

Central–peripheral differences in audiovisual and visuotactile event perception

**Yi-Chuan Chen, Daphne Maurer,
Terri L. Lewis, Charles Spence & David
I. Shore**

**Attention, Perception, &
Psychophysics**

ISSN 1943-3921
Volume 79
Number 8

Atten Percept Psychophys (2017)
79:2552–2563
DOI 10.3758/s13414-017-1396-4

Attention, Perception, & Psychophysics

VOLUME 79, NUMBER 8 ■ NOVEMBER 2017

AP&P

EDITOR

Michael D. Dodd, *University of Nebraska-Lincoln*

ASSOCIATE EDITORS

Wendy Jo Adams, *University of Southampton*

Meghan Clayards, *McGill University*

Frank H. Durgin, *Swarthmore College*

Karla Evans, *The University of York*

Michael C. Hout, *New Mexico State University*

Liqiang Huang, *The Chinese University of Hong Kong*

Todd A. Kahan, *Bates College*

Ashleigh M. Maxcey, *The Ohio State University*

Lisa Sanders, *University of Massachusetts Amherst*

Stefan Van der Stigchel, *Utrecht University*

A PSYCHONOMIC SOCIETY PUBLICATION

www.psychonomic.org

ISSN 1943-3921



Your article is protected by copyright and all rights are held exclusively by The Psychonomic Society, Inc.. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".

Central–peripheral differences in audiovisual and visuotactile event perception

Yi-Chuan Chen^{1,2} · Daphne Maurer¹ · Terri L. Lewis¹ · Charles Spence² · David I. Shore¹

Published online: 16 August 2017
© The Psychonomic Society, Inc. 2017

Abstract We examined audiovisual and visuotactile integration in the central and peripheral visual field using visual fission and fusion illusions induced by sounds or taps. The fission illusion occurs when a single flash is perceived as two flashes if paired with two beeps or taps; the fusion illusion, by contrast, occurs when two flashes are perceived as a single flash if the flashes are paired with a single beep or tap. Beeps and taps induced similar patterns of illusions: the fission illusion was larger in the periphery than in the center, whereas the fusion illusion was larger in the center than in the periphery. An analysis based on signal detection theory revealed that both a decline in discriminability and a shift in criterion were associated with the more pronounced fission induced by sounds in the periphery. In contrast, only a shift in criterion was associated with the larger fission induced by taps in the periphery, and the larger fusion induced by a sound or tap in the center. To accommodate these findings, two accounts are proposed: audiovisual signals are more likely to be integrated in peripheral than in central vision, and the interpretation of visual signals favors discontinuous percepts, especially in the periphery.

Keywords Vision · Audition · Touch · Eccentricity · Multisensory integration · Illusion

✉ David I. Shore
dshore@mcmaster.ca

¹ Department of Psychology, Neuroscience & Behaviour, McMaster University, 1280 Main St West, Hamilton, ON L8S 4K1, Canada

² Crossmodal Research Laboratory, Department of Experimental Psychology, University of Oxford, Oxford, UK

Introduction

In everyday life, perceptual events often provide information to multiple sensory modalities. For example, each time a ball bounces, it produces both a visual and an auditory signal at the moment of impact. In order to perceive the location of the ball and predict its future trajectory optimally, an observer typically integrates signals from both senses. The process of multisensory integration is nevertheless complicated: not only the nature of the sensory signals (such as the travel time of distal information) and the characteristics of each sensory system (such as the different neural transduction times and spatial reference frames) have to be taken into account, but also the observers' knowledge and previous experience (see Welch & Warren, 1980, for an early review). According to recent computational models of multisensory integration, information from the different senses is integrated in a statistically optimal fashion; that is, each sensory signal is weighted according to its reliability, or variance, during integration (see Alais & Burr, 2004; Andersen, Tiippana, & Sams, 2005; Ernst & Banks, 2002; Wozny, Beierholm, & Shams, 2008).

The present study examined a multisensory illusion induced by either auditory or tactile signals (Bresciani, Dammeier, & Ernst, 2006, 2008; Shams, Kamitani, & Shimojo, 2000, 2002; Violentyev, Shimojo, & Shams, 2005) in central and peripheral vision. We chose to examine the impact of visual eccentricity on multisensory integration because the spatial and temporal resolution in central and peripheral vision differ (e.g., Hartmann, Lachenmayr, & Brettel, 1979; Wright & Johnson, 1983). Furthermore, crossmodal neural connectivity between sensory-dominant areas responsible for the peripheral visual field is more extensive than that for the central visual field (at least for audiovisual integration; see Falchier, Clavagnier, Barone, & Kennedy, 2002), a pattern leading to the prediction of increasing multisensory

integration with increasing visual eccentricity (e.g., Charbonneau, Veronneau, Boudrias-Fournier, Lepore, & Collignon, 2013; Gleiss & Kayser, 2013; though see Fiebelkorn, Foxe, Butler, & Molholm, 2011).

Two types of event illusion induced by sound have been reported: in the *fission* illusion, a single flash is perceived as two flashes when accompanied by two beeps (Shams et al., 2000, 2002); in the *fusion* illusion, two flashes are perceived as a single flash when accompanied by a single beep (Andersen, Tiippana, & Sams, 2004). The fission illusion occurs more frequently in the periphery (5° eccentricity) than in the center; by contrast, the fusion illusion appears to be unaffected by eccentricity (Tremblay et al., 2007). These results suggest that the fission and fusion illusions plausibly have different underlying mechanisms rather than simply reflecting two facets of a single underlying phenomenon. Although both fission and fusion illusions have also been observed with a tactile inducer (Bresciani et al., 2006; Violentyev et al., 2005; Wozny et al., 2008), the effect of visual eccentricity has not yet been examined.

In the present study, we compared the fission and fusion illusions induced by beeps (Experiment 1) or taps (Experiment 2) in the center and periphery. One or two flashes were presented in the center or at 10° into the periphery and paired with zero, one, or two beeps (or taps). The accuracy of participants' reports concerning the number of flashes and the magnitude of the illusions were calculated. We further computed the participant's discriminability and criterion when responding to the flashes based on signal detection theory in order to better understand the mechanisms underlying the fission and fusion illusions.

Experiment 1: sound-induced flash illusion

Method

Participants

The final sample consisted of 20 McMaster University undergraduates (mean age = 20.3 years, six males) who took part in the study in exchange for course credit. All of the participants had normal or corrected-to-normal vision (20/20 vision in each eye on the Lighthouse eye chart) and normal hearing by self-report. All of them were naïve regarding the purpose of the study. Two additional participants were excluded: one did not finish the study, and the other did not reach 80% accuracy on the catch trials presented in the center during the main experiment (see Design), suggesting that the participant might report the number of beeps rather than flashes on some trials. All of the participants signed a written consent form after the procedures were explained. The study was cleared by the McMaster University Research Ethics Board

(2001028) and conformed to the Tri-Council Statement on Ethics (Canada).

Apparatus and stimuli

The stimuli were presented on a LED monitor (60-Hz refresh rate) with speakers at the left and right of the monitor. Both monitor and speakers were connected to a Mac Mini computer controlled by Matlab (MathWorks Inc., Natick, MA, USA) and the Psychtoolbox extension (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997). The background display on the monitor consisted of three grey 2° rings (inner diameter, with a thickness of 0.06°) on a black background, one ring in the center and the others placed with their centers at 10° to the left or right of center. Targets consisted of white flashes presented in one of the three rings (see Fig. 1a). The flash was a 2° disc with a duration of 17 ms. The auditory stimulus consisted of a 10 ms white noise burst (plus 1-ms ramps at onset and offset) with a loudness of 57.5 dB SPL (the room had a background noise level of 40 dB SPL). Participants sat with their head on a chinrest located 50 cm from the monitor and entered their responses on a standard Apple keyboard.

Design

Three within-participant factors were manipulated: Flash (1F or 2F), Congruency (unimodal, congruent, incongruent, or catch), and Location (center or periphery). On each trial, one flash (1F) or two flashes (2F) were presented, and the flash(es) could be presented in silence (0B), with one beep (1B), or with two beeps (2B). These combinations created three types of trials: unimodal (1F0B and 2F0B), congruent (1F1B and 2F2B), and incongruent (1F2B and 2F1B). The fission illusion would be observed if two flashes were reported in the 1F2B condition, and the fusion illusion would be observed if one flash was reported in the 2F1B condition. The stimulus onset asynchrony (SOA) between the paired flash and beep was 0 ms, and the SOA between two successive flashes or beeps was 83 ms (see Fig. 1b). In half of the trials, the flash(es) was/were presented in the central ring where the participants fixated, and for the other half, in the periphery (equally often in the left or right ring). Both flashes, when presented, always occurred in the same location.

Two types of catch trials were presented. The stimuli were the same as the incongruent trials (i.e., 1F2B and 2F1B), but the SOA between the closest beep and flash was set to 300 ms (auditory leading). With this timing, the incongruent audiovisual stimuli would be unlikely to induce the fission illusion (see Shams et al., 2002). We used the data from catch trials with one flash in the center paired with two beeps preceding by 300 ms to establish an exclusion criterion for participants who might not have been attentive or who might have incorrectly reported the number of beeps instead of the number of

flashes. We chose this condition because the single flash occurred in the center where the participants were instructed to fixate, and hence the flash was unlikely to be missed or misperceived as two flashes.¹

There were 20 trials for each Flash (2) × Congruency (3) × Location (2) condition, and 10 trials for each type of catch trial (1F2B or 2F1B) in each Location (2), giving rise to a total of 280 trials. These trials were divided into 10 blocks with balanced conditions. The participants were given a short break between blocks. It took participants roughly 30 min to complete the experiment.

Procedure

The participants were instructed to fixate the central ring. Each trial was initiated by pressing the space bar on the keyboard. After a randomly selected interval (ranging from 500 to 1500 ms in 200-ms increments), the first visual stimulus was presented either with or without an accompanying auditory stimulus. The participants' task was to report the number of flashes appearing on the screen (one or two) while ignoring the beeps. The participants keyed their answer into the computer by pressing the '1,' '2,' or '0' key if they saw one flash, two flashes, or missed the flashes, respectively.² They could take as long as needed to respond, and pressed the space bar to start the next trial.

Two practice sessions were conducted prior to the main experiment. The first practice session consisted of eight unimodal trials, half with one flash and the other half with two flashes. The aim was to ensure that the participants perceived the brief flashes accurately. The SOAs between the two flashes were 117 ms in the first round and 83 ms in the second round. At least 85% accuracy was required on each round in order to be included in the final analysis. The second practice session consisted of 16 trials (eight unimodal, four congruent, and four catch trials; half of each type with one flash and the other half with two flashes), which was designed to ensure that the participants reported the number of flashes rather than the number of beeps. Given that there should be no illusion on these trials, participants needed at least 85% accuracy to be included in the final analysis. All of the participants passed

¹ One participant in Experiment 1 and two in Experiment 2 did not reach 80% accuracy on these catch trials and their data were excluded from the analyses. Had we included them, the accuracy results would have been identical to those reported in the text.

² When the participants missed the flash(es) completely by pressing the '0' key, the trial was discarded and not included in the calculation of accuracy or in further analyses. In Experiments 1 (sound-induced flash illusion) and 2 (tap-induced flash illusion), there were 17 and 32 trials of such types of responses, respectively, among 5,600 trials (20 participants × 280 trials – i.e., 0.3% and 0.6%, respectively, of the total trials). Excluding this small proportion of missing flash(es) responses should not influence the results reported in the current study.

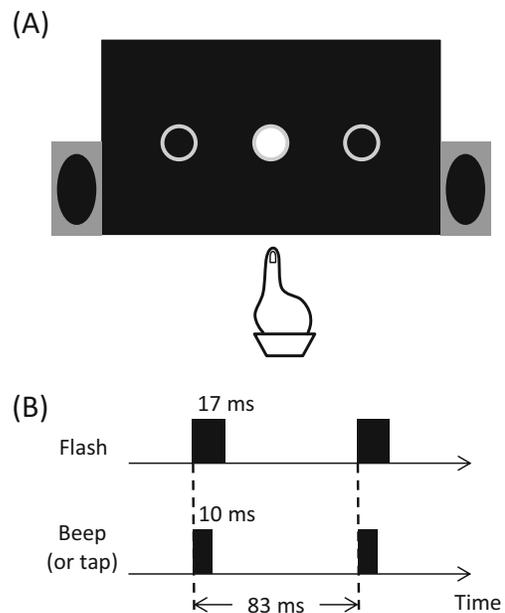


Fig. 1 Schematic display of the experiment setting. **(a)** Three rings (2° inner diameter) were presented on the monitor, one in the center and the other two 10° to the left and right. Visual targets (one or two flashes) were presented randomly in one of the rings. In Experiment 1, the auditory stimuli were presented from the loudspeakers placed on either side of the monitor. In Experiment 2, the tactile stimuli were presented on the participants' right index finger aligned with the central ring and body midline. **(b)** The duration and SOA of the visual and auditory stimuli (Experiment 1) or visual and tactile stimuli (Experiment 2) when two flashes and two beeps (or taps) were presented

these criteria. The 280 experimental trials followed these two practice runs.

Results

In both experiments, the data were submitted to repeated measures analyses of variance (ANOVA). A Greenhouse-Geisser correction was used when the sphericity assumption was violated. *Post-hoc* tests were pairwise t-tests with Bonferroni correction.

Accuracy

We calculated the percentage of correct responses for each participant in each condition. The accuracy data (Fig. 2a) were submitted to a three-way ANOVA with the factors of Flash (1F or 2F), Congruency (unimodal, congruent, incongruent, or catch), and Location (center or periphery). The main effect of Congruency was significant ($F(1.22,23.10) = 65.95$, $MSE = 0.12$, $p < .001$). Two two-way interactions were significant: Flash × Congruency ($F(1.95,36.99) = 15.96$, $MSE = 0.03$, $p < .001$) and Flash × Location ($F(1,19) = 9.81$, $MSE = 0.04$, $p < .01$). Critically, the three-way interaction was also significant ($F(3,57) = 7.43$, $MSE = 0.01$, $p < .001$).

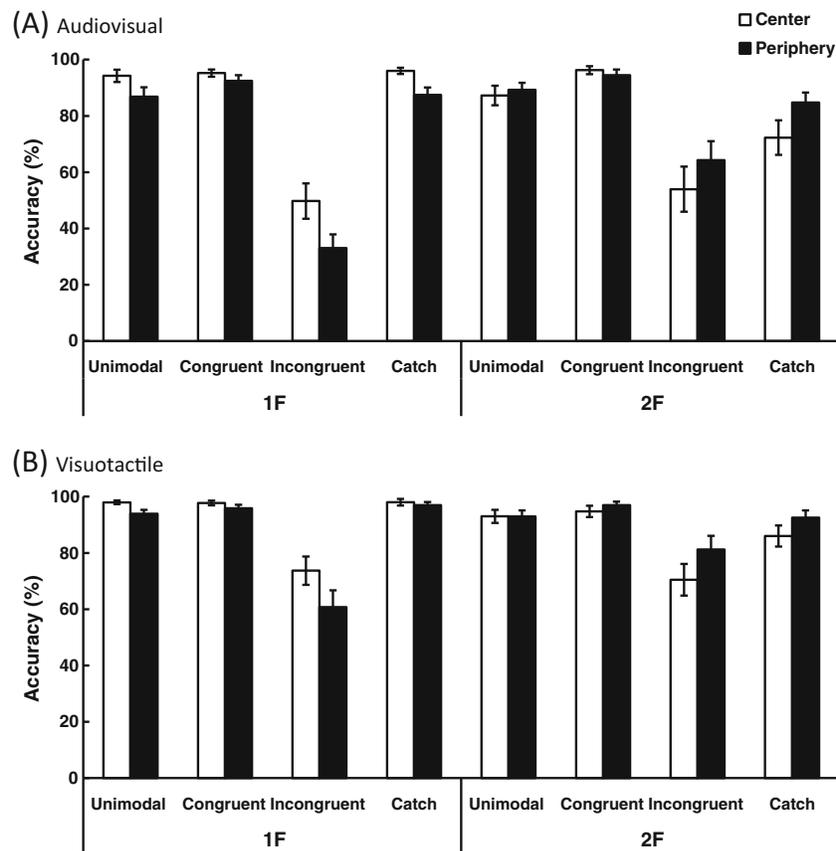


Fig. 2 The mean accuracy of reporting the number of flashes when paired with beeps in Experiment 1 (**Panel a**) or taps in Experiment 2 (**Panel b**). The error bars represent ± 1 standard error of the mean. *1F* one flash, *2F* two flashes

To understand the three-way interaction, the Flash factor was broken down into two two-way ANOVAs with the factors of Congruency and Location. In the 1F condition (left panel in Fig. 2a), both main effects were significant. For the Congruency factor ($F(1.32, 24.99) = 79.39$, $MSE = 0.07$, $p < .001$), pairwise t-tests demonstrated that the accuracy in the incongruent condition was significantly lower than in the other three conditions ($ps < .001$), with no significant difference among the other three conditions ($ps > .49$). These results suggest a significant fission illusion for the incongruent condition but not for the catch trials. For the Location factor ($F(1, 19) = 15.06$, $MSE = 0.02$, $p < .005$), accuracy was higher in the center than in the periphery. Critically, the interaction between Congruency and Location was significant ($F(1.84, 35.04) = 5.43$, $MSE = 0.01$, $p < .05$). Paired t-tests revealed that accuracy was *higher* in the center than in the periphery for the incongruent condition ($t(19) = 3.60$, $p < .01$) and for the catch trials ($t(19) = 3.49$, $p < .01$).

In the 2F condition (right panel in Fig. 2a), the main effect of Congruency was significant ($F(1.32, 24.99) = 26.63$, $MSE = 0.09$, $p < .001$). Pairwise t-tests of all comparisons demonstrated that accuracy was highest in the congruent condition and next highest in the unimodal condition, both of which were higher than accuracy on the catch trials and in the incongruent

condition, which had the lowest accuracy (all $ps < .05$). This result suggests a significant fusion illusion in the incongruent condition as well as in the catch trials when compared to the congruent condition. The interaction between Congruency and Location was significant ($F(3, 57) = 4.58$, $MSE = 0.01$, $p < .01$). Paired t-tests did not identify the source of this interaction, but inspection of the graph and a statistical trend indicate that accuracy tended to be *lower* in the center than in the periphery in the incongruent condition ($t(19) = 2.20$, $p = .16$).

Magnitude of the illusion

The magnitude of the fission illusion was calculated by subtracting accuracy in the 1F/incongruent from the 1F/congruent condition. Similarly, the magnitude of the fusion illusion was calculated by subtracting accuracy in the 2F/incongruent from the 2F/congruent condition. Each was calculated separately for the center and periphery. These values (see Fig. 3a) were submitted to a two-way ANOVA with the factors of Illusion (fission or fusion) and Location (center or periphery). The magnitude of the illusion was larger for the fission illusion (52.5%) than for the fusion illusion (36.3%; $F(1, 19) = 8.87$, $MSE = 0.06$, $p < .01$). The two-way interaction was significant ($F(1, 19) = 14.15$,

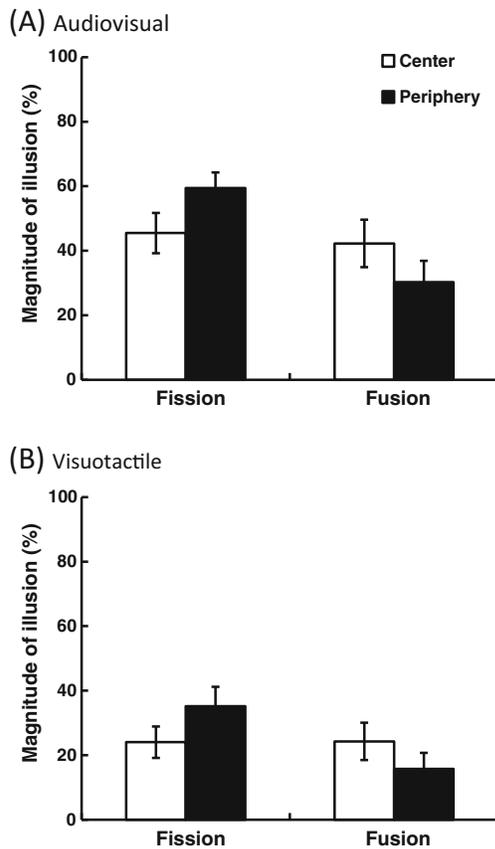


Fig. 3 The mean magnitude of the fission illusion (accuracy in the 1F/incongruent condition subtracted from the 1F/congruent condition) and the fusion illusion (accuracy in the 2F/incongruent subtracting from 2F/congruent condition) for the audiovisual stimuli in Experiment 1 (**Panel a**) and the visuotactile stimuli in Experiment 2 (**Panel b**). The error bars represent ± 1 standard error of the mean

$MSE = 0.02, p < .005$). The results of paired t-tests demonstrated that the fission illusion was larger in the periphery than in the center ($t(19) = 3.23, p < .01$); by contrast, the fusion illusion was larger in the center than in the periphery ($t(19) = 2.75, p < .05$).

Discriminability and criterion

In order to determine whether the different magnitudes of the fission and fusion illusions in the center versus periphery were attributable to a change in the discriminability of the flashes and/or criterion for reporting the number of flashes induced by the presentation of the beeps, the data were analyzed in terms of signal detection theory (see Fig. 4a; McCormick & Mamassian, 2008; Violentyev et al., 2005). The presentation of two flashes was defined as the target, whereas the presentation of one flash was the noise. Hence, the d' represents the discriminability of whether one or two flashes were presented, while the criterion (c in the present study) represents the tendency to report “one flash” or “two flashes” (see Table 2 in Witt, Taylor, Sugovic, & Wixted, 2015).

The accuracy in the two-flash conditions, therefore, constituted the *hit rate*, while the error rate in the one-flash condition (i.e., reporting two flashes when there was only one flash presented) constituted the *false alarm* (FA) rate (see Fig. 4b). The hit and FA rates were then transformed to z scores. The d' and c were calculated using the following equations (Macmillan & Creelman, 2005):

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

$$c = -0.5 * [z(\text{Hit rate}) + z(\text{FA rate})]$$

The d' and c were calculated separately when no sound (0B), or else one beep (1B) or two beeps (2B) were presented (Rosenthal, Shimojo, & Shams, 2009; Violentyev et al., 2005).

When computing d' and c in the 2B condition, the participants' accuracy in the 2F/congruent condition was used as the *hit rate*, while their error rate in the 1F/incongruent condition (i.e., two flashes were reported) where the fission illusion may occur was used as the *FA rate*. In the 1B condition, the participants' accuracy in the 2F/incongruent condition (i.e., two flashes were reported correctly) where the fusion illusion may occur was used as the *hit rate*, while their error rate in the 1F/congruent condition was used as the *FA rate*. Therefore, d' and c in the 2B and 1B conditions can be used to represent the participant's performance for the fission and fusion illusions, respectively. In the 0B condition, the participant's accuracy in the 2F/unimodal was used as the *hit rate*, and the error rate in the 1F/unimodal, the *FA rate*.

The d' values (see Fig. 4c) were submitted to a two-way ANOVA with the factors of Sound (0B, 1B, and 2B) and Location (center or periphery). The main effect of Sound was significant ($F(2,38) = 33.56, MSE = 0.59, p < .001$). Pairwise t-tests demonstrated that d' decreased significantly with an increasing number of beeps (0B > 1B > 2B; all $ps < .05$). The interaction between Sound and Location was also significant ($F(2,38) = 12.64, MSE = 0.17, p < .001$). A post-hoc t-test demonstrated that, in the 2B condition, d' was *lower* in the periphery than in the center ($t(19) = 5.10, p < .001$). This may explain the fact that the fission illusion was larger in the periphery than in the center (i.e., lower discriminability, larger illusion). However, there was no significant difference of Location in the 0B and 1B conditions ($ts < 1.9, ps > .24$).

The c values (see Fig. 4d) were submitted to a two-way ANOVA with the factors of Sound (0B, 1B, and 2B) and Location (center or periphery). The main effect of Sound was significant ($F(1.31,24.97) = 68.77, MSE = 0.59, p < .001$), and the pairwise t-tests demonstrated that c decreased in the order of 1B (positive values suggesting a tendency of reporting one flash), 0B (close to 0, thus suggesting no specific tendency), and 2B (negative values suggesting a

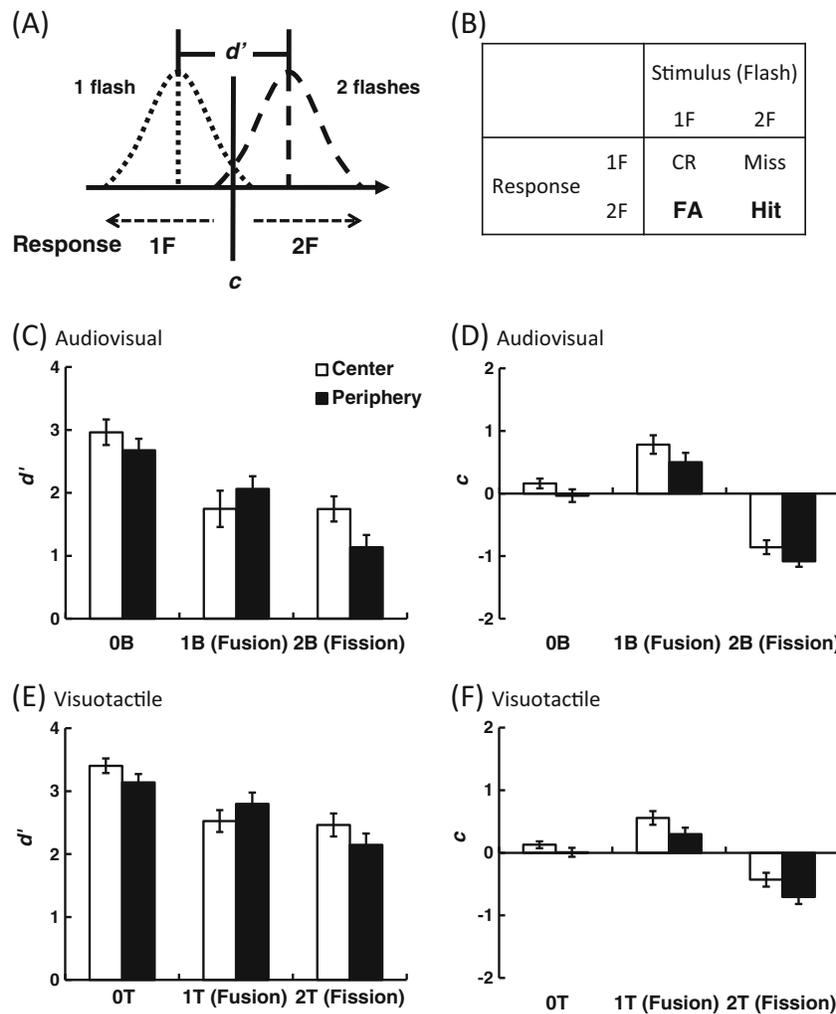


Fig. 4 (a) A schematic representation of the design in terms of signal detection theory. The distribution highlighted by the dashed line represents the target (two-flash condition) and the distribution highlighted by the dotted line represents the noise (one-flash condition). (b) Accordingly, the *hit* rate is the accuracy in the two-flash (2F) condition, and the counterpart, *miss* rate, is the error rate in the two-flash (2F) condition. The *false alarm* (FA) rate is the error rate in the one-flash (1F) condition, so the counterpart, *correct rejection* (CR) rate, is the accuracy in the one-flash (1F) condition. The hit and FA rates were

calculated separately for the no beep (0B), one beep (1B), and two beeps (2B) conditions. (c) and (d) represent the mean discriminability (d') and criterion (c) for the audiovisual stimuli in Experiment 1. (e) and (f) represent the mean d' and c for the visuotactile stimuli in Experiment 2 in which the comparable tap conditions replaced the beep conditions. In (d) and (f), the shift of c toward positive values indicates that the participants tended to report “one flash”, whereas the shift toward negative values indicates the tendency to report “two flashes”. The error bars represent ± 1 standard error of the mean. *B* beeps, *T* taps

tendency to report two flashes), and the three conditions were all significantly different from each other ($ps < .001$). Hence, the occurrence of the fission and fusion illusions can be attributed partly to a shift in criterion because of the numbers of beeps. The c value was also lower in the periphery (-0.21) than in the center (0.03 ; $F(1,19) = 7.22$, $MSE = 0.23$, $p < .05$), indicating that participants were more likely to report two flashes when the flashes were presented in the periphery than in the center. This result suggests that the larger fission illusion (i.e., reporting two flashes) in the periphery than in the center, as well as the larger fusion illusion (i.e., reporting one flash) in the center than in the periphery, can be attributed partly to the shift of criterion. The interaction between Sound and Location was not significant ($F < 1$, $p > .70$).

Discussion

In Experiment 1, we replicated both the visual fission and fusion illusions induced by sounds (Andersen et al., 2004; Shams et al., 2000, 2002; Wozny et al., 2008). A significant fission illusion was observed in the 1F/incongruent condition, while a significant fusion illusion was observed in both the 2F/incongruent and 2F/catch conditions. In the analysis based on signal detection theory, the presentation of sound reduced participants’ discriminability (d') to the flashes, as well as shifted their criterion (c) toward reporting the number of beeps.

Comparing the illusions in the center and periphery further clarify the possible mechanisms of the illusions. For the fission illusion, accuracy in the 1F/incongruent condition

was lower in the periphery than in the center, leading to a larger fission illusion in the periphery than in the center. This finding replicates the results reported by Shams et al. (2000) and Tremblay et al. (2007). We attribute this pattern of fission illusion to a lower discriminability (d') in the periphery than in the center, as well as to a lower criterion (c , i.e., the participants were more likely to report two flashes) in the periphery than in the center. On the other hand, the pattern was reversed for the fusion illusion: accuracy in the 2F/incongruent condition was lower in the center than in the periphery, leading to a larger fusion illusion in the center than in the periphery. The only other study to test the effect of eccentricity on the fusion illusion found no difference between the center and the periphery (Tremblay et al., 2007); however, those researchers presented the visual stimuli at an eccentricity of 5° rather than 10° as tested in the present study. The drop-off of the fusion illusion at 10° that we observed was attributable only to a lower criterion (c) in the periphery than in the center and not to a change in discriminability (d'). The meaning of the criterion shift will be discussed further for both audiovisual and visuotactile illusions in the [General discussion](#).

A number of differences between the fission and fusion illusions emerged in the present study. First, a reversed pattern in the center and periphery was observed: the fission illusion was larger in the periphery whereas the fusion illusion was larger in the center. Second, the larger fission illusion in the periphery than center was attributable to a change of both discriminability and criterion, whereas the larger fusion illusion in the center than periphery was attributable only to a change of criterion. Third, in catch trials where the closest beep led the flash by 300 ms, a temporal interval that is beyond the window of audiovisual simultaneity (Chen, Shore, Lewis, & Maurer, 2016; Zampini, Guest, Shore, & Spence, 2005), the fission illusion was eliminated (see also Shams et al., 2002); however, the fusion illusion remained.

Taken together, these differences suggest that the fission and fusion illusions have different underlying mechanisms. Specifically, the narrower temporal window for fission than fusion suggests that fission is plausibly the result of audiovisual integration. In addition, only the fission illusion, rather than the fusion illusion, is associated with a change of discriminability to visual flashes when comparing center and periphery. In summary, the fission illusion is a phenomenon with a solid perceptual basis (see also Berger, Martelli, & Pelli, 2003; McCormick & Mamassian, 2008; Shams et al., 2002), whereas the fusion illusion is likely associated with the stage of post-perceptual processing.

It is worth noting that in the condition where no sound was presented (unimodal condition), the participants' accuracy and discriminability (d') to the flashes were similar in the center and periphery. However, their criterion (c) was lower in the periphery than in the center; that is, our participants tended to

report two flashes rather than one in the periphery even when no sound was presented.

Experiment 2: tap-induced flash illusion

Method

The final sample consisted of a new group of 20 volunteers similar to those described in Experiment 1 (mean age = 18.4 years, nine males) except that all were right-handed as confirmed by a handedness questionnaire. An additional three participants were excluded: one had only 80% accuracy in the tap practice session (see below), while the other two did not reach 80% accuracy in the catch trials where one flash paired with two taps presented in the center during the main experiment.

The visual stimuli were the same as in Experiment 1. The auditory beeps were replaced by tactile taps (T). Participants were asked to place the index finger of their right hand on top of the tapping device where a small metallic pin could move upward and then downward, giving rise to a brief tactile tap. The tapping device was placed in front of the monitor and 20° below the central ring so as to align with participant's body midline. A single tap lasted 10 ms, and in the double-tap condition, the second tap was presented with an SOA of 83 ms. White noise was presented continuously through closed-ear headphones in order to mask the clicking noise produced by the tapping device. Participants were instructed to report verbally the number of flashes that they perceived while ignoring the number of taps. An experimenter, unaware of the actual number of flashes, keyed participants' answers into the computer.

Before the unimodal and bimodal practice sessions, participants completed an additional tap practice to verify that they felt the number of taps accurately. The participants had to reach 90% accuracy on the five 1T and five 2T practice trials in order to be included in the final sample.

Results

Accuracy

We calculated the percentage of correct responses for each participant in each condition. The accuracy data (see Fig. 2b) were submitted to a three-way ANOVA with the factors of Flash (1F or 2F), Congruency (unimodal, congruent, incongruent, or catch), and Location (center or periphery). There was a significant main effect of Congruency ($F(1.25, 23.82) = 40.17$, $MSE = 0.07$, $p < .001$). Two two-way interactions were significant: Flash \times Congruency ($F(1.60, 30.30) = 5.16$, $MSE = 0.04$, $p < .05$) and Flash \times Location ($F(1, 19) = 12.06$, $MSE = 0.02$, $p < .005$).

Critically, the three-way interaction was also significant ($F(1.73,32.78) = 6.78$, $MSE = 0.01$, $p < .01$).

To further investigate the source of the three-way interaction, the results for each flash condition were analyzed separately. In the 1F condition (see left panel in Fig. 2b), both main effects were significant: for the Congruency factor ($F(1.10,20.97) = 36.27$, $MSE = 0.07$, $p < .001$), the pairwise t-tests demonstrated that accuracy in the incongruent condition was significantly lower than in the other three conditions ($ps < .001$), and the latter three conditions were not significantly different from one another (all $ps > .77$). These results suggest a significant fission illusion in the incongruent condition. For the Location factor ($F(1,19) = 9.73$, $MSE = 0.01$, $p < .01$), accuracy was higher in the center than in the periphery. Critically, the two-way interaction was significant ($F(1.45,27.56) = 6.84$, $MSE = 0.01$, $p < .01$). Paired t-tests suggest that accuracy was *higher* in the center than in the periphery only in the incongruent condition ($t(19) = 3.14$, $p < .05$).

In the 2F condition (see right panel in Fig. 2b), the main effect of Congruency was significant ($F(1.60,30.40) = 14.09$, $MSE = 0.04$, $p < .001$). Accuracy was significantly lower in the incongruent condition than in the other three conditions ($ps < .05$), and there was no difference between the latter three conditions (all $ps > .23$), suggesting a significant fusion illusion in the incongruent condition. The main effect of Location was also significant ($F(1,19) = 7.87$, $MSE = 0.01$, $p < .05$), with accuracy being lower in the center than in the periphery. The two-way interaction was significant ($F(2.04,38.83) = 4.08$, $MSE = 0.01$, $p < .05$). Paired t-tests demonstrated that the *lower* accuracy in the center than in the periphery was significant only in the incongruent condition ($t(19) = 2.88$, $p < .05$).

Magnitude of illusion

The magnitudes of the fission and fusion illusions (see Fig. 3b) were calculated and submitted to a two-way ANOVA with the factors of Illusion (fission or fusion) and Location (center or periphery). Neither main effect was significant ($Fs < 2.59$, $ps > .12$); however, their interaction was ($F(1,19) = 7.99$, $MSE = 0.02$, $p < .05$). The fission illusion was larger in the periphery than in the center ($t(19) = 2.61$, $p < .05$); in contrast, the fusion illusion was marginally larger in the center than in the periphery ($t(19) = 2.21$, $p = .08$).

Discriminability and criterion

The d' values (see Fig. 4e) were submitted to a two-way ANOVA with the factors of Tap (0T, 1T, and 2T) and Location (center or periphery). The analysis revealed a significant main effect of Tap ($F(1.39,26.44) = 24.06$, $MSE = 0.57$, $p < .001$). Pairwise t-tests demonstrated that d' was

significantly lower in the 1T and 2T conditions than in the 0T condition (both $ps < .001$), while the 1T and 2T conditions were not significantly different ($p = .17$). This result suggests that the participants' discriminability of the flashes was lower when taps were presented than when no tap was presented. The interaction between Tap and Location was significant ($F(1.50,28.52) = 5.49$, $MSE = 0.26$, $p < .05$). However, paired t-tests revealed no significant difference in the center and periphery in all three tap conditions (0T: $t(19) = 2.05$, $p = .16$; 1T: $t(19) = 1.63$, $p = .36$; 2T: $t(19) = 2.08$, $p = .16$). Thus, unlike the sound-induced fission illusion (Experiment 1), the larger tap-induced fission illusion in the periphery than in the center cannot be attributed to any significant change in discriminability to the flashes.

The c values (see Fig. 4f) were submitted to a two-way ANOVA on the factors of Tap (0T, 1T, and 2T) and Location (center or periphery). The main effect of Tap was significant ($F(1.33,25.19) = 38.65$, $MSE = 0.40$, $p < .001$), and pairwise t-tests indicated that c was decreased in the order of 1T (0.43), 0T (0.07), and 2T (−0.57), and they were all significantly different from each other ($ps < .001$). That the presentation of taps induced the shift of criterion partly explains the patterns of the fission and fusion illusions. Participants were more likely to report two flashes rather than one flash when the flashes were presented in the periphery ($c = -0.13$) compared to the center ($c = 0.09$; $F(1,19) = 14.29$, $MSE = 0.10$, $p < .005$). The two-way interaction was not significant ($F(2,38) = 1.77$, $MSE = 0.04$, $p = .18$).

Discussion

The visual fission and fusion illusions induced by the taps generally demonstrated similar patterns to those illusions induced by sounds (see Figs. 2, 3, and 4). That is, the fission and fusion illusions observed in the 1F/incongruent and 2F/incongruent conditions were attributable both to lower discriminability (d') to the flash and a shift of criterion (c) induced by the presentation of taps. In addition, accuracy in the 1F/incongruent condition was lower in the periphery than in the center, leading to a larger fission illusion in the periphery than in the center; on the other hand, in the 2F/incongruent condition, accuracy was lower in the center than in the periphery, leading to a marginally larger fusion illusion in the center than in the periphery. These patterns were attributable mainly to a shift in criterion (c), rather than a change of discriminability (d'). This difference from the sound-induced fission illusion reported in Experiment 1 (i.e., lower d' in the periphery than in the center) will be discussed in the [General discussion](#).

Another difference between Experiments 1 and 2 was the results in the 2F/catch trials: when the sound (or tap) led the flashes by 300 ms, a significant fusion illusion was still induced by the beep, but the tap did not induce a significant illusion (see Fig. 2). This suggests that the temporal window

for a fusion illusion induced by a sound is wider than that induced by a tap. Nevertheless, this explanation is confounded by the fact that the location of the stimuli differed: in particular, the sounds in Experiment 1 were perceived as having been presented somewhere in the center of the monitor (or ventriloquized to the location of flashes in the eccentricity of 10°; see Jackson, 1953). However, the taps in Experiment 2 were perceived on the right index finger placed 20° below the monitor. Previous research has demonstrated that temporal and spatial modulations on multisensory integration are highly correlated: the window of temporal simultaneity is wider when multisensory signals are presented close in space rather than more separated (see Spence, Baddeley, Zampini, James, & Shore, 2003; Spence, Shore, & Klein, 2001; Zampini et al., 2005). Hence, it is premature to draw any conclusion about whether differences in catch trials for the fusion illusion induced by a sound versus a tap were attributable to a wider temporal integration window or to spatial modulation.

General discussion

In the present study, we compared the fission and fusion illusions induced by sounds (Experiment 1) or taps (Experiment 2) in the central and peripheral visual field. The results can be summarized in terms of three outcomes (see Table 1). First, participants' accuracy was lower in the incongruent conditions than in the congruent conditions either when one flash (1F) or two flashes (2F) were presented, thus suggesting significant fission and fusion illusions, respectively. The occurrence of the fission and fusion illusions was attributable to the presentation of beeps or taps giving rise to a lower discriminability (d') as well as a shift of criterion (c) when responding to the flashes. Second, in the 1F/incongruent condition, the lower accuracy in the periphery than in the center led to a larger fission illusion in the periphery than in the center. This pattern of fission illusions induced by sounds has been reported in previous studies (Shams et al., 2000; Tremblay et al., 2007), and our analyses indicate that such a pattern can be attributed to both a lower d' (i.e., a lower discriminability between one vs. two flashes) and a lower c (i.e., a tendency to report two flashes) in the periphery than in the center. A similar pattern for the fission illusion induced by taps is reported here for the first time; nevertheless, it was attributed mainly to the lower c in the periphery than in the center without a significant change of d' . Third, we observed a larger fusion illusion in the center than in the periphery either induced by a sound or tap, which has not been reported previously, perhaps because of the larger eccentricity used in the current (10°) than the previous study (e.g., 5°; Tremblay et al., 2007). However, the larger fusion illusion in the center was attributed only to the higher c (i.e., a tendency to report one

flash) in the center than periphery, with no difference in d' between the two locations.

Taken together, the similar center/periphery patterns for sound- and tap-induced flash illusions support a common effect of visual eccentricity on the illusions, regardless of the modality of the inducer. However, the analyses of changes in discriminability from center to periphery suggest that only the audiovisual fission illusion is underpinned by a perceptual mechanism: its increased magnitude in the periphery was associated with a decreased ability to discriminate one from two flashes. All other changes in the magnitude of the illusions with eccentricity (larger visuotactile fission in the periphery than center, and larger audiovisual and visuotactile fusions in the center than periphery) correlated only with changes in criterion, a pattern suggesting a tendency to report two flashes in the periphery but one flash in the center.

The results of the present study therefore suggest that *audiovisual* fission and fusion illusions have different underlying mechanisms. Only the larger fission illusion in the periphery is a result of a change in discriminability to the flashes, whereas the larger fusion illusion in the center is not. In addition, when the sound(s) led the flash(es) by 300 ms, the fission illusion was eliminated whereas the fusion illusion remained. Together, the evidence suggests that the audiovisual fission illusion is a result of audiovisual integration. Hence, audiovisual fission and fusion illusions cannot be taken as the two facets of a single phenomenon (e.g., Bolognini et al., 2016); nor can they be accommodated easily using a single mechanism of multisensory integration (see Andersen et al., 2004).

The perceptual component of the *sound-induced* fission illusion is consistent with previous behavioral studies suggesting that the illusory flash in the fission illusion appears perceptually real (Berger et al., 2003; Fiedler, O'Sullivan, Schroter, Miller, & Ulrich, 2011; McCormick & Mamassian, 2008; though see van Erp, Philippi, & Werkhoven, 2013). In addition, neuropsychological studies suggest that the neural activity associated with the illusory flash in the fission illusion can be observed 30–60 ms after the onset of the second beep, fast enough, in other words, to reflect feedforward processing in the primary visual cortex (Mishra, Martinez, Sejnowski, & Hillyard, 2007; Shams, Iwaki, Chawla, & Bhattacharya, 2005; Watkins, Shams, Tanaka, Haynes, & Rees, 2006; see also Bolognini et al., 2016 for evidence from patients with visual field deficits caused by brain damage). Consistent with this view, the fission illusion can be modulated by transcranial direct current stimulation (tDCS) applied over occipital and temporal cortices (Bolognini, Ressetti, Casati, Mancini, & Vallar, 2011). On the other hand, the fusion illusion is associated with neural activity at a later time point (Mishra, Martinez, & Hillyard, 2008) and is not modulated by tDCS over occipital and temporal cortices (Bolognini et al., 2011), consistent with a more post-perceptual (or decisional) account

Table 1 A summary of the patterns of fission and fusion illusions induced by sounds (Experiment 1) or taps (Experiment 2) when the flashes were presented in the central versus peripheral visual field. d' and c represent the sensitivity and criterion, respectively, based on signal detection theory

	Fission illusion				Fusion illusion			
	1 Flash, 2 Beeps		1 Flash, 2 Taps		2 Flashes, 1 Beep		2 Flashes, 1 Tap	
	Center	Periphery	Center	Periphery	Center	Periphery	Center	Periphery
Accuracy	>		>		<		<	
Magnitude of the illusion	<		<		>		>	
d'	>		=		=		=	
c	>		>		>		>	

for this illusion (though see Watkins, Shams, Josephs, & Rees, 2007).

The mechanism underlying the shift of criterion (c) in the fission and fusion illusion, nevertheless, has been debated recently. Witt et al. (2015; see also Witt, Taylor, Sugovic, & Wixted, 2016) proposed that such a criterion shift when comparing fission and fusion illusions should be attributed sometimes to “perceptual bias” rather than “decisional bias.” The former is based on the hypothesis that the presentation of the sound may induce equal shifts of both signal and noise distributions, leading to the shift of calculated criterion, even though the actual decisional criterion may not change (see Fig. 1 in Witt et al., 2016). Indeed, the current results demonstrate that both the hit and FA rates were higher in the 2B condition (hit rate: 95.4%; FA rate: 58.6%) than in the 1B condition (hit rate: 59.1%; FA rate: 6.1%). Hence, for the comparison of fission and fusion illusions, it is not possible to disentangle whether the shift of c is a result of perceptual bias or decisional bias. Nevertheless, it should be noted that discriminability to the flashes was lower in the audiovisual fission illusion (in the 2B condition) than in the audiovisual fusion illusion (in the 1B condition), consistent with a genuine perceptual effect (see Knotts & Shams, 2016).

It should be further noted that Witt et al. (2015, p. 297) stated that “The claim that c can measure perceptual biases applies only when the manipulation causes a shift in both distributions. If the manipulation shifts only one distribution, then a perceptual bias should reveal itself in d' , and c will reflect decisional processes alone.” The main focus in the current study was the comparison between the fission (or fusion) illusion in the center vs. periphery. When estimating the discriminability and criterion in the center and periphery, only either the hit or FA rate was changed, but they did not change together. Specifically, when comparing the *audiovisual* fission illusion in the central and peripheral conditions, only the FA rates were different (50.3% vs. 66.9%), but the hit rates were similar (96.3% vs. 94.5%). For the *audiovisual* fusion illusion, only the hit rates were different (54.0% vs. 64.3%), but not the FA rates (4.8% vs. 7.5%). This was also true for *visuotactile*

illusions when comparing the central and peripheral conditions: for the fission illusion, the FA rates were different (26.3% vs. 39.3%), while the hit rates were similar (94.8% vs. 97.0%); for the fusion illusion, only the hit rates were different (70.5% vs. 81.2%), while the FA rates were similar (2.3% vs. 4.1%). Consequently, the shift of criterion between central and peripheral conditions may be attributable to the decisional bias in the current study.³

Based on the signal detection analysis, we propose two distinct mechanisms, one perceptual and the other decisional, to account for the results. The perceptual mechanism, for greater *audiovisual* integration in the periphery, may arise because of the larger number of neural projections from auditory cortex to visual cortex in the periphery (Falchier et al., 2002; Rockland & Ojima, 2003). This perceptual mechanism in terms of early audiovisual interactions can partly explain the lower discriminability to the flashes in the periphery than center in the sound-induced fission illusion. By contrast, to our knowledge, there is no empirical evidence for similar linkages between vision and touch. Consistently, we did not observe a significant modulation by the presentation of taps on discriminability to the flashes in the periphery. One account of these results proposes that the mapping between visual and auditory spatial representations remains relatively constant across time – the eyes and ears are fixed in the head. Hence, the hardwired connections likely facilitate the rapid audiovisual integration and modulate peripheral vision. The mapping between vision and touch, on the other hand, requires the coordination of both external and proprioceptive reference frames (e.g., Shore,

³ We acknowledge, however, that, as pointed out by a reviewer, these arguments rest on assumptions about the underlying distributions; specifically, multisensory integration might lead to a more reliable response in the congruent condition (i.e., a distribution with smaller variance, see Witt et al., 2016, p. 5). Whether this assumption is true can be verified using computational modeling of people’s responses in the sound- or tap-induced illusions (e.g., Adams, 2016; Bresciani et al., 2006). If only the response variance in the congruent rather than in the incongruent condition was reduced, then a definitive test would require plotting the whole receiver operating characteristic (ROC) curve by creating different hit rates and measuring the accompanying false alarm rates.

Spry & Spence, 2002), most likely mediated by the multisensory coordination network in the parietal lobe (Azañón, Long, Soto-Faraco & Haggard, 2010; Gallace & Spence, 2014; Lloyd, Shore, Spence & Calvert, 2003). Thus, spatial coordination between vision and touch appears to occur in higher multisensory association areas rather than depending on direct connections as has been suggested for audiovisual interactions.

The second proposed mechanism is decisional: stimuli in the periphery are interpreted as dynamic rather than static. This is supported by the tendency to report two flashes (i.e., a lower criterion) for all visual stimuli in this location, even when presented unimodally. This mechanism is perhaps underpinned by the fact that the temporal resolution (measured by the critical flicker fusion frequency) is higher in peripheral than central vision (Hartmann et al., 1979; Rovamo, & Raninen, 1984; see Strasburger, Rentschler, & Jüttner, 2011, for a review). Hence, ambiguous stimuli (e.g., in the incongruent condition in the current study) may be resolved as dynamic events in the periphery since this is more likely to be correct. We therefore suggest that our findings of larger fission and smaller fusion in the periphery than in the center arise from a general assumption of dynamic rather than static stimuli in the periphery. This decisional mechanism in peripheral vision is consistent with the explanation proposed by Shams et al. (2002) stating that discontinuous information would modulate the other modality containing continuous information (i.e., in the fission case), but not vice versa (i.e., in the fusion case) in the periphery. Consistent with this notion, the unimodal fission illusion (i.e., one flash perceived as two when presented with second visual stimulus with two flashes) is stronger when the target flash is presented 4° in the periphery compared to in the center (see Chatterjee, Wu, & Sheth, 2011, Experiment 2; see also Wilson & Singer, 1981).

In summary, the results of the present study demonstrate that both audiovisual and visuotactile processing are influenced by visual eccentricity: a larger fission illusion in the periphery than in the center is attributable to changes in both discriminability and criterion when induced by sounds while only by changes in criterion when induced by taps. In contrast, a larger fusion illusion in the center than in the periphery is attributable to a change only in criterion when induced either by a sound or tap. These results can be explained by combining a perceptual mechanism that results in audiovisual (but not visuotactile) integration being more pronounced in the periphery than in the center, and a decisional mechanism that visual events in the periphery are plausibly dynamic rather than static irrespective of the modality of the inducers.

Acknowledgements This study was funded by the James S. McDonnell Foundation and NSERC to DM, and Discovery grant from NSERC to DIS. YCC and CS are currently supported by AHRC Rethinking the Senses grant (AH/L007053/1). We thank Sally Stafford,

Shalini Seneviratne, Savneet Uppal, and Miriam Mokhtar for their great help in data collection.

References

- Adams, W. J. (2016). The development of audio-visual integration for temporal judgements. *PLoS Computational Biology*, *12*(4), e1004865. doi:10.1371/journal.pcbi.1004865
- Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, *14*, 257–262.
- Andersen, T. S., Tiippana, K., & Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Cognitive Brain Research*, *21*, 301–308.
- Andersen, T. S., Tiippana, K., & Sams, M. (2005). Maximum Likelihood Integration of rapid flashes and beeps. *Neuroscience Letters*, *380*, 155–160.
- Azañón, E., Longo, M. R., Soto-Faraco, S., & Haggard, P. (2010). The posterior parietal cortex remaps touch into external space. *Current Biology*, *20*, 1304–1309.
- Berger, T. D., Martelli, M., & Pelli, D. G. (2003). Flicker flutter: Is an illusory event as good as the real thing? *Journal of Vision*, *3*, 406–412.
- Bolognini, N., Rossetti, A., Casati, C., Mancini, F., & Vallar, G. (2011). Neuromodulation of multisensory perception: A tDCS study of the sound-induced flash illusion. *Neuropsychologia*, *49*, 231–237.
- Bolognini, N., Convento, S., Casati, C., Mancini, F., Brighina, F., & Vallar, G. (2016). Multisensory integration in hemianopia and unilateral spatial neglect: Evidence from the sound induced flash illusion. *Neuropsychologia*, *87*, 134–143.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.
- Bresciani, J. P., Dammeier, F., & Ernst, M. O. (2006). Vision and touch are automatically integrated for the perception of sequences of events. *Journal of Vision*, *6*, 554–564.
- 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*, 753–760.
- Charbonneau, G., Véronneau, M., Boudrias-Fournier, C., Lepore, F., & Collignon, O. (2013). The ventriloquist in periphery: Impact of eccentricity-related reliability on audio-visual localization. *Journal of Vision*, *13*(12):20, 1–14.
- Chatterjee, G., Wu, D. A., & Sheth, B. R. (2011). Phantom flashes caused by interactions across visual space. *Journal of Vision*, *11*(2):14, 1–14.
- Chen, Y.-C., Shore, D. I., Lewis, T. L., & Maurer, D. (2016). The development of the perception of audiovisual simultaneity. *Journal of Experimental Child Psychology*, *146*, 17–33.
- Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*, 429–433.
- Falchier, A., Clavagnier, S., Barone, P., & Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. *The Journal of Neuroscience*, *22*, 5749–5759.
- Fiebelkorn, I. C., Foxe, J. J., Butler, J. S., & Molholm, S. (2011). Auditory facilitation of visual-target detection persists regardless of retinal eccentricity and despite wide audiovisual misalignments. *Experimental Brain Research*, *213*, 167–174.
- Fiedler, A., O'Sullivan, J. L., Schröter, H., Miller, J., & Ulrich, R. (2011). Illusory double flashes can speed up responses like physical ones: Evidence from the sound-induced flash illusion. *Experimental Brain Research*, *214*, 113–119.
- Gallace, A., & Spence, C. (2014). *In touch with the future: The sense of touch from cognitive neuroscience to virtual reality*. Oxford, UK: Oxford University Press.

- Gleiss, S., & Kayser, C. (2013). Eccentricity dependent auditory enhancement of visual stimulus detection but not discrimination. *Frontiers in Integrative Neuroscience*, 7, 52.
- Hartmann, E., Lachenmayr, B., & Brettel, H. (1979). The peripheral critical flicker frequency. *Vision Research*, 19(9), 1019–1023.
- Jackson, C. V. (1953). Visual factors in auditory localization. *Quarterly Journal of Experimental Psychology*, 5, 52–65.
- Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception*, 36 ECVPAbstract Supplement, 14.
- Knotts, J. D., & Shams, L. (2016). Clarifying signal detection theoretic interpretations of the Müller-Lyer and sound-induced flash illusions. *Journal of Vision*, 16(11):18, 1–4.
- Lloyd, D. M., Shore, D. I., Spence, C., & Calvert, G. A. (2003). Multisensory representation of limb position in human premotor cortex. *Nature Neuroscience*, 6, 17–18.
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide*. Mahwah, NJ: Lawrence Erlbaum Associates.
- McCormick, D., & Mamassian, P. (2008). What does the illusory-flash look like? *Vision Research*, 48, 63–69.
- 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, 4120–4131.
- Mishra, J., Martinez, A., & Hillyard, S. A. (2008). Cortical processes underlying sound-induced flash fusion. *Brain Research*, 1242, 102–115.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442.
- Rockland, K. S., & Ojima, H. (2003). Multisensory convergence in calcarine visual areas in macaque monkey. *International Journal of Psychophysiology*, 50, 19–26.
- Rosenthal, O., Shimojo, S., & Shams, L. (2009). Sound-induced flash illusion is resistant to feedback training. *Brain Topography*, 21, 185–192.
- Rovamo, J., & Raninen, A. (1984). Critical flicker frequency and M-scaling of stimulus size and retinal illuminance. *Vision Research*, 24, 1127–1131.
- Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature*, 408, 788.
- Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research*, 14, 147–152.
- Shams, L., Iwaki, S., Chawla, A., & Bhattacharya, J. (2005). Early modulation of visual cortex by sound: An MEG study. *Neuroscience Letters*, 378, 76–81.
- Shore, D. I., Spry, E., & Spence, C. (2002). Confusing the mind by crossing the hands. *Cognitive Brain Research*, 14, 153–163.
- Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. *Journal of Experimental Psychology: General*, 130, 799–832.
- Spence, C., Baddeley, R., Zampini, M., James, R., & Shore, D. I. (2003). Multisensory temporal order judgments: When two locations are better than one. *Perception & Psychophysics*, 65, 318–328.
- Strasburger, H., Rentschler, I., & Jüttner, M. (2011). Peripheral vision and pattern recognition: A review. *Journal of Vision*, 11(5):13, 1–82.
- Tremblay, C., Champoux, F., Voss, P., Bacon, B. A., Lepore, F., & Théoret, H. (2007). Speech and non-speech audio-visual illusions: A developmental study. *PLoS One*, 2(8), e742.
- van Erp, J. B., Philippi, T. G., & Werkhoven, P. (2013). Observers can reliably identify illusory flashes in the illusory flash paradigm. *Experimental Brain Research*, 226, 73–79.
- Violentyev, A., Shimojo, S., & Shams, L. (2005). Touch-induced visual illusion. *Neuroreport*, 16, 1107–1110.
- Watkins, S., Shams, L., Tanaka, S., Haynes, J. D., & Rees, G. (2006). Sound alters activity in human V1 in association with illusory visual perception. *NeuroImage*, 31, 1247–1256.
- Watkins, S., Shams, L., Josephs, O., & Rees, G. (2007). Activity in human V1 follows multisensory perception. *NeuroImage*, 37, 572–578.
- Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, 88, 638–667.
- Wilson, J. T., & Singer, W. (1981). Simultaneous visual events show a long-range spatial interaction. *Perception & Psychophysics*, 30, 353–356.
- Witt, J. K., Taylor, J. E. T., Sugovic, M., & Wixted, J. T. (2015). Signal detection measures cannot distinguish perceptual biases from response biases. *Perception*, 44, 289–300.
- Witt, J. K., Taylor, J. E. T., Sugovic, M., & Wixted, J. T. (2016). Further clarifying signal detection theoretic interpretations of the Müller-Lyer and sound-induced flash illusions. *Journal of Vision*, 16(11): 19, 1–7.
- Wozny, D. R., Beierholm, U. R., & Shams, L. (2008). Human trimodal perception follows optimal statistical inference. *Journal of Vision*, 8(3):24, 1–11.
- Wright, M. J., & Johnston, A. (1983). Spatiotemporal contrast sensitivity and visual field locus. *Vision Research*, 23(10), 983–989.
- Zampini, M., Guest, S., Shore, D. I., & Spence, C. (2005). Audio-visual simultaneity judgments. *Perception & Psychophysics*, 67, 531–544.