BRILL
Multisensory Research 31 (2018) 251–272
brill.com/msr

Aging Impairs Audiovisual Facilitation of Object Motion Within Self-Motion

Eugenie Roudaia [1,*], **Finnegan J. Calabro** [2,4], **Lucia M. Vaina** [2,3] **and Fiona N. Newell** [1]

[1] School of Psychology and Institute of Neuroscience, Trinity College Dublin, Dublin, Ireland
[2] Brain and Vision Research Laboratory, Department of Biomedical Engineering, Boston University, Boston, MA, USA
[3] Department of Neurology, Harvard Medical School, Boston, MA, USA
[4] Department of Psychiatry and Department of Bioengineering, University of Pittsburgh, Pittsburgh, PA, USA

Received 10 March 2017; accepted 27 July 2017

Abstract

The presence of a moving sound has been shown to facilitate the detection of an independently moving visual target embedded among an array of identical moving objects simulating forward self-motion (Calabro *et al.*, *Proc. R. Soc. B*, 2011). Given that the perception of object motion within self-motion declines with aging, we investigated whether older adults can also benefit from the presence of a congruent dynamic sound when detecting object motion within self-motion. Visual stimuli consisted of nine identical spheres randomly distributed inside a virtual rectangular prism. For 1 s, all the spheres expanded outward simulating forward observer translation at a constant speed. One of the spheres (the target) had independent motion either approaching or moving away from the observer at three different speeds. In the visual condition, stimuli contained no sound. In the audiovisual condition, the visual stimulus was accompanied by a broadband noise sound co-localized with the target, whose loudness increased or decreased congruent with the target's direction. Participants reported which of the spheres had independent motion. Younger participants showed higher target detection accuracy in the audiovisual compared to the visual condition at the slowest speed level. Older participants showed overall poorer target detection accuracy than the younger participants, but the presence of the sound had no effect on older participants' target detection accuracy at either speed level. These results indicate that aging may impair cross-modal integration in some contexts. Potential reasons for the absence of auditory facilitation in older adults are discussed.

* To whom correspondence should be addressed. E-mail: eugenie.roudaia@umontreal.ca

DOI:10.1163/22134808-00002600

Keywords
Optic flow, motion, audiovisual integration, aging

1. Introduction

When we move about in the world, the characteristic pattern of motion generated on our retinae as a result of our self-motion is termed optic flow. When dynamic objects are present in the scene while we are moving, our perceptual system needs to disambiguate the pattern of retinal motion to extract information about our own heading direction and speed, as well as the direction and speed of objects in the scene. The flow-parsing hypothesis (Warren and Rushton, 2007, 2009) proposes that the visual system extracts independent object motion from the pattern of retinal motion by subtracting, or parsing out, the optic flow components that are detected by specialized optic flow detectors sensitive to global motion patterns (for a review, see Lappe *et al.*, 1999).

When an event stimulates multiple sensory modalities, it is often beneficial to integrate information across the senses as opposed to treating information from each sense in isolation, especially when unisensory information is unreliable (e.g., Ernst and Banks, 2002; for review, see Alais *et al.*, 2010). In the case of self-motion perception, several studies have shown that visual and vestibular cues are integrated when making heading direction judgments (Butler *et al.*, 2010, 2011). The presence of vestibular information congruent with visual optic flow has also been shown to improve the perception of the direction of object motion (Macneilage *et al.*, 2012). Interestingly, auditory motion signals have also been found to improve the detection of independent object motion within self-motion (Calabro *et al.*, 2011). In that task, participants had to detect a sphere that was looming toward or receding away from the observer amid eight identical spheres that were expanding outward so as to simulate forward observer motion. The presentation of a co-localized dynamic sound (whose amplitude was increased or decreased congruent with the looming or receding visual target motion) led to more accurate detection of the target sphere compared to performance in the visual-only condition, or performance in a condition that presented a co-localized, but static (not changing in amplitude) sound. Using MEG and Dynamic Granger causality connectivity analyses, Vaina *et al.* (2010) showed that, in the visual condition, there was evidence for reciprocal connectivity between the prefrontal cortex and visual motion area MT+. In contrast, in the audiovisual condition, the prefrontal cortex was connected with the right superior temporal polysensory area, an area that also received incoming inputs from auditory cortex, but not with area MT+. These results are consistent with the hypothesis put forth by Vaina *et* *al.* (2010) that the target in the audiovisual condition was treated as a coherent audiovisual object.

Healthy aging is known to affect certain types of motion perception (Allard *et al.*, 2013; Bennett *et al.*, 2007; Billino *et al.*, 2008; Pilz *et al.*, 2010; Roudaia *et al.*, 2010; Snowden and Kavanagh, 2006). Studies examining self-motion perception in older adults have found reductions in the sensitivity to heading direction based on optic flow (Lich and Bremmer, 2014; Warren *et al.*, 1989) and reduced sensitivity when detecting coherent lamellar optic flow at high speeds (Atchley and Andersen, 1998). Older observers are also worse at judging whether a looming object will collide or pass by the observer, especially in the context of simulated self-motion (Andersen and Enriquez, 2006). These studies indicate that aging impacts the ability to parse retinal motion into observer-generated and object-generated motion. These declines in motion processing have been suggested to contribute to slower and less accurate decisions in the context of walking, street-crossing, driving, or intercepting objects (Berard *et al.*, 2009; DeLucia *et al.*, 2003; Dommes and Cavallo, 2011; Dommes *et al.*, 2013; François *et al.*, 2011).

Given that both visual and auditory sensory processing deteriorate with aging (Faubert, 2002; Hutchinson *et al.*, 2012; Owsley, 2010; Schneider *et al.*, 2010; Spear, 1993), it has been suggested that older adults may derive a greater benefit from combining information across the senses than younger adults. Early studies examined this question by comparing reaction times to brief, isolated stimuli presented in one or multiple modalities in younger and older adults. The findings confirmed this hypothesis, as older adults typically showed greater facilitation in reaction time in the multisensory conditions compared to unisensory conditions than younger adults (Diederich*et al.*, 2008; Laurienti *et al.*, 2006; Mahoney *et al.*, 2011; Peiffer *et al.*, 2007). An MEG study that examined cross-modal facilitation of response times to semantically congruent visual and auditory stimuli found that greater response time facilitation with congruent audiovisual stimuli was related to activation of posterior parietal and medial prefrontal regions specifically in the older group, pointing to potential differences in the way multisensory information is combined with aging (Diaconescu *et al.*, 2012). Older adults have also been shown to benefit from congruent auditory–visual stimulation in the context of speech processing (Winneke and Phillips, 2011), especially when the semantic information was unpredictable (Maguinness *et al.*, 2011). Thus, a number of studies have found that older adults benefit from multisensory congruent stimulation to the same or greater degree than younger adults (for reviews, see Freiherr *et al.*, 2013; Mozolic *et al.*, 2012).

At the same time, a number of other studies have shown that multisensory integration mechanisms are altered in aging. For example, older adults tend to integrate auditory and visual signal across a wider time window than younger adults (e.g., Diederich *et al.*, 2008; Setti *et al.*, 2011), although a growing number of studies have shown that this observation depends on the type of task or the specific multisensory mechanisms solicited (Bedard and Barnett-Cowan, 2015; McGovern *et al.*, 2014). Only a few studies have examined how aging affects multisensory processing in the context of object motion or self-motion. Roudaia and colleagues (2013) examined the effect of aging on the bounce-stream illusion (Sekuler *et al.*, 2001), in which participants judge whether two discs moving along two intersecting paths either streamed past one another or bounce off of one another. Results in that study showed that the presentation of a brief click sound at the same time as the discs overlapped was less effective at biasing the perception of the discs' motion in older compared to younger participants, suggesting that audiovisual integration in the context of motion may be reduced in older age, at least as assessed by the bounce-stream illusion. In contrast, another study found that the addition of auditory information resulted in a greater benefit for older compared to younger participants in a task where they were required to maintain a constant driving speed in a driving simulator (Ramkhalawansingh *et al.*, 2016). This result suggested aging does not impair the ability to combine auditory and visual information to estimate the speed of self-motion based on optic flow. Thus, whether cross-modal effects are enhanced or attenuated in aging appears to be highly task-dependent.

The aim of the current study was to examine whether older adults can use dynamic auditory information to enhance the detection of visual object motion in the context of self-motion. We used an adaptation of the same paradigm in which younger adults had demonstrated auditory facilitation of visual object motion detection (Calabro *et al.*, 2011). In that study, participants who showed the worst target detection in the visual condition had shown larger improvements in performance with the co-localized dynamic sound, suggesting that sounds may be especially beneficial to those who show difficulty in the task (Calabro *et al.*, 2011). Given that we expected older adults to show worse accuracy in the current task overall and given previous reports of enhanced multisensory integration in aging (Mahoney *et al.*, 2011), we hypothesized that older adults would show greater auditory facilitation than the younger group in this task.

2. Methods

The research protocol was approved by the research ethics committee of Trinity College Dublin, which ensured adherence to the tenets of the Declaration of Helsinki. All participants gave written informed consent to participate in the study.

2.1. Participants

Seventeen younger (*M*: 26.0 years, range: 19–31 years, 11 females) and seventeen older (*M*: 67.5 years, range: 61–76 years, 10 females) adults took part in this study. Younger participants were university students recruited at Trinity College Dublin and older participants were independently living residents of Dublin recruited through advertisements in community centres and local newspapers. A general health questionnaire was used to screen for previous history of neurological disorders, visual abnormalities, or hearing impairment. To screen for cognitive impairment, the Montreal Cognitive Assessment (Nasreddine *et al.*, 2005) was administered to older participants and participants who scored 23 or less were not recruited (Luis, Keegan and Mullan, 2009). Younger and older participants had normal or corrected-to-normal visual acuity (near and far Snellen acuity better than 6/9) and normal contrast sensitivity ($M = 1.99$, $\mathrm{SD} = 0.1$ log contrast on the Pelli–Robson Contrast Sensitivity Test). Older participants were also screened with the Hughson Westlake pure-tone audiometry test and those who were classified as having hearing impairment in at least one ear were not recruited.

2.2. Apparatus

The experiment was programmed in the Matlab environment using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) and Open GL libraries. Stimulus generation and presentation were controlled by a MacBook running OSX 10.6. Stimuli were presented on an HP L1710 17″ LCD monitor with a resolution of 1280×800 and refresh rate of 60 Hz. The display subtended $33 \times 21°$ at a viewing distance of 60 cm. Auditory stimuli were presented with Sennheiser HD 201 headphones. The experiment was conducted in a quiet room. A small table lamp provided low illumination.

2.3. Stimuli

The stimuli were the same as described in the study by Calabro *et al.* (2011) and examples are shown in Fig. 1. Visual stimuli consisted of nine textured, purple spheres (mean luminance: 28 cd/m^2) presented against a black background (mean luminance: 0.3 cd/m^2). The spheres were distributed inside a simulated virtual rectangular prism measuring 25 cm wide by 25 cm high by 60 cm deep. To avoid overlap between spheres, the frontoparallel plane was divided into nine equally sized (invisible) wedges extending from the fixation point and one sphere was placed in each wedge at a random eccentricity and at a random depth ranging from 25 to 35 cm. Objects had a mean diameter of 1.58°, but their size scaled with distance. The stimuli were viewed binocularly but did not contain any stereo disparity, so that change in the spheres’

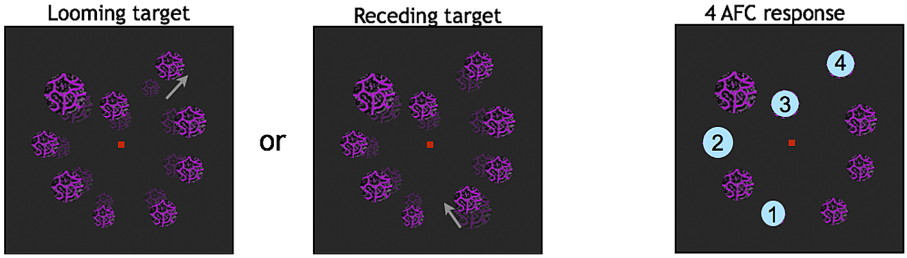

Figure 1. Schematic of experimental trials and task: Observers viewed nine spheres that moved towards them simulating looming motion for 1 s. One of the nine spheres contained independent motion, either looming or receding relative to the other spheres. In the audiovisual trials, a sound was also presented. The sound was co-localized with the target and its loudness increased for a looming target and decreased for receding target. The response screen required observers to identify the target in a four-alternative forced choice task.

size and texture details were the only two cues available to motion in depth. A red fixation square (0.1°) was placed at the centre of the display. For 1 s, eight spheres comprising the background scene moved outward, simulating observer self-motion toward the fixation point at 3 cm/s, or equivalently, simulating a rigid scene of eight objects approaching a stationary observer. The ninth sphere (the target) moved according to the sum of the scene motion vector and an independent motion vector either in the same or in the opposite direction relative to the scene, with a speed of 4, 6, or 8 cm/s. Participants were asked to detect the sphere that had independent motion relative to the scene, either looming toward the observer faster than the scene, or receding away from the observer. After the 1 s motion, the spheres disappeared from the screen for 0.25 s and reappeared at their last coordinates, but set to the average sphere size to ensure participants would not be biased to select the largest or smallest of all the spheres as their response. The target and three randomly-selected non-targets were selected in each trial and assigned labels 1–4 in random order. Participants reported which of those four spheres was the target by pressing one of four keyboard buttons.

In the audiovisual condition, the visual target was accompanied by a moving sound that was either looming or receding, congruent with the direction of the target object motion. The sound was a broadband noise filtered between 0.3 and 12 kHz co-localized with the approximate azimuth location of the target sphere using inter-aural time and intensity cues. Sound motion was simulated by increasing or decreasing the sound amplitude by ∼10 dB SPL, from a starting sound level of ∼65 dB SPL. The rate of loudness change was consistent with a sound motion of ∼3.5 cm/s. This auditory motion speed was constant across visual speed conditions.

2.4. Procedure

Prior to the main experiment, participants performed a sound localization task for the looming and receding sounds in two separate blocks, in counterbalanced order. On each trial of the localization task, a sound stimulus was presented at one of 5 azimuths ($-10°$, $-5°$, $0°$, $5°$, $10°$) for 1 s at the same time as five vertical bars appeared on the screen at those locations. Participants were asked to fixate on the vertical bar in the middle of the display ($0°$) while listening to the sound and press one of five keys on the keyboard to report the location that best matched their perceived sound location. No feedback was given. All locations were presented in random order with 10 repetitions per location. At this time, the experimenter also informed participants that sounds were looming in one block and receding in another block and ensured that participants heard the difference between the two types of sound by presenting two types of sounds as a demo. All the participants reported being able to hear the increasing and decreasing loudness change in the looming and receding stimuli, respectively.

After the auditory localization task, participants completed one visual and one audiovisual practice block to become familiarized with the target object search task. The practice stimuli contained six spheres instead of nine and target speeds were ± 6 cm/s or ± 8 cm/s. There were five trials of each speed all presented in randomly interleaved order. Participants were told that they would see several spheres approaching them at a constant speed and that one of the spheres would be either approaching faster than the rest of the scene, or would be receding away from the scene. They were asked to detect the target sphere. Each trial began with a red fixation square that participants were asked to fixate throughout the trial. The visual or audiovisual was displayed for 1 s, followed by a blank screen containing the fixation square for 0.25 s, followed by the response screen containing all the spheres at their x–y locations on the last frame (see Fig. 1 for an illustration). Four of the spheres were labelled with numbers 1–4 and participants responded by pressing one of four corresponding keys on the keyboard to indicate which sphere they believed to be the target. The duration of the blank interval between the stimulus offset and the response screen was kept brief to reduce the memory load of the response phase, while also being long enough to ensure that the response screen did not perceptually group with the motion sequence or disrupt the processing of the motion sequence. Participants were told to take their time to respond as accurately as possible. There was no time limit to complete the response and no feedback as to the accuracy of the response was given. The next trial began 0.3 s after a response was made.

The practice phase was followed by visual and audiovisual experimental blocks of trials where participants detected the target among nine spheres.

Each block of trials contained looming and receding targets moving at 4, 6, and 8 cm/s. All speeds and directions were presented in randomly interleaved order, with 15–20 repetitions of each type per block. One younger and three older participants only completed one block of each condition. All the other participants completed two visual and two audiovisual blocks, whose order was chosen randomly from the following orders with equal probability: ABBA, ABAB, BABA, BAAB.

2.5. Analysis

Data from the auditory localization experiment were analyzed using signal detection theory by calculating the sensitivity d' for discriminating each pair of adjacent locations ($-10°$ and $-5°$, $-5°$ and $0°$, $0°$ and $5°$, $5°$ and $10°$), and then computing the sum of all d' to obtain a measure of global sensitivity for discriminating sound locations (Macmillan and Creelman, 2005). Data from the visual target detection task were expressed as a proportion of correct trials. Median response times were also calculated for all conditions, after excluding very slow outliers (>10 s). Response times were defined as the time between the stimulus onset and the button press.

Statistical analyses and data visualisation were performed in R using the ggplot2, car, ez, wrs2, and TOSTER packages (Fox and Weisberg, 2011; Lakens, 2017; Lawrence, 2015; R Core Team, 2015; Wickham, 2009; Wilcox, 2016). Mauchy's Test of Sphericity was used to test for violations of the sphericity assumption for all effects and interactions containing within-subject factors. When multiple comparisons were performed, the Holm–Bonferroni procedure maintained a family-wise alpha at 0.05. When sphericity was violated, the degrees of freedom were adjusted using the Greenhouse–Geisser epsilon, $\hat{\varepsilon}$. The generalised eta-squared, $\hat{\eta}^2_G$, effect size measure is reported for all effects (Olejnik and Algina, 2003). This measure represents the proportion of variance associated with the effect, relative to variance due to individual differences. The general guidelines for the effect size magnitudes of $\hat{\eta}^2_G$ are 0.02, 0.13, and 0.26 for small, medium, and large effects, respectively (Bakeman, 2005).

3. Results

3.1. Auditory Localization Performance

Figure 2 shows individuals' average of perceived azimuth location and the total d' sensitivity for discriminating the azimuth locations for the looming and receding sounds separately. High positive d' values indicate good localization performance. The localization performance was similar for the two motion directions (receding or looming) and showed high variability across participants in both age groups. Comparing average d' for both groups and motion directions in a 2 (age group) $\times$ 2 (direction) split-plot ANOVA revealed a

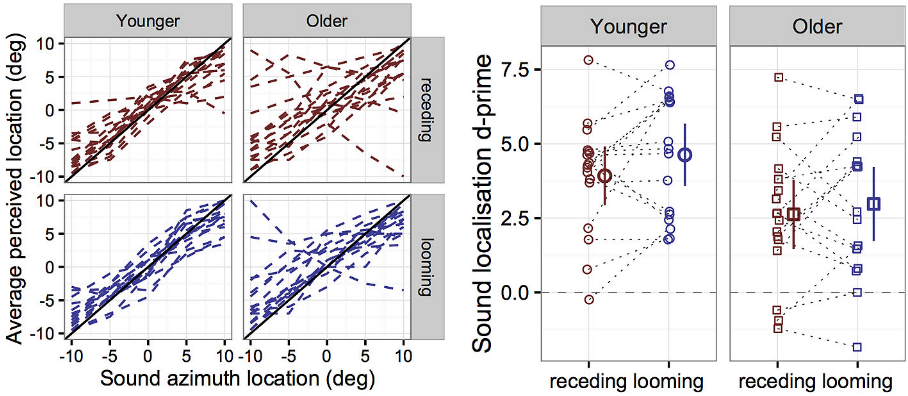

Figure 2. The left panel shows average perceived location (deg) plotted against presented azimuth location for individual participants (dashed lines). The right panel shows total sensitivity (d') for discriminating the sound locations. Small circle and square symbols show individual younger and older participants' data, respectively. Group averages and their bootstrapped 95% CI are shown in larger symbols offset to the right.

medium-size effect of age group, $F(1, 32) = 4.76$, $p = 0.04$, $\hat{\eta}^2_G = 0.11$, no effect of direction, $F(1, 32) = 2.81$, $p = 0.10$, $\hat{\eta}^2_G = 0.02$, and no interaction between age and direction, $F(1, 32) = 0.32$, $p = 0.57$, $\hat{\eta}^2_G = 0$.

Figure 2 shows that one younger and three older participants showed negative d' for at least one direction, with one participant showing highly negative d' for both directions. Negative d' would indicate that perceived azimuth locations were reversed, such that left locations were perceived on the right and vice versa. This reversal indicates a potential mistake either in response mapping or incorrect placement of headphones. Excluding data from the four participants who showed negative d' values, the effect of age group was no longer statistically significant, $F(1, 28) = 2.8$, $p = 0.11$, $\hat{\eta}^2_G = 0.07$. In the following figures and statistical analyses, data from these four participants were excluded. However, the general pattern of results remained unchanged even when all participants were included.

3.2. *Visual Target Detection Accuracy*

The top panel of Fig. 3 shows the accuracy of the younger and older adults in correctly identifying the location of the target object in the visual and audiovisual conditions, for looming and receding speeds. A mixed-model ANOVA with age group as the between-subject factor and sound condition, target speed, and direction as within-subject factors revealed a main effect of age group, $F(1, 28) = 53.1$, $p < 0.001$, $\hat{\eta}^2_G = 0.46$, reflecting the fact that older participants showed overall lower accuracy in all conditions. This result is consistent with previous studies showing age-related declines in the perception of

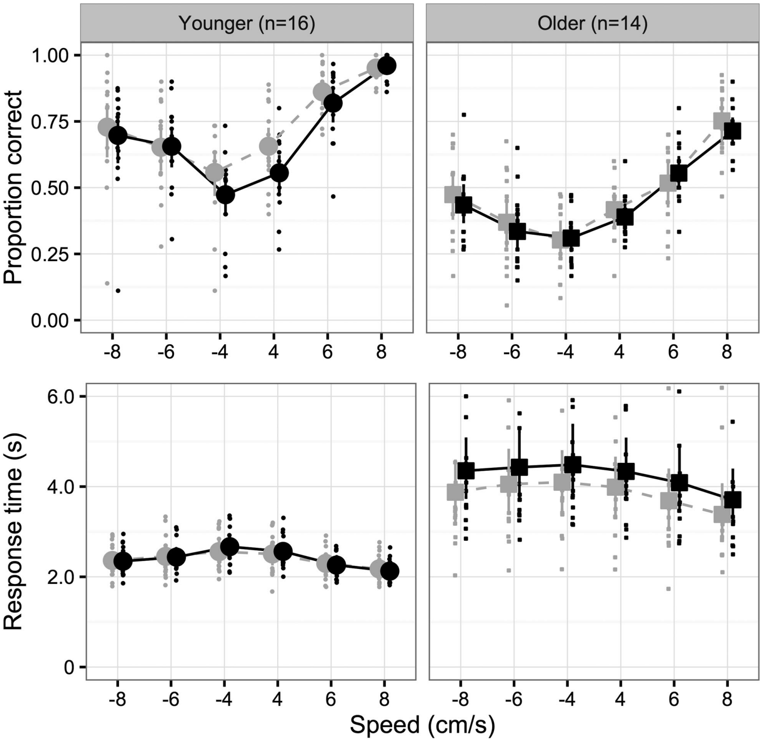

Figure 3. Accuracy (top panels) and response times (bottom panels) for reporting the target object in the visual (black) and audiovisual (grey) conditions for younger (circles) and older (squares) participants across six speeds (negative values correspond to receding motion and positive values correspond to looming motion). Large symbols represent group averages and small symbols show individual data points. Error bars show the bootstrapped 95% CIs of the group averages.

object motion within self-motion (Andersen and Enriquez, 2006). There was a main effect of target speed, $F(2, 56) = 137.2$, $p < 0.001$, $\hat{\eta}^2_G = 0.30$, and an Age group × Speed interaction, $F(2, 56) = 5.04$, $p = 0.01$, $\hat{\eta}^2_G = 0.02$, reflecting the fact that target detection accuracy improved with increasing speed and that the effect of speed differed between the two age groups. There was also a main effect of target direction, $F(1, 28) = 85.5$, $p < 0.001$, $\hat{\eta}^2_G = 0.30$, and a Direction × Speed interaction, $F(2, 56) = 31.3$, $p < 0.001$, $\hat{\eta}^2_G = 0.06$, suggesting that performance was better for looming than receding targets, and that increasing speed had a greater effect on detecting looming compared to receding targets. Regarding the effect of sound condition, the main effect of sound was not significant, $F(1, 28) = 3.76$, $p = 0.06$, $\hat{\eta}^2_G = 0.01$, but there was a significant three-way Age group × Sound × Speed interaction, $F(2, 56) = 3.30$, $p = 0.04$, $\hat{\eta}^2_G = 0.007$. See Fig. 3 for an illustration of these results. The presence of the three-way interaction indicates that the effect of the presence of co-localised sound on target detection depended on target speed and age group. None of the other interactions were significant.

To decompose the three-way interaction, we tested the effect of sound condition at each speed level, collapsed across motion direction, for each group separately. In the younger group, the effect of sound was statistically significant at the slowest speed levels, ±4 cm/s, $F(1, 15) = 13.4$, $p = 0.002$, $\hat{\eta}^2_G = 0.13$, reflecting the fact that accuracy was higher in the audiovisual condition ($M = 0.61$, SD = 0.13) compared to the visual condition ($M = 0.51$, SD = 0.11). The effect of sound was not significant at the medium and high speeds for younger adults [±6 cm/s: $F(1, 15) = 1.2$, $p = 0.29$; ±8 cm/s: $F(1, 15) = 0.56$, $p = 0.46$]. In the older group, the effect of sound was not statistically significant at any of the speed levels [±4 cm/s: $F(1, 13) = 0.10$, $p = 0.76$; ±6 cm/s: $F(1, 13) = 0.001$, $p = 0.96$; ±8 cm/s: $F(1, 13) = 1.92$, $p = 0.19$].

To better illustrate these effects, we plotted the difference in accuracy between the audiovisual and visual conditions across speeds for younger and older groups in Fig. 4. The crossbars in the figure represent the average difference score and its 95% CI. Consistent with the statistical results reported above, the difference scores at the slowest speed in the younger group lie mostly above the zero line and the 95% CI of the average difference score does not overlap zero, indicating reliably better performance for the audiovisual over the visual only condition. In contrast, the 95% CIs of the average difference scores overlap zero at the faster speeds in the younger group and at

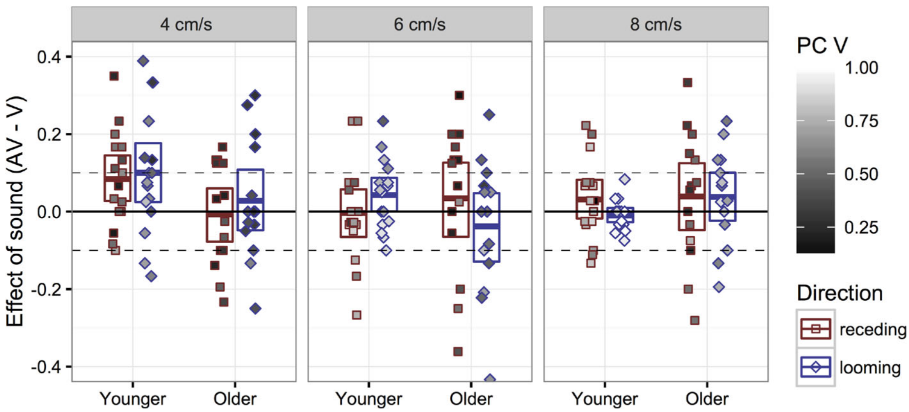

Figure 4. Difference scores of target detection accuracy in the audiovisual and visual conditions are shown for all individual participants for all speeds and directions. Red square symbols show data for receding targets and blue diamond symbols show data for looming targets. Crossbars show group averages and their bootstrapped 95% CI. The grey-scale fill colour (PC V) shows the proportion correct in the visual condition.

all speeds in the older group, suggesting equivalent performance across both sensory conditions.

The above results suggested that unlike younger participants, older participants showed no evidence of auditory facilitation on visual target detection in either speed condition. However, given that a failure to reject the null hypothesis of no difference between conditions is not sufficient to conclude that performance in two conditions is the same (as that is the null hypothesis itself), we conducted equivalence tests using the two-sided t-tests approach (Schuirmann, 1987) to examine whether performance in the visual and audiovisual conditions in the older group was indeed not different. An equivalence test allows us to choose the smallest effect size of interest and tests the null hypothesis that the true effect is greater than the one specified. If the equivalence test is statistically significant, it allows us to conclude that the two conditions are unlikely to differ by more than the effect size specified and can thus be considered to be equivalent (Lakens, 2017). We chose to test for a difference score of 10% or greater, as this ensured the equivalence test had 80% power to detect an effect, given the variability of the difference scores, our sample size, and the family-wise alpha of 0.05. The equivalence test was not significant for ±4 cm/s for younger participants $\{t(15) = 0.09$, $p = 0.53$, mean difference = 0.92, 98% CI {0.02–0.16]}, which is consistent with the statistical analysis reported above and indicates that the difference between the visual and audiovisual conditions is not smaller than 10%. The equivalence test was statistically significant for the other two speeds in the younger group [±6 cm/s: $t(15) = -4.5$, $p < 0.001$, 8 cm/s: $t(15) = -6.24$, $p < 0.001$] and for all speeds in the older group [(±4 cm/s, $t(13) = -2.76$, $p = 0.008$; ±6 cm/s: $t(13) = 2.4$, $p = 0.02$; ±8 cm/s: $t(13) = -2.20$, $p = 0.02$]. These results allow us to conclude that accuracy in the visual and audiovisual conditions was equivalent (within 10%) for all speeds in the older group and the middle and fastest speeds in the younger group.

3.3. Visual Target Detection Response Times

For completeness, we also present the data for response times for the visual target detection in Fig. 3. Note that participants were not required to respond quickly and responses were made only after the response screen appeared 0.25 s after stimulus offset. Therefore, this measure does not only reflect perceptual and motor processing, but is also highly influenced by the speed of stimulus–response mapping and response confidence. Given that response times were timed from the stimulus onset to the button press, response time could not be shorter than 1.25 s.

The split-plot ANOVA on response times revealed a main effect of age group, $F(1, 28) = 25.0$, $p < 0.001$, $\hat{\eta}^2_G = 0.44$, with older participants showing much slower responses than younger participants. The ANOVA also revealed main effects of target speed, $F(2, 56) = 31.3$, $p < 0.001$, $\hat{\eta}^2_G = 0.02$, direction, $F(1, 28) = 45.7$, $p < 0.001$, $\hat{\eta}^2_G = 0.02$, as well as a target Direction × Speed interaction, $F(2, 56) = 26.1$, $p < 0.001$, $\hat{\eta}^2_G = 0.003$. These results reflect the fact that responses were overall faster for looming than receding targets and became faster with increasing speed, but that the rate of decrease with speed was greater for looming targets. There was also a significant Age group × Direction interaction, $F(1, 28) = 7.41$, $p = 0.01$, $\hat{\eta}^2_G = 0.003$, and an Age group × Direction × Speed interaction, $F(2, 56) = 8.14$, $p < 0.001$, $\hat{\eta}^2_G = 0.001$). These interactions indicate that the magnitude of the speed and direction effects differed across the two age groups.

Comparing performance for the visual and audiovisual condition revealed a significant main effect of sound, $F(1, 28) = 7.6$, $p = 0.01$, $\hat{\eta}^2_G = 0.01$, and a significant Age group × Sound interaction, $F(1, 28) = 7.0$, $p = 0.01$, $\hat{\eta}^2_G = 0.01$. None of the other interactions with sound were statistically significant. Examining the effect of sound in each group, collapsed across all speed conditions, revealed that older participants' responses were 39 ms faster in the audiovisual condition ($M = 3.85$, SD = 1.25) compared to the visual condition ($M = 4.23$, SD = 1.35), $F(1, 13) = 6.93$, $p = 0.02$, $\hat{\eta}^2_G = 0.02$. In contrast, younger participants showed similar response times across the audiovisual ($M = 2.39$, SD = 0.38) and visual ($M = 2.40$, SD = 0.30) conditions, $F(1, 15) = 0.04$, $p = 0.83$, $\hat{\eta}^2_G = 0.001$. The absence of an effect of sound in the younger group may be due to the fact that responses were already very quick in the visual condition and there was no room for response time facilitation in the audiovisual condition. The absence of significant interactions between the effect of sound and target speed levels or motion direction indicates that response time facilitation in older adults was constant across all speeds.

3.4. Relationship Between the Effect of Sound and Sound Localization Ability

One potential explanation for the large inter-subject variability in the effect of sound on target detection is the variability in the sensitivity to the sound azimuth location. Participants who show good sound localization may have been able to narrow down their target search to fewer items by directing their attention to that location, which would also improve the probability of audiovisual binding with the correct target. We examined this question by testing for a positive correlation between the difference scores (AV–V), averaged across target speed levels, and sound localization d' for each age group and sound direction separately.

We used a one-tailed, rank-sum Spearman test and calculated 90% CI using the bootstrap percentile method. In the older group, there was a positive correlation for receding sounds, $r_s(15) = 0.48$, $p = 0.03$, 90% CI [0.13, 0.73], but a correlation near zero for looming sounds, $r_s(15) = -0.01$, $p = 0.51$, 90% CI [−0.46, 0.44]. In the younger group, both correlations were near zero [receding, $r_s(15) = -0.12$, $p = 0.68$; 90% CI [−0.54, 0.34]; looming, $r_s(15) = 0.04$, $p = 0.44$, 90% CI [−0.44, 0.51]]. No significant correlations were found between the RT facilitation and sound localization in either group. Thus, differences in sound localization ability did not account for significant variability in auditory facilitation of visual target detection in young adults, in either motion direction of the target. Good auditory localization was associated with a greater benefit in the audiovisual condition in older adults, but only in the case of receding sounds. However, given the wide confidence interval around the correlation coefficient and the noisy nature of small sample correlations, this finding should be treated with caution, especially given that there is no theoretical reason why auditory localization would provide a benefit on target detection for the receding, but not the looming direction of the target.

3.5. Relationship Between the Effect of Sound and Visual Target Detection

Calabro *et al.* (2011) reported a negative correlation between the magnitude of acoustic facilitation due to a co-localized, moving sound relative to a static sound and overall visual target performance, indicating that participants who showed poor visual target detection performance benefited more from the dynamic, congruent sound. Here, we examined whether low target detection accuracy in the visual-only condition was associated with a high AV–V difference score for each group and target direction separately, averaged across all speed levels. The Spearman one-tailed, rank-sum test showed a significant negative correlation in the older group for both looming targets, $r_s = -0.44$, $p = 0.04$, 90% CI [−0.74, −0.04], and receding targets, $r_s = -0.42$, $p = 0.04$, 90% CI [−0.74, −0.01]. Results from the younger group also showed a negative correlation for looming targets, $r_s = -0.44$, $p = 0.04$, 90% CI [−0.76, −0.04], but a correlation near zero for receding targets, $r_s = -0.07$, $p = 0.39$, 90% CI [−0.48, 0.38]. The negative correlations suggest that participants who showed worse target detection in the visual-only condition also benefited more from the co-localized sound. Here again, however, the confidence intervals for the correlation coefficients were very wide, making it difficult to know the magnitude of the true relationship. Correlation estimates based on small sample sizes are known to be particularly unstable (Schönbrodt and Perugini, 2013).

4. Discussion

The goal of the current study was to examine whether older adults may benefit from congruent auditory information when performing a complex visual motion perception task. We employed the task developed by Calabro *et al.* (2011), where participants detected an independently-moving target sphere embedded among eight other identical spheres moving radially outward so as to simulate forward observer self motion. We compared target detection accuracy with and without a spatially co-localized dynamic auditory broadband noise moving congruent with the direction of the target. We found that younger participants showed facilitation of target detection in the presence of the dynamic sound when visual targets were looming or receding at 4 cm/s, but not at the two faster speeds. In contrast, older participants showed no improvement in target detection accuracy in the presence of the sound at any speed level in this task. The lack of a beneficial effect on target detection accuracy in the presence of congruent cross-modal information in older adults is surprising, given that most previous studies have found equivalent or enhanced multisensory integration with aging (Freiherr *et al.*, 2013; Mozolic *et al.*, 2012).

The current study also revealed that aging significantly impairs performance in this task, as older adults showed substantially poorer accuracy than the younger group at all speeds. Given that the response screen appeared shortly after the stimulus offset, this difference cannot be due to age-differences in visual memory. Instead, this finding adds to a growing number of studies showing age-related reductions in sensitivity to optic flow and declines in the perception of approaching objects in the context of self-motion (Andersen and Enriquez, 2006; Lich and Bremmer, 2014; Mapstone *et al.*, 2006; Warren *et al.*, 1989). Previous studies have shown that target detection in the current task relies on global motion processing as opposed to processing of relative motion of adjacent objects (Calabro *et al.*, 2011), consistent with flow-parsing (Warren and Rushton, 2009). An fMRI study with Granger causality connectivity analyses showed that detection of the independent object motion relative to viewing optic flow only is accompanied by increases in the connectivity within a network of areas consisting of areas LO, V3A, KO, and hMT bilateraly, and areas in the intraparietal sulcus and the medial region of the dorsal intraparietal sulcus (Calabro and Vaina, 2012). The detection of object motion within simulated self-motion in younger observers has been shown to improve with increasing number of objects comprising the optic flow and with increasing stimulus duration (Royden and Connors, 2010). Thus, the effect of aging seen here may be especially large due to the relative sparseness of the visual stimulus and the duration of the stimulus display.

The results from younger participants only partially replicate the findings of Calabro *et al.* (2011), who found auditory facilitation of a similar magnitude but at all the target speeds they tested (2–8 cm/s, looming and receding). What might explain the lack of improvement in target detection at the medium and faster speeds in the current study? Given that we used the same visual and auditory stimuli, this discrepancy in the results may be due to differences in the testing procedure [we block-randomized visual only and audiovisual moving conditions, while Calabro *et al.* (2011) intermixed auditory-static and audio-moving sounds in the same block and compared those conditions directly; we used three speeds here, while they tested four speeds], or due to random differences in participant samples related to variation in individual binding tendency (Odegaard and Shams, 2016) or experience-related factors such as driving, action video game habits or experience with psychophysical studies.

What processes drive the auditory facilitation in the current task? Given that the sound was co-localized with the visual target, facilitation may result from cross-modal spatial cueing of the target location, or it may be due to audiovisual binding of the sound and visual target, or both. Given that multisensory integration and cross-modal spatial attention are highly interdependent (Talsma *et al.*, 2010), it is often difficult to disambiguate these possibilties. Calabro *et al.* (2011) examined this question by comparing the effect of static and dynamic sounds that were either co-localized with the target or non-spatially informative. They found that facilitation was most robust when the dynamic sound was co-localized with the target compared to being presented in the centre of the screen. At the same time, the presentation of a spatially-congruent but static sound had a much milder effect compared to the co-localized sound that also moved in a direction congruent with the target. Those results suggested that although spatial co-localization is necessary to observe the auditory facilitation, cross-modal spatial cueing alone does not fully account for the benefit in target detection. One possibility is that spatial attention to the location of the sound increases the likelihood of audiovisual binding within the target by increasing motion processing at that location and by reducing the set size of potential candidates for binding.

Could the absence of auditory facilitation in older adults be due to an inability to use the auditory spatial information to improve visual target detection? Our auditory localization experiment showed that younger and older adults had similar sensitivity to sound azimuth location. Therefore, older participants should have been able to use the sound's azimuth location to direct their attention to the approximate spatial location of the target or to the correct hemifield. However, although previous studies have reported preserved crossmodal attention in aging (e.g., Hugenschmidt *et al.*, 2009), it could be that older adults may require more time to localize the sound or to shift their attention to the sound's location than younger adults. Another possibility is that they were indeed able to use the sound's location to shift their attention to the correct azimuth position, but were less successful at binding the sound with the correct visual target. The analysis of response times revealed that older participants responded significantly faster in the audiovisual condition compared to the visual condition, even though no emphasis was put on responding quickly. The fact that the magnitude of response time facilitation was constant across speeds and motion directions may be a sign of a non-specific effect related to overall attention to the task or arousal, but it can also reflect increased response confidence in the audiovisual condition. The higher response confidence in the absence of improvements in accuracy may be the result of audiovisual bindings with non-targets. One way to determine if this was likely to be the case would be to examine the spatial pattern of responses. Unfortunately, we did not record the spatial location of the response sphere and thus are unable to determine from the present study whether participants were more likely to choose spheres on the same hemifield as the sound location. Future studies can shed light on this issue by examining the spatial pattern of incorrect responses, or by tracking the shifts of spatial attention in the current task using eyemovement recording or electrophysiology.

Could age-related differences in the overall accuracy level explain the lack of auditory facilitation in older adults? By using a range of speed levels, we ensured that our stimuli produced a range of accuracy levels in both groups. If auditory facilitation is only observed at a certain level of accuracy, then we should have observed auditory facilitation in older adults at the medium speed, where accuracy in the older group matched the accuracy level of the younger group at the slowest speed. Thus, it is unlikely that overall age differences in accuracy at detecting visual targets can explain the lack of audiovisual interactions in older adults.

Yet another possibility is that audiovisual binding fails in older adults due to age-related declines in integration of global motion or perception of optic flow (Atchley and Andersen, 1998; Mapstone *et al.*, 2006). Previous research indicates that visual perceptual grouping can occur prior to audiovisual integration (Kawachi *et al.*, 2014; Sanabria *et al.*, 2005). In our stimulus, the eight spheres forming the visual scene must be grouped together to extract the self-motion heading and speed which, in turn, can help to segregate the independently-moving target sphere. It is possible that the stronger the background elements are grouped together, the easier it becomes to bind the sound with the independently moving target. If older adults are slower or less efficient at integrating the background spheres into a coherent whole, this may reduce the likelihood for cross-modal binding of the sound with the independently moving object. Studies using transient sounds paired with visual search displays have shown that audiovisual integration is less likely to occur when multiple visual events are equally viable candidates for binding with a single auditory event (Van der Burg *et al.*, 2014). It is possible that by facilitating the extraction of optic flow from the background spheres in different ways we may increase the chances of observing audiovisual facilitation in older adults. This may be done by increasing the density of background spheres (a counterintuitive suggestion, given that it would increase the number of potential 'distracters') or increasing the stimulus duration to allow more time for global motion grouping to occur.

Similarly, it may be that aging affected the ability to perceive 3D motion from the monocular depth cues contained in the current stimulus, resulting in a percept of spheres expanding radially outward on a flat frontoparallel plane instead of as static objects moving forward in depth. This type of visual percept should not be expected to bind with a sound moving in depth. The stimulus used in the experiment contained no stereo disparity and used only two monocular cues to depth. A recent study reported that a sound moving in depth had a greater impact on detection of visual motion in depth when the visual stimulus contained stereo disparity (Harrison *et al.*, 2015). Similarly, congruent vestibular cues had a more robust effect on perception of heading from optic flow when visual stimuli contained stereo disparity, compared to only 2D information (Butler *et al.*, 2011). Future studies should examine whether auditory facilitation of object motion detection can be recovered in older adults when additional depth or vestibular self-motion cues are available to help the perception of motion in depth.

The results of this study indicate that spatially informative, congruent auditory information may not always be of benefit to older adults in a complex visual motion task. Many other types of audiovisual interactions have been reported in motion perception in younger adults, such as cross-modal dynamic capture (Soto-Faraco *et al.*, 2003, 2004, 2005), cross-modal motion aftereffects (Jain *et al.*, 2008; Kitagawa and Ichihara, 2002), and disambiguation of walking direction of a point-light-walker with directional footstep sounds (Schouten *et al.*, 2011). Future studies should examine whether aging may affect audiovisual integration in other motion perception tasks, as the underlying mechanisms in the above-mentioned interactions are likely to differ from those involved in the current complex motion task.

5. Conclusion

The capacity to encode the motion direction and speed of moving objects in the scene while walking or driving is an important ability that is known to decline in older age. The current study showed that whereas the presence of a dynamic broadband noise moving coherently with an independently-moving visual target can enhance its detection at some speed levels in younger adults, older adults do not benefit from this additional auditory information. This lack of multisensory benefit indicates that aging may impair audiovisual integration in cluttered, dynamic scenes.

Acknowledgements

This research was funded by the European Commission FP7 'VERVE' Project, Grant No. 288914 awarded to FNN. We thank Dr David McGovern for many helpful discussions and all the participants for their time.

References

Alais, D., Newell, F. N. and Mamassian, P. (2010). Multisensory processing in review: from physiology to behaviour, *Seeing Perceiv.* **23**, 3–38.

Allard, R., Lagacé-Nadon, S. and Faubert, J. (2013). Feature tracking and aging, *Front. Psychol.* **4**, 427. DOI:10.3389/fpsyg.2013.00427.

Andersen, G. J. and Enriquez, A. (2006). Aging and the detection of observer and moving object collisions, *Psychol. Aging* **21**, 74–85.

Atchley, P. and Andersen, G. J. (1998). The effect of age, retinal eccentricity, and speed on the detection of optic flow components, *Psychol. Aging* **13**, 297–308.

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs, *Behav. Res. Meth.* **37**, 379–384.

Bedard, G. and Barnett-Cowan, M. (2015). Impaired timing of audiovisual events in the elderly, *Exp. Brain Res.* **234**, 331–340.

Bennett, P. J., Sekuler, R. and Sekuler, A. B. (2007). The effects of aging on motion detection and direction identification, *Vision Res.* **47**, 799–809.

Berard, J. R., Fung, J., McFadyen, B. J. and Lamontagne, A. (2009). Aging affects the ability to use optic flow in the control of heading during locomotion, *Exp. Brain Res.* **194**, 183–190.

Billino, J., Bremmer, F. and Gegenfurtner, K. R. (2008). Differential aging of motion processing mechanisms: evidence against general perceptual decline, *Vision Res.* **48**, 1254–1261.

Brainard, D. H. (1997). The psychophysics toolbox, *Spat. Vis.* **10**, 433–436.

Butler, J. S., Smith, S. T., Campos, J. L. and Bülthoff, H. H. (2010). Bayesian integration of visual and vestibular signals for heading, *J. Vis.* **10**, 23.

Butler, J. S., Campos, J. L., Bülthoff, H. H. and Smith, S. T. (2011). The role of stereo vision in visual-vestibular integration, *Seeing Perceiv.* **24**, 453–470.

Calabro, F. J. and Vaina, L. M. (2012). Interaction of cortical networks mediating object motion detection by moving observers, *Exp. Brain Res.* **221**, 177–189.

Calabro, F. J., Soto-Faraco, S. and Vaina, L. M. (2011). Acoustic facilitation of object movement detection during self-motion, *Proc. R. Soc. B Biol. Sci.* **278**, 2840–2847.

DeLucia, P. R., Kathryn Bleckley, M., Meyer, L. E. and Bush, J. M. (2003). Judgments about collision in younger and older drivers, *Transp. Res. Part F Traffic Psychol. Behav.* **6**, 63–80.

Diaconescu, A. O., Hasher, L. and McIntosh, A. R. (2012). Visual dominance and multisensory integration changes with age, *NeuroImage* **65C**, 152–166.

Diederich, A., Colonius, H. and Schomburg, A. (2008). Assessing age-related multisensory enhancement with the time-window-of-integration model, *Neuropsychologia* **46**, 2556–2562.

Dommes, A. and Cavallo, V. (2011). The role of perceptual, cognitive, and motor abilities in street-crossing decisions of young and older pedestrians, *Ophthalmic Physiol. Opt.* **31**, 292–301.

Dommes, A., Cavallo, V. and Oxley, J. (2013). Functional declines as predictors of risky street-crossing decisions in older pedestrians, *Accid. Anal. Prev.* **59**, 135–143.

Ernst, M. O. and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion, *Nature* **415**(6870), 429–433.

Faubert, J. (2002). Visual perception and aging, *Can. J. Exp. Psychol.* **56**, 164–176.

Fox, J. and Weisberg, S. (2011). *An {R} Companion to Applied Regression*, 2nd edn. SAGE, Los Angeles, CA, USA.

François, M., Morice, A. H., Blouin, J. and Montagne, G. (2011). Age-related decline in sensory processing for locomotion and interception, *Neuroscience* **172**, 366–378.

Freiherr, J., Lundstrom, J. N., Habel, U., Reetz, K., Lundström, J. N., Habel, U. and Reetz, K. (2013). Multisensory integration mechanisms during aging, *Front. Hum. Neurosci.* **7**, 863. DOI:10.3389/fnhum.2013.00863.

Harrison, N. R., Witheridge, S., Makin, A., Wuerger, S. M., Pegna, A. J. and Meyer, G. F. (2015). The effects of stereo disparity on the behavioural and electrophysiological correlates of perception of audio-visual motion in depth, *Neuropsychologia* **78**, 51–62.

Hugenschmidt, C., Peiffer, A. and McCoy, T. (2009). Preservation of crossmodal selective attention in healthy aging, *Exp. Brain Res.* **198**, 273–285.

Hutchinson, C. V., Arena, A., Allen, H. A. and Ledgeway, T. (2012). Psychophysical correlates of global motion processing in the aging visual system: a critical review, *Neurosci. Biobehav. Rev.* **36**, 1266–1272.

Jain, A., Sally, S. L. and Papathomas, T. V. (2008). Audiovisual short-term influences and aftereffects in motion: examination across three sets of directional pairings, *J. Vis.* **8**, 1–13.

Kawachi, Y., Grove, P. and Sakurai, K. (2014). A single auditory tone alters the perception of multiple visual events, *J. Vis.* **14**, 1–13.

Kitagawa, N. and Ichihara, S. (2002). Hearing visual motion in depth, *Nature* **416**(6877), 172–174.

Lakens, D. (2017). Equivalence tests: a practical primer for t-tests, correlations, and meta-analyses, *Soc. Psychol. Personal. Sci.* **1**, 1–8.

Lappe, M., Bremmer, F. and van den Berg, A. V. (1999). Perception of self-motion from visual flow, *Trends Cognitive Sci.* **3**, 329–336.

Laurienti, P. J., Burdette, J. H., Maldjian, J. A. and Wallace, M. T. (2006). Enhanced multisensory integration in older adults, *Neurobiol. Aging* **27**, 1155–1163.

Lawrence, M. A. (2015). ez: easy analysis and visualization of factorial experiments, R package version 4.3. Available at: http://cran.r-project.org/package=ez.

Lich, M. and Bremmer, F. (2014). Self-motion perception in the elderly, *Front. Hum. Neurosci.* **8**, 681. DOI:10.3389/fnhum.2014.00681.

Luis, C. A., Keegan, A. P. and Mullan, M. (2009). Cross validation of the Montreal Cognitive Assessment in community dwelling older adults residing in the Southeastern US, *Int. J. Geriatr. Psychiatry* **24**, 197–201.

Macmillan, N. A. and Creelman, C. D. (2005). *Detection Theory: a User's Guide*, 2nd edn. Lawrence Erlbaum Associates, Inc., Mahwah, NJ, USA.

Macneilage, P. R., Zhang, Z., Deangelis, G. C. and Angelaki, D. E. (2012). Vestibular facilitation of optic flow parsing, *PloS One* **7**, e40264. DOI:10.1371/journal.pone.0040264.

Maguinness, C., Setti, A., Burke, K. E., Kenny, R. A. and Newell, F. N. (2011). The effect of combined sensory and semantic components on audio-visual speech perception in older adults, *Front. Aging Neurosci.* **3**, 19. DOI:10.3389/fnagi.2011.00019.

Mahoney, J. R., Li, P. C. C., Oh-Park, M., Verghese, J. and Holtzer, R. (2011). Multisensory integration across the senses in young and old adults, *Brain Res.* **1426**, 43–53.

Mapstone, M., Logan, D. and Duffy, C. J. (2006). Cue integration for the perception and control of self-movement in ageing and Alzheimer's disease, *Brain* **129**, 2931–2944.

McGovern, D. P., Roudaia, E., Stapleton, J., McGinnity, T. M. and Newell, F. N. (2014). The sound-induced flash illusion reveals dissociable age-related effects in multisensory integration, *Front. Aging Neurosci.* **6**, 250. DOI:10.3389/fnagi.2014.00250.

Mozolic, J. L., Hugenschmidt, C. E., Peiffer, A. M. and Laurienti, P. J. (2012). Multisensory integration and aging, in: *The Neural Bases of Multisensory Processes*, M. M. Murray and M. T. Wallace (Eds). CRC Press, Boca Raton, FL, USA.

Nasreddine, Z. S., Phillips, N. A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L. and Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening, *J. Am. Ger. Soc.*, 695–699.

Odegaard, B. and Shams, L. (2016). The brain's tendency to bind audiovisual signals is stable but not general, *Psychol. Sci.* **27**, 583–591.

Olejnik, S. and Algina, J. (2003). Generalized eta and omega squared statistics: measures of effect size for some common research designs, *Psychol. Meth.* **8**, 434–447.

Owsley, C. (2010). Aging and vision, *Vision Res.* **51**, 1610–1622.

Peiffer, A. M., Mozolic, J. L., Hugenschmidt, C. E. and Laurienti, P. J. (2007). Age-related multisensory enhancement in a simple audiovisual detection task, *Neuroreport* **18**, 1077–1081.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies, *Spat. Vis.* **10**, 437–442. DOI:10.1163/156856897X00366.

Pilz, K. S., Bennett, P. J. and Sekuler, A. B. (2010). Effects of aging on biological motion discrimination, *Vision Res.* **50**, 211–219.

R Core Team (2015). *R: a Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Ramkhalawansingh, R., Keshavarz, B., Haycock, B., Shahab, S. and Campos, J. L. (2016). Age differences in visual-auditory self-motion perception during a simulated driving task, *Front. Psychol.* **7**, 595. DOI:10.3389/fpsyg.2016.00595.

Roudaia, E., Bennett, P. J., Sekuler, A. B. and Pilz, K. S. (2010). Spatiotemporal properties of apparent motion perception and aging, *J. Vis.* **10**, 5. DOI:10.1167/10.14.5.

Roudaia, E., Sekuler, A. B., Bennett, P. J. and Sekuler, R. (2013). Aging and audio-visual and multi-cue integration in motion, *Front. Psychol.* **4**, 267. DOI:10.3389/fpsyg.2013.00267.

Royden, C. S. and Connors, E. M. (2010). The detection of moving objects by moving observers, *Vision Res.* **50**, 1014–1024.

Sanabria, D., Soto-Faraco, S., Chan, J. and Spence, C. (2005). Intramodal perceptual grouping modulates multisensory integration: evidence from the crossmodal dynamic capture task, *Neurosci. Lett.* **377**, 59–64.

Schneider, B. A., Pichora-Fuller, K. and Daneman, M. (2010). Effects of senescent changes in audition and cognition on spoken language comprehension, in: *The Aging Auditory System*, S. Gordon-Salant, R. D. Frisina, A. N. Popper and R. R. Fay (Eds), pp. 167–210. Springer New York, New York, NY, USA.

Schönbrodt, F. D. and Perugini, M. (2013). At what sample size do correlations stabilize? *J. Res. Pers.* **47**, 609–612.

Schouten, B., Troje, N. F., Vroomen, J. and Verfaillie, K. (2011). The effect of looming and receding sounds on the perceived in-depth orientation of depth-ambiguous biological motion figures, *PloS One* **6**, e14725. DOI:10.1371/journal.pone.0014725.

Schuirmann, D. J. (1987). A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability, *J. Pharmacokinet. Biopharm.* **15**, 657–680.

Sekuler, R., Sekuler, A. and Lau, R. (2001). Sound alters visual motion perception, *Nature* **385**(6614), 308.

Setti, A., Burke, K. E., Kenny, R. A. and Newell, F. N. (2011). Is inefficient multisensory processing associated with falls in older people? *Exp. Brain Res.* **209**, 375–384.

Snowden, R. J. and Kavanagh, E. (2006). Motion perception in the ageing visual system: minimum motion, motion coherence, and speed discrimination thresholds, *Perception* **35**, 9–24.

Soto-Faraco, S., Kingstone, A. and Spence, C. (2003). Multisensory contributions to the perception of motion, *Neuropsychologia* **41**, 1847–1862.

Soto-Faraco, S., Spence, C. and Kingstone, A. (2005). Assessing automaticity in the audiovisual integration of motion, *Acta Psychol.* **118**, 71–92.

Soto-Faraco, S., Spence, C., Lloyd, D. and Kingstone, A. (2004). Moving multisensory research along: motion perception across sensory modalities, *Curr. Dir. Psychol. Sci.* **13**, 29–32.

Spear, P. D. (1993). Neural bases of visual deficits during aging, *Vision Res.* **33**, 2589–2609.

Talsma, D., Senkowski, D., Soto-Faraco, S. and Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration, *Trends Cogn. Sci.* **14**, 400–410.

Vaina, L. M., Calabro, F., Lin, F.-H. and Hämäläinen, M. S. (2010). Long-range coupling of prefrontal cortex and visual (MT) or polysensory (STP) cortical areas in motion perception, in: *IFMBE Proceedings*, Vol. 28, pp. 298–301.

Van der Burg, E., Cass, J. and Alais, D. (2014). Window of audio-visual simultaneity is unaffected by spatio-temporal visual clutter, *Sci. Rep.* **4**, 5098. DOI:10.1038/srep05098.

Warren, P. A. and Rushton, S. K. (2007). Perception of object trajectory: parsing retinal motion into self and object movement components, *J. Vis.* **7**, 1–11.

Warren, P. A. and Rushton, S. K. (2009). Optic flow processing for the assessment of object movement during ego movement, *Curr. Biol.* **19**, 1555–1560.

Warren, W. H., Blackwell, A. W. and Morris, M. W. (1989). Age differences in perceiving the direction of self-motion from optical flow, *J. Gerontol.* **44**, 147–153.

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, New York, NY, USA.

Wilcox, L. M. (2016). Representation and measurement of stereoscopic volumes, *J. Vis.* **16**, 16. DOI:10.1167/16.11.16.

Winneke, A. H. and Phillips, N. A. (2011). Does audiovisual speech offer a fountain of youth for old ears? An event-related brain potential study of age differences in audiovisual speech perception, *Psychol. Aging* **26**, 427–438.