_{Num} | df_{Den} | $\tilde{\epsilon}$ | F | p | η^2_p |
|--|------------|------------|--------------------|--------|--------|------------|
| Presentation | 1 | 24 | | 14.49 | .001 | .38 |
| Velocity | 1 | 24 | | 40.29 | < .001 | .63 |
| Presentation \times velocity | 1 | 24 | | 0.01 | .916 | .00 |
| Actual TTC | 3 | 72 | 0.39 | 198.22 | < .001 | .89 |
| Presentation \times actual TTC | 3 | 72 | 0.77 | 1.05 | .364 | .04 |
| Actual TTC \times velocity | 3 | 72 | 0.87 | 13.92 | < .001 | .37 |
| Presentation \times actual TTC \times velocity | 3 | 72 | 0.87 | 1.47 | .234 | .06 |

Displayed are uncorrected numerator degrees of freedom (df_{Num}), denominator degrees of freedom (df_{Den}), the Huynh-Feldt multiplier for sphericity correction ($\tilde{\epsilon}$), F -values, p -values and partial η^2 (η^2_p)

as a function of actual TTC at the higher velocity. Taken together, the final distance or the final optical size may have been a central heuristic in the TTC estimation for vehicles at constant velocities.

Discussion

We investigated the role of auditory information in TTC judgements for accelerating vehicles from a pedestrian's perspective. Without vehicle sound, participants estimated the TTC for accelerating vehicles substantially longer than it actually was. More precisely, TTC estimates in the visual-only (V-only) condition showed a pattern similar to first-order TTC estimation, which does not consider the acceleration. That is, TTC was increasingly overestimated with increasing actual TTC and decreasing initial velocity, resulting in large estimation errors. It is conceivable that the participants did not perceive the acceleration because the human visual system is quite insensitive to acceleration (Gottsdanker et al., 1961; Snowden & Braddick, 1991; Werkhoven et al., 1992). In contrast, the added vehicle sound mostly removed the first-order pattern so that the TTC estimates in the audiovisual (AV) condition were close to the veridical values. Hence, the auditory information helped the participants to judge the TTC more accurately for accelerating vehicles, and thus played a major role.

It is an interesting and open question how exactly the vehicle sound improved the accuracy of TTC estimation for accelerating vehicles. Although the present study was not designed to answer this question, we briefly outline possible explanations. First, it is important to note that auditory information about the vehicle's motion in space was available from the dynamic spatial sound field generated by the approaching vehicle, just as in the visual domain. For instance, interaural time and level differences auditorily indicate the angle between the left curbside and the vehicle's right front tire from the observer's perspective (Middlebrooks & Green, 1991). As the vehicle approached, this dynamically changing lateral angle provided information

about the distance, velocity and acceleration, analogous to the optical angle. In principle, participants might be sensitive to the second-order information (acceleration) provided by the motion of the acoustic object in space than second-order motion information provided by the motion of the visual object. In such a scenario, in the AV condition, participants could simply base their TTC estimates on only the auditory information and ignore the visual information. Because auditory acceleration detection is not substantially better than in the visual domain (Locke et al., 2016; Perrott et al., 1993), we expect that an auditory-only (A-only) presentation would not result in more accurate TTC estimates for accelerating vehicles than the V-only condition. However, to investigate whether people are better able to use second-order motion information in the auditory than in the visual modality, additional experiments comparing TTC estimates for accelerating objects in an A-only to those in a V-only presentation condition are needed. Further comparisons with an AV condition could then reveal how the auditory and visual second-order information is weighted against each other.

Alternatively, as an anonymous reviewer suggested, rather than reflecting a better use of second-order motion information, the largely accurate mean TTC estimates in the AV condition might be the result of a rather simple averaging process between underestimated auditory and overestimated visual TTCs. In the V-only condition, the TTC overestimation increased with longer actual TTCs. If now the auditory TTC estimates show the opposite pattern, that is, a TTC underestimation increasing with actual TTC, then the average between the visual and the auditory TTC estimates will again be accurate. We consider it unlikely that the auditory TTC estimates show exactly the opposite pattern as visual TTC estimates, but again data on A-only TTC estimation for accelerating objects would be necessary to answer this question.

Instead, we suppose that the vehicle's salient sound profile, which provides information beyond the vehicle's spatial motion, played a central role. During acceleration, the sound spectrum of a vehicle with internal combustion engine shifts to higher frequencies as the engine speed increases, and the

sound level rises as the engine load increases (Zeller, 2018). Thus, the sound profile conveyed acceleration cues, which were visually not available. Therefore, participants might have noticed the acceleration due to the characteristic and probably familiar sound profile. In line with current theoretical models of interceptive actions (e.g., Rust & Palmer, 2021; Zago et al., 2009), the familiar sound of an accelerating vehicle might have served as an heuristic cue that triggered the retrieval of prior traffic experiences and thus resulted in updated expectations about the vehicle's motion, considering acceleration. In turn, this might have triggered either a) a rather unspecific correction of the visual first-order TTC estimates, or b) an attention shift towards the relevant second-order information in the dynamic approach. The former case assumes that only first-order information (velocity) can be used in TTC estimation. Hence, participants could still have made first-order TTC estimates in the AV condition but then shortened their estimates by an absolute amount of time (e.g., 1 s) as a consequence of the detected (potentially dangerous) acceleration. This would be compatible with a safety strategy previously observed with, e.g., blurred vision (Hecht et al., 2021), and threatening sounds (Braly et al., 2021). When using such a strategy, the mean TTC estimates in the AV condition in Fig. 2 would show a pattern parallel to the estimates in the V-only condition, but shifted towards the x -axis by a constant amount, independent of the actual TTC. However, the data in Fig. 2 are not compatible with such a pattern. Alternatively, participants could have shortened their estimates pro rata, e.g., by 50%. At an initial velocity of 10 km/h and an actual TTC of 5 s, participants estimated the TTC of the accelerating vehicle to be on average 8.28 s in the V-only condition, and 4.87 s for the same actual TTC in the AV condition, which represents a reduction of approximately 40%. For an actual TTC of 1.25 s, however, the mean TTC estimates were 1.61 s in the V-only and 1.24 s in the AV condition, respectively, which represents a reduction of approximately 20%. Therefore, the data are not entirely consistent with a constant pro rata reduction of the first-order TTC estimates. In contrast, case b) assumes that second-order information (provided by the auditory and/or visual motion in space) can indeed be used in TTC estimation, but only if attention is directed to it. That is, the available second-order auditory and/or visual information is disregarded until participants become aware of the current acceleration signaled by the sound profile.

In sum, the characteristic sound profile of the accelerating vehicle may have been the key information to compensate for the erroneous visual TTC estimates, but further research is needed to identify the perceptual cues and mechanisms underlying the beneficial role of the additional vehicle sound. In particular, it would be of high relevance to include an A-only presentation condition to investigate

whether the largely accurate TTC estimates in the AV condition are due to the exclusive use of auditory motion information or whether visual information is still incorporated. Also, the following aspects limit the generalizability of our findings: First, the present study only investigated the sound of a conventional vehicle with internal combustion engine. For other vehicles, such as hybrid or electric vehicles, changes in vehicle sound during acceleration are likely to be less salient than for conventional vehicles. This could reduce the benefit of the vehicle sound to pedestrians estimating TTC. Second, we only presented vehicle approaches at an acceleration level of 2 m/s^2 . Since auditory acceleration cues might be much more subtle during lower acceleration, the improvement in TTC estimation when the vehicle sound is added to the visual representation could be significantly reduced at lower acceleration levels. Nonetheless, we can also expect that the visual TTC estimates for vehicle approaches at a lower acceleration differ less from the actual TTC than at a higher acceleration because the difference between first- and second-order estimation decreases with decreasing object acceleration. For accelerations above 2 m/s^2 , this difference would be more prominent, but so would be the auditory acceleration cues possibly correcting a visual first-order estimation. Third, in an older sample, sensory functioning/cognitive performance might be substantially different than in our sample, which comprises rather young people. This could affect the use of perceptual cues and mechanisms underlying the TTC estimation, as was already observed for approaches at constant velocities (Keshavarz et al., 2017). Finally, the simulated scene might have offered a fairly large number of visual distance cues (road markings house fronts). In an environment with fewer visual distance cues, participants might rely even more heavily on auditory distance cues during their estimation. Taken together, future research should include an A-only presentation condition, and shed light on different vehicle types, acceleration levels, groups with different levels of sensory functioning, and environments to gain a deeper understanding of the role of auditory information during TTC estimation for accelerating objects.

For constant velocity approaches, we observed, on average, mostly accurate TTC estimates both with and without added vehicle sound. The data also indicated a significant difference between the presentation conditions, with slightly less underestimated TTCs in the AV than in the V-only condition. Thus, the vehicle sound had an effect on the TTC estimates although visual information was available, but did not provide a particularly large benefit in the TTC estimation for constant velocity approaches, which is consistent with the previous literature (DeLucia et al., 2016; Hassan, 2012; Keshavarz et al., 2017; Schiff & Oldak, 1990; Zhou et al., 2007).

Practical Implications

Our data show that pedestrians commit perceptual errors when the vehicle sound is not available to them - at least in situations where they are confronted with an accelerating vehicle. Since these perceptual errors manifest as TTC overestimation, pedestrians are likely to cross the road in a risky manner, i.e. (too) close in front of an accelerating vehicles. By listening to loud music that masks the vehicle sounds or wearing headphones with noise cancellation, pedestrians might thus jeopardize their own safety on the road not only due to an impaired acoustic detection of vehicles, but also in relation to risky TTC overestimations for accelerating vehicles. This aspect should be included in educational programs on traffic safety. In the same line of reasoning, if the vehicles emit hardly any noise, our results imply negative consequences for pedestrian safety in interaction with accelerating vehicles, in this context without pedestrians having any influence. The latter could be the case with quieter electric vehicles, which would mean that pedestrian safety and the goal of reducing noise pollution in cities by increasing the number of electric vehicles are in critical tension with each other (Mendonça et al., 2013).

Conclusions

Our results clearly demonstrate that the sound emitted by an accelerating vehicle provides a pronounced benefit for pedestrians' TTC estimations. Without vehicle sound, the TTC estimates were quite erroneous and showed a first-order pattern indicating that the vehicle's acceleration was not adequately considered, compatible with previous data. The added vehicle sound helped to substantially increase the accuracy of TTC estimates for accelerating vehicles, presumably due to the improved consideration of acceleration information. We conclude that the added sound of an accelerating vehicle can compensate for erroneous visual TTC estimates in a highly effective way, which has important implications for traffic safety.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12144-022-03375-6>.

Acknowledgements We thank Agnes Münch for assistance in programming the traffic scenario and Heiko Hecht for comments on an initial draft.

Author Contributions Marlene Wessels: Conceptualization, Methodology, Software, Formal Analysis, Resources, Data curation, Writing – Original Draft, Writing – review & editing, Visualization, Supervision, Project administration.

Carolina Zähme: Conceptualization, Investigation, Writing – review & editing.

Daniel Oberfeld: Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Software, Supervision, Writing – review & editing.

Funding Open Access funding enabled and organized by Projekt DEAL. This work was supported by funding from Deutsche Forschungsgemeinschaft (www.dfg.de; OB 346/8–1) granted to Daniel Oberfeld.

Data Availability The data are available in an online repository (Open Science Framework, https://osf.io/gqtxn/?view_only=7a429d24a0cb4e9e8faa6517a63cd62e).

Declarations

Ethics Approval The experimental procedure was in accordance with principles outlined in the Declaration of Helsinki and the local ethics committee of the Johannes Gutenberg-University Mainz approved the study (approval number: 2019-JGU-psychEK-S011).

Consent Prior to testing, all participants received experimental information and gave written informed consent to participate and to publish their data.

Competing Interests The authors have no conflicts or competing interests to declare.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ahrens, J., Rabenstein, R., & Spors, S. (2014). Sound field synthesis for audio presentation. *Acoustics Today*, 10(2), 15–25.
- Bach, M. (1996). The Freiburg visual acuity test-automatic measurement of visual acuity. *Optometry and Vision Science*, 73(1), 49–53. <https://doi.org/10.1097/00006324-199601000-00008>
- Benguigui, N., & Bennett, S. J. (2010). Ocular pursuit and the estimation of time-to-contact with accelerating objects in prediction motion are controlled independently based on first-order estimates. *Experimental Brain Research*, 202(2), 327–339. <https://doi.org/10.1007/s00221-009-2139-0>
- Benguigui, N., Ripoll, H., & Broderick, M. P. (2003). Time-to-contact estimation of accelerated stimuli is based on first-order information. *Journal of Experimental Psychology: Human Perception and Performance*, 29(6), 1083–1101. <https://doi.org/10.1037/0096-1523.29.6.1083>
- Bennett, S. J., & Benguigui, N. (2016). Spatial estimation of accelerated stimuli is based on a linear extrapolation of first-order information. *Experimental Psychology*, 63(2), 98–106. <https://doi.org/10.1027/1618-3169/a000318>

- Bennett, A. G., & Rabbetts, R. B. (1998). *Clinical visual optics* (3rd ed.). Butterworth-Heinemann.
- Braly, A. M., DeLucia, P., & Oberfeld, D. (2021). Does affective content of sounds affect auditory time-to-collision estimation? *Auditory Perception & Cognition*, 1–23. <https://doi.org/10.1080/25742442.2021.1997064>.
- Butler, A. A., Lord, S. R., & Fitzpatrick, R. C. (2016). Perceptions of speed and risk: Experimental studies of road crossing by older people. *PLoS One*, 11(4), e0152617. <https://doi.org/10.1371/journal.pone.0152617>
- Calderone, J. B., & Kaiser, M. K. (1989). Visual acceleration detection: Effect of sign and motion orientation. *Perception & Psychophysics*, 45(5), 391–394. <https://doi.org/10.3758/BF03210711>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). L. Erlbaum Associates.
- DeLucia, P. R. (1991). Pictorial and motion-based information for depth perception. *Journal of Experimental Psychology: Human Perception and Performance*, 17(3), 738–748.
- DeLucia, P. R. (2013). Effects of size on collision perception and implications for perceptual theory and transportation safety. *Current Directions in Psychological Science*, 22(3), 199–204. <https://doi.org/10.1177/0963721412471679>
- DeLucia, P. R., & Warren, R. (1994). Pictorial and motion-based depth information during active control of self-motion: Size-arrival effects on collision avoidance. *Journal of Experimental Psychology: Human Perception and Performance*, 20(4), 783–798.
- DeLucia, P. R., Preddy, D., & Oberfeld, D. (2016). Audiovisual integration of time-to-contact information for approaching objects. *Multisensory Research*, 29(4–5), 365–395. <https://doi.org/10.1163/22134808-00002520>
- DeLucia, P. R., Braly, A. M., & Savoy, B. R. (2020). Does the size-arrival effect occur with an active collision-avoidance task in an immersive 3D virtual reality environment? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 64(1), 1094–1094. <https://doi.org/10.1177/1071181320641263>
- Dommes, A., Cavallo, V., & Oxley, J. (2013). Functional declines as predictors of risky street-crossing decisions in older pedestrians. *Accident Analysis & Prevention*, 59, 135–143. <https://doi.org/10.1016/j.aap.2013.05.017>
- Gottsdanker, R., Frick, J. W., & Lockard, R. (1961). Identifying the acceleration of visual targets. *British Journal of Psychology*, 52(1), 31–42.
- 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(3), 566–578. <https://doi.org/10.3813/AAA.919337>
- Hassan, S. E. (2012). Are normally sighted, visually impaired, and blind pedestrians accurate and reliable at making street crossing decisions? *Investigative Ophthalmology & Visual Science*, 53(6), 2593. <https://doi.org/10.1167/iovs.11-9340>
- Hecht, H., Brendel, E., Wessels, M., & Bernhard, C. (2021). Estimating time-to-contact when vision is impaired. *Scientific Reports*, 11(21213), 14. <https://doi.org/10.1038/s41598-021-00331-5>
- Hofbauer, M., Wuerger, S. M., Meyer, G. F., Roehrbein, F., Schill, K., & Zetsche, C. (2004). Catching audiovisual mice: Predicting the arrival time of auditory-visual motion signals. *Cognitive, Affective, & Behavioral Neuroscience*, 4(2), 241–250. <https://doi.org/10.3758/CABN.4.2.241>
- Huynh, H., & Feldt, L. S. (1976). Estimation of the box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational Statistics*, 1(1), 69–82.
- Kaiser, M. K., & Hecht, H. (1995). Time-to-passage judgments in non-constant optical flow fields. *Perception & Psychophysics*, 57(6), 817–825. <https://doi.org/10.3758/BF03206797>
- Keshavarz, B., & Hecht, H. (2011). Validating an efficient method to quantify motion sickness. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 53(4), 415–426. <https://doi.org/10.1177/0018720811403736>
- Keshavarz, B., Campos, J. L., DeLucia, P. R., & 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(3), 929–944. <https://doi.org/10.3758/s13414-016-1270-9>
- Law, D. J., Pellegrino, J. W., Mitchell, S. R., Fischer, S. C., McDonald, T. P., & Hunt, E. B. (1993). Perceptual and cognitive factors governing performance in comparative arrival-time judgments. *Journal of Experimental Psychology: Human Perception and Performance*, 19(6), 1183–1199.
- Lee, D. N., & Young, D. S. (1985). *Visual timing of interceptive action. In brain mechanisms and spatial vision* (Vol. 21, pp. 1–30). Springer.
- Lee, D. N., Young, D. S., Reddish, P. E., Lough, S., & Clayton, T. M. H. (1983). Visual timing in hitting an accelerating ball. *The Quarterly Journal of Experimental Psychology Section A*, 35(2), 333–346. <https://doi.org/10.1080/14640748308402138>
- Locke, S. M., Leung, J., & Carlile, S. (2016). Sensitivity to auditory velocity contrast. *Scientific Reports*, 6(1), 27725. <https://doi.org/10.1038/srep27725>
- López-Moliner, J., Maiche, A., & Estaún, S. (2003). Perception of acceleration in motion-in-depth with only monocular and both monocular and binocular information. *Psicológica*, 24, 93–108.
- Mendonça, C., Freitas, E., Ferreira, J. P., Raimundo, I. D., & Santos, J. A. (2013). Noise abatement and traffic safety: The trade-off of quieter engines and pavements on vehicle detection. *Accident Analysis & Prevention*, 51, 11–17. <https://doi.org/10.1016/j.aap.2012.10.018>
- Middlebrooks, J. C., & Green, D. M. (1991). Sound localization by human listeners. *Annual Review of Psychology*, 42, 135–159.
- 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 and Prevention*. (in press)
- Perrott, D. R., Costantino, B., & Ball, J. (1993). Discrimination of moving events which accelerate or decelerate over the listening interval. *The Journal of the Acoustical Society of America*, 93(2), 1053–1057. <https://doi.org/10.1121/1.405553>
- Petzoldt, T. (2014). On the relationship between pedestrian gap acceptance and time to arrival estimates. *Accident Analysis & Prevention*, 72, 127–133. <https://doi.org/10.1016/j.aap.2014.06.019>
- Prime, S. L., & Harris, L. R. (2010). Predicting the position of moving audiovisual stimuli. *Experimental Brain Research*, 203(2), 249–260. <https://doi.org/10.1007/s00221-010-2224-4>
- Rosenbaum, D. A. (1975). Perception and extrapolation of velocity and acceleration. *Journal of Experimental Psychology: Human Perception and Performance*, 1(4), 395–403. <https://doi.org/10.1037/0096-1523.1.4.395>
- Rust, N. C., & Palmer, S. E. (2021). Remembering the past to see the future. *Annual Review of Vision Science*, 7, 349–365. <https://doi.org/10.1146/annurev-vision-093019-112249>
- Schiff, W., & Detwiler, M. L. (1979). Information used in judging impending collision. *Perception*, 8(6), 647–658.
- Schiff, W., & Oldak, R. (1990). Accuracy of judging time to arrival: Effects of modality, trajectory, and gender. *Journal of Experimental Psychology: Human Perception and Performance*, 16(2), 303–316.
- Senot, P., Prévost, P., & McIntyre, J. (2003). Estimating time to contact and impact velocity when catching an accelerating object with the hand. *Journal of Experimental Psychology: Human Perception*

- and Performance, 29(1), 219–237. <https://doi.org/10.1037/0096-1523.29.1.219>
- Snowden, R. J., & Braddick, O. J. (1991). The temporal integration and resolution of velocity signals. *Vision Research*, 31(5), 907–914. [https://doi.org/10.1016/0042-6989\(91\)90156-Y](https://doi.org/10.1016/0042-6989(91)90156-Y)
- Tharanathan, A. (2009). Effects of constant and non-constant velocity motion on judgments of collision-avoidance action gap. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 53(23), 1762–1765. <https://doi.org/10.1177/154193120905302304>
- Tharanathan, A., & DeLucia, P. R. (2006). Time-to-collision judgments of constant and non-constant velocities: Implications for rear-end collisions. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50, 2463–2467.
- Tresilian, J. R. (1995). Perceptual and cognitive processes in time-to-contact estimation: Analysis of prediction-motion and relative judgment tasks. *Perception & Psychophysics*, 57(2), 231–245. <https://doi.org/10.3758/BF03206510>
- Werkhoven, P., Snippe, H. P., & Alexander, T. (1992). Visual processing of optic acceleration. *Vision Research*, 32(12), 2313–2329. [https://doi.org/10.1016/0042-6989\(92\)90095-Z](https://doi.org/10.1016/0042-6989(92)90095-Z)
- Zago, M., Bosco, G., Maffei, V., Iosa, M., Ivanenko, Y. P., & Lacquaniti, F. (2004). Internal models of target motion: Expected dynamics overrides measured kinematics in timing manual interceptions. *Journal of Neurophysiology*, 91(4), 1620–1634. <https://doi.org/10.1152/jn.00862.2003>
- Zago, M., McIntyre, J., Senot, P., & Lacquaniti, F. (2009). Visuo-motor coordination and internal models for object interception. *Experimental Brain Research*, 192(4), 571–604. <https://doi.org/10.1007/s00221-008-1691-3>
- Zeller, P. (Ed.). (2018). *Handbuch Fahrzeugakustik*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-18520-6>
- Zhou, L., Yan, J., Liu, Q., Li, H., Xie, C., Wang, Y., Campos, J. L., & Sun, H. (2007). Visual and Auditory Information Specifying an Impending Collision of an Approaching Object. In J. A. Jacko (Ed.), *Human-Computer Interaction. Interaction Platforms and Techniques* (Vol. 4551, pp. 720–729). Springer. https://doi.org/10.1007/978-3-540-73107-8_80
- Zotter, F., & Frank, M. (2019). *Ambisonics: A practical 3D audio theory for recording, studio production, sound reinforcement, and virtual reality* (Vol. 19). Springer International Publishing. <https://doi.org/10.1007/978-3-030-17207-7>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.