



# Sound-induced flash illusion is modulated by the depth of auditory stimuli: Evidence from younger and older adults

Yawen Sun<sup>3,4</sup> · Heng Zhou<sup>1</sup> · Chunmei Liu<sup>2</sup> · Aijun Wang<sup>1</sup> · Chunlin Yue<sup>5</sup> · Ming Zhang<sup>1</sup>

Accepted: 5 July 2022 / Published online: 18 July 2022

© The Psychonomic Society, Inc. 2022

## Abstract

Sound-induced flash illusion (SiFI) refers to an auditory dominance effect in which observers misperceive the number of visual flashes due to the simultaneous presentation of a different number of auditory beeps. Although the SiFI remains robust despite many factors, little is known about how the dominance of audition over vision changes as a function of spatial location in 3D space and how this corresponds with aging. In the present study, we recruited younger (18–26 years old) and older adults (60–76 years old) and presented visual stimuli in near space and auditory stimuli in far space (Vnear\_Afar) or visual stimuli in far space and auditory stimuli in near space (Vfar\_Anear), with the aim of exploring the effect of aging on spatial dominance in SiFI. The results showed that both participant groups exhibited a SiFI auditory dominance effect; however, older adults were more susceptible to illusions than younger adults. More importantly, for younger adults, when auditory stimuli were presented in far space, the accuracy and  $d'$  indicated that SiFI was significantly enhanced in the fission illusion, and the reaction time (RT) results showed that RT became longer in the fusion illusion. However, for older adults, the SiFI was not affected by the spatial location. Therefore, the present results reveal for the first time that the SiFI effect is modulated by the depth of auditory stimuli. Moreover, older adults were not susceptible to the depth dimension, thus showing age-related sensory decline and specifically showing a decreased ability to localize visual and auditory stimuli.

**Keywords** Sound-induced flash illusion (SiFI) · Auditory dominance · Near and far spaces · Fission and fusion illusions

## Introduction

Our bodies receive various types of information simultaneously, and most of this information is obtained from multisensory modalities. Two main behavioral outcomes of multisensory integration have been identified. One of the two outcomes is the redundancy effect, which involves the merging of information across senses. However, in some cases, one sensory modality from multisensory information also competes for preference for consciousness, which involves the sensory

dominance effect. Specifically, one modality dominates the others to receive preferential processing and ultimately dominates the awareness and behavior of observers (Chen & Zhou, 2013; Driver & Noesselt, 2008; Hirst et al., 2018; Huang et al., 2015; Koelewijn et al., 2010; Spence, 2011; Stein & Meredith, 1993; Talsma et al., 2010).

In most cases involving spatial information, visual information more frequently dominates other sensory modalities. However, especially in regard to temporal information, auditory information also dominates other sensory modalities

Yawen Sun, Heng Zhou and Chunmei Liu contributed equally to this work.

✉ Aijun Wang  
ajwang@suda.edu.cn

✉ Chunlin Yue  
y\_chunlin002@163.com

✉ Ming Zhang  
psyzm@suda.edu.cn

<sup>1</sup> Department of Psychology, Research Center for Psychology and Behavioral Sciences, Soochow University, Suzhou, China

<sup>2</sup> Jiangsu Provincial Key Constructive Laboratory for Big Data of Psychology and Cognitive Science, Yancheng Teachers University, Yancheng, China

<sup>3</sup> Institute of Psychology, Chinese Academy of Sciences, Beijing, China

<sup>4</sup> Department of Psychology, University of Chinese Academy of Sciences, Beijing, China

<sup>5</sup> School of Physical Education and Sport Science, Soochow University, Suzhou, China

(Bresciani et al., 2008; Repp, 2000, 2002; Wang et al., 2019). Sound-induced flash illusion (SiFI) is a typical phenomenon of information competition defined by Shams et al. (2000). When a certain number of visual flashes are presented with an unequal number of auditory beeps that are either presented successively or simultaneously within 100 ms, observers will misperceive the number of visual flashes due to the successive or simultaneous presentation of different numbers of auditory beeps (Hirst et al., 2020; Keil, 2020). Auditory information receives preferential processing over visual information (Shams et al., 2000, 2002). SiFI has been divided into fission and fusion illusions (Andersen et al., 2004; Shams et al., 2000, 2002). The fission illusion occurs when two visual flashes are perceived if a single visual flash is presented accompanied by two auditory stimuli (Shams et al., 2000, 2002); the fusion illusion occurs when one flash is perceived if two flashes are presented with one auditory stimulus (Andersen et al., 2004). Some studies have suggested that a greater SiFI effect corresponds to a lower visual perceptual sensitivity, which is calculated by  $d'$ , and looser subjective judgment criteria, which are calculated by  $\beta$  (Keil, 2020; Kumpik et al., 2014; McCormick & Mamassian, 2008).

Existing studies have shown that SiFI perception varies across individuals (Kaiser et al., 2019; Keil & Senkowski, 2018), ranging from 3% to 86% (Mishra et al., 2007). Studies have also suggested that neural oscillations orchestrate the individual perception of the SiFI effect (Kaiser et al., 2019; Keil & Senkowski, 2018; Lange et al., 2014). Moreover, some studies have illustrated that perceptual processing changes significantly with aging (Cabeza et al., 2004; Grady, 2010; Salat et al., 2009), which involves multisensory perception. Older adults display increased multisensory integration compared to younger adults (Hirst et al., 2020). Setti et al. (2011) first indicated a greater effect of beeps on older adults judging the number of visual flashes than on younger adults, and subsequent studies have confirmed the results (DeLoss et al., 2013). In addition, older adults are susceptible to SiFI (Hernández et al., 2019), especially fission illusions, across a much wider range of temporal asynchronies in SiFI (McGovern et al., 2014; Setti et al., 2011), presumably due to older adults' enlarged temporal window of integration (TWI) (Diederich et al., 2008; Laurienti et al., 2006; Peiffer et al., 2007; Wu et al., 2012). However, one recent study indicated a lower perceptual sensitivity of auditory stimuli in older adults; consequently, they were more dependent on visual stimuli (Sun et al., 2020). At the same time, age-related differences in integration effects have been reported to be due to sensory modality and measurements. Mahoney et al. (2011) found that younger adults but not older adults exhibited a significant increase in the magnitude of coactivation for auditory-somatosensory (AS) and auditory-visual (AV) pairings (Mahoney et al., 2011), whereas older adults presented a significantly greater reaction time (RT) benefit when

processing concurrent visual-somatosensory (VS) coactivation using a race model. In general, the evidence for audiovisual integration in older adults is controversial (Chan et al., 2018; DeLoss et al., 2013; DeLoss & Andersen, 2015; McGovern et al., 2014; Sun et al., 2020).

SiFI has been revealed to be robust, irrespective of a variety of experimental manipulations, such as the interstimulus time interval (beep SOA (stimulus onset asynchrony) and flash SOA) (Shams et al., 2000, 2002; Shams et al., 2005), stimulus properties (Andersen et al., 2004; Setti & Chan, 2011), spatial location (Abadi & Murphy, 2014), feedback mode (Rosenthal et al., 2009), and the amount of attention allocated (Mishra et al., 2008, 2010). However, all studies on SiFI have been conducted in two-dimensional (2D) space, while we live in a real 3D world with depth. Therefore, a critical factor that has been ignored in current studies is the "depth" in the real environment. Another key factor that has been ignored is the spatial distance from the perceived sensory information to the observers along the depth dimension of the 3D world, which would help observers evaluate a potentially threatening/rewarding stimulus. In real life, the stimuli that appear close might capture observers' attention to ensure survival; conversely, stimuli far from observers may require less attention (Graziano & Cooke, 2006; Previc, 1998). More importantly, although studies have found that SiFI auditory dominance is not affected when visual and auditory stimuli are presented in the same or different spatial locations in 2D space (Abadi & Murphy, 2014; DeLoss & Andersen, 2015; Innes-Brown & Crewther, 2009), researchers do not clearly understand whether sensory dominance changes with the depth of sensory input.

Early evidence indicated that attention resources decrease with the distance from the observers' position in 3D space, with the maximal attentional focus in the visual space around the observers' body (Maringelli et al., 2001). Therefore, fewer attentional resources are allocated to far space than to near space. When attentional resources are limited, the competitive advantage of the dominant modality may increase, leading to a higher cost of the non-dominant modality (Desimone et al., 1990; Desimone & Duncan, 1995). As Michail and Keil (2018) recently reported, participants' SiFI perception increases when attentional resources are limited. Therefore, because SiFI is an illusion effect in which audition dominates vision, if far-away stimuli are allocated fewer attentional resources than near stimuli, a plausible hypothesis is that auditory stimuli in far space will obtain more competitive advantages than vision; thus, the frequency of the SiFI will increase compared to near-space auditory stimuli combined with far-space visual stimuli. Meanwhile, according to the perception/action model (Goodale & Milner, 1992), the representation of objects in the near space is achieved through movement, which activates the dorsal visual pathway, while the representation of objects in the far space is achieved through

perception, which activates the ventral visual pathway. Thus, sensory dominance may also be influenced by 3D space, as multisensory gain (tendency to integrate stimuli) is greater when dominant stimuli are presented in the far space compared to the near space (Noel et al., 2018; Previc, 1998; Van der Stoep et al., 2016a, b). Recent research has discovered sensory dominance changes along the depth dimension in the 3D world, as reported in a study on the Colavita visual dominance effect (Yue et al., 2015). The authors found that the magnitude of the visual dominance effect was significantly increased as long as the visual stimuli were presented in far space, irrespective of whether the auditory stimuli were presented in near or far space. However, researchers have not determined how auditory stimuli in far space affect visual flashes in near space when attention is focused on near space and how auditory stimuli in near space affect visual flashes in far space when attention is focused on far space.

Based on these results, in the present study, by presenting visual and audiovisual stimuli either in near or in far space, we hypothesized that the dominance of audition over vision would change as a function of the spatial location, i.e., attention resources in far space would be less than those in near space, the competitive advantage of auditory modality may increase, and the visual modality would have more difficulty capturing attention resources, which would result in a larger auditory advantage in far space than in near space. In the experiment, visual unimodal stimuli were presented in far or near space, and audiovisual bimodal stimuli were always presented at different distances. If the SiFI auditory dominance effect was affected by spatial dominance in the depth dimension, then we would find that when the auditory stimuli were presented in far space and the visual stimuli were in near space, observers would be more susceptible to SiFI. Moreover, an age-related decline in localized visual and auditory stimuli (Abel et al., 2000; Andersen & Atchley, 1995; Norman et al., 2000) seemed to predict the difference in performance compared to younger adults in this study. Studies have reported older adults' degradation in vertical and front-back discrimination of auditory space (Abel et al., 2000; Dobrova et al., 2011; Noble et al., 1994) and the depth of visual spatial attention may be somewhat less for older adults (Pierce et al., 2011). Overall, we assumed that when the auditory stimuli were presented in far space, increased auditory dominance would be observed compared to near space, namely, more SiFI. Additionally, due to the decline in older adults' visual and auditory spatial localization abilities, older adults might be less susceptible to spatial location factors.

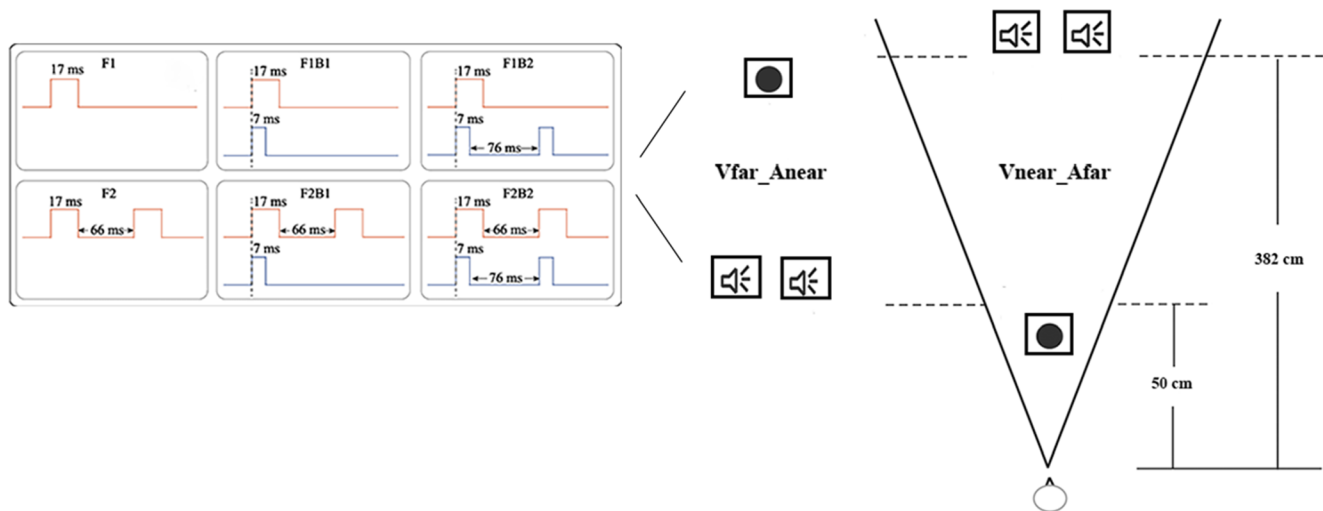
## Materials and methods

**Participants** Thirty-nine older adults and 29 younger adults participated in the present experiment. If the accuracy of any

of the non-illusion conditions (F1, F2, F1B1, and F2B2) was less than 60%, the participant was considered to be unable to identify the stimulus properly, and the data from the participant were rejected. Based on this criterion, 14 older adults and five younger adults were excluded from the experiment. Therefore, ultimately, 25 older adults (11 males and 14 females) aged 60–76 years (mean age = 64 years) and 24 college students (eight males and 16 females) aged 18–26 years (mean age = 22 years) were included in the analysis. All participants were naive to the experimental procedure and were paid for their participation in the experiment. All participants were rescreened for self-reported eye disease, neurological disorders (e.g., Alzheimer's disease, Parkinson's disease, or stroke), and any significant hearing loss. All participants had normal or corrected-to-normal visual acuity. All participants provided written informed consent according to the standards of the Declaration of Helsinki. The study was approved by the Ethics Committee of the Department of Psychology, Soochow University.

We used G\*Power 3.1.9.2 to conduct a sensitivity analysis of an *F* test and to evaluate the statistical power in the present study (Faul et al., 2007, 2009). The input parameters were  $\alpha$  err prob = 0.05, power ( $1 - \beta$  err prob) = 0.80, total sample size = 49, number of groups = 2, and number of measurements = 2, and the output was an effect size  $f = 0.29$ , which indicated sufficient power to detect moderately sized effects (Cohen, 1988). We used SPSS 226.0 to perform ANOVA and *t* tests for accuracy and RTs. Additionally, we also conducted Bayesian tests (Wetzels & Wagenmakers, 2012) using JASP 0.16.

**Stimuli and apparatus** As shown in Fig. 1, all stimuli were presented at two distances from participants: 50 cm for the near condition and 382 cm for the far condition. In near space, all visual stimuli were presented on a ThinkPad X390 laptop computer with a screen resolution of  $1,024 \times 768$  pixels and a refresh rate of 60 Hz. The visual stimuli were white discs (with a radius of view of  $2^\circ$ ) presented at a  $5^\circ$  viewing angle below the central fixation point for 17 ms by Presentation software (Neurobehavioral Systems Inc.). The visual stimuli were presented at a  $5^\circ$  angle of view below the central fixation point because the visual stimuli had the greatest illusion in the peripheral field of view when presented with the accompanying auditory stimuli (Shams et al., 2002). The auditory target was a pure tone (3.5 kHz, 75 dB, for 7 ms, from Shams et al., 2002) that was presented at 50 cm via two loudspeakers on either side of the monitor. In far space, visual stimuli were presented on the screen by a projector (EPSON CB-X06E), while speakers were placed on either side of the monitor, both of which were located 382 cm away from the participants. The angle of view ( $2^\circ$ ) and luminance ( $9.5 \text{ cd/m}^2$ ) were the same in both



**Fig. 1** Flow chart of each trial. The left side of the figure shows the time flow chart of F1, F2, F1B1, F2B2, F1B2, and F2B1, and the right side shows the diagram of the experimental setup (using F1B2 as an example). The behavioral targets included a 7-ms auditory pure tone (blue lines) and a 17-ms visual circle (red lines). Two types of stimuli were used: unimodal visual stimuli (F1 and F2) in which only the visual target was presented and bimodal stimuli in which the auditory and visual targets

were simultaneously presented (F1, F2, F1B1, F1B2, F2B1, and F2B2). The visual and auditory stimuli were presented at different spatial distances (50 cm for the near condition and 382 cm for the far condition). Two types of spatial locations were used: Vfar\_Anear refers to visual stimuli presented in near space and auditory stimuli presented in far space, and Vnear\_Anear refers to visual stimuli presented in far space and auditory stimuli presented in near space

near and far spaces. Similarly, the loudness of the auditory stimuli (75 dB) in the near and far space was matched (the decibel meter was placed at the participants' head position).

**Experimental design and procedure** The experiment was a 2 (participant group: older vs. younger adults) × 2 (spatial location: Vfar\_Anear vs. Vnear\_Anear) × 2 (illusion type: fission illusion vs. fusion illusion) mixed design, the participant group was the between-group variable, and others were considered within-group variables. For the sake of discussion, trials are named using this unified expression. For example, F = flash, B = beep, F2 means only two visual stimuli were presented, and F2B1 means that two visual stimuli were accompanied by one auditory stimulus. We were more interested in the two types of illusions (fission and fusion illusions); therefore, in our comprehensive analysis, we only focused on the F1B2 and F2B1 conditions.

At the beginning of the experiment, participants were required to pass a test to determine whether they understood the task and were able to discriminate the beeps or flashes in isolation. The formal experimental procedure is shown in Fig. 1. Stimuli were presented 50 cm (near) or 382 cm (far) away from the monitor and the loudspeakers in a dimly lit room. Participants were instructed to place their chin on a chinrest and maintain central fixation throughout the experiment. Two types of stimuli were used: unimodal visual stimuli (F1 and F2) and bimodal audiovisual stimuli (F1B1, F1B2, F2B1, and F2B2). In each trial, one or two visual target stimuli (the duration was 17 ms) accompanied by no, one, or two

auditory stimuli (the duration was 7 ms) were presented. The formal experiment consisted of two sections, including the Vnear\_Anear condition (visual stimuli were presented 50 cm (near) and auditory stimuli were presented 382 cm (far) away from the monitor and the loudspeakers) and the Vfar\_Anear condition (visual stimuli were presented 382 cm (far) and auditory stimuli were presented 50 cm (near) away from the monitor and the loudspeakers). The participants' task was to determine if they perceived one or two visual stimuli by pressing the left or right mouse button within 1,500 ms after the stimuli were presented. We balanced the button response, with half of the participants pressing the left button and half pressing the right button if they perceived one flash. Each participant was required to complete 384 trials (96 trials per block, four blocks in total) in each space condition, 64 trials under each experimental condition, and the intertrial interval (ITI) between trials was randomized from 400 ms to 700 ms in steps of 100 ms. Half of the participants performed the Vnear\_Anear condition first, and the other half of the participants performed the opposite condition first.

auditory stimuli (the duration was 7 ms) were presented. The formal experiment consisted of two sections, including the Vnear\_Anear condition (visual stimuli were presented 50 cm (near) and auditory stimuli were presented 382 cm (far) away from the monitor and the loudspeakers) and the Vfar\_Anear condition (visual stimuli were presented 382 cm (far) and auditory stimuli were presented 50 cm (near) away from the monitor and the loudspeakers). The participants' task was to determine if they perceived one or two visual stimuli by pressing the left or right mouse button within 1,500 ms after the stimuli were presented. We balanced the button response, with half of the participants pressing the left button and half pressing the right button if they perceived one flash. Each participant was required to complete 384 trials (96 trials per block, four blocks in total) in each space condition, 64 trials under each experimental condition, and the intertrial interval (ITI) between trials was randomized from 400 ms to 700 ms in steps of 100 ms. Half of the participants performed the Vnear\_Anear condition first, and the other half of the participants performed the opposite condition first.

## Results

**Accuracy** For conditions in which the visual flash was presented alone, the accuracy of F1 and F2 were greater than 80% in both participant groups and spatial locations (see Table 1); these results were significantly larger than a random level 50% ( $p < 0.001$ ), which indicated that all participants had the ability to correctly identify the number of flashes. We

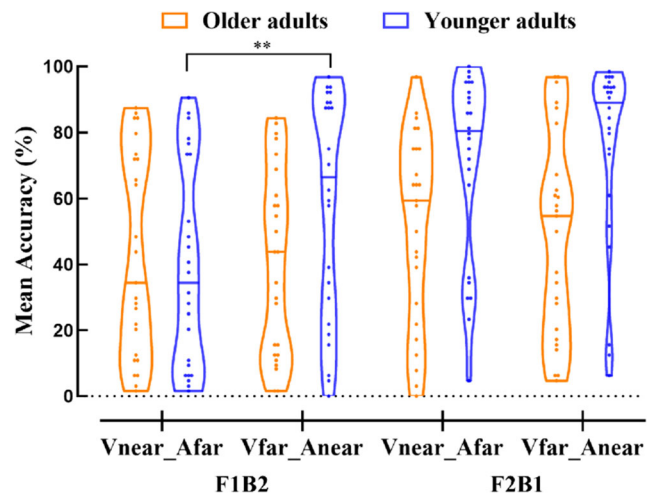


**Table 1** Mean accuracy of older and younger adults under illusion conditions (M ± SD)

	Older adults		Younger adults	
	Vnear_Afar	Vfar_Anear	Vnear_Afar	Vfar_Anear
F1	90.69 ± 6.80	92.65 ± 6.32	94.33 ± 5.13	94.66 ± 4.80
F2	87.37 ± 10.94	84.14 ± 10.00	92.13 ± 6.76	94.60 ± 5.18
F1B1	94.40 ± 6.57	95.54 ± 4.36	95.05 ± 7.49	95.95 ± 4.74
F1B2	42.94 ± 30.87	40.65 ± 28.06	39.19 ± 30.95	58.55 ± 34.04
F2B1	51.89 ± 28.80	49.30 ± 30.85	70.90 ± 28.43	75.33 ± 28.54
F2B2	92.95 ± 8.33	93.50 ± 9.13	97.52 ± 3.91	97.20 ± 3.02

conducted a 2 (participant group: older vs. younger adults) × 2 (spatial location: Vfar\_Anear vs. Vnear\_Afar) × 4 (condition: F1B1 vs. F1B2 vs. F2B1 vs. F2B2) repeated-measures ANOVA to further analyze the participants' performance. The main effect of the participant group was significant,  $F(1, 47) = 6.83, p = 0.012, \eta_p^2 = 0.13, BF_{10} = 39.45$ , indicating that the accuracy of younger adults (79%) was significantly greater than that of older adults (70%). The main effect of spatial location was significant,  $F(1, 47) = 4.32, p = 0.043, \eta_p^2 = 0.08, BF_{10} = 0.15$ , indicating that the accuracy of Vfar\_Anear (76%) was significantly greater than that of Vnear\_Afar (73%). The main effect of the condition was significant,  $F(3, 141) = 81.68, p < 0.001, \eta_p^2 = 0.64, BF_{10} = 3.249e+13$ . Post hoc tests showed that the accuracy in the F1B1 condition (95%) was significantly higher than the accuracy in the F1B2 condition (45%), and the accuracy in the F2B2 condition (95%) was significantly higher than the accuracy in the F2B1 condition (61%), indicating that the auditory dominance effect occurred, namely, the fission and the fusion illusion. The interaction between the participant group and spatial location was significant,  $F(1, 47) = 8.70, p = 0.005, \eta_p^2 = 0.16, BF_{10} = 0.30$ . The interaction between the participant group and condition was significant,  $F(3, 141) = 3.00, p = 0.033, \eta_p^2 = 0.06, BF_{10} = 55.59$ . The interaction between the spatial location and condition was significant,  $F(3, 141) = 3.64, p = 0.014, \eta_p^2 = 0.07, BF_{10} = 0.06$ . The interaction among the spatial location, condition, and participant group was significant,  $F(3, 141) = 5.90, p < 0.001, \eta_p^2 = 0.11, BF_{10} = 0.07$ . Although some interactions of the Bayesian factors produced weak evidence in support of the null hypothesis, the appearance of the fission and fusion illusion was supported by the Bayesian factors.

We performed additional analyses to determine how the fission and fusion illusions were affected by the participant group and the spatial location, respectively. For the F1B2 condition (Fig. 2 left), we conducted a 2 (participant group: older vs. younger adults) × 2 (spatial location: Vfar\_Anear vs. Vnear\_Afar) repeated-measures ANOVA. The main effect of the participant group was not significant,  $F < 1, BF_{10} = 2.63$ .



**Fig. 2** Mean accuracy of the younger and older adults for the fission (left data) and fusion (right data) illusion conditions under the Vfar\_Anear and Vnear\_Anear conditions. Two types of illusions were shown: fission illusion (F1B2) and fusion illusion (F2B1). F1B2 refers to a visual stimulus accompanied by two auditory stimuli. F2B1 refers to two visual stimuli accompanied by an auditory stimulus. The conditions Vfar\_Anear and Vnear\_Afar are the same as those described in Fig. 1

The main effect of the spatial location was significant,  $F(1, 47) = 5.89, p = 0.019, \eta_p^2 = 0.11, BF_{10} = 4.98$ , indicating that the accuracy of Vfar\_Anear (49%) was significantly greater than that of Vnear\_Afar (41%). The interaction between the participant group and condition was significant,  $F(1, 47) = 9.50, p = 0.003, \eta_p^2 = 0.17, BF_{10} = 9.14$ . A simple analysis of the interaction effects was performed. No significant effects were observed in the Vfar\_Anear,  $t(47) = 2.02, p = 0.29, BF_{10} = 1.44$ , and Vnear\_Afar conditions,  $t < 1, BF_{10} = 0.31$ . For the older adults, no significant effect was observed,  $t < 1, BF_{10} = 0.24$ . However, for the younger adults, the accuracy of Vfar\_Anear (59%) was significantly greater than that of Vnear\_Afar (39%),  $t(23) = 3.86, p = 0.002, Cohen's d = 0.71, CI = [0.08, 0.31], BF_{10} = 19.12$ . The results suggested that for older adults, the fission illusion was not affected by the spatial location. For younger adults, the fission illusion was affected by the spatial location, and the fission illusion was much greater in the Vnear\_Afar condition than in the Vfar\_Anear condition.

For the F2B1 condition (Fig. 2 right), we conducted a 2 (participant group: older vs. younger adults) × 2 (spatial location: Vfar\_Anear vs. Vnear\_Afar) repeated-measures ANOVA. The main effect of the participant group was significant,  $F(1, 47) = 7.92, p = 0.007, \eta_p^2 = 0.14, BF_{10} = 4.35$ , indicating that the accuracy of younger adults (73%) was significantly greater than that of older adults (51%). The main effect of spatial location was not significant,  $F < 1, BF_{10} = 0.25$ . The interaction between the participant group and condition was not significant,  $F(1, 47) = 2.23, p = 0.142, \eta_p^2 = 0.05, BF_{10} = 0.47$ . Based on these results, the fusion illusion

was not affected by the spatial location, and the fusion illusion was much greater in older adults than in younger adults.

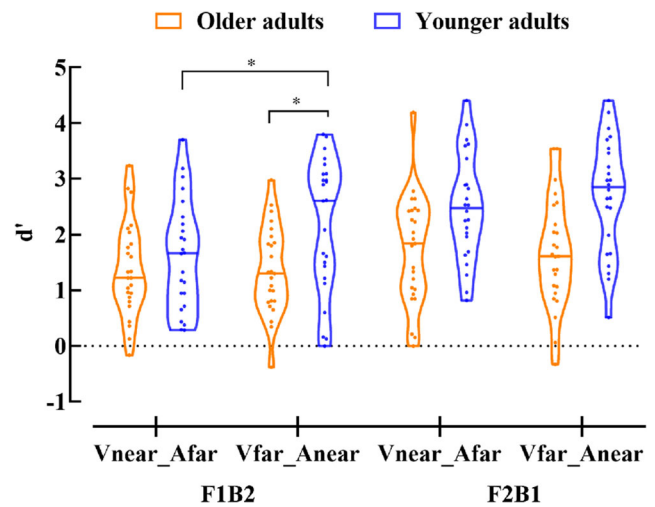
**Signal detection theory analysis** We next determined whether the older and younger adults' different magnitudes of fission and fusion illusion at different spatial locations were related to changes in flash discriminability and/or criteria for reporting the number of flashes caused by beeps. The data were analyzed according to signal detection theory (Chen et al., 2017; Vanes et al., 2016). Conditions were divided into fission conditions (F1B2 and F2B2) and fusion conditions (F1B1 and F2B1). For fission conditions, two flashes were defined as the signal; therefore, in F2B2, the reaction of "2 flash" was a hit, while the reaction of "1 flash" was a miss. In contrast, in F1B2, the "2 flash" response was a false alarm (FA), and the "1 flash" response was a correct rejection. For fusion conditions, one flash was defined as the signal; therefore, in F1B1, the reaction of "1 flash" was a hit, while the reaction of "2 flash" was a miss. In contrast, in F2B1, the "1 flash" response was a false alarm (FA), and the "2 flash" response was a correct rejection. The parameter  $d'$  represents the discriminability of whether one or two flashes were presented, and  $\ln(\beta)$  represents the response criterion or tendency, which was calculated using the following equations (Macmillan & Creelman, 2004; Rosenthal et al., 2009):

$$d' = z(\text{Hit rate}) - z(\text{FA})$$

$$\ln(\beta) = \frac{z(\text{FA})^2 - z(\text{Hit rate})^2}{2}$$

For  $p = 0$  and  $p = 1$  event considerations, log-linear transformations were applied to calculate the hit and false alarm rates (adding 0.5 each to the number of hits and false alarms and adding 1 each to the total number of signal trials and no-signal trials) to avoid extreme  $d'$  values and  $\ln(\beta)$  (e.g., Vanes et al., 2016).

**Sensitivity ( $d'$ )** We analyzed  $d'$  in the illusion condition (F1B2 and F2B1) to supplement the accuracy results. For the F1B2 condition (Fig. 3, left panel), we conducted a 2 (participant group: older vs. younger adults)  $\times$  2 (spatial location: Vfar\_Anear vs. Vnear\_Afar) repeated-measures ANOVA. The main effect of the participant group was significant,  $F(1, 47) = 5.05, p = 0.029, \eta_p^2 = 0.10, \text{BF}_{10} = 3.18$ , indicating that the  $d'$  of younger adults (1.91) was significantly greater than that of older adults (1.36). The main effect of the spatial location was significant,  $F(1, 47) = 4.32, p = 0.043, \eta_p^2 = 0.08, \text{BF}_{10} = 1.78$ , indicating that the  $d'$  of Vfar\_Anear (1.77) was significantly greater than that of Vnear\_Afar (1.50). The interaction between the participant group and condition was significant,  $F(1, 47) = 5.07, p = 0.029, \eta_p^2 = 0.10, \text{BF}_{10} =$



**Fig. 3** Mean  $d'$  under illusion conditions for older (left data) and younger (right data) adults. Two types of illusions were shown: fission illusion (F1B2) and fusion illusion (F2B1). The conditions Vfar\_Anear and Vnear\_Afar are the same as those described in Fig. 1

2.86. A simple effect analysis was conducted. In the Vfar\_Anear condition, the  $d'$  of younger adults (2.19) was significantly greater than that of older adults (1.35),  $t(46) = 3.04, p = 0.020$ , Cohen's  $d = 0.83, CI = [0.09, 1.58], \text{BF}_{10} = 7.53$ . In the Vnear\_Afar condition, a significant effect was not observed,  $t < 1, \text{BF}_{10} = 0.43$ . For the older adults, no significant effect was observed,  $t < 1, \text{BF}_{10} = 0.33$ . However, for the younger adults, the  $d'$  of Vfar\_Anear (2.19) was significantly greater than that of Vnear\_Afar (1.63),  $t(23) = 3.03, p = 0.024$ , Cohen's  $d = 0.56, CI = [0.12, 0.99], \text{BF}_{10} = 4.31$ . The results suggested that older adults' judgment of flashes was not affected by the spatial location in the fission illusion condition. For younger adults, the participants' judgment of flashes was affected by the spatial location, and the fission illusion was much greater under the Vnear\_Afar condition than under the Vfar\_Anear condition. In addition, the fission illusion of older adults was greater than that of the younger adults under the Vfar\_Anear condition, but no difference was observed under the Vnear\_Afar condition.

For the F2B1 condition (Fig. 3, right panel), we conducted a 2 (participant group: older vs. younger adults)  $\times$  2 (spatial location: Vfar\_Anear vs. Vnear\_Afar) repeated-measures ANOVA. The main effect of the participant group was significant,  $F(1, 47) = 12.35, p = 0.001, \eta_p^2 = 0.21, \text{BF}_{10} = 21.65$ , indicating that the  $d'$  of younger adults (2.60) was significantly greater than that of older adults (1.70). The main effect of the spatial location was not significant,  $F < 1, \text{BF}_{10} = 0.26$ . The interaction between the participant group and condition was not significant,  $F(1, 47) = 1.48, p = 0.23, \eta_p^2 = 0.03, \text{BF}_{10} = 0.41$ . Based on these results, the participants' judgment of flashes was not affected by the spatial location in the fusion illusion condition, and the fusion illusion was much greater in older adults than younger adults.

**Criterion  $\ln(\beta)$**  We analyzed the criterion  $\ln(\beta)$  according to the illusion condition (F1B2 and F2B1). For the F1B2 condition, we conducted a 2 (participant group: older vs. younger adults)  $\times$  2 (spatial location: Vfar\_Anear vs. Vnear\_Afar) repeated-measures ANOVA. The main effect of the participant group was not significant,  $F(1, 47) = 2.84, p = 0.098, \eta_p^2 = 0.06, BF_{10} = 0.74$ . The main effect of the spatial location was not significant,  $F < 1, BF_{10} = 0.24$ . The interaction between the participant group and condition was not significant,  $F(1, 47) = 2.89, p = 0.096, \eta_p^2 = 0.06, BF_{10} = 0.34$ . For the F2B1 condition, we conducted a 2 (participant group: older vs. younger adults)  $\times$  2 (spatial location: Vfar\_Anear vs. Vnear\_Afar) repeated-measures ANOVA. The main effect of the participant group was not significant,  $F < 1, BF_{10} = 0.27$ . The main effect of the spatial location was not significant,  $F < 1, BF_{10} = 0.18$ . The interaction between the participant group and condition was not significant,  $F < 1, BF_{10} = 0.08$ . Thus, all participants' response criterion  $\ln(\beta)$  was consistent.

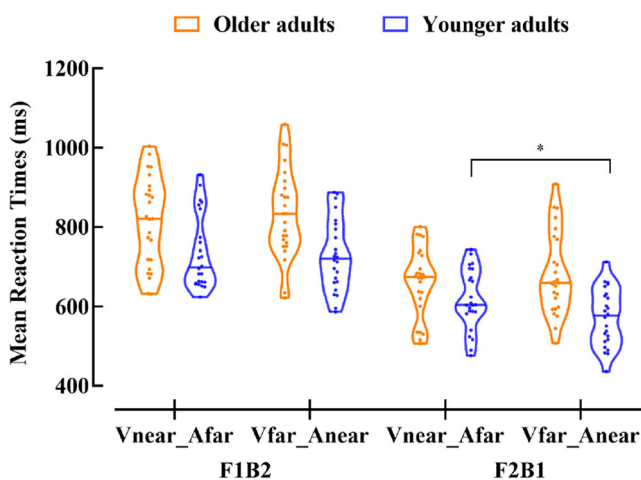
**Reaction times** We analyzed the RT of the correct trials under the illusion condition (F1B2 and F2B1). For the F1B2 condition (Fig. 4, left panel), we conducted a 2 (participant group: older vs. younger adults)  $\times$  2 (spatial location: Vfar\_Anear vs. Vnear\_Afar) repeated-measures ANOVA. The main effect of the participant group was significant,  $F(1, 47) = 11.43, p = 0.001, \eta_p^2 = 0.20, BF_{10} = 15.58$ , indicating that the RT of older adults (824 ms) was significantly longer than that of younger adults (732 ms). The main effect of the spatial location was not significant,  $F < 1, BF_{10} = 0.26$ . The interaction between the participant group and condition was not significant,  $F(1, 47) = 1.38, p = 0.25, \eta_p^2 = 0.03, BF_{10} = 0.38$ . The results suggested that older adults took longer to make correct

judgments than younger adults under the fission illusion condition.

For the F2B1 condition (Fig. 4, right panel), we conducted a 2 (participant group: older vs. younger adults)  $\times$  2 (spatial location: Vfar\_Anear vs. Vnear\_Afar) repeated-measures ANOVA. The main effect of the participant group was significant,  $F(1, 47) = 12.69, p = 0.001, \eta_p^2 = 0.21, BF_{10} = 69.04$ , indicating that the RT of older adults (673 ms) was significantly longer than that of younger adults (595 ms). The main effect of the spatial location was not significant,  $F(1, 47) = 1.33, p = 0.25, \eta_p^2 = 0.03, BF_{10} = 1.78$ . The interaction between the participant group and condition was significant,  $F(1, 47) = 8.48, p = 0.005, \eta_p^2 = 0.15, BF_{10} = 7.05$ . A simple effect analysis was conducted. Under the Vfar\_Anear condition, the RT of older adults (683 ms) was significantly longer than that of younger adults (572 ms),  $t(46) = 4.50, p < 0.001$ , Cohen's  $d = 1.23, CI = [43.81, 117.15], BF_{10} = 245.19$ , and under the Vnear\_Afar condition, no significant effect was observed,  $t(46) = 1.84, p = 0.422, BF_{10} = 1.29$ . For the older adults, no significant effect was observed,  $t(24) = 1.10, p = 0.28, BF_{10} = 0.36$ . However, for the younger adults, the RT of Vnear\_Afar (618 ms) was significantly greater than that of Vfar\_Anear (572 ms),  $t(23) = 2.85, p = 0.039$ , Cohen's  $d = 0.71, CI = [18.37, 72.88], BF_{10} = 18.29$ . Therefore, under the fusion illusion condition, older adults spent more time than younger adults only in the Vfar\_Anear condition, but not the Vnear\_Afar condition. In addition, the response time of older adults was not affected by the spatial location, while the response time of younger adults was longer in Vnear\_Afar than in Vfar\_Anear.

## Discussion

The present study adopted the classic SiFI paradigm (Shams et al., 2000, 2002) to investigate how auditory stimuli presented in far space affected visual flashes presented in near space when attention was focused on the near space, how auditory stimuli presented in near space affected visual flashes presented in far space when attention was focused on the far space, and how the effect interacted with aging when audiovisual bimodal stimuli were presented in either near or in far space. Although both groups showed the same tendency to respond, older adults were more susceptible (lower accuracy and  $d'$ ) to the fusion illusion than younger adults. Moreover, the magnitude of the fission illusion was enhanced when auditory stimuli were presented in far versus near space to younger adults but not among older adults (Fig. 2). In terms of the RT results, older adults consistently showed a slower response to the fission illusion than younger adults. However, older adults were slower to respond to the fusion illusion than younger adults only under the Vfar\_Anear condition. Younger adults were slower to respond to Vnear\_Afar than to Vfar\_Anear (Fig.



**Fig. 4** Mean reaction times of older (left data) and younger (right data) adults under illusion conditions. Two types of illusions were showed: fission illusion (F1B2) and fusion illusion (F2B1). The conditions Vfar\_Anear and Vnear\_Afar are the same as those described in Fig. 1

3). Therefore, the SiFI auditory dominance of younger adults was enhanced under the *Vnear\_Afar* condition, while the older adults' SiFI was not affected by the spatial location.

The classic SiFI effect (Shams et al., 2000) has been documented in both older and younger adults in both the present study and previous studies. Notably, older adults have shown a larger fusion illusion and longer RTs than younger adults in the fission illusion, consistent with previous studies reporting that older adults were more susceptible to the SiFI, showing a larger magnitude of the illusion (Deloss et al., 2013; McGovern et al., 2014; Setti et al., 2011). Lower perceptual sensitivity (lower  $d'$ ) might be the cause of a greater SiFI and could therefore be a measure of susceptibility to illusions (Kumpik et al., 2014; McCormick & Mamassian, 2008). Therefore, the results indicated that older adults' perceptual sensitivity to visual stimuli was lower under the fusion illusion condition, resulting in older adults being more dependent on auditory stimuli and experiencing a greater SiFI. However, the present study did not reveal differences in fission illusion conditions between older and younger adults, which was inconsistent with previous studies (Deloss et al., 2013; Setti et al., 2011) and may be due to the effect of spatial location on the fission illusion.

According to the accuracy results, auditory information was more dominant in *Vnear\_Afar* than in *Vfar\_Anear* for younger adults in F1B2; the signal detection theory analysis also indicated that younger adults had a lower  $d'$  when the auditory stimuli were presented in far space, suggesting that younger adults were more sensitive to the fission illusion. As shown in previous studies, attentional resources are mostly centered around the observer (Downing & Pinker, 1985; Gawryszewski et al., 1987; Maringelli et al., 2001); therefore, fewer attentional resources are allocated to far space than to near space. Attentional resources are shared across sensory modalities (Klemen et al., 2009; Sinnett et al., 2007); as a result, while attention resources are limited, the competitive advantage of the dominant modality potentially increases (Desimone et al., 1990; Desimone & Duncan, 1995). Sinnett et al. (2007) found that the magnitude of the visual dominance effect decreased when available attentional resources were manipulated directly by reducing the perceptual load in the auditory modality; the magnitude of the visual dominance effect was also reduced when the direction of attention was biased to the auditory modality (Sinnett et al., 2007), indicating that the overall attentional resources are limited and stimuli from visual and auditory modalities competed for attentional resources. Taken together, regarding the SiFI, overall attentional resources decreased in far space compared to near space, resulting in the competitive advantages of the auditory stimuli increasing in far space compared to near space and a greater SiFI under the *Vnear\_Afar* condition. Yue et al. (2015) found that the visual dominance effect (Colavita effect) increased when visual stimuli were presented in far space. Our

study further clarifies that the increase in the illusion effect on far space reported by Yue et al. (2015) is not due to the visual stimuli being presented in far space but rather to the dominant sensory stimuli.

Furthermore, different neuroanatomical systems seem to be involved in the processing of near and far space. The perception/action model of the ventral/dorsal stream proposes that the ventral stream transforms visual or auditory information to a perception, which plays a crucial role in object recognition; the dorsal stream, on the other hand, processes visual or auditory information to prepare the observer to execute actions (Goodale et al., 1991; Goodale & Milner, 1992; Rauschecker, 2012; Shmuelof & Zohary, 2005; Vuilleumier et al., 1998; Wise, 2003). This phenomenon might explain the differences in the dorsal and ventral streams, supporting near and far space processing. Since a direct action cannot be performed in far space, the ventral stream is activated; in contrast, the dorsal stream subserves near-space processing because objects can be directly manipulated in near space (Menemeyer et al., 1992; Previc, 1990, 1998; Romanski et al., 2000; Weiss et al., 2003). Thus, the evidence suggests sensory information processing patterns of ventral/perception/far space and dorsal/movement/near space. Visual and auditory information processing occur along both dorsal/where and ventral/what pathways (Belin & Zatorre, 2000; Mishkin et al., 1983). In addition, studies have shown that in 2D experimental environments, the dorsal visual stream exhibited significantly increased neural activity when vision dominated audition, which indicates that the ventral/dorsal pathway was related to multisensory competition (Huang et al., 2015). Therefore, we suspect that in regard to the auditory dominant SiFI effect, when auditory stimuli are presented in far space, the neural activity of the auditory ventral stream that is mainly responsible for processing perception/far space would increase. As a result, the competitive advantages of the auditory stimuli were strengthened in far space, and the magnitude of the SiFI became larger. The present results of the fission illusion support this hypothesis.

For older adults, the fission illusion was not affected by the "depth" factor. One possible explanation was that older adults had increased multisensory integration, which helped counteract the age-related decline in sensory processing according to the compensation mechanism (Hairston et al., 2003; Reuter-Lorenz & Cappell, 2008). Multisensory illusion of older adults was so strong that the effect of spatial location on SiFI was ignored. Another possibility was that the ability of age-related spatial perception declines (Andersen & Atchley, 1995; He et al., 1998; Owsley, 2011). In particular, older adults' decreased ability to localize visual and auditory stimuli in the environment might result in the lack of a significant difference in SiFI between the *Vfar\_Anear* and the *Vnear\_Afar* conditions for older adults. Atchley and Kramer (2000) conducted an experiment that examined the extent to



which age influenced 3D spatial attention. Depth had no effect on older adults, potentially because older adults fail to tightly focus their attention along the depth axis. A subsequent study added to the evidence that the breadth of spatial attention of older drivers is constant across various depths (Pierce et al., 2011). Regarding sound location capability, both peripheral and central age-related auditory deficits, such as high-frequency hearing loss or a decline in temporal processing, might affect the utilization of the auditory spatial cues involved in sound localization. For instance, high-frequency peripheral hearing loss in older listeners has been associated with degradation in vertical and front-back discrimination in auditory space (Abel et al., 2000; Noble et al., 1994; Rakerd et al., 1998). Therefore, we predicted that auditory and visual localization would generally deteriorate in older adults due to age-related central spatial processing deficits. In conclusion, older adults were not susceptible to “depth”; therefore, we suspected that one of the reasons why older adults were susceptible to SiFI but were not affected by the space location is an age-related reduction in the spatial localization capacity.

More importantly, spatial location affected fission and fusion illusions in older and younger adults differently. For the fission illusion, accuracy and signal detection theory results indicated that younger adults were substantially affected by the spatial location, as a significant difference was observed between the  $V_{far\_A_{near}}$  and  $V_{near\_A_{far}}$  conditions. However, for the fusion illusion, neither participant group was affected by the spatial location, which may be due to the instability of the fusion illusion (Andersen et al., 2004; Shams et al., 2000). Thus, the effect of spatial dominance on the SiFI varied with age, and thus the fission and fusion illusion effects differed between the groups. However, for the RT results, the fusion illusion was also affected by spatial location; namely, when the auditory stimuli were presented in far space, the younger adults’ RT was longer. This result indicated that the auditory competitive advantages will be strengthened and that the illusion will increase, leading to a longer RT. Therefore, the RT results also indicated that the fusion illusion was also affected by spatial location, suggesting that although accuracy is the main outcome of the SiFI (Andersen et al., 2004; DeLoss et al., 2013; Keil, 2020), future studies should focus on RT results (DeLoss & Andersen, 2015). Nevertheless, the present study reported that an enhanced auditory dominance effect was observed in far space in younger adults, which is critical for the effective perception of unsafe objects or events in far space for better adaptation to the world and an increased chance of survival (Graziano & Cooke, 2006).

## Conclusions

In summary, by positioning auditory and visual stimuli in near and far space separately, the present study

showed that the fission illusion of younger adults was modulated by the depth of auditory stimuli. Although a significant fission illusion effect was observed regardless of whether the auditory stimuli appeared in near or far space, the fission illusion was significantly enhanced in far space. The effect of the spatial location was only observed on younger adults, not on older adults, which indicated a decline in attention and spatial localization capacity due to aging.

**Authors’ contributions** AW and MZ were responsible for the experimental design, interpretation of the data, and revision of the manuscript. AW, CY, and CL performed the experiments. YS and HZ were responsible for analyzing the data. YS, HZ, and AW drafted the manuscript.

**Funding** This research was supported by grants from the Jiangsu Provincial Key Constructive Laboratory for Big Data of Psychology and Cognitive Science (72592162005G), the 14th Five-year Plan of Jiangsu Province Education Science (B/2021/01/87), the Humanities and Social Sciences Research Project of Soochow University (22XM0017), the Interdisciplinary Research Team of Humanities and Social Sciences of Soochow University (2022), the Japan Society for the Promotion of Science KAKENHI (20K04381) and the National Natural Science Foundation of China (31700939 and 31871092).

## Declarations

**Conflicting interests** The author(s) have no potential conflicts of interest to declare with respect to the research, authorship, and/or publication of this article.

**Informed consent** Informed consent was obtained from all individual participants included in the study.

**Ethical approval** All experiments were performed in accordance with the principles of the Declaration of Helsinki. This study was approved by the Academic Committee of the Department of Psychology, Soochow University, China.

## References

- Abadi, R. V., & Murphy, J. S. (2014). Phenomenology of the sound-induced flash illusion. *Experimental Brain Research*, 232(7), 2207–2220.
- Abel, S. M., Giguère, C., Consoli, A., & Papsin, B. C. (2000). The effect of aging on horizontal plane sound localization. *The Journal of the Acoustical Society of America*, 108(2), 743–752.
- Andersen, G. J., & Atchley, P. (1995). Age-Related differences in the detection of three-dimensional surfaces from optic flow. *Psychology and Aging*, 10(4), 650–658.
- Andersen, T. S., Tiippana, K., & Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Cognitive Brain Research*, 21(3), 301–308.
- Atchley, P., & Kramer, A. F. (2000). Age related changes in the control of attention in depth. *Psychology and Aging*, 15, 78–87.
- Belin, P., & Zatorre, R. J. (2000). ‘What’, ‘where’ and ‘how’ in auditory cortex. *Nature Neuroscience*, 3(10), 965–966.

- Bresciani, J. P., Dammeier, F., & Ernst, M. O. (2008). Tri-modal integration of visual, tactile and auditory signals for the perception of sequences of events. *Brain Research Bulletin*, *75*(6), 753–760.
- Cabeza, R., Daselaar, S. M., Dolcos, F., Prince, S. E., Budde, M., & Nyberg, L. (2004). Task-independent and task-specific age effects on brain activity during working memory visual attention and episodic retrieval. *Cerebral Cortex*, *14*(4), 364–375.
- Chan, J. S., Connolly, S. K., & Setti, A. (2018). The Number of Stimulus-Onset Asynchronies Affects the Perception of the Sound-Induced Flash Illusion in Young and Older Adults. *Multisensory Research*, *31*(3–4), 175–190.
- Chen, Q., & Zhou, X. (2013). Vision dominates at the prereponse level and audition dominates at the response level in cross-modal interaction: behavioral and neural evidence. *The Journal of Neuroscience*, *33*(17), 7109–7121.
- Chen, Y. C., Maurer, D., Lewis, T. L., Spence, C., & Shore, D. I. (2017). Central-peripheral differences in audiovisual and visuotactile event perception. *Attention, Perception, & Psychophysics*, *79*(8), 2552–2563.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.
- DeLoss, D. J., & Andersen, G. J. (2015). Aging, spatial disparity, and the sound-induced flash illusion. *PLoS One*, *10*(11), e0143773.
- DeLoss, J. D., Russell, S. P., & George, J. A. (2013). Multisensory integration, aging, and the sound-induced flash illusion. *Psychology and Aging*, *28*(3), 802–812.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*(1), 193–222.
- Desimone, R., Wessinger, M., Thomas, L., & Schneider, W. (1990). Attentional control of visual perception: Cortical and subcortical mechanisms. *Cold Spring Harbor Symposia on Quantitative Biology*, *55*, 963–971.
- Diederich, A., Colonius, H., & Schomburg, A. (2008). Assessing age-related multisensory enhancement with the time-window-of-integration model. *Neuropsychologia*, *46*(10), 2556–2562.
- Dobrev, M. S., O'Neill, W. E., & Paige, G. D. (2011). Influence of aging on human sound localization. *Journal of Neurophysiology*, *105*(5), 2471–2486.
- Downing, C., & Pinker, S. (1985). The spatial structure of visual attention. In M. I. Posner & O. S. M. Marin (Eds.), *Attention and performance* (Vol. 11, pp. 171–187). Erlbaum.
- Driver, J., & Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on 'sensory-specific' brain regions, neural responses, and judgments. *Neuron*, *57*(1), 11–23.
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G\*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G\*Power 3.1: tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160.
- Gawryszewski, L. G., Riggio, L., Rizzolatti, G., & Umiltà, C. (1987). Movements of attention in the three spatial dimensions and the meaning of "neutral" cues. *Neuropsychologia*, *25*(1), 19–29.
- Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15*(1), 20–25.
- Goodale, M. A., Milner, A. D., Jakobson, L. S., & Carey, D. P. (1991). A neurological dissociation between perceiving objects and grasping them. *Nature*, *349*(6305), 154–156.
- Grady, C. L. (2010). Cognitive neuroscience of aging. *Annals of the New York Academy of Sciences*, *1124*(1), 127–144.
- Graziano, M. S. A., & Cooke, D. F. (2006). Parieto-frontal interactions, personal space, and defensive behavior. *Neuropsychologia*, *44*(13), 2621–2635.
- Hairston, W. D., Laurienti, P. J., Mishra, G., Burdette, J. H., & Wallace, M. T. (2003). Multisensory enhancement of localization under conditions of induced myopia. *Experimental Brain Research*, *152*(3), 404–408.
- He, N., Dubno, J. R., & Mills, J. H. (1998). Frequency and intensity discrimination measured in a maximum-likelihood procedure from young and aged normal hearing subjects. *The Journal of the Acoustical Society of America*, *103*(1), 553–565.
- Hernández, B., Setti, A., Kenny, R. A., & Newell, F. N. (2019). Individual differences in ageing, cognitive status, and sex on susceptibility to the sound-induced flash illusion: a large-scale study. *Psychology and Aging*, *34*(7), 978–990.
- Hirst, R. J., Cragg, L., & Allen, H. A. (2018). Vision dominates audition in adults but not children: A meta-analysis of the Colavita effect. *Neuroscience and Biobehavioral Reviews*, *94*, 286–301.
- Hirst, R. J., McGovern, D. P., Setti, A., Shams, L., & Newell, F. N. (2020). What you see is what you hear: Twenty years of research using the Sound-Induced Flash Illusion. *Neuroscience & Biobehavioral Reviews*, *118*, 759–774.
- Huang, S., Li, Y., Zhang, W., Zhang, B., Liu, X., Mo, L., & Chen, Q. (2015). Multisensory competition is modulated by sensory pathway interactions with fronto-sensorimotor and default-mode network regions. *The Journal of Neuroscience*, *35*(24), 9064–9077.
- Innes-Brown, H., & Crewther, D. (2009). The impact of spatial incongruence on an auditory-visual illusion. *PLoS One*, *4*(7), e6450.
- Kaiser, M., Senkowski, D., Busch, N. A., Balz, J., & Keil, J. (2019). Single trial prestimulus oscillations predict perception of the sound-induced flash illusion. *Scientific Reports*, *9*(5983), 1–8.
- Keil, J. (2020). Double flash illusions: current findings and future directions. *Frontiers in Neuroscience*, *14*, 298–304.
- Keil, J., & Senkowski, D. (2018). Neural oscillations orchestrate multisensory processing. *The Neuroscientist*, *24*(6), 609–626.
- Klemen, J., Büchel, C., & Rose, M. (2009). Perceptual load interacts with stimulus processing across sensory modalities. *European Journal of Neuroscience*, *29*(12), 2426–2434.
- Koelewijn, T., Bronkhorst, A., & Theeuwes, J. (2010). Attention and the multiple stages of multisensory integration: A review of audiovisual studies. *Acta Psychologica*, *134*(3), 372–384.
- Kumpik, D. P., Roberts, H. E., King, A. J., & Bizley, K. B. (2014). Visual sensitivity is a stronger determinant of illusory processes than auditory cue parameters in the sound-induced flash illusion. *Journal of Vision*, *14*(7), 60–65.
- Lange, J., Keil, J., Schnitzler, A., Van Dijk, H., & Weisz, N. (2014). The role of alpha oscillations for illusory perception. *Behavioural Brain Research*, *271*, 294–301.
- Laurienti, P. J., Burdette, J. H., Maldjian, J. A., & Wallace, M. T. (2006). Enhanced multisensory integration in older adults. *Neurobiology of Aging*, *27*(8), 1155–1163.
- Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. Psychology Press.
- Mahoney, J. R., Li, P. C., Oh-Park, M., Verghese, J., & Holtzer, R. (2011). Multisensory integration across the senses in young and older adults. *Brain Research*, *1426*(17), 43–53.
- Maringelli, F., McCarthy, J., Steed, A., Slater, M., & Umiltà, C. (2001). Shifting visuo-spatial attention in a virtual three-dimensional space. *Cognitive Brain Research*, *10*(3), 317–322.
- McCormick, D., & Mamassian, P. (2008). What does the illusory-flash look like? *Vision Research*, *48*(1), 63–69.
- McGovern, D. P., Eugenie, R., John, S., Martin, M. G. T., & Newell, F. N. (2014). The sound-induced flash illusion reveals dissociable age-related effects in multisensory integration. *Frontiers in Aging Neuroscience*, *6*(250), 1–9.
- Mennemeier, M., Wertman, E., & Heilman, K. M. (1992). Neglect of near peripersonal space. Evidence for multidirectional attentional systems in humans. *Brain*, *115*(1), 37–50.
- Michail, G., & Keil, J. (2018). High cognitive load enhances the susceptibility to non-speech audiovisual illusions. *Scientific Reports*, *8*(1), 11530.

- Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences*, 6, 414–417.
- Mishra, J., Martinez, A., Sejnowski, T. J., & Hillyard, S. A. (2007). Early cross-modal interactions in auditory and visual cortex underlie a sound-induced visual illusion. *The Journal of Neuroscience*, 27(15), 4120–4131.
- Mishra, J., Martinez, A., & Hillyard, S. A. (2008). Cortical processes underlying sound-induced flash fusion. *Brain Research*, 1242(4), 102–115.
- Mishra, J., Martinez, A., & Hillyard, S. A. (2010). Effect of attention on early cortical processes associated with the sound-induced extra flash illusion. *Journal of Cognitive Neuroscience*, 22(8), 1714–1729.
- Noble, W., Byrne, D., & Lepage, B. (1994). Effects on sound localization of configuration and type of hearing impairment. *The Journal of the Acoustical Society of America*, 95(2), 992–1005.
- Noel, J. P., Modi, K., Wallace, M. T., & Nathan, V. D. S. (2018). Audiovisual integration in depth: multisensory binding and gain as a function of distance. *Experimental Brain Research*, 236(7), 1939–1951.
- Norman, J. F., Dawson, T. E., & Butler, A. K. (2000). The effects of age upon the perception of depth and 3-D shape from differential motion and binocular disparity. *Perception*, 29(11), 1335–1159.
- Owsley, C. (2011). Aging and vision. *Vision Research*, 51(13), 1610–1622.
- Peiffer, A. M., Mozolic, J. L., Hugenschmidt, C. E., & Laurienti, P. J. (2007). Age-related multisensory enhancement in a simple audiovisual detection task. *Neuroreport*, 18(10), 1077–1081.
- Pierce, R., Bian, Z., & Andersen, G. (2011). Age-Related Limits of 3D Spatial Attention in Dual-Task Driving. *Driving Assessment Conference*, 6(2011), 314–320.
- Previc, F. H. (1990). Functional specialization in the lower and upper visual fields in humans: its ecological origins and neurophysiological implications. *The Behavioral and Brain Sciences*, 13(3), 519–542.
- Previc, F. H. (1998). The neuropsychology of 3-D space. *Psychological Bulletin*, 124(2), 123–164.
- Rakerd, B., Vander Velde, T. J., & Hartmann, W. M. (1998). Sound localization in the median sagittal plane by listeners with presbycusis. *Journal of the American Academy of Audiology*, 9(6), 466–479.
- Rauschecker, J. P. (2012). Ventral and dorsal streams in the evolution of speech and language. *Frontiers in Evolutionary Neuroscience*, 4(7), 1–4.
- Repp, B. H. (2000). Compensation for subliminal timing perturbations in perceptual-motor synchronization. *Psychological Research Psychologische Forschung*, 63(2), 106–128.
- Repp, B. H. (2002). Perception of timing is more context sensitive than sensorimotor synchronization. *Perception & Psychophysics*, 64(5), 703–716.
- Reuter-Lorenz, P., & Cappell, K. A. (2008). Neurocognitive aging and the compensation hypothesis. *Current Directions in Psychological Science*, 17(3), 177–182.
- Romanski, L. M., Tian, B., Fritz, J. B., Mishkin, M., & Rauschecker, J. P. (2000). Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nature Neuroscience*, 2(12), 1131–1136.
- Rosenthal, O., Shimojo, S., & Shams, L. (2009). Sound-induced flash illusion is resistant to feedback training. *Brain Topography*, 21(3–4), 185–192.
- Salat, D. H., Lee, S. Y., Van der Kouwe, A. J., & Greve, D. N. (2009). Age-associated alterations in cortical gray and white matter signal intensity and gray to white matter contrast. *NeuroImage*, 48(1), 21–28.
- Setti, A., & Chan, J. S. (2011). Familiarity of objects affects susceptibility to the sound-induced flash illusion. *Neuroscience Letters*, 492(1), 19–22.
- Setti, A., Burke, K. E., Kenny, R. A., & Newell, F. N. (2011). Is inefficient multisensory processing associated with falls in old people? *Experimental Brain Research*, 209(3), 375–384.
- Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature*, 408(12), 2670–2671.
- Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research*, 14(1), 147–152.
- Shams, L., Ma, W. J., & Beierholm, U. (2005). Sound-induced flash illusion as an optimal percept. *Neuroreport*, 16(17), 1923–1927.
- Shmuelof, L., & Zohary, E. (2005). Dissociation between ventral and dorsal fmri activation during object and action recognition. *Neuron*, 47(3), 457–470.
- Sinnett, S., Spence, C., & Soto-Faraco, S. (2007). Visual dominance and attention: the colavita effect revisited. *Perception & Psychophysics*, 69(5), 673–686.
- Spence, C. (2011). Crossmodal correspondences: a tutorial review. *Attention, Perception, & Psychophysics*, 73(4), 971–995.
- Stein, B. E., & Meredith, M. A. (1993). Multisensory integration. (book reviews: the merging of the senses). *Science*, 261, 928–929.
- Sun, Y., Liu, X., Li, B., Sava-Segal, C., & Zhang, M. (2020). Effects of repetition suppression on sound induced flash illusion with aging. *Frontiers in Psychology*, 11, 216.
- Talsma, D., Senkowski, D., Sotofaraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences*, 14(9), 400–410.
- Van der Stoep, N., Serino, A., Farnè, A., Serino, C., Di Luca, M., & Spenc, C. (2016a). Depth: the forgotten dimension in multisensory research. *Multisensory Research*, 29, 493–524.
- Van der Stoep, N., Van der Stigchel, S., Nijboer, T. C. W., & Van der Smagt, M. J. (2016b). Audiovisual integration in near and far space: effects of changes in distance and stimulus effectiveness. *Experimental Brain Research*, 234(5), 1175–1188.
- Vanes, L. D., White, T. P., Wigton, R. L., Joyce, D., Collier, T., & Shergill, S. S. (2016). Reduced susceptibility to the sound-induced flash fusion illusion in schizophrenia. *Psychiatry Research*, 245, 58–65.
- Vuilleumier, P., Valenza, N., Mayer, E., Reverdin, A., & Landis, T. (1998). Near and far visual space in unilateral neglect. *Annals of Neurology*, 43(3), 406–410.
- Wang, A., Sang, H., He, J., Sava-Segal, C., Tang, X., & Zhang, M. (2019). Effects of cognitive expectation on sound-induced flash illusion. *Perception*, 48(12), 12140–12134.
- Weiss, P. H., Marshall, J. C., Zilles, K., & Fink, G. R. (2003). Are action and perception in near and far space additive or interactive factors? *Neuroimage*, 18(4), 837–846.
- Wetzels, R., & Wagenmakers, E. J. (2012). A default Bayesian hypothesis test for correlations and partial correlations. *Psychonomic Bulletin and Review*, 19, 1057–1064.
- Wise, R. J. S. (2003). Language systems in normal and aphasic human subjects: functional imaging studies and inferences from animal studies. *British Medical Bulletin*, 65, 95–119.
- Wu, J., Yang, W., Gao, Y., & Kimura, T. (2012). Age-related multisensory integration elicited by peripherally presented audiovisual stimuli. *Neuroreport*, 23(10), 616–620.
- Yue, Z., Jiang, Y., Li, Y., Wang, P., & Chen, Q. (2015). Enhanced visual dominance in far space. *Experimental Brain Research*, 233(10), 2833–2843.

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