

# 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

Behrang Keshavarz<sup>1,2</sup> · Jennifer L. Campos<sup>1,3</sup> · Patricia R. DeLucia<sup>4</sup> · Daniel Oberfeld<sup>5</sup>

Published online: 17 January 2017  
© The Psychonomic Society, Inc. 2017

**Abstract** Estimating time to contact (TTC) involves multiple sensory systems, including vision and audition. Previous findings suggested that the ratio of an object's instantaneous optical size/sound intensity to its instantaneous rate of change in optical size/sound intensity ( $\tau$ ) drives TTC judgments. Other evidence has shown that heuristic-based cues are used, including final optical size or final sound pressure level. Most previous studies have used decontextualized and unfamiliar stimuli (e.g., geometric shapes on a blank background). Here we evaluated TTC estimates by using a traffic scene with an approaching vehicle to evaluate the weights of visual and auditory TTC cues under more realistic conditions. Younger (18–39 years) and older (65+ years) participants made TTC estimates in three sensory conditions: visual-only, auditory-only, and audio–visual. Stimuli were presented within an immersive virtual-reality environment, and cue weights were calculated for both visual cues (e.g., visual  $\tau$ , final optical size) and auditory cues (e.g., auditory  $\tau$ , final sound pressure level). The results demonstrated the use of visual  $\tau$  as well as heuristic

cues in the visual-only condition. TTC estimates in the auditory-only condition, however, were primarily based on an auditory heuristic cue (final sound pressure level), rather than on auditory  $\tau$ . In the audio–visual condition, the visual cues dominated overall, with the highest weight being assigned to visual  $\tau$  by younger adults, and a more equal weighting of visual  $\tau$  and heuristic cues in older adults. Overall, better characterizing the effects of combined sensory inputs, stimulus characteristics, and age on the cues used to estimate TTC will provide important insights into how these factors may affect everyday behavior.

**Keywords** Time to contact · Relative cue weights · Multisensory integration · Age · Visual · Auditory · Looming · Virtual environments

To identify the potential for collision while moving through the environment, various cues can be used both within and among sensory modalities, such as vision and audition. Once the possibility of a collision is recognized, it is important to estimate the time remaining until the collision would occur (time to contact or time to collision: TTC), so that effective collision avoidance responses can be made. In the visual domain, accurate TTC information is reliably provided by the optical invariant *tau* ( $\tau$ ) when certain assumptions are met (e.g., objects are rigid and move at a constant velocity; see Lee, 1976).  $\tau$  is defined as the ratio of an object's instantaneous optical size to its instantaneous rate of optical expansion (Lee, 1976). Observers use or are sensitive to  $\tau$  in a variety of tasks such as absolute and relative TTC judgments (e.g., (e.g., Bootsma & Oudejans, 1993; Schiff & Oldak, 1990; Todd, 1981).  $\tau$  does not require perceptual estimates of the incoming object's absolute size, distance, or velocity. However, some work has demonstrated that even when  $\tau$  is available, other

---

✉ Daniel Oberfeld  
oberfeld@uni-mainz.de

<sup>1</sup> Department of Research, iDAPT, Toronto Rehabilitation Institute - University Health Network, Toronto, Ontario, Canada

<sup>2</sup> Department of Psychology, Ryerson University, Toronto, Ontario, Canada

<sup>3</sup> Department of Psychology, University of Toronto, Toronto, Ontario, Canada

<sup>4</sup> Department of Psychological Sciences, Texas Tech University, Lubbock, TX, USA

<sup>5</sup> Institute of Psychology, Section Experimental Psychology, Johannes Gutenberg-Universität, 55099 Mainz, Germany

visual characteristics of the object can influence TTC judgments by providing heuristic cues. For instance, studies by DeLucia and colleagues (DeLucia, 1991, 2004; DeLucia, Preddy, & Oberfeld, 2016) showed that an approaching object's final optical size before it disappears (i.e., optical size on the final frame in degrees of visual angle,  $\theta_{final}$ ) affects judgments such that larger final optical sizes are associated with earlier arrival estimates. This “size-arrival effect” was reported not only in experiments using laboratory-type stimuli (e.g., a disc increasing in size presented on a blank background), but also in more naturalistic traffic-related settings (Caird & Hancock, 1994; Horswill, Helman, Ardiles, & Wann, 2005; Petzoldt, 2014; Schleinitz, Petzoldt, Krems, & Gehlert, 2016). Relative size may be considered a heuristic that does not reliably provide accurate TTC information (Braunstein, 1976; Cutting & Wang, 2000; DeLucia, 2004; Hosking & Crassini, 2011). It does not guarantee the correct solution or define a single event in 3-D space.

In natural scenes, apart from  $\tau$  and final optical size, additional visual heuristic cues could potentially be used in TTC judgments. These include the final rate of optical expansion ( $\dot{\theta}_{final}$ ; Gray and Regan, 1998), the change in visual angle across the presentation duration ( $\Delta\theta = \theta_{final} - \theta_{initial}$ ; Gray & Regan, 1998), and the final visual distance of the stimulus ( $D_{final\ v}$ ; Yan, Lorv, Li, & Sun, 2011). In the present study, we measured the influence of each of these three visual heuristic cues on TTC judgments. Specifically, we performed analyses to determine the extents to which visual  $\tau$ ,  $\theta_{final}$ ,  $\dot{\theta}_{final}$ ,  $\Delta\theta$ , and  $D_{final\ v}$  were used in such judgments.

It is known that individuals can use auditory information alone to make TTC estimates and that blind individuals can use acoustical TTC information with accuracy comparable to sighted individuals' abilities to use visual TTC information (Schiff & Oldak, 1990). In the auditory domain, accurate TTC information is reliably provided about objects travelling on a straight path at a constant velocity by a  $\tau$ -like ratio of the objects' instantaneous acoustic intensity to its instantaneous rate of change in intensity (Jenison, 1997; Shaw, McGowan, & Turvey, 1991). This is often referred to as *auditory*  $\tau$ . However, as in vision, other auditory heuristic cues could potentially be used to estimate TTC in addition to auditory  $\tau$ . Our previous study (DeLucia et al., 2016) was the first to show that auditory judgments of TTC are significantly influenced by an auditory heuristic cue. Specifically, final sound pressure level ( $SPL_{final}$ ), a heuristic cue analogous to the final optical size in visual TTC estimation, was used to make TTC judgments. In the present study, we measured the influence of  $SPL_{final}$ , the final rate of change in sound pressure ( $\dot{p}_{final}$ ), and the change in sound pressure level across the presentation duration ( $\Delta SPL = SPL_{final} - SPL_{initial}$ ). We performed analyses to determine the extent to which auditory  $\tau$ ,  $SPL_{final}$ ,  $\dot{p}_{final}$ , and  $\Delta SPL$  were used in the TTC judgments.

The vast majority of previous TTC studies have examined either visual inputs alone or auditory inputs alone without considering how these cues are combined when both are available. Importantly, experiencing combined visual and auditory inputs is arguably most comparable to typical real-world conditions. In our previous study (DeLucia et al., 2016), we measured TTC judgments of an approaching object presented with visual inputs only (solid textured square on a blank background), auditory inputs only (1.0-kHz tone), or both concurrently. The relative weights of auditory and visual information, and of different sources of information within each modality (i.e., auditory or visual  $\tau$  vs. final optical size or final sound pressure level), were determined using multiple regression techniques. Several results are noteworthy. First, TTC estimates were shorter and closer to the veridical value in the auditory-only than in the visual-only condition. Second, combining auditory and visual inputs did not result in higher accuracy (deviation of estimates from the veridical value) or greater precision (variability of the estimates) than in the unimodal conditions. This result is consistent with Schiff and Oldak (1990; but see Zhou et al., 2007). However, the regression analyses showed that both auditory and visual information was used (Prime & Harris, 2010), although visual information was weighted more heavily than auditory information. Third, the relative use of  $\tau$  versus two heuristic-based cues (i.e., final sound pressure level and final optical size) differed between auditory and visual TTC judgments. Specifically, the use of heuristics was greater in the auditory-only condition than in the visual-only condition, demonstrating an auditory analog to the size-arrival effect.

In the present study, our experimental design allowed for the analysis of weights for additional heuristic cues including the final rate of optical expansion/sound level pressure and the change in visual angle and sound pressure level across the stimulus presentation. Some of these cues have been studied using visual stimuli (Gray & Regan, 1998; Yan et al., 2011), but not using auditory stimuli, or combined visual and auditory stimuli. Further, prior studies have used relatively simple stimuli and thus, it is not known whether the relative weighting of visual and auditory cues in TTC judgments, and the relative use of  $\tau$ -based information and heuristics cues, would be different under more naturalistic conditions. We will return to this point later.

The age of an observer is another important consideration when understanding the role of different sensory inputs in TTC judgments. Healthy older adults often experience age-related sensory, cognitive, and motor declines that may affect TTC estimates and may result in relative weightings of visual and auditory sensory inputs that differ from the relative weightings in younger adults. Indeed, differences in *visual* TTC judgments have been previously observed between younger and older adults (e.g., DeLucia, Bleckley, Meyer, & Bush, 2003; Hancock & Manser, 1997; Schiff, Oldak, &

Shah, 1992). However, age differences have not been considered in TTC judgments when auditory or concurrent visual and auditory information is presented.

When performing other simple auditory–visual detection and discrimination tasks, healthy older adults have been shown to exhibit heightened sensory integration demonstrated by, for instance, proportionally faster response times in bimodal than in unimodal stimulus presentations than do younger adults (Diederich & Colonius, 2004; Laurienti, Burdette, Maldjian, & Wallace, 2006; Mozolic, Hugenschmidt, Peiffer, & Laurienti, 2012). These age-related differences in multisensory integration may, however, depend on the nature and complexity of the task and the types of sensory inputs available (e.g., Freiherr, Lundström, Habel, & Reetz, 2013; McGovern, Roudaia, Stapleton, McGinnity, & Newell, 2014). Importantly, general functional implications are associated with possible age-related changes in TTC judgments with respect to navigation and collision avoidance during typical, daily mobility-related tasks such as walking and driving. Of particular note is that older pedestrians show less flexible judgments when making realistic street crossing decisions (e.g., Lobjois & Cavallo, 2009), and older drivers are overrepresented in serious collision statistics that occur due to driver error when making left hand turns across traffic (Cicchino & McCartt, 2015).

In prior research, quantitative estimates of the weighting of different sources of visual information in TTC estimates were obtained with only very basic visual stimuli (e.g., simple geometric shapes) without contextual or depth cues to scale these visual inputs. To address this issue in the present study, participants visually estimated the TTC of a smaller or a larger object (car vs. fire truck) that approached within a realistic simulated street scene. Similarly, auditory TTC estimates might be influenced by familiar acoustical properties of the stimulus (Jenison, 1997; Kolarik, Moore, Zahorik, Cirstea, & Pardhan, 2016), which were not varied in our previous TTC study that presented a pure tone (DeLucia et al., 2016). In the present study, we used familiar vehicle engine sounds to add more realism and additional acoustical information. Due to technical limitations, we varied acoustic intensity (but not reflected sound, spatial audio, or contextual sounds like traffic noise), which is probably the most important cue to TTC for sound sources approaching at a constant velocity on a straight path (Jenison, 1997).

The present study had two main objectives. First, we wanted to determine whether the pattern of results in our earlier study of auditory–visual integration in TTC judgments (DeLucia et al., 2016) would be observed with more complex, contextualized, and realistic stimuli. Toward this aim, we measured TTC estimates with visual, auditory, and combined audio–visual conditions using simulations of a traffic scene in downtown Toronto. We then used these TTC estimates to quantify the relative weights of visual and auditory inputs,

and to quantify the use of auditory  $\tau$ , visual  $\tau$  and several visual and auditory heuristic cues (e.g., final optical size/sound pressure level), including heuristic cues not considered in our earlier study (i.e., final rate of optical expansion/change in sound pressure and change in visual angle/sound pressure level across the presentation duration). Second, we examined age-related effects on TTC judgments by determining whether older adults demonstrate differences (as compared to younger adults) in (a) overall TTC judgments (e.g., over- vs. underestimation of TTC) and (b) the use of particular TTC cues ( $\tau$  vs. different visual and auditory heuristic cues).

## Method

### Participants

Forty-two adult volunteers participated (26 female, 16 male). Twenty-one participants were older (13 female,  $M_{\text{age}} = 70.31$  years,  $SD = 4.40$  years; eight male,  $M_{\text{age}} = 70.13$ ,  $SD = 2.53$ ), and 21 participants were younger (13 female,  $M_{\text{age}} = 27.31$ ,  $SD = 6.07$ ; eight male,  $M_{\text{age}} = 26.36$ ,  $SD = 2.83$ ). All were healthy (i.e., no self-reported vestibular disorders, recent history of brain injury, musculoskeletal disorder, or acute psychiatric disorders). Although all participants reported having normal or corrected-to-normal vision, 10% of younger participants and 23% of older participants still scored below average on a Snellen test for visual acuity (lower than 20/25) while wearing their corrective lenses. Despite poorer visual acuity in some cases, all participants were able to detect the incoming object at its farthest position. Overall, 90% of the older participants and 57% of the younger participants wore glasses or contact lenses.

Four of the older participants (19%) wore hearing aids, but they were not excluded from the data analyses.<sup>1</sup> A short auditory detection task prior to the test session ensured that older and younger participants could hear the sounds used in the experiment at the lowest volumes and at the same frequencies presented during the experiment. Specifically, during this detection task, each of the two quietest sounds (70 and 75 dB SPL) from the experiment was presented five times in a randomized order, and participants were asked to press a button as soon as they heard the sound. All of the participants successfully did so, but the older adults ( $M = 0.43$  s,  $SEM = 0.07$ ) showed significantly slower reaction times in general than did

<sup>1</sup> All statistical analyses on estimated TTCs and regression weights were also performed with participants with hearing aids excluded. The means were very similar to the means when the four hearing-aid users were included. Also, the analysis of variance results did not change considerably, except for a significant interaction between age group and TTC level in the analysis of estimated TTC that was present when hearing-aid users were excluded, but not when they were included. We therefore decided to keep the participants with hearing aids in the data analyses.

the younger adults ( $M = 0.35$  s,  $SEM = 0.08$ ),  $F(1, 39) = 12.808$ ,  $p = .001$ ,  $\eta_p^2 = .407$ . The study protocol was approved by University Health Network’s research ethics board and participants gave written informed consent prior to the experiment. Participants were compensated with \$10 per hour. Participants were naïve with respect to the purpose of the study.

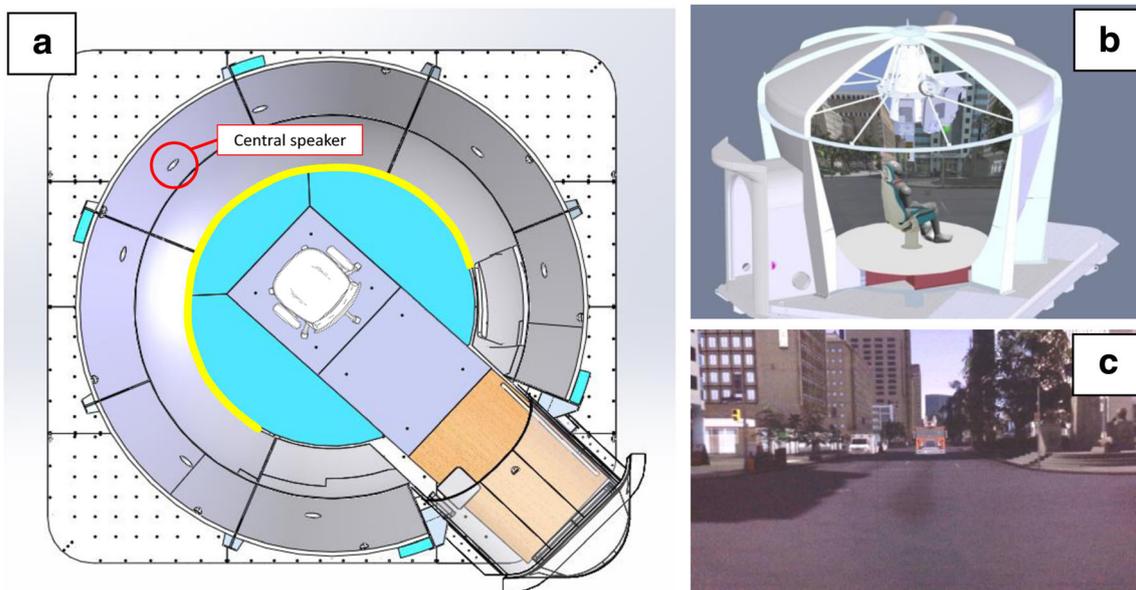
### Apparatus and stimuli

The study was conducted in StreetLab, a virtual-reality dome-shaped laboratory located at the Toronto Rehabilitation Institute’s Challenging Environment and Assessment Laboratory (see Fig. 1). Participants sat 100 cm in front of StreetLabs’s large, immersive, curved projection screen. Visual stimuli were created using OpenScene Graph and were displayed using six calibrated projectors (Eyevis ESP-LED series with LED technology), resulting in a field of view of  $240^\circ$  horizontally and  $105^\circ$  vertically. Picture resolution accounted for 6.5 arcmin per optical line pair. Auditory stimuli were presented from a speaker located at 0 deg azimuth (Meyersound MP-4XP) located behind the projection screen. The height of the chair was individually adjusted, resulting in an eye height of approximately 140 cm above the floor, which aligned the participant’s ear level with the height of the speaker and the eye level with the center of the image.

The virtual scene consisted of a simulation of a three-lane road in downtown Toronto (see Fig. 1, right panel) without environmental sounds (e.g., no other ambient traffic noise, etc.). While the participant remained in a stationary position, three different sensory conditions were used to simulate a

vehicle approaching at a constant speed on a collision course toward the participant: a visually simulated vehicle (visual-only), an auditory simulated vehicle (auditory-only), or a vehicle that was simulated with both auditory and visual inputs (audio-visual).

In the *visual-only condition*, the stimulus was either a smaller vehicle (Audi A4) or a larger vehicle (fire truck). The dimensions of the vehicles were  $2.0\text{ m} \times 1.3\text{ m}$  (Width  $\times$  Height) and  $2.4\text{ m} \times 2.9\text{ m}$  for the Audi and fire truck, respectively. The optical size of the vehicle increased as it approached, providing information about the vehicle’s TTC through the optical invariant, tau ( $\tau$ ). Before the vehicle arrived at the participant’s location (i.e., before it virtually “collided” with the participant), it disappeared from the display. The time between the vehicle’s disappearance and the collision with the participant defined the actual *time to contact* (TTC), which varied across trials. Given that TTCs in the range of 0.5–3.5 s are typically encountered in real driving situations (car following) at speeds slower than 40 mph (Kusano, Chen, Montgomery, & Gabler, 2015), nine different visual TTCs were tested here that covered that range (0.7, 1.0, 1.3, 1.6, 2.0, 2.4, 2.8, 3.2, and 3.6 s). Vehicle size and presentation duration were varied in order to reduce the correlations between such cues as final optical size and final rate of expansion (cf. Gray & Regan, 1998; Oberfeld, Hecht, & Landwehr, 2011), so that the observer’s weightings of these cues could be more easily estimated in the analyses. Consequently, the starting distance of the vehicle varied across trials on the basis of presentation duration, vehicle speed, and TTC level (see the Study Design section for details). Other features of the virtual scene (e.g., other cars or people) did not move.



**Fig. 1** (A) Bird’s eye view of the laboratory. The thick yellow line represents the screen. The center speaker behind the screen is circled. The chair is positioned in front of the screen. (B) 3-D view of

StreetLab, with the chair in the center of the lab. (C) Screenshot of the visual stimulus, showing the larger vehicle (fire truck) approaching the observer on a collision course.

In the *auditory-only condition*, the projector screens were all blackened and participants kept their eyes closed. The auditory stimulus was a monaural recording of a car's engine sound that increased in acoustic intensity upon approach, simulating a vehicle that approached the observer head-on at a constant speed in the acoustic far field and in the absence of reflected sound in the simulated scene. The sound pressure of the auditory object as a function of its relative distance from the listener followed an inverse law, with a loss of approximately 6 dB ( $20 \log_{10} 2$ ) in sound pressure for each doubling of distance (Zahorik, 2002). Thus, information about the TTC of the auditory object was available in terms of the  $\tau$ -like ratio  $I(t)/\dot{I}(t)$ , where  $I(t)$  is the instantaneous acoustic intensity at time  $t$ , and  $\dot{I}(t)$  is the derivative of  $I(t)$  with respect to time (Shaw et al., 1991). As in the visual condition, the sound of the car stopped before arriving at the participant's location (i.e., before virtual collision). Again, two different engine sound levels (quieter [70 dB SPL] vs. louder [85 dB SPL]) were selected to reduce the correlation between the auditory TTC and final SPL (analogous to the variations in vehicle size in the visual-only condition). This resulted in engine sounds with either 70 or 85 dB SPL (including ambient noise of approximately 45 dB SPL) at a virtual distance of 4 m from the observer. There were nine different auditory TTCs (0.7, 1.0, 1.3, 1.6, 2.0, 2.4, 2.8, 3.2, and 3.6 s). Again, the starting distances of the sound varied depending on presentation duration, vehicle speed, and the different TTC levels.

In the *audio-visual condition*, the visual and auditory cues described above were presented simultaneously and provided combined information about the simulated vehicle. In some trials the auditory and visual stimuli provided congruent information, whereas in other trials those stimuli were offset from each other. Specifically, on some trials the visual stimulus arrived ahead of the auditory stimulus, and on other trials the visual stimulus arrived after the auditory stimulus (see the details in the Study Design section). This allowed us to use regression analyses to estimate the relative weights of the visual and auditory information during bimodal trials (DeLucia et al., 2016).

## Study design

In the *visual-only condition*, 72 unique visual stimuli were created by factorially by crossing nine values of TTC (0.7, 1.0, 1.3, 1.6, 2.0, 2.4, 2.8, 3.2, and 3.6 s), two values of vehicle speed (25 and 40 km/h [15 and 25 mph]), two vehicle sizes (small and large), and two presentation durations (1.5 and 3.0 s). Each stimulus was presented two times, resulting in 144 trials for the visual-only condition.

In the *auditory-only condition*, 72 unique auditory stimuli were created by factorially crossing nine values of TTC (0.7, 1.0, 1.3, 1.6, 2.0, 2.4, 2.8, 3.2, and 3.6 s), two values of vehicle

speed (25 and 40 km/h [15 and 25 mph]), two values of sound source intensity (70 and 85 dB SPL presented 4 m from the participant), and two presentation durations (1.5 and 3.0 s). Each unique stimulus was presented two times, resulting in 144 trials for the auditory-only condition.

In the *audio-visual condition*, when the auditory ( $TTC_a$ ) and the visual ( $TTC_v$ ) inputs were *congruent*, the same nine levels of TTC were included (0.7, 1.0, 1.3, 1.6, 2.0, 2.4, 2.8, 3.2, and 3.6 s). For *incongruent* trials,  $TTC_a$  and  $TTC_v$  were varied independently, resulting in the visual TTC being either shorter than the auditory TTC ( $TTC_v < TTC_a$ ) or longer than the auditory TTC ( $TTC_v > TTC_a$ ). In other words, across different trials the visual stimulus either "arrived" earlier or later than the auditory stimulus. Three of the nine TTCs were combined to three TTC levels (TTC level 1.0 s = TTCs 0.7, 1.0, and 1.3 s; TTC level 2.0 s = TTCs 1.6, 2.0, and 2.4 s; TTC level 3.2 s = TTCs 2.8, 3.2, and 3.6 s). Only the auditory and visual TTCs from a single TTC level were combined, so that the auditory and visual stimuli were perceived as belonging to a single object rather than to two separate objects, which might occur at large differences between auditory and visual TTC. Because the sensitivity for discriminating TTCs is slightly lower at shorter than at longer TTCs (Oberfeld & Hecht, 2008), a higher proportional difference in visual versus auditory arrival times was selected for the shorter than for the longer TTC levels (i.e., 30% deviation at TTC level 1.0, 20% deviation at TTC level 2.0, and 12.5% deviation at TTC level 3.2).

Consequently, in the audio-visual condition, 432 unique stimuli were created by factorially crossing three values of TTC level (1.0, 2.0, and 3.2 s), nine combinations of  $TTC_v$  and  $TTC_a$  within each TTC level (see Table 1), two values of vehicle speed (25 and 40 km/h [15 and 25 mph]), two vehicle sizes (small and large), two values of sound source intensity (70 and 85 dB SPL presented 4 m from the participant), and two presentation durations (1.5 and 3.0 s). Each unique stimulus was presented once. Table 1 shows all combinations of visual and auditory TTCs.

The trials were divided into ten blocks, each with 72 trials. Each block contained only trials from a single sensory condition (i.e., visual-only, auditory-only, or audio-visual). This resulted in two blocks for the visual-only stimuli, two blocks for the auditory-only stimuli, and six blocks for the audio-visual stimuli. Two fixed block orders were chosen, and participants were randomly assigned to one of the two block orders, with the same number of participants in each age group assigned to the same block order. Due to the large number of trials (720), participants were tested across two different days with at least 24 h between the test sessions, to prevent fatigue and carryover effects. All participants completed a total of 720 trials across all three sensory conditions (visual-only, auditory-only, and audio-visual). Age Group was a between-subjects factor (younger [18–39] and older [65+]).

**Table 1** Time-to-contact (TTC) differences between the visual TTC ( $TTC_v$ ) and the auditory TTC ( $TTC_a$ ) in the audio–visual condition (i.e.,  $\Delta TTC = TTC_v - TTC_a$ )

| TTC level | $TTC_v$ | $TTC_a$ | $TTC_v$ vs. $TTC_a$ | $\Delta TTC$ (s) |
|-----------|---------|---------|---------------------|------------------|
| 1.0s      | 0.7s    | 0.7s    | $TTC_v = TTC_a$     | 0.0s             |
|           |         | 1.0s    | $TTC_v < TTC_a$     | -0.3s            |
|           |         | 1.3s    | $TTC_v < TTC_a$     | -0.6s            |
|           | 1.0s    | 0.7s    | $TTC_v > TTC_a$     | +0.3s            |
|           |         | 1.0s    | $TTC_v = TTC_a$     | 0.0s             |
|           |         | 1.3s    | $TTC_v < TTC_a$     | -0.3s            |
|           | 1.3s    | 0.7s    | $TTC_v > TTC_a$     | +0.6s            |
|           |         | 1.0s    | $TTC_v > TTC_a$     | +0.3s            |
|           |         | 1.3s    | $TTC_v = TTC_a$     | 0.0s             |
| 2.0s      | 1.6s    | 1.6s    | $TTC_v = TTC_a$     | 0.0s             |
|           |         | 2.0s    | $TTC_v < TTC_a$     | -0.4s            |
|           |         | 2.4s    | $TTC_v < TTC_a$     | -0.8s            |
|           | 2.0s    | 1.6s    | $TTC_v > TTC_a$     | +0.4s            |
|           |         | 2.0s    | $TTC_v = TTC_a$     | 0.0s             |
|           |         | 2.4s    | $TTC_v < TTC_a$     | -0.4s            |
|           | 2.4s    | 1.6s    | $TTC_v > TTC_a$     | +0.8s            |
|           |         | 2.0s    | $TTC_v > TTC_a$     | +0.4s            |
|           |         | 2.4s    | $TTC_v = TTC_a$     | 0.0s             |
| 3.2s      | 2.8s    | 2.8s    | $TTC_v = TTC_a$     | 0.0s             |
|           |         | 3.2s    | $TTC_v < TTC_a$     | -0.4s            |
|           |         | 3.6s    | $TTC_v < TTC_a$     | -0.8s            |
|           | 3.2s    | 2.8s    | $TTC_v > TTC_a$     | +0.4s            |
|           |         | 3.2s    | $TTC_v = TTC_a$     | 0.0s             |
|           |         | 3.6s    | $TTC_v < TTC_a$     | -0.4s            |
|           | 3.6s    | 2.8s    | $TTC_v > TTC_a$     | +0.8s            |
|           |         | 3.2s    | $TTC_v > TTC_a$     | +0.4s            |
|           |         | 3.6s    | $TTC_v = TTC_a$     | 0.0s             |

## Procedure

Participants completed a prediction-motion task (Schiff & Detwiler, 1979). Specifically, they pressed a joystick button to indicate the point in time at which the virtually approaching vehicle (visual or auditory) would “collide” with them, had it continued to approach them with a constant speed after it disappeared. The TTC estimation was defined as the time between the disappearance of the vehicle and the participant’s button response. If the participant did not press the button, the trial was terminated by the program after 10 s, and the trial was excluded from the data analysis and not repeated.

At the beginning of the first test session, participants performed the visual acuity test and the baseline auditory detection task. A practice block was used to familiarize participants with the experimental procedure. The practice block contained 24 randomly selected trials from each of the visual-only, auditory-only, and audio–visual conditions, resulting in a total of 72 practice trials. Feedback about participants’ TTC judgments was not given during either the practice or the

experimental blocks. After practice, four of the ten experimental blocks were presented in Session 1. In the second test session, participants finished the six experimental blocks that had not been presented in the first session. Again, a brief practice block, consisting of four trials from each sensory condition (total of 12 trials), was presented prior to the beginning of the second session block.

## Results

Our analyses focused on the relative weightings of auditory and visual  $\tau$  and of each of the three auditory and four visual heuristic cues. We will report regression analyses for the two auditory and visual unimodal conditions and for the audio–visual condition. However, in the first section of the Results, we will present analysis-of-variance (ANOVA) results comparing the average TTC estimates in the unimodal (visual-only, auditory-only) and the congruent audio–visual conditions for younger and older adults. This allowed for

comparisons of TTC judgment accuracy across age groups in each sensory condition.

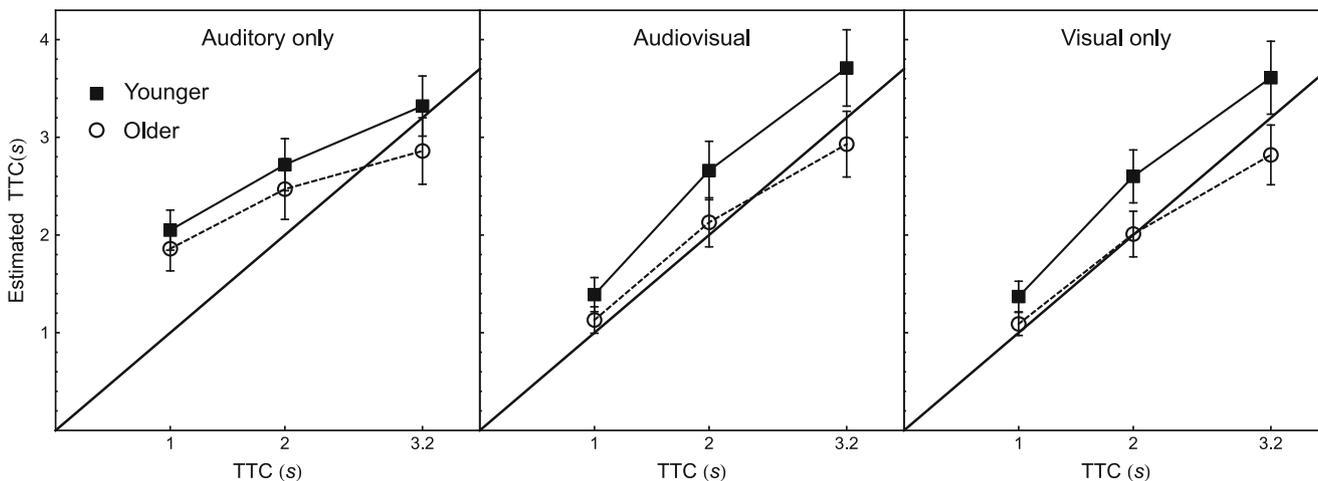
**Average TTC estimates**

Figure 2 shows the average TTC estimates ( $TTC_{est}$ ) for each age group at each TTC level. TTC estimates were averaged across visual vehicle sizes and sound intensities to allow for direct comparisons among the three sensory conditions. Within each TTC level (1.0, 2.0, and 3.2 s), the TTC estimates were also averaged across the variations of each TTC level (e.g., 0.7, 1.0, and 1.3 s for the TTC level of 1.0 s; see Table 1). Only those trials with congruent visual and auditory TTCs were selected for the analysis of the audio–visual condition, so that we could directly compare the TTC estimates among the three sensory conditions. A repeated-measures ANOVA with a univariate approach was conducted, including the within-subjects factors Sensory Condition (visual-only, auditory-only, and audio–visual), TTC Level (1.0, 2.0, and 3.2 s), Presentation Duration (1.5 and 3.0 s), and Speed (25 and 40 km/h). Age Group (younger and older) was a between-subjects factor. The degrees of freedom were Huynh–Feldt corrected to account for violations of sphericity (Huynh & Feldt, 1976; Lecoutre, 1991), and the correction factor  $\tilde{\epsilon}$  is reported.

**Sensory condition** We found a significant effect of TTC level,  $F(2, 80) = 168.73, \tilde{\epsilon} = .530, p < .001, \eta_p^2 = .808$ , such that the mean TTC estimates increased approximately linearly with actual TTC level (see Fig. 2). The effect of age group was not significant,  $F(1, 40) = 1.621, p = .210, \eta_p^2 = .039$ , but a significant effect of sensory condition was observed,  $F(2, 80) = 10.15, \tilde{\epsilon} = .620, p = .001, \eta_p^2 = .202$ . Simple contrast comparisons revealed longer estimated TTCs for the auditory-only condition than for the audio–visual condition,  $F(1, 40) = 7.396, p = .010, \eta_p^2 = .156$ , and the visual-only condition,

$F(1, 40) = 13.489, p = .001, \eta_p^2 = .252$ . Mean TTC estimates were significantly shorter for the visual-only than for the audio–visual condition,  $F(1, 40) = 6.361, p = .016, \eta_p^2 = .137$ . Additionally, the interaction between TTC level and sensory condition was significant,  $F(4, 160) = 63.790, \tilde{\epsilon} = .444, p < .001, \eta_p^2 = .615$ . To further disentangle this interaction, repeated measures ANOVAs (rmANOVAs) with post-hoc single comparisons (Bonferroni corrected) were calculated for each TTC level separately (1.0, 2.0, and 3.2 s) to compare TTC estimates between the sensory conditions (auditory-only vs. visual-only vs. audio–visual). The results showed that for the TTC level of 1.0 s, TTC estimates were larger in the auditory-only condition than in either the visual-only condition ( $p < .001$ ) or the audio–visual condition ( $p < .001$ ), but the difference between the visual-only and audio–visual conditions was not significant ( $p = .709$ ). For the TTC level of 2.0 s, TTC estimates were larger in the auditory-only than in the visual-only condition ( $p = .007$ ), but differences were not significant between the auditory-only and audio–visual conditions ( $p = .090$ ), or between the visual-only and audio–visual conditions ( $p = .075$ ). For the TTC level of 3.2 s, no significant differences were apparent among the sensory conditions, meaning that TTC estimates were significantly larger in the auditory-only condition than in the visual-only and audio–visual conditions only for the TTC levels of 1.0 and 2.0 s, but not for the TTC level of 3.2 s.

**Speed** A significant effect of speed emerged in the ANOVA,  $F(1, 40) = 215.237, p < .001, \eta_p^2 = .843$ , indicating shorter mean TTC estimates when the vehicles approached the observer with slower speed. This finding is consistent with previous studies (e.g., Oberfeld & Hecht, 2008; Petzoldt, 2014), and can be attributed to either a size-arrival effect (DeLucia, 1991) or a distance bias (Law et al., 1993), because the slower speed condition had a larger final optical vehicle



**Fig. 2** Mean time-to-contact (TTC) estimates for each TTC level, separated by sensory condition for older (circles) and younger (boxes) adults. Error bars indicate plus and minus one standard error of the mean (SEM).

size and a closer final distance than did the faster condition. A significant interaction between TTC level and speed,  $F(2, 80) = 27.397$ ,  $\tilde{\varepsilon} = 1.000$ ,  $p < .001$ ,  $\eta_p^2 = .407$ , indicated larger differences in TTC estimates between the two speeds with increasing actual TTC (see Fig. 3).

**Presentation duration** The interaction between sensory condition and presentation duration was significant,  $F(2, 80) = 19.082$ ,  $\tilde{\varepsilon} = .958$ ,  $p < .001$ ,  $\eta_p^2 = .323$ , demonstrating that a longer presentation duration resulted in shorter TTC estimates only in the auditory-only condition (see Fig. 4), and not in the visual-only and audio–visual conditions (but see McLeod & Ross, 1983). Finally, a significant four-way interaction among sensory condition, TTC level, speed, and age group was found,  $F(4, 160) = 2.668$ ,  $\tilde{\varepsilon} = 1.000$ ,  $p = .034$ ,  $\eta_p^2 = .067$ , for which we have no explanation. No other effects were significant ( $p$  values ranging from .067 to .903). In particular, there were no other significant interactions involving age group.

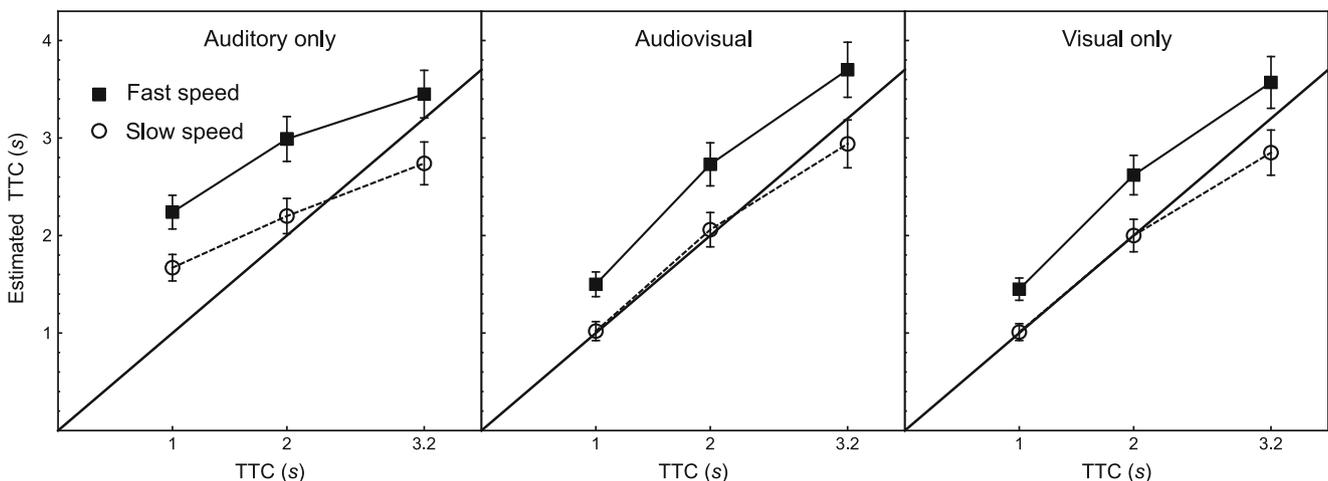
### Cue weights

In each of the three sensory conditions, the display contained several different sources of information that may have been used to make TTC judgments, including both accurate cues to TTC (i.e., the  $\tau$  variables) as well as other, heuristic cues (e.g., final optical size or final SPL). We estimated the degree to which the different cues contributed to the TTC estimates by means of multiple linear regression analyses conducted for each of the three sensory conditions independently. The estimated TTC was entered as the criterion (dependent variable). The multiple linear regression model contained an intercept term and the predictors  $TTC_a$  and  $TTC_v$ . Note that the acoustic  $\tau$ - variable ( $I(t)/\dot{I}(t)$ ) is proportional to  $TTC_a$

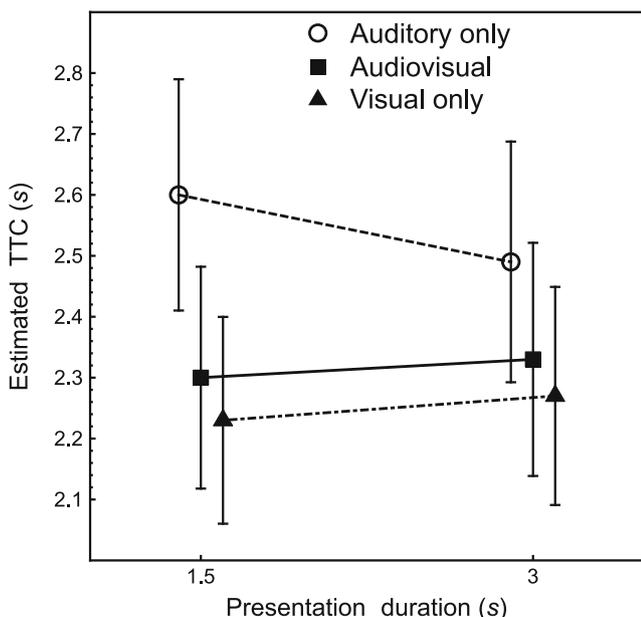
(i.e., perfectly correlated with  $TTC_a$ ) (Jenison, 1997). Also, the optical  $\tau$ - variable ( $\theta_{final}/\dot{\theta}_{final}$ ) was almost perfectly correlated with  $TTC_v$  for our stimuli. As additional visual predictors, the regression model also contained the inverse of the final optical size,  $\theta_{final}^{-1}$  (i.e., optical size on the final frame in degrees of visual angle) (e.g., DeLucia, 1991), the inverse of the final rate of optical expansion,  $\dot{\theta}_{final}^{-1}$ , and the inverse of the change in visual angle across the presentation duration,  $\Delta\theta^{-1} = 1/(\theta_{final} - \theta_{initial})$ . Because the simulated visual scene contained many monocular depth cues (e.g., occlusion, relative height, and familiar size; see the lower right panel of Fig. 1), the final visual distance of the stimulus ( $D_{final,v}$ ) was also included as a predictor (Yan et al., 2011). As additional auditory predictors, the model contained the inverse of the final sound pressure level ( $SPL_{final}^{-1}$ ; as was done in DeLucia et al., 2016), the inverse of the final rate of change in sound pressure,  $\dot{p}_{final}^{-1}$ , and the inverse of the change in sound pressure level across the presentation duration,  $\Delta SPL^{-1} = 1/(SPL_{final} - SPL_{initial})$ . The inverse values were used for several cues because they are positively and more linearly related to the actual TTCs. In the visual-only and auditory-only conditions, only the predictors representing the corresponding sensory modality were included, whereas for the audio–visual condition, both auditory and visual predictors were included. All predictors were  $z$ -standardized and entered simultaneously. Separate regression models were fitted for each combination of participant and sensory condition. The individual estimates were then analyzed using repeated-measures ANOVAs and related tests.

### Visual-only condition

In the condition providing only visual information, the five predictors included in the multiple regression model were



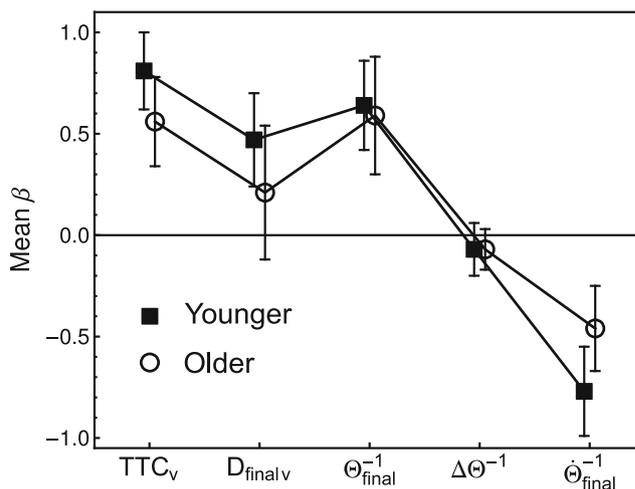
**Fig. 3** Mean TTC estimates for each TTC level, shown separated by sensory condition and speed (squares: fast, circles: slow). Error bars indicate  $\pm 1$  SEM.



**Fig. 4** Mean time-to-contact (TTC) estimates, shown separated by sensory condition and presentation duration. The lines are shifted horizontally to prevent overlap of the error bars, which show  $\pm 1$  SEM. circles, auditory only; squares, audiovisual; triangles, visual only

$TTC_v$ ,  $D_{final\ v}$ ,  $\theta_{final}^{-1}$ ,  $\dot{\theta}_{final}^{-1}$ , and  $\Delta\theta^{-1}$ . Following the recommendations by Belsley, Kuh, and Welsch (1980), we analyzed the externally studentized residuals and the DFFITS index proposed by Belsley et al. (1980) as a measure of the influence of an observation. For each individual regression model, externally studentized residuals with an absolute value exceeding 1.96 or with an absolute DFFITS value exceeding  $2\sqrt{p/N}$  were defined as outliers (Belsley et al., 1980), where  $N$  is the number of trials collected in the visual condition and  $p = 5$  is the number of predictors. Across participants, this resulted in the exclusion of between 6% and 18% of the trials ( $M = 12.9\%$ ,  $SD = 2.6\%$ ). Q-Q plots of the residuals showed no systematic deviations from normality, and plots of  $TTC_{est}$  as a function of the predictors showed no severe deviations from linearity. Thus, linear multiple regression was an appropriate method to assess the influence of the predictors on the TTC estimates. Across participants, the regression models showed a good fit, with  $R^2$  ranging from .63 to .93 ( $M = .83$ ,  $SD = .08$ ) for the younger group and from .49 to .94 ( $M = .81$ ,  $SD = .13$ ) for the older group.

Figure 5 shows the mean regression coefficients for the five visual cues. As can be seen from the confidence intervals (CIs), the average regression coefficients for all visual cues except those for the change in visual angle across presentation durations (older and younger groups) and final distance (older group only) were significantly different from 0. Thus, as expected, both the accurate cue ( $TTC_v$ ) and heuristic cues such as the final optical size were significantly associated with the TTC estimates. This finding is compatible with data from DeLucia et al. (2016), obtained in both an audio–visual and



**Fig. 5** Visual-only condition. Mean estimated regression coefficients for the five visual cues. Error bars show 95% confidence intervals (CIs).  $TTC_v$ , visual TTC;  $D_{final\ v}$ , final distance;  $\theta_{final}^{-1}$ , inverse of the final optical size;  $\Delta\theta^{-1}$ , inverse of the change in visual angle across the presentation duration;  $\dot{\theta}_{final}^{-1}$ , inverse of the final rate of expansion; squares, younger adults; circles, older adults

a visual-only condition, whereas Gray and Regan (1998) reported that the final optical size, final rate of optical expansion and the change in visual angle across the presentation duration showed no relation to TTC estimates in a visual-only condition. As expected, the TTC estimates increased with the actual visual TTC, with the final visual distance (Law et al., 1993), and with the inverse of the final optical size (DeLucia, 1991).

A surprising result was that, on average, a negative coefficient was estimated for the inverse final optical rate of expansion. Thus, according to the regression model, fast final rates of optical expansion were associated with longer TTC estimates. We believe that this counterintuitive result is due to the correlations among the predictors. In such a case, so-called sign reversals can occur in a regression model (e.g., DeCarlo & Cross, 1990; Greene, 2008, p. 59). Because we presented simulations of realistic traffic situations, there were substantial correlations between some of the visual cues (see Table 2). In contrast, in experiments using more abstract, laboratory types of stimuli, these correlations can be reduced by providing very limited monocular depth cues or scaling information (Gray & Regan, 1998; Regan & Hamstra, 1993).

To gain insight into the conditions in which a negative regression coefficient for the inverse final rate of optical expansion occurred, we fitted models containing all possible subsets of the five predictors. In models containing only the inverse final rate of optical expansion, the regression coefficient was not negative for any of the 42 participants. In fact, the average bivariate correlation between inverse final rate of optical expansion and estimated TTC was positive ( $M = .77$ ,  $SD = .10$ ,  $p < .001$ ). If the model contained only  $TTC_v$  and the inverse final rate of optical expansion, a negative coefficient was found for only two of the 42 participants. Also, if  $TTC_v$

**Table 2** Visual-only condition: Pairwise Pearson correlation coefficients between the five predictors ( $z$ -standardized), across the 144 trials

|                                    | $D_{\text{final } v}$ | $\theta_{\text{final}}^{-1}$ | $\dot{\theta}_{\text{final}}^{-1}$ | $\Delta\theta^{-1}$ |
|------------------------------------|-----------------------|------------------------------|------------------------------------|---------------------|
| TTC <sub>v</sub>                   | .781                  | .769                         | .859                               | .794                |
| $D_{\text{final}}$                 |                       | .984                         | .953                               | .924                |
| $\theta_{\text{final}}^{-1}$       |                       |                              | .961                               | .934                |
| $\dot{\theta}_{\text{final}}^{-1}$ |                       |                              |                                    | .946                |

All coefficients are significantly different from 0 ( $p < .001$ ).

was not included in the model, a negative coefficient for inverse final rate of optical expansion was present in only 29 of the 294 models (10%). In contrast, if the models contained the inverse final rate of optical expansion, TTC<sub>v</sub>, and at least one additional predictor, a negative sign for the inverse final rate of optical expansion occurred in 83.3% of the cases. The negative sign occurred most frequently when the model contained the inverse final rate of optical expansion and TTC<sub>v</sub>, plus either the inverse final rate of optical expansion, or final visual distance, or both. In conclusion, additional research using a design with reduced correlations between final rate of optical expansion and the other cues seems necessary to answer the question of whether the negative relation between the inverse final rate of optical expansion and the TTC estimate reflects a systematic TTC estimation strategy rather than just a statistical phenomenon.

What can be concluded about the extents to which participants relied on the different visual cues when making their TTC estimates? In the case of correlated predictors, it can be misleading to gauge relative importance of the predictors by considering the squared standardized regression coefficients (cf. Tonidandel & LeBreton, 2011). For this reason, we used the “dominance analysis” approach proposed by Budescu (1993), which has been shown to be a useful measure of the relative importances of predictors in a regression model, both on theoretical grounds and in simulation studies (LeBreton, Ployhart, & Ladd, 2004; Thomas, Zumbo, Kwan, & Schweitzer, 2014; Tonidandel & LeBreton, 2011). Dominance analysis provides a quantitative measure of relative importance through an examination of changes in the variance accounted for ( $\Delta R^2$ ) resulting from adding a predictor to all possible regression models containing subsets of the predictors. For example, if there are three predictors (A, B, and C), then predictor C can be added to four possible subset models (i.e., containing only the intercept term, intercept and predictor A, intercept and predictor B, and intercept and predictors A and B, respectively). A predictor’s *general dominance weight* (GDW; Azen & Budescu, 2003) is found by averaging the squared semipartial correlations across all of

the possible subset models. This measure indexes a variable’s contribution to the prediction of the dependent variable, by itself and in combination with the other predictors. The sum of the GDWs is the total proportion of variance explained by the regression model,  $R^2$ .

As is shown in Fig. 6, the GDW was highest for TTC<sub>v</sub>, as expected. However, the heuristic cues also explained a significant portion of the variance, compatible with the dominance analysis results by DeLucia et al. (2016). The change in visual angle across the presentation duration was the least important predictor. Because the total proportions of variance accounted for ( $R^2$ ), and thus also the mean GDWs, differed slightly between the two age groups, the five GDWs were normalized to a sum of 1.0 for each participant for the comparison of relative importances. An rmANOVA with the within-subjects factor Cue and the between-subjects factor Age Group showed a significant effect of cue on the normalized GDWs,  $F(4, 156) = 24.60$ ,  $\tilde{\epsilon} = .325$ ,  $p < .001$ ,  $\eta_p^2 = .387$ , but no significant Group  $\times$  Cue interaction,  $F(4, 156) = 0.70$ ,  $p = .44$ . Thus, the relative importances of the five visual cues for the TTC estimates did not differ significantly between the two age groups.

#### Auditory-only condition

In the auditory-only condition, the multiple regression model contained the predictors TTC<sub>a</sub>,  $SPL_{\text{final}}^{-1}$ ,  $\dot{p}_{\text{final}}^{-1}$ , and  $\Delta SPL^{-1}$ . Applying the same outlier criteria as described above resulted in the exclusion of between 5% and 14% of the trials ( $M = 11\%$ ,  $SD = 2\%$ ). Across participants, the regression models showed a somewhat inferior fit as compared to the visual-only condition, with  $R^2$ s ranging between .41 and .87 ( $M = .70$ ,  $SD = .14$ ) for the younger group and between .24 and .90 ( $M = .70$ ,  $SD = .19$ ) for the older group.

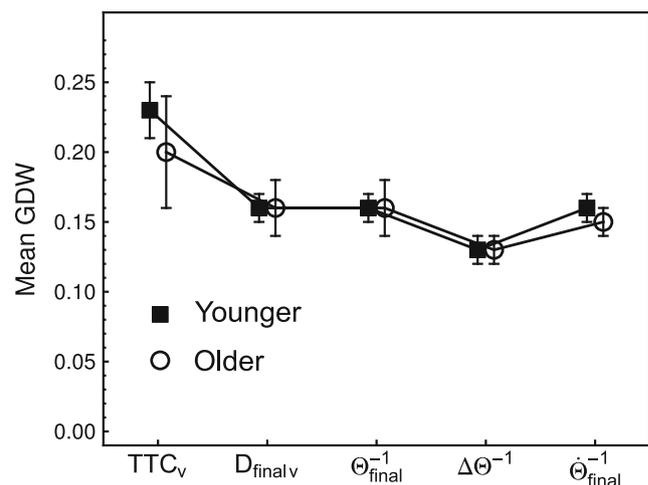
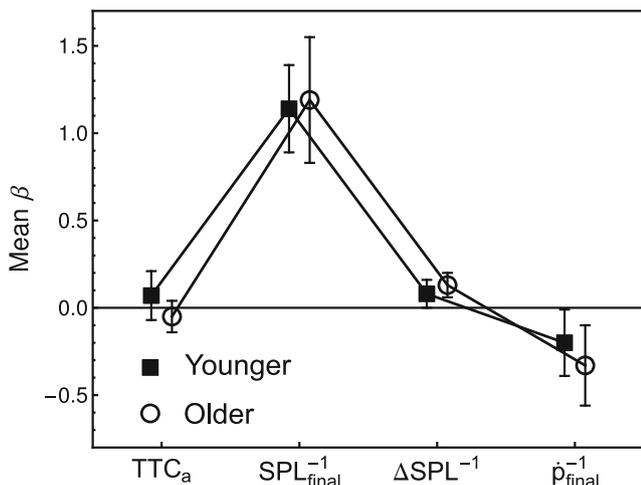
**Fig. 6** Visual-only condition: Mean general dominance weights (GDWs) for the five visual cues. Error bars show 95% confidence intervals (CIs). squares, younger adults; circles, older adults

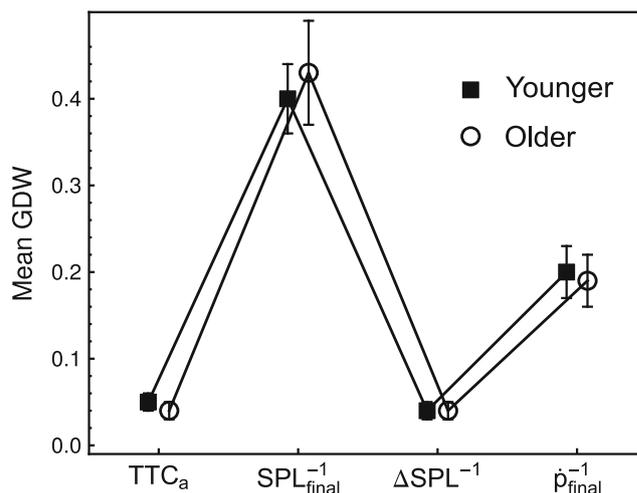
Figure 7 shows the mean regression coefficients for the four predictors. The data show a striking difference from the visual-only condition, because the mean regression coefficient for the exact cue,  $TTC_a$ , was not significant. Instead, we observed a high positive regression coefficient for the final sound pressure level, compatible with the data from DeLucia et al. (2016). The regression coefficient for change in sound pressure level was small (although significantly different from 0 for the older group), and the final rate of change in sound pressure received a comparably small but significant negative weight, similar to what was found for the inverse final expansion rate in the visual-only condition. The two age groups showed very similar patterns of regression weights, as is shown in Fig. 7.

The GDWs, displayed in Fig. 8, clearly demonstrate that the participants relied primarily on the final sound pressure level when making the TTC estimates for the auditory object. The mean GDW for  $TTC_a$  was close to 0, and the change in SPL also played no important role in the TTC estimates. Only the final rate of change in sound pressure explained an additional portion of the variance.

To compare the relative uses of the four cues between the two age groups, the GDWs were again normalized as described above. An rmANOVA with the within-subjects factor Cue and the between-subjects factor Age Group showed a significant effect of cue on the normalized GDWs,  $F(3, 120) = 687.7, \hat{\epsilon} = .745, p < .001, \eta_p^2 = .95$ , but no significant Age Group  $\times$  Cue interaction,  $F(3, 120) = 2.25, p = .105$ . Thus, the relative importance of the four auditory cues for the TTC estimates did not differ significantly between the two age groups.



**Fig. 7** Auditory-only condition: Mean regression coefficients for the four auditory cues. Error bars show 95% confidence intervals (CIs).  $TTC_a$ , auditory TTC;  $SPL_{final}^{-1}$ , inverse of the final sound pressure level;  $\Delta SPL^{-1}$ , inverse of the change in SPL across the presentation duration;  $\dot{p}_{final}^{-1}$ , inverse of the final rate of change in sound pressure; squares, younger adults; circles, older adults.

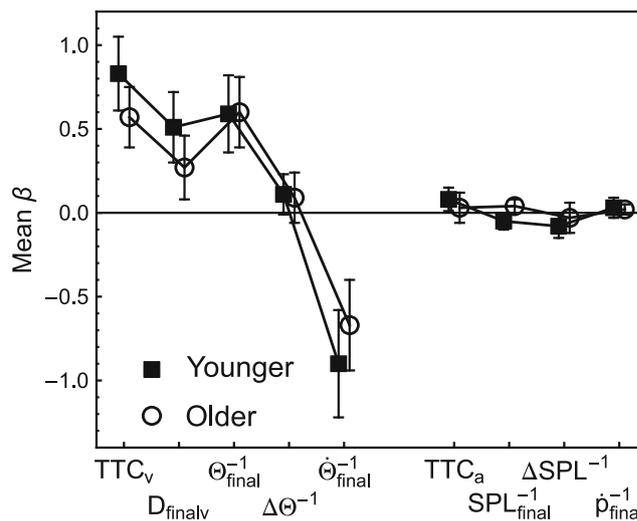


**Fig. 8** Auditory-only condition: Mean general dominance weights (GDWs) for the four auditory cues. Error bars show 95% confidence intervals (CIs). squares, younger adults; circles, older adults

**Audio–visual condition**

How do older and younger adults integrate auditory and visual cues when making TTC estimates in a realistic traffic scenario? To answer this question, the regression analyses for the audio–visual condition contained the five visual and four auditory cues as predictors. Applying the same outlier criteria as above resulted in the exclusion of between 2% and 14% of the trials ( $M = 10\%, SD = 2\%$ ). Across participants, the regression models showed a good fit, with  $R^2$  ranging between .54 and .93 ( $M = .81, SD = .09$ ) for the younger group, and between .53 and .92 ( $M = .78, SD = .12$ ) for the older group.

Figure 9 shows the mean regression coefficients for the nine predictors. As can be seen from the CIs, on average the

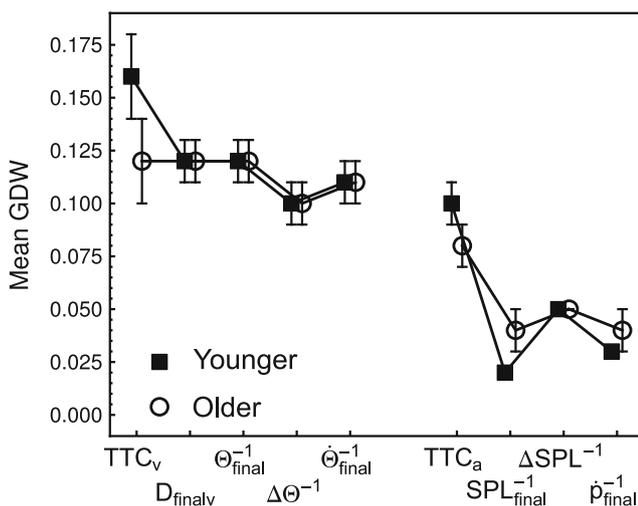


**Fig. 9** Audio–visual condition: Mean estimated regression coefficients for the five visual and four auditory predictors of the TTC estimates. Error bars show 95% confidence intervals (CIs). squares, younger adults; circles, older adults

regression coefficients for all visual cues except the change in visual angle were significantly different from 0 and showed a pattern similar to the one in the visual-only condition. In contrast, the regression coefficients for the auditory cues were all close to 0, although in the younger group the coefficients for  $TTC_a$  and the regression coefficients for final sound pressure level and change in sound pressure level across the presentation duration were significant, whereas for the older group none of the auditory predictors were significant. This pattern differs markedly from the cue weights observed in the audio–visual condition of our previous study (DeLucia et al., 2016), in which the association between estimated TTC and final sound pressure level was much stronger. As we will discuss below, one likely explanation for these diverging results is that in the present study we combined relatively sophisticated visual simulations with very basic auditory simulations containing no spatial audio, simulated reflected sound, or contextual environmental sounds.

The GDWs, displayed in Fig. 10, confirm the conclusion that the TTC estimates in the audio–visual condition were dominated by visual cues. Upon visual inspection, among the auditory cues, only the GDW for  $TTC_a$  was comparable in level to the GDWs of the visual cues, which is somewhat surprising, given the higher reliance on final SPL than on  $TTC_a$  that we found in the auditory-only condition.

An rmANOVA showed a significant effect of cue on the normalized GDWs,  $F(8, 320) = 170.6$ ,  $\tilde{\epsilon} = .357$ ,  $p < .001$ ,  $\eta_p^2 = .81$ . Unlike in the auditory-only and visual-only conditions, the Age Group  $\times$  Cue interaction was also significant,  $F(8, 320) = 5.20$ ,  $p = .002$ ,  $\eta_p^2 = .115$ . This was presumably due to the GDWs for the accurate cues ( $TTC_a$  and  $TTC_v$ ) appearing to be smaller in the older than in the younger age group, meaning that the older adults relied less on accurate cues and more on heuristic-based cues than did the younger adults.



**Fig. 10** Audio–visual condition: Mean GDWs for the five visual and four auditory predictors of the TTC estimates. Error bars show 95% confidence intervals (CIs). squares, younger adults; circles, older adults

## Discussion

The objectives of the present study were to quantify the relative use of different visual and auditory cues when estimating TTC, including  $\tau$  variables and heuristic cues. Here, we investigate this for the first time under more realistic, familiar, and contextualized conditions than in previous studies. The weights of each cue were estimated during both unimodal conditions (visual-only and auditory-only) and a combined cue condition (audio–visual). Age-related differences were evaluated by comparing the average TTC estimates and estimated cue weights of older adults and younger adults.

When comparing the average TTC estimates, it was shown that, for both age groups, the average estimates were the longest during the auditory-only condition and shortest in the visual-only condition, with estimates in the audio–visual condition falling between the two unisensory estimates (when considered across all TTC levels). This finding is somewhat inconsistent with the results of our prior study that used less realistic and decontextualized stimuli (DeLucia et al., 2016), which showed no significant difference between visual-only and audio–visual estimates. It is therefore possible, that the auditory information provided through the more realistic scene in the present study resulted in these inputs being used to a greater extent during combined cue conditions than when the auditory and visual inputs were less realistic or familiar (i.e., DeLucia et al., 2016). It may be that the two sensory inputs were interpreted as more likely to be causally related in the familiar scene and thus they were more readily combined than under conditions for which no a priori knowledge about the known associations between the visual image and sound cues were available. However, it should also be noted that in the present study, just as in DeLucia et al. (2016), the mean TTC estimates in the audio–visual condition were more similar to the estimates in the visual-only condition than in the auditory-only condition. This pattern is also consistent with the weight analyses that showed a much stronger reliance on visual than on auditory cues in the audio–visual condition. On a more general level, Ashmead, Guth, Wa, Long, and Ponchillia (2005) and Guth, Long, Emerson, Ponchillia, and Ashmead (2013) reported that in a real traffic situation blind individuals made more frequent risky street crossing decisions than sighted individuals. In principle, the overestimation of the TTC of an approaching vehicle based only on the auditory, as compared to both the auditory and visual, information shown in our data might contribute to people accepting a too-short gap when crossing a street. However, the approximate time needed to cross one lane is 4 s (Guth et al., 2013), and our data showed that the relative TTC overestimation in the auditory-only condition decreased with increases in actual TTC (see Fig. 2).

Another interesting finding is that participants mostly overestimated TTC, whereas many prior studies of TTC have

reported underestimations, particularly as actual TTC increased above 1 s (e.g., Heuer, 1993; Oberfeld & Hecht, 2008; Schiff & Detwiler, 1979). In our earlier study (DeLucia et al., 2016), we also observed overestimations. We attributed them to our simple displays, in accordance with accounts in other studies (Geri, Gray, & Grutzmacher, 2010; Gray & Regan, 1998). For example, Geri et al. noted the importance of context to whether participants overestimate or underestimate TTC. The results of the present study do not support the idea that overestimations occur only when the displays are simple and lack context. The present visual displays were highly complex, with substantial context, but still resulted in overestimations, although the magnitude of overestimation appeared smaller in the visual-only and audio–visual conditions of the present study than in our prior study. Further research will be needed to determine the factors that underlie these observed TTC overestimations versus underestimations. We also observed a significant effect of vehicle speed on TTC estimates, indicating that estimates were shorter when vehicle speed was slower. This finding is consistent with previous studies (e.g., Oberfeld & Hecht, 2008; Petzoldt, 2014) and can be attributed to a size-arrival effect (DeLucia, 1991) or to a distance bias (Law et al., 1993). Specifically, at slower speeds, when the screen goes blank, the final optical vehicle size is larger, and the vehicle is closer in terms of final distance, than in the faster condition.

In addition to comparing average estimates across groups and sensory conditions, the design of this study allowed us to estimate the relative cue weights within each of the unimodal conditions and in the bimodal condition. These cues included five visual cues (visual  $\tau$ , final optical size, final rate of optical expansion, change in visual angle across the presentation duration, final visual distance) and four auditory cues (auditory  $\tau$ , final sound pressure level, final rate of change in sound pressure, change in sound pressure level across the presentation duration).

The cue-weighting results were generally comparable with those from our previous study (DeLucia et al., 2016). Specifically, when *only visual information* was available, estimates were consistent with the use of visual  $\tau$ , final optical size, and final rate of optical expansion. Analyses of relative importance supported the idea that visual  $\tau$  was the most dominant cue, with other heuristic cues explaining a significant portion of the variance, and with the change in visual angle across the presentation duration being the least important predictor. One of the few other studies that had examined the use of heuristic visual cues reported no association between TTC estimates and final optical size, final rate of optical expansion, and the change in visual angle across the presentation duration (Gray & Regan, 1998). This might be due to differences in the experimental designs, such that, in the study by Gray and Regan, the TTC and the final rate of optical expansion were independent by varying the starting size about the mean by

40%. Although this was possible with the laboratory type of stimuli they presented, our visual scene included realistic vehicle types (car vs. fire truck) that varied in size, but much less so than in Gray and Regan (1998). In another study, Yan et al. (2011) examined the use of visual heuristic cues and found an association between the TTC estimates and final visual distance, as is consistent with our results. Those authors also reported a correlation between the velocity of the object and the TTC estimates, but they did not analyze the final rates of optical expansion.

When *only auditory information* was available, the weightings of the cues were not the same as in the visual-only condition. Instead, auditory  $\tau$  was virtually ignored, whereas the final sound pressure level was weighted highest (a result also observed in DeLucia et al., 2016). Thus, the visual and auditory versions of the size-arrival effect were observed in the present study, as they were in our previous study, even when other realistic contextual cues were provided.

When *both auditory and visual information* were available, the estimates appeared to be dominated by visual cues, with visual  $\tau$  being weighted the highest. The possibility should be considered that the visual information was weighted more than the auditory information due to differences in the reliabilities of the visual and auditory information—for example, from the perspective of optimal cue integration (e.g., Green, 1958; Landy, Maloney, Johnston, & Young, 1995). However, our study was not designed to address this issue, which deserves further investigation. Among the auditory cues, only the relative importance of auditory  $\tau$  was somewhat comparable to any of the visual cues. This is very interesting, given that in the auditory-only condition, auditory  $\tau$  was not used at all. A speculative possibility is that the use of auditory  $\tau$  was enhanced by the processing of visual  $\tau$ , because both cues are measured on the same scale and can be averaged; when presented alone, auditory  $\tau$  cannot be averaged with the other (heuristic) cues, which are measured on different scales. This pattern differs markedly from the cue weights observed in the audio–visual condition of our previous study (DeLucia et al., 2016), in which the relative importance of the final sound pressure level cue was much higher. We also found an interesting age group by cue interaction in this condition, whereby older adults were less likely to weight visual  $\tau$  (and, to a lesser extent, auditory  $\tau$ ) as highly as younger adults did, relative to the heuristic-based cues.

The present study was the first to evaluate whether relative visual and auditory cue weights differ between younger and older adults. Overall, the results demonstrated very few differences with respect to age, apart from the stronger reliance on heuristic cues by the older group in the audio–visual condition. These results are somewhat inconsistent with other studies of visual TTC, which have demonstrated more conservative estimates (i.e., underestimations of TTC) by older adults

(e.g., DeLucia et al., 2003; Hancock & Manser, 1997; Schiff et al., 1992).

The main difference between these previous studies and the present study was the inclusion of familiar contextual information in the stimuli used in the present study, as well as the addition of sound cues. This could mean that older adults benefit more from these added cues, thereby minimizing any observable age differences. Because older adults often experience declines in sensory function, they may exploit the use of contextual cues, when available, to a greater extent than do younger adults (e.g., Pichora-Fuller, 2008), thereby leading to performance that is more comparable to that of younger adults.

A limitation of the present study is that the auditory representation of the vehicle changed only in acoustic intensity during approach. The combination of rather sophisticated visual simulations with very basic auditory simulations containing no spatial audio, simulated reflected sound, or contextual environmental sounds very likely contributed to the near-zero weights for auditory cues in the audio–visual condition and these weights were lower than in our previous study. Future studies should combine visual and auditory simulation with comparable levels of fidelity. Furthermore, the frequency content of the auditory stimulus may also be an important factor to consider for auditory  $\tau$ -based estimates (Gordon, Russo, & MacDonald, 2013), particularly when evaluating age effects, given that older adults often lose sensitivity at the higher frequencies. It is also important to note that the older adults in the present study were screened for sensory and cognitive impairments. There was no evidence of visual or vestibular impairments, and the few participants with known hearing loss were hearing-aid users and were able to perform the baseline auditory detection task. Therefore, this group (a) may not be representative of the older adult population more generally speaking, or (b) may have shown less evidence of age-related effects than in past studies in which these factors may not have been directly accounted for. To this point, Butler, Lord, and Fitzpatrick (2016) have demonstrated that older adults with reduced physical and cognitive function made more unsafe traffic-crossing decisions under certain task constraints than did those who were higher functioning. Future research should better consider and account for the effects of these age-related declines on the weighting and combination of different sensory cues when making TTC judgments.

Overall, the present study demonstrates that under realistic, contextualized conditions, the TTC estimates of both older and younger adults reflect the use of both visual and auditory cues, and of both accurate  $\tau$ -like cues and such heuristic cues as final size or sound pressure level. Visual  $\tau$  was the highest-weighted cue across both the visual-only and audio–visual conditions. These results are, in general, consistent with previous findings; however, there is evidence that the inclusion of a more realistic scene and familiar visual and auditory stimuli

had some effects on the TTC estimates. Future research should provide a more comprehensive evaluation of which aspects of the more realistic scenarios contributed to these observed differences.

**Author note** We thank Bruce Haycock and Susan Gorski for technical support. We are grateful to Hengameh Nazemroaya for her help with data collection, and to Marius Frenken for his assistance in organizing the references. Steve Prime and an anonymous reviewer provided helpful comments on an earlier version of the manuscript.

## References

- Ashmead, D. H., Guth, D., Wa, R. S., Long, R. G., & Ponchillia, P. E. (2005). Street crossing by sighted and blind pedestrians at a modern roundabout. *Journal of Transportation Engineering*, *131*, 812–821. doi:10.1061/(asce)0733-947x(2005)131:11(812)
- Azen, R., & Budescu, D. V. (2003). The dominance analysis approach for comparing predictors in multiple regression. *Psychological Methods*, *8*, 129–148. doi:10.1037/1082-989x.8.2.129
- Belsley, D. A., Kuh, E., & Welsch, R. A. (1980). *Regression diagnostics: Identifying influential data and sources of collinearity*. Hoboken, NJ: Wiley.
- Bootsma, R. J., & Oudejans, R. R. D. (1993). Visual information about time-to-collision between two objects. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 1041–1052. doi:10.1037/0096-1523.19.5.1041
- Braunstein, M. L. (1976). *Depth perception through motion*. New York, NY: Academic Press.
- Budescu, D. V. (1993). Dominance analysis: A new approach to the problem of relative importance of predictors in multiple regression. *Psychological Bulletin*, *114*, 542–551.
- 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*, e152617. doi:10.1371/journal.pone.0152617
- Caird, J. K., & Hancock, P. A. (1994). The perception of arrival time for different oncoming vehicles at an intersection. *Ecological Psychology*, *6*, 83–109. doi:10.1207/s15326969eco0602\_1
- Cicchino, J. B., & McCart, A. T. (2015). Critical older driver errors in a national sample of serious US crashes. *Accident Analysis and Prevention*, *80*, 211–219. doi:10.1016/j.aap.2015.04.015
- Cutting, J. E., & Wang, R. F. (2000). Heading judgments in minimal environments: The value of a heuristic when invariants are rare. *Perception & Psychophysics*, *62*, 1146–1159. doi:10.3758/bf03212119
- DeCarlo, L. T., & Cross, D. V. (1990). Sequential effects in magnitude scaling: Models and theory. *Journal of Experimental Psychology: General*, *119*, 375–396. doi:10.1037/0096-3445.119.4.375
- DeLucia, P. R. (1991). Pictorial and motion-based information for depth perception. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 738–748.
- DeLucia, P. R. (2004). Multiple sources of information influence time-to-contact judgments: Do heuristics accommodate limits in sensory and cognitive processes? In H. Hecht & G. J. P. Savelsbergh (Eds.), *Time-to-contact* (pp. 243–286). Amsterdam, The Netherlands: Elsevier Science.
- DeLucia, P. R., Bleckley, M. K., Meyer, L. E., & Bush, J. M. (2003). Judgments about collision in younger and older drivers. *Transportation Research Part F*, *6*, 63–80. doi:10.1016/S1369-8478(02)00047-5

- DeLucia, P. R., Preddy, D., & Oberfeld, D. (2016). Audiovisual integration of time-to-contact information for approaching objects. *Multisensory Research*, 29. doi:10.1163/22134808-00002520
- Diederich, A., & Colonius, H. (2004). Bimodal and trimodal multisensory enhancement: Effects of stimulus onset and intensity on reaction time. *Perception & Psychophysics*, 66, 1388–1404. doi:10.3758/bf03195006
- Freiherr, J., Lundström, J. N., Habel, U., & Reetz, K. (2013). Multisensory integration mechanisms during aging. *Frontiers in Human Neuroscience*, 7, 863. doi:10.3389/fnhum.2013.00863
- Geri, G. A., Gray, R., & Grutzmacher, R. (2010). Simulating time-to-contact when both target and observer are in motion. *Displays*, 31, 59–66. doi:10.1016/j.displa.2009.11.002
- Gordon, M. S., Russo, F. A., & MacDonald, E. (2013). Spectral information for detection of acoustic time to arrival. *Attention, Perception, & Psychophysics*, 75, 738–750. doi:10.3758/s13414-013-0424-2
- Gray, R., & Regan, D. (1998). Accuracy of estimating time to collision using binocular and monocular information. *Vision Research*, 38, 499–512.
- Green, D. M. (1958). Detection of multiple component signals in noise. *Journal of the Acoustical Society of America*, 30, 904–911.
- Greene, W. H. (2008). *Econometric analysis* (6th ed.). Upper Saddle River, NJ: Pearson.
- Guth, D. A., Long, R. G., Emerson, R. S. W., Ponchillia, P. E., & Ashmead, D. H. (2013). Blind and sighted pedestrians' road-crossing judgments at a single-lane roundabout. *Human Factors*, 55, 632–642. doi:10.1177/0018720812459884
- Hancock, P. A., & Manser, M. P. (1997). Time-to-contact: More than Tau alone. *Ecological Psychology*, 9, 265–297.
- Heuer, H. (1993). Estimates of time to contact based on changing size and changing target vergence. *Perception*, 22, 549–563.
- Horswill, M. S., Helman, S., Ardiles, P., & Wann, J. P. (2005). Motorcycle accident risk could be inflated by a time to arrival illusion. *Optometry and Vision Science*, 82, 740–746. doi:10.1097/01.opx.0000175563.21423.50
- Hosking, S. G., & Crassini, B. (2011). The influence of optic expansion rates when judging the relative time to contact of familiar objects. *Journal of Vision*, 11(6), 20. doi:10.1167/11.6.20
- 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, 69–82. doi:10.2307/1164736
- Jenison, R. L. (1997). On acoustic information for motion. *Ecological Psychology*, 9, 131–151.
- Kolarik, A. J., Moore, B. C. J., Zahorik, P., Cirstea, S., & Pardhan, S. (2016). Auditory distance perception in humans: A review of cues, development, neuronal bases, and effects of sensory loss. *Attention, Perception, & Psychophysics*, 78, 373–395. doi:10.3758/s13414-015-1015-1
- Kusano, K. D., Chen, R., Montgomery, J., & Gabler, H. C. (2015). Population distributions of time to collision at brake application during car following from naturalistic driving data. *Journal of Safety Research*, 54, 95–104. doi:10.1016/j.jsr.2015.06.011
- Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision Research*, 35, 389–412.
- Laurienti, P. J., Burdette, J. H., Maldjian, J. A., & Wallace, M. T. (2006). Enhanced multisensory integration in older adults. *Neurobiology of Aging*, 27, 1155–1163. doi:10.1016/j.neurobiolaging.2005.05.024
- 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, 1183–1199.
- LeBreton, J. M., Ployhart, R. E., & Ladd, R. T. (2004). A Monte Carlo comparison of relative importance methodologies. *Organizational Research Methods*, 7, 258–282. doi:10.1177/1094428104266017
- Lecoutre, B. (1991). A correction for the epsilon approximate test in repeated measures designs with two or more independent groups. *Journal of Educational Statistics*, 16, 371–372.
- Lee, D. N. (1976). Theory of visual control of braking based on information about time-to-collision. *Perception*, 5, 437–459.
- Lobjois, R., & Cavallo, V. (2009). The effects of aging on street-crossing behavior: From estimation to actual crossing. *Accident Analysis and Prevention*, 41, 259–267. doi:10.1016/j.aap.2008.12.001
- McGovern, D. P., Roudaia, E., Stapleton, J., McGinnity, T. M., & Newell, F. N. (2014). The sound-induced flash illusion reveals dissociable age-related effects in multisensory integration. *Frontiers in Aging Neuroscience*, 6, 250. doi:10.3389/fnagi.2014.00250
- McLeod, R. W., & Ross, H. E. (1983). Optic flow and cognitive factors in time-to-collision estimates. *Perception*, 12, 417–423.
- Mozolic, J. L., Hugenschmidt, C. E., Peiffer, A. M., & Laurienti, P. J. (2012). Multisensory integration and aging. In M. M. Murray & M. T. Wallace (Eds.), *The neural bases of multisensory processes* (pp. 381–392). Boca Raton, FL: CRC Press.
- Oberfeld, D., & Hecht, H. (2008). Effects of a moving distractor object on time-to-contact judgments. *Journal of Experimental Psychology: Human Perception and Performance*, 34, 605–623. doi:10.1037/0096-1523.34.3.605
- Oberfeld, D., Hecht, H., & Landwehr, K. (2011). Effects of task-irrelevant texture motion on time-to-contact judgments. *Attention, Perception, & Psychophysics*, 73, 581–596. doi:10.3758/s13414-010-0040-3
- Petzoldt, T. (2014). On the relationship between pedestrian gap acceptance and time to arrival estimates. *Accident Analysis and Prevention*, 72, 127–133. doi:10.1016/j.aap.2014.06.019
- Pichora-Fuller, M. K. (2008). Use of supportive context by younger and older adult listeners: Balancing bottom-up and top-down information processing. *International Journal of Audiology*, 47, S72–S82. doi:10.1080/14992020802307404
- Prime, S. L., & Harris, L. R. (2010). Predicting the position of moving audiovisual stimuli. *Experimental Brain Research*, 203, 249–260. doi:10.1007/s00221-010-2224-4
- Regan, D., & Hamstra, S. J. (1993). Dissociation of discrimination thresholds for time to contact and for rate of angular expansion. *Vision Research*, 33, 447–462.
- Schiff, W., & Detwiler, M. L. (1979). Information used in judging impending collision. *Perception*, 8, 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, 303–316.
- Schiff, W., Oldak, R., & Shah, V. (1992). Aging persons estimates of vehicular motion. *Psychology and Aging*, 7, 518–525. doi:10.1037/0882-7974.7.4.518
- Schleinitz, K., Petzoldt, T., Krems, J. F., & Gehlert, T. (2016). The influence of speed, cyclists' age, pedaling frequency, and observer age on observers' time to arrival judgments of approaching bicycles and e-bikes. *Accident Analysis and Prevention*, 92, 113–121. doi:10.1016/j.aap.2016.03.020
- Shaw, B. K., McGowan, R. S., & Turvey, M. T. (1991). An acoustic variable specifying time-to-contact. *Ecological Psychology*, 3, 253–261.
- Thomas, D. R., Zumbo, B. D., Kwan, E., & Schweitzer, L. (2014). On Johnson's (2000) relative weights method for assessing variable importance: A reanalysis. *Multivariate Behavioral Research*, 49, 329–338. doi:10.1080/00273171.2014.905766
- Todd, J. T. (1981). Visual information about moving objects. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 795–810.

- Tonidandel, S., & LeBreton, J. M. (2011). Relative importance analysis: A useful supplement to regression analysis. *Journal of Business and Psychology, 26*, 1–9. doi:10.1007/s10869-010-9204-3
- Yan, J. J., Lorv, B., Li, H., & Sun, H. J. (2011). Visual processing of the impending collision of a looming object: Time to collision revisited. *Journal of Vision, 11*(12), 7. doi:10.1167/11.12.7
- Zahorik, P. (2002). Assessing auditory distance perception using virtual acoustics. *Journal of the Acoustical Society of America, 111*, 1832–1846. doi:10.1121/1.1458027
- Zhou, L., Yan, J., Liu, Q., Li, H., Xie, C., Wang, Y., . . . Sun, H.-J. (2007). Visual and auditory information specifying an impending collision of an approaching object. In J. A. Jacko (Eds.), *Human–computer interaction: Interaction platforms and techniques. Proceedings of the 12th international conference* (pp. 720–729). Berlin, Germany: Springer.