RESEARCH ARTICLE

Temporal dynamics of the flash-induced bouncing effect

Hui Zhong | Song Zhao (ORCID) | Tingji Chen | Wanlu Yang | Xinyin Huang | Wenfeng Feng (ORCID)

Department of Psychology, School of Education, Soochow University, Suzhou, Jiangsu, China

Correspondence
Xinyin Huang and Wenfeng Feng, Department of Psychology, School of Education, Soochow University, Suzhou 215123, Jiangsu, China.
Email: hxy5123@163.com (X. Y. H.) and fengwfly@gmail.com (W. F. F.)

Funding information
National Natural Science Foundation of China, Grant/Award Number: 31771200

Abstract

Two identical visual disks moving toward each other on a two-dimensional (2D) display are more likely to be perceived as "streaming through" than "bouncing off" each other after their coincidence. However, either a brief auditory tone or visual flash presented at the coincident moment of the disks can strikingly increase the incidence of the bouncing percept. Despite the neural substrates underlying the sound-induced bouncing effect have been widely investigated, little is known about the neural mechanisms underlying the flash-induced bouncing effect. The present study used event-related potential recordings to explore the temporal dynamics of the flash-induced bouncing effect. The results showed that the amplitude of the post-coincidence parietooccipital P2 component (190–230 ms after coincidence) elicited by the visual motion was significantly smaller on bouncing relative to streaming trials only when the flash was presented but not when absent. In addition, the parietal P3 component (330–430 ms) was found to be larger on bouncing than streaming trials when the flash was presented, but the opposite was true when no flash was presented. These electrophysiological findings suggest that the flash-induced bouncing effect may occur at both perceptual and postperceptual stages of processing.

KEYWORDS
bistable perception, ERPs, flash-induced bouncing effect, stream/bounce

1 | INTRODUCTION

Consider two identical visual disks moving toward each other along the same horizontal line with equal speed in a 2D display: two disks start their motion from opposite sides, coincide (i.e., completely overlap) at the center of the screen, move apart, and stop at each other's starting point. Observers typically perceived the two disks after their coincidence as either "streaming through" or "bouncing off" each other (Metzger, 1934), whereas the incidence of streaming percept was much higher than bouncing percept (Bertenthal, Banton, & Bradbury, 1993; Sekuler & Sekuler, 1999). Interestingly, many previous studies have found that when delivering a salient sound that was task-irrelevant at the moment of two disks' coincidence, the dominant percept of this bistable motion display could be strikingly reversed from streaming toward bouncing (Adams & Grove, 2018; Dufour, Touzalin, Moessinger, Brochard, & Després, 2008; Grassi & Casco, 2010; Grove, Ashton, Kawachi, & Sakurai, 2012; Kawabe & Miura, 2006; Remijn, Ito, & Nakajima, 2004; Sekuler, Sekuler, & Lau, 1997; Watanabe & Shimojo, 2001a). Moreover, several studies have shown that this bouncing effect could be induced not only by the auditory transient, but also by a unimodal visual flash presented at the coincident moment of the two disks (Adams & Grove, 2018; Burns & Zanker, 2000; Kawabe & Miura, 2006; Watanabe & Shimojo, 1998, 2001a).

Hui Zhong and Song Zhao contributed equally to this work and are cofirst authors.

Although these transient-induced bouncing effects have been consistently replicated, there are still considerable controversies about their underlying processing mechanisms. One representative hypothesis held that these transient-induced bouncing effects occur at the perceptual stage of processing (e.g., Berger & Ehrsson, 2017; Dufour et al., 2008; Remijn et al., 2004; Watanabe & Shimojo, 2001b). Specifically, this hypothesis proposed that the salient visual or auditory transient presented at the moment of the two disks' coincidence makes the observers genuinely see the two disks reverse their motion after coincidence and return to their original starting position. Contrary to the perceptual hypothesis, the inference hypothesis proposed that the transient-induced bouncing effects originate from shifted cognitive bias by the transient at late decision-making stage (e.g., Adams & Grove, 2018; Grove et al., 2012; Grove, Robertson, & Harris, 2016; Grove & Sakurai, 2009; Zeljko & Grove, 2016). That is, the transient appearing at the coincident moment of the two disks imitates the causal consequence (i.e., a release of sound or light energy) when two objects collide in the natural environment. According to this real-life experience, observers can infer in a post hoc manner that the two disks were more likely to bounce off each other after the coincident transient onset (for details, see Adams & Grove, 2018).

A growing number of neuroscientific studies have explored the neural mechanisms responsible for the sound-induced bouncing effect (Bushara et al., 2003; Hipp, Engel, & Siegel, 2011; Maniglia, Grassi, Casco, & Campana, 2012; Zhao, Wang, Feng, & Feng, 2020; Zhao, Wang, Xu, Feng, & Feng, 2018; Zvyagintsev, Nikolaev, Sachs, & Mathiak, 2011). For instance, a functional magnetic resonance imaging study (Bushara et al., 2003) found that the coincident sound elicited higher activation in a series of multisensory brain areas (e.g., prefrontal and posterior parietal cortices) but lower activation in unisensory auditory and visual cortices on the audiovisual bouncing trials than streaming trials. Besides, using transcranial magnetic stimulation (TMS) to temporally interrupt the function of the posterior parietal cortex, Maniglia et al. (2012) showed that the sound-induced bouncing effect was significantly diminished relative to the condition without TMS, indicating a causal role the multisensory posterior parietal cortex plays in the sound-induced bouncing effect. More recently, Zhao et al. (2018, 2020) investigated the temporal dynamics of the sound-induced bouncing effect using event-related potential (ERP) recordings, and found that early cross-modal ERP components within 100–200 ms after sound onset over frontocentral and occipital regions were closely associated with the occurrence of this bouncing effect. These results demonstrated that early cross-modal interactions occurring at the perceptual stage of processing contribute to the sound-induced bouncing effect, thereby providing evidence for the perceptual hypothesis mentioned above.

In contrast to the sound-induced bouncing effect whose neural substrates have been widely investigated, little is known, however, about the neural mechanisms underlying the similar bouncing effect induced by the intramodal salient visual flash. In particular, it is currently unclear whether the intramodal visual flash would also influence the streaming/bouncing motion perception at the perceptual stage of processing as the effect of sound, or instead, at the postperceptual stages (e.g., the decision-making stage) as proposed by the inference

hypothesis mentioned above. To shed light on these issues, the current study examined the electrophysiological time course of the flash-induced bouncing effect using ERP recordings. Based on a percept-based analysis, the present study showed that the amplitude of the parietooccipital P2 component (190–230 ms after coincidence) elicited by the visual motion was significantly decreased on bouncing relative to streaming trials only when the flash was presented but not when absent. In addition, the parietal P3 component (330–430 ms) was found to be larger on bouncing than streaming trials when the flash was presented, but the opposite was true when no flash was presented. These electrophysiological findings suggest that the flash-induced bouncing effect occurs at both the perceptual stage and the postperceptual stage.

2 | METHODS

2.1 | Participants

A total of 32 healthy undergraduates (23 females, mean age of 19.2 years) participated in this experiment after giving written informed consent as approved by the Human Research Protections Program of Soochow University. All participants reported normal or corrected-to-normal visual acuity as well as normal hearing. They were all unaware of the hypothesis of the experiment. All experimental procedures were in agreement with the Declaration of Helsinki.

2.2 | Stimuli and procedure

The experiment was conducted in a dark and sound attenuated booth. All stimuli were generated and scripted using "Presentation" software (Neurobehavioral Systems, version 18.0). Visual motion stimuli were presented on a 27-in. LCD monitor (ASUS VG278HE, resolution $1,920 \times 1,080$, refresh rate 120 Hz) on which the background color was set to gray. Participants sat in front of the monitor with a viewing distance of 60 cm and held fixation on a red dot ($0.20 \times 0.20°$ of visual angle), which was presented at the center of the monitor throughout each block.

As illustrated in Figure 1, four different stimulus conditions were designed for the experiment, which were labeled as "V," "VF," "Catch," and "F" conditions for convenience.

1 The *V condition* referred to the classic streaming/bouncing display (see Figure 1b). Specifically, two identical black disks (each 1.49° in diameter) were originally presented at the opposite edges of the background on the first frame, separated by 14.9° horizontally and placed at the same vertical height as the fixation dot. From Frames 2–10, the two disks gradually moved toward each other along the same horizontal path. Each frame appeared immediately after the offset of the preceding frame, and the duration of each frame was 50 ms (i.e., frame to frame SOA was 50 ms). At the onset moment of Frame 11 (500 ms after the onset of Frame 1), the two disks visually overlapped each other completely (i.e., coincidence) at the
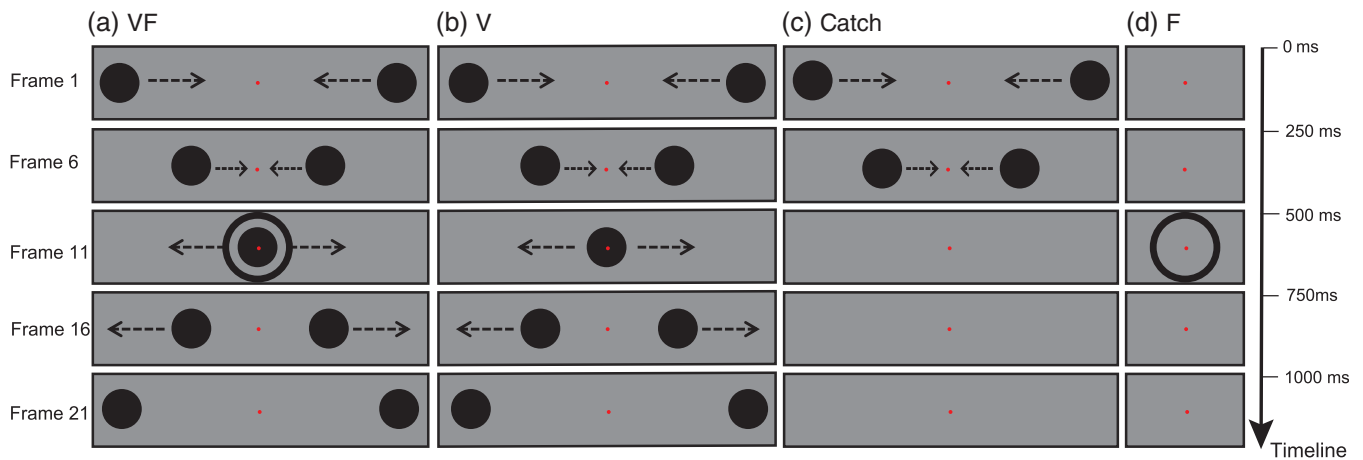
FIGURE 1 Schematic diagram of the four different stimulus conditions designed for the experiment, which were labeled as V, VF, Catch, and F, respectively (see Section 2 for details). (a) Stimuli sequences illustrated for VF condition, in which two identical visual disks started their motion from opposite sides, moved toward one another, coincided with the flash of a visual circle simultaneously, moved apart, and stopped at each other's starting positions and then disappeared. The long solid axis on the right (timeline) indicates the onset moments of the key frames shown here and each frame lasted for 50 ms. The dashed arrows near the disks indicate the direction of two disks' movement in the next frame and were not shown in the real experimental displays. (b–d) Stimuli sequences illustrated for V, Catch and F conditions, respectively

center of the monitor. From Frames 12–21, the two disks gradually moved apart from each other and eventually stopped at each other's starting point. Given the initial separation of 14.9° and the duration of 50 ms for each frame as well as a total of 21 frames, the motion of the two disks took 1,000 ms with a constant speed of 14.9°/s (i.e., 0.745° per frame).

2 The *VF condition* consisted of the same visual motion sequences as those in V condition, as well as a visual flash (black circle, 3.88° in diameter, 50 ms duration) presented at the same time and location as the two visual disks' coincidence, without occluding the coincidence of two disks (see Figure 1a; cf. Watanabe & Shimojo, 1998, 2001b).

3 In the *Catch condition*, the visual stimuli presented from Frames 1–8 were exactly the same as those in V condition, but no stimulus was presented from Frames 9–21 (Figure 1c). That is, two visual disks moved toward one another and then suddenly disappeared just before their coincidence, eliciting neither streaming nor bouncing percept. These catch trials were designed to ensure that participants were responding veridically based on their perceptual outcomes after two disks' coincidence event happened (Zhao et al., 2017, 2018; Zhao et al., 2020).

4 In the *F condition*, no visual motion was presented from Frames 1–21, but the flash (50 ms duration) was presented at the onset of Frame 11 and the center of the monitor (Figure 1d). This F condition was arranged to prevent subjects from guessing the relationship between transient flash and the visual motion.

All stimulus conditions were presented randomly with equal probability (25% for each of the four conditions) in each block, and the intertrial intervals ranged from 1,200 to 1,600 ms. Participants were required to indicate whether the two visual disks appeared to "stream through" or "bounce off" each other after their coincidence as rapidly and accurately as possible in V and VF conditions, by pressing one of

two buttons on a keyboard. The response buttons for "streaming" and "bouncing" percepts were counterbalanced across participants. No responses were required for the other two stimulus conditions. The whole experiment consisted of 30 blocks of 60 trials each and participants were allowed to have a rest after finishing each block.

2.3 | Electrophysiological recording and analysis

The electroencephalogram (EEG) was recorded continuously from 57 tin electrodes mounted in an elastic cap (Electro-Cap International, Inc.) using a NeuroScan SynAmps system (NeuroScan, Inc.). These electrode sites (FPz, FP1, FP2, Fz, F1, F2, F3, F4, F7, F8, FCz, FC1, FC2, FC3, FC4, FC5, FC6, Cz, C1, C2, C3, C4, C5, C6, T7, T8, CPz, CP1, CP2, CP3, CP4, CP5, CP6, TP7, TP8, Pz, P1, P2, P3, P4, P5, P6, P7, P8, POz, PO3, PO4, PO7, PO8, Oz, O1, O2, I3, I4, SI3, SI4, and the right mastoid) were positioned according to a modified 10–10 system montage (McDonald, Teder-Salejarvi, Russo, & Hillyard, 2003). Horizontal eye movements were monitored via a bipolar pair of electrodes positioned at the left and right outer canthi (horizontal EOG). Vertical eye movements and blinks were monitored bipolarly by two electrodes above and below the left eye (vertical EOG). All electrodes were referenced to the left mastoid electrode during recording, and all electrode impedances were kept below 5 kΩ. The online EEG and EOG signals were amplified with a gain of 10,000, band-pass filtered from 0.05 to 100 Hz, and were continuously digitized with a sampling rate of 1,000 Hz.

In offline processing, the continuous EEG signals were first low-pass filtered digitally (30 Hz, 24 dB/octave) using a zero phase-shift FIR filter to attenuate high-frequency noise triggered by muscle activities or external electrical sources. The filtered EEG data were rereferenced to the algebraic mean of the left and right mastoid electrodes. EEG signals were then divided into 600-ms epochs

time-locked to the onset of Frame 11 (i.e., the moment of two disks' coincidence) with a 100-ms precoincidence baseline and were baseline-corrected. Automatic artifact rejection was performed based on a threshold of ±80 μV for both EEG and EOG channels, in order to eliminate epochs contaminated by eye movements, eye blinks, and muscle activities. The artifact-free epochs were then averaged separately according to both stimulus condition (V, VF) and perceptual outcome (stream, bounce). On average, there were 127 artifact-free epochs on V_bounce trials, 242 on V_stream trials, 244 on VF_bounce trials, and 121 on VF_stream trials, respectively.

Because unimodal visual stimuli were used in the current study, a series of classic visual-evoked ERP components should have been expected, such as P1, N1, P2, and P3 (Anllo-Vento & Hillyard, 1996). However, since similar dynamic visual display as the present visual motion sequence has been shown to attenuate substantially the amplitudes of P1 and N1 components when ERP waveforms were time-locked to the critical event that occurred after the initial onset of the dynamic visual display (Leblanc, Prime, & Jolicoeur, 2008; Martens, Munneke, Smid, & Johnson, 2007; Van der Burg, Talsma, Olivers, Hickey, & Theeuwes, 2011), no obvious P1 and N1 components were expected in the present V and VF conditions. Therefore, the present study focused mainly on the postcoincidence P2 and P3 components. The time windows and electrode sites for measuring the post-coincidence P2 and P3 components were selected based *primarily* on previous ERP studies and *secondarily* on the grand-averaged ERP waveforms. First, prior investigations have shown that the P2 component was typically prominent from about 180 to 250 ms after stimulus onset (e.g., Anllo-Vento & Hillyard, 1996; Di Russo, Aprile, Spitoni, & Spinelli, 2008; Federmeier & Kutas, 2002; Freunberger, Klimesch, Doppelmayr, & Höller, 2007; Luck & Hillyard, 1994; Omoto et al., 2010), and the posterior P2 was typically measured over bilateral parietooccipital channels (e.g., Anllo-Vento & Hillyard, 1996; Federmeier & Kutas, 2002; Freunberger et al., 2007; Omoto et al., 2010). Therefore, the present study quantified the postcoincidence P2 component (time-locked to the coincidence of the two disks, i.e., Frame 11) as mean amplitude during a time window of 190–230 ms after two disks' coincidence over four homologous pairs of bilateral parietooccipital electrodes consisting of P7/P8, P5/P6, PO3/PO4, and PO7/PO8. These parameters fitted well with the actual timings and scalp distributions of postcoincidence P2 components in all conditions (see Figure 4). Second, as the P3 component is known to be maximal from approximately 300 to 600 ms after stimulus onset over middle centroparietal electrodes (Anllo-Vento & Hillyard, 1996; Johnson, 1986; Luck & Hillyard, 1994; Polich, 2007; Rik, Martijn, Jongsma, Kessels, & Francesco, 2014; Sergent, Baillet, & Dehaene, 2005), the current study measured the post-coincidence P3 component (time-locked to the coincidence of the two disks, i.e., Frame 11) as mean voltage within a broad time window of 330–430 ms over six neighboring middle centroparietal sites including CP1, CPz, CP2, P1, Pz, and P2. These parameters fitted well with the actual timings and scalp distributions of postcoincidence P3 component in all conditions (see Figure 5).

For statistical analysis, the postcoincidence P2 amplitudes were subjected to two-way repeated-measure analysis of variance (ANOVAs) with factors of perceptual outcome (stream, bounce) and brain hemisphere (left, right) separately for V and VF conditions. The P3 amplitudes were then subjected to one-way ANOVAs with a single factor of perceptual outcome (stream, bounce) also separately for V and VF conditions. Using this percept-based ERP analysis, the neural basis of the flash-induced bouncing effect (if there were) would be manifested as different patterns of the bounce verse stream comparison between VF and V conditions. Note that we did not include the stimulus condition (V vs. VF) as a factor in the above ANOVAs because the physical properties of stimuli at the coincident moment of two disks were inherently distinct between V and VF conditions (see Figure 1a,b, Frame 11). All reported *p*-values for ANOVA results were corrected using the Greenhouse–Geisser method (the same below).

Finally, the ERPs before the coincidence of the two visual disks were also analyzed, since the precoincidence P2 amplitude has been found to be predictive of subsequent streaming or bouncing percept (Zhao et al., 2017). More importantly, a recent study found that a brief sound presented at the moment of two disks' coincidence could interrupt the predictiveness of the precoincidence P2 amplitude on subsequent perceptual outcomes (Zhao et al., 2018). Given that the visual flash presented at the moment of two disks' coincidence has similar behavioral effect as the sound (i.e., increasing the incidence of bouncing percept), it is reasonable to expect that the predictiveness of the precoincidence P2 amplitude on subsequent perceptual outcomes could also be eliminated by the intermediate appearance of visual flash. To this end, the EEG was averaged in 600-ms epochs time-locked to the onset of Frame 1 (i.e., the initial onset of two visual disks, see Figure 1a,b), including a 100-ms prestimulus baseline. The precoincidence P2 component was measured as mean amplitude within a time interval of 250–300 ms after Frame 1 onset over a cluster of five parietal electrode sites (CPz, P1, Pz, P2, POz). These parietal electrodes, rather than the frontocentral electrodes reported previously (Zhao et al., 2017, 2018), were chosen because the current precoincidence P2 components were consistently evident over these parietal electrodes but not over frontocentral electrodes (see Figure 3, right) under *all* conditions (i.e., V_stream, V_bounce, VF_stream, and VF_bounce trials), which is logically equivalent to using the "*collapsed localizer*" method (Luck & Gaspelin, 2017) to select electrodes when the analysis parameters cannot be determined based on previous studies. The precoincidence P2 amplitudes were then subjected to a two-way repeated-measure ANOVA with factors of stimulus condition (V, VF) and perceptual outcome (stream, bounce). After finding a significant two-way interaction, the P2 amplitudes were further analyzed with specific comparisons between streaming and bouncing trials under each stimulus condition.

3 | RESULTS

3.1 | Behavioral results

A one-way ANOVA with a factor of stimulus condition (V vs. VF) on the percentage of bouncing percept showed a significantly sharp increase of bouncing responses for VF condition relative to

V condition (V: 34.3 ± 2.6% (*M* ± *SEM*); VF: 66.6 ± 2.6%; *F*(1, 31) = 58.61, *p* < .0001, η^2_p = 0.65; see Figure 2, left), which verified the classic flash-induced bouncing effect introduced by Watanabe and Shimojo (1998). Then, a stimulus condition (V vs. VF) × perceptual outcome (stream vs. bounce) two-way repeated-measure ANOVA on reaction times (RTs) did not show a significant main effect of either stimulus condition (*F*(1, 31) = 1.91, *p* = .176, η^2_p = 0.58; V: 540.1 ± 25.6 ms; VF: 548.9 ± 27.5 ms) or perceptual outcome (*F*(1, 31) = 0.25, *p* = .619, η^2_p = 0.08; stream: 542.3 ± 28.2 ms;

bounce: 546.8 ± 25.2 ms), and the stimulus condition × perceptual outcome interaction was also not significant (*F*(1, 31) = 0.20, *p* = .662, η^2_p = 0.06).

3.2 | ERP results

3.2.1 | The coincident flash eliminated the predictiveness of precoincidence P2 amplitude on perceptual outcome

ERP waveforms before the coincidence of the two visual disks are shown in Figure 3. The stimulus condition (V vs. VF) × perceptual outcome (stream vs. bounce) two-way ANOVA on the precoincidence P2 amplitudes (250–300 ms, time-locked to the initial onset of the two visual disks, i.e., Frame 1) revealed that the main effects of stimulus condition (*F*(1, 31) = 0.14, *p* = .713, η^2_p < 0.01) and perceptual outcome (*F*(1, 31) = 0.89, *p* = .354, η^2_p = 0.03) were not significant. However, the stimulus condition × perceptual outcome interaction was significant as expected (*F*(1, 31) = 4.76, *p* = .037, η^2_p = 0.13). Specific contrasts revealed that the precoincidence P2 amplitude was significantly larger on streaming than bouncing trials only in V condition (*F*(1, 31) = 4.96, *p* = .033, η^2_p = 0.14; V_stream: 2.66 ± 0.48 μV (*M* ± *SEM*); V_bounce: 2.06 ± 0.38 μV; see Figure 3a) but not in VF condition (*F*(1, 31) = 0.23, *p* = .637, η^2_p = 0.01; VF_stream: 2.36 ± 0.54 μV; VF_bounce: 2.52 ± 0.44 μV; see Figure 3b). These findings suggest that introducing a brief visual flash at the moment of two disks' coincidence might interrupt the predictiveness of the precoincidence P2 amplitude on subsequent perceptual outcome.
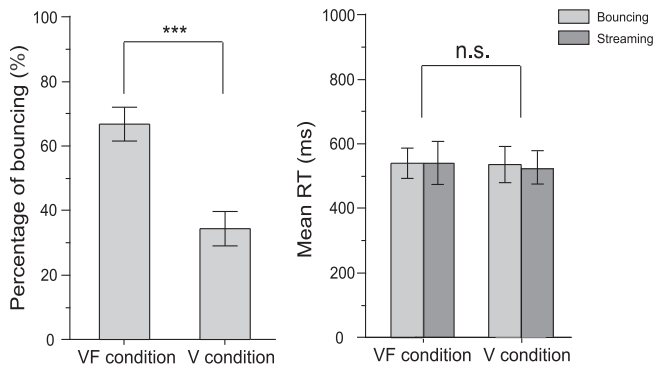

FIGURE 2 Behavioral results. The left graph depicted group mean (*n* = 32) percentage of bouncing responses under V and VF conditions. A highly significant increase of bouncing responses was found in VF condition relative to V condition. The right graph depicted group mean reaction times (RTs) as functions of stimulus condition (V vs. VF) and perceptual response (streaming vs. bouncing). Error bars in both graphs indicate ±1 *SEM*. ***: *p* < .001; n. s.: nonsignificant
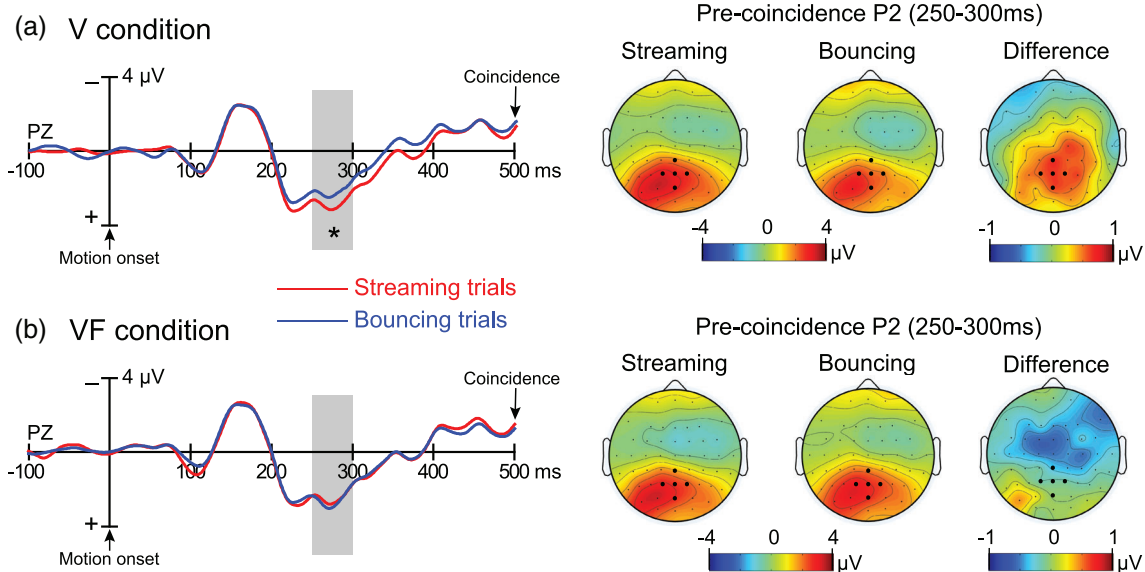

FIGURE 3 Grand-average event-related potential (ERP) waveforms elicited by the visual motion *before* the coincidence of the two disks (time-lock to the initial onset of two disks, i.e., the onset of Frame 1) are shown on the left for streaming and bouncing trials in V condition (a) and VF condition (b). Example waveforms shown were recorded from Pz electrode. Shaded areas show the time interval (250–300 ms) during which the precoincidence P2 component was quantified. Scalp topographies within the precoincidence P2 interval are shown on the right for streaming and bouncing trials as well as their differences (streaming minus bouncing) in V condition (a) and VF condition (b). *: *p* < .05
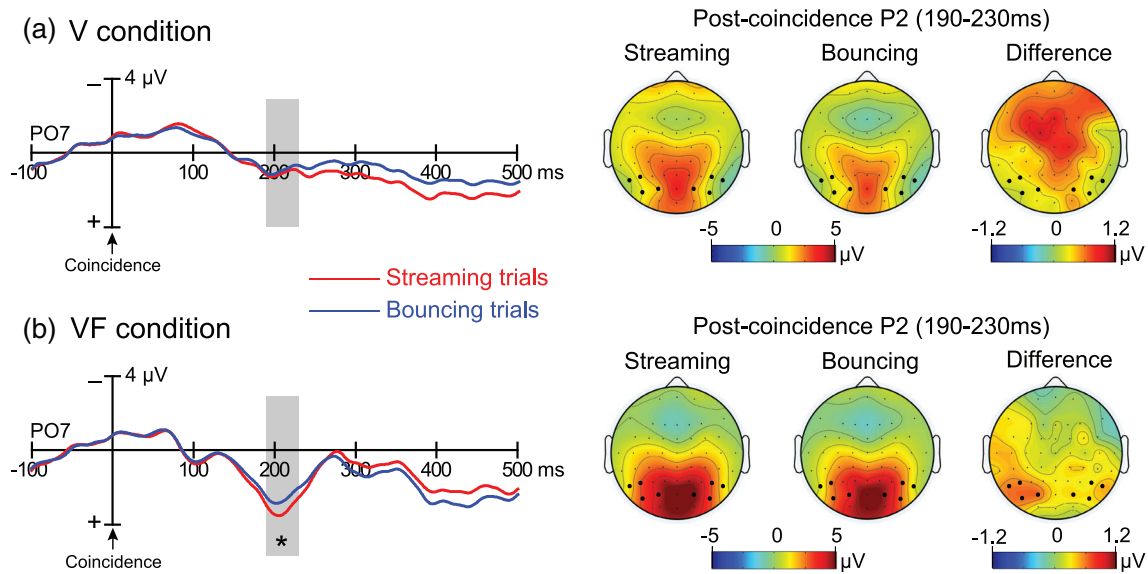
FIGURE 4 Grand-average event-related potential (ERP) waveforms elicited by the visual motion *after* the two disks' coincidence (time-lock to the coincident moment of two disks, i.e., the onset of Frame 11) are shown on the left for streaming and bouncing trials in V condition (a) and VF condition (b). Example waveforms shown were recorded from PO7 electrode. Shaded areas show the time interval (190–230 ms) within which the postcoincidence P2 components were quantified. Scalp topographies within the postcoincidence P2 interval are shown on the right for streaming and bouncing trials as well as their differences (streaming minus bouncing) in V condition (a) and VF condition (b). *: $p < .05$
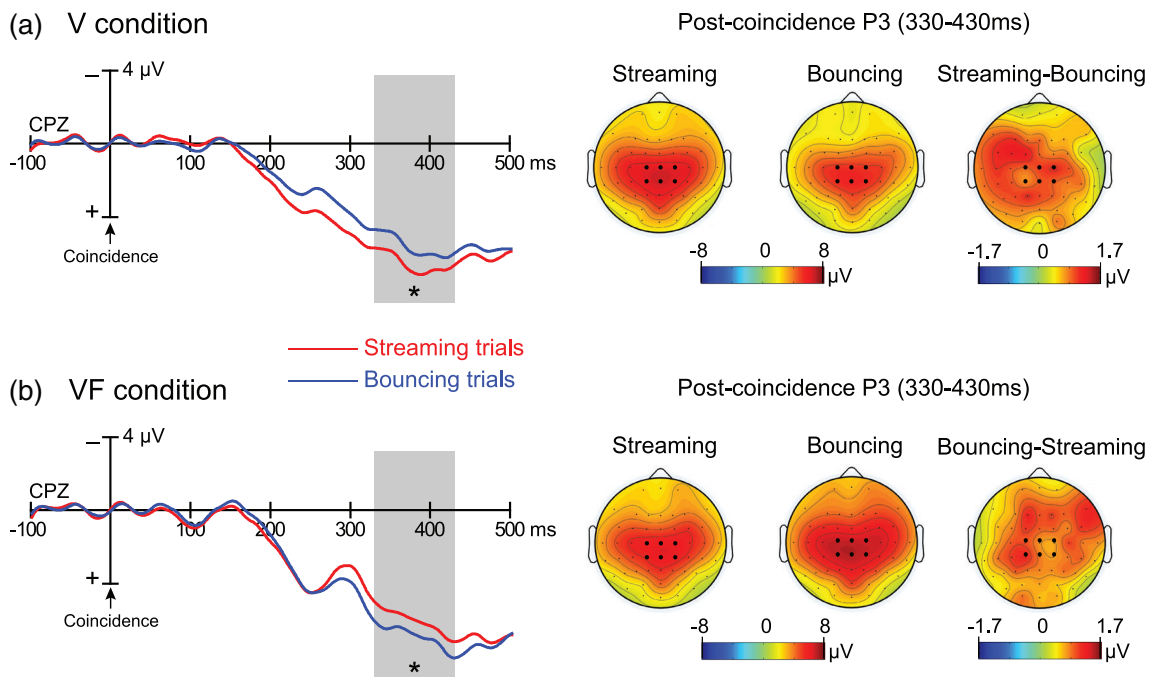

FIGURE 5 Similar to Figure 4 but for event-related potential (ERP) waveforms recorded from CPz electrode and scalp topographies during the time window of postcoincidence P3 component (330–430 ms, shaded areas). *: $p < .05$

3.2.2 | The flash-induced bouncing effect was associated with ERP modulations at both perceptual and postperceptual stages

After the coincidence of the two visual disks, there was first a clear positivity during 190–230 ms with a parietooccipital scalp topography

(Figure 4), which was most likely to be the visual P2 component (Boylan, Clement, & Harding, 1984; Lee et al., 2006; Luck & Hillyard, 1994). The perceptual outcome (stream vs. bounce) × brain hemisphere (left vs. right) ANOVA conducted for the postcoincidence P2 (time-locked to the coincidence of the two disks, i.e., Frame 11) amplitudes in V condition (Figure 4a) did not show a significant main effect

of perceptual outcome ($F(1, 31) = 2.18$, $p = .150$, $\eta^2_p = 0.01$; stream: 1.00 ± 0.43 µV; bounce: 0.75 ± 0.46 µV) or brain hemisphere ($F(1, 31) = 1.967$, $p = .171$, $\eta^2_p = 0.06$; left: 1.10 ± 0.45 µV; right: 0.65 ± 0.49 µV), and the two-way interaction was also nonsignificant ($F(1, 31) = 0.02$, $p = .903$, $\eta^2_p < 0.01$). In contrast, the postcoincidence P2 amplitudes in VF condition (Figure 4b) showed a significant main effect of perceptual outcome ($F(1, 31) = 6.35$, $p = .017$, $\eta^2_p = 0.17$), with postcoincidence P2 being smaller on bouncing trials (2.50 ± 0.56 µV) than streaming trials (3.00 ± 0.63 µV), but the main effect of brain hemisphere ($F(1, 31) = 0.89$, $p = .354$, $\eta^2_p < 0.01$; left: 2.68 ± 0.57 µV; right: 2.82 ± 0.67 µV) and the perceptual response × brain hemisphere interaction were not significant ($F(1, 31) = 1.63$, $p = 0.211$, $\eta^2_p = 0.05$). These percept-based analyses indicated that neural activity at a relatively early stage of perceptual processing contributes to the generation of the flash-induced bouncing effect. Of note, the observed postcoincidence P2 modulation in VF condition could not be accounted by postcoincidence P2 difference between streaming and bouncing percepts in V condition because there was no significant postcoincidence P2 difference in V condition.

Following the postcoincidence P2 component, there was a broad positive wave maximal within 330–430 ms over centroparietal scalp (Figure 5). Given its late latency and scalp distribution, this slow positive wave was most likely to be the typical P3/P300 component associated with postperceptual evaluation of task-defined events (Johnson, 1986; Polich, 2007). The one-way ANOVA conducted for P3 amplitudes in V condition (Figure 5a) showed that the P3 was significantly larger on streaming than bouncing trials (V_stream: 6.30 ± 0.76 µV; V_bounce: 5.34 ± 0.69 µV; $F(1, 31) = 4.87$, $p = .035$, $\eta^2_p = 0.14$). Instead, the opposite was true in VF condition (Figure 5b), with greater P3 amplitude on bouncing versus streaming trials (VF_stream: 5.92 ± 0.75 µV; VF_bounce: 6.78 ± 0.76 µV; $F(1, 31) = 4.94$, $p = .034$, $\eta^2_p = 0.14$). These results suggest that higher level cognitive processes also appear to play a role in the flash-induced bouncing effect.

4 | DISCUSSION

The current study investigated the neural mechanisms by which the salient visual flash presented at the coincident moment of two disks influenced the streaming/bouncing motion perception. Consistent with the classic flash-induced bouncing effect (Watanabe & Shimojo, 1998), the behavioral results showed that the visual flash presented synchronously with the coincidence of two disks (i.e., VF condition) significantly increased the percentage of bouncing percept relative to that when no flash was presented (i.e., V condition), to the extent that the perceptual dominance was reversed from streaming to bouncing. The reliable flash-induced bouncing effect replicated in the present study ensured the validity of comparing ERPs on bouncing trials with those on streaming trials separately for VF condition and V condition. Using this percept-based ERP analysis, the neural basis of the flash-induced bouncing effect would be manifested as different patterns of the bounce verse stream comparison between VF and V conditions.

The ERP results first showed that the precoincidence P2 component (250–300 ms, time-locked to the initial onset of two visual disks) was significantly larger on streaming than bouncing trials under V condition. This result seems to provide further evidence for the finding of Zhao et al. (2017, 2018), which indicated that the precoincidence P2 amplitude was predictive of subsequent postcoincidence perceptual outcome in visual streaming/bouncing display. Previous studies of other bistable perceptions have also found that brain activities prior to stimulus onset could influence observers' subsequent perceptual outcome (Britz, Landis, & Michel, 2009; Hesselmann, Kell, Eger, & Kleinschmidt, 2008; Hesselmann, Kell, & Kleinschmidt, 2008; Williams, Elfar, Eskandar, Toth, & Assad, 2003). The precoincidence P2 effect seems to be in line with these findings and implies that the fluctuations of brain activities before stimulus-driven processing play an important role when presented with ambiguous visual signals. More importantly, the precoincidence P2 effect was found to be absent under the VF condition in the present study. This result seems to extend the recent finding of Zhao et al. (2018) that the precoincidence P2 amplitude was not predictive of the subsequent perceptual outcome when a transient auditory tone was delivered synchronously with the coincidence of two disks. It should be noted that these findings *do not* indicate the presence of a salient stimulus (visual or auditory) impact the precoincidence brain activity per se, because the salient stimulus was presented after the coincidence. Instead, given that the transient-induced bouncing effects were manifested as streaming or bouncing percept being redetermined by the coincident transient, the absence of precoincidence P2 effect when the coincident transient was presented thus suggests that the transient stimulus, irrespective of its modality, presented synchronously with the coincidence of two disks might substantially attenuate/eliminate the predictability of the precoincidence brain activity on the final perceptual outcome.

However, the interpretation above should be treated with caution given the difference in scalp topography between the current precoincidence P2 component (parietal maximum) and that reported in previous studies (frontocentral maximum; Zhao et al., 2017, 2018). The reason for this disparity, as far as we can speculate, seems be a slightly difference in stimulus parameter. Specifically, in the studies of Zhao et al. (2017, 2018), the two moving visual disks were placed above the fixation (i.e., at upper visual field). In contrast, the present study presented the visual motion at the same vertical height as the fixation in order to ensure that when the two disks coincided, the transient visual flash could appear at identical location as the two visual disks' coincidence. This speculation might be supported, at least in part, by a previous finding that P2 amplitude varied with the location (upper/lower visual field) of stimulus (Ao, Fan, He, & Chen, 2000). Hence, future ERP studies are needed to test this speculation by manipulating the height of streaming/bouncing visual motion relative to the fixation systematically.

The ERP results after the coincidence of two disks showed that the amplitude of the parietooccipital postcoincidence P2 component (190–230 ms, time-locked to the coincidence of the two disks) elicited by the visual motion was significantly decreased on bouncing relative to streaming trials in the VF condition. In contrast, this P2 difference was not found between bouncing and streaming trials in the V condition,

suggesting the decreased P2 amplitude on bouncing relative to streaming trials in the VF condition resulted from the presentation of the coincident visual flash. Previous studies have found evidence that the visual-evoked P2 component is involved in visual search processes (Luck & Hillyard, 1994), memory processes (Dunn, Dunn, Languis, & Andrews, 1998), and language processes (Federmeier & Kutas, 2002; Wlotko & Federmeier, 2007). Moreover, it has been suggested that the P2 component is related to the reactivation of occipital brain areas (Di Russo et al., 2008). More relevant to the current study, the bilateral parietooccipital P2 component has been shown to be sensitive to depth perception (Liu et al., 2009; Omoto et al., 2010). For instance, Omoto et al. (2010) found that the P2 amplitude elicited by images yielding three-dimensional (3D) impression was significantly larger than images yielding 2D impression after minimizing the difference in physical properties between the two kinds of images. It is worth mentioning that several pioneering studies of streaming/bouncing motion (e.g., Sekuler & Sekuler, 1999; Watanabe & Shimojo, 2001a, 2001b) have proposed that since multiple moving objects in the real world are typically located at different depth planes, our perceptual system might utilize this 3D depth experience when perceiving the 2D streaming/bouncing motion, which leads to the streaming percept becoming dominant. Combining the sensitivity of P2 component to depth perception with this proposal, the present finding of smaller postcoincidence P2 amplitude on bouncing than streaming trials in VF condition may suggest that the role the coincident visual flash plays on inducing bouncing percept is to bind the two disks into the same depth plane.

In addition, the timing of the present postcoincidence P2 effect (i.e., ~200 ms after flash onset) implies that the above-mentioned binding process that might be reflected by the postcoincidence P2 effect occurs at an early stage of perceptual processing. It is noteworthy that two recent ERP studies on the sound-induced bouncing effect consistently observed an occipitally distributed cross-modal positivity also around 200 ms after coincidence (Zhao et al., 2018, 2020). However, this cross-modal positivity was found to be larger, instead of smaller, on audiovisual bouncing than streaming trials. Furthermore, before this occipital positivity, there was an earlier cross-modal positivity (around 150 ms after coincidence) over the frontocentral scalp that was also responsible for the sound-induced bouncing effect (Zhao et al., 2018, 2020). These disparities in brain dynamics between the flash-induced and the sound-induced bouncing effects thus suggest that these two kinds of transient-induced bouncing effects might rely on distinct neural pathways at the perceptual stage of processing. This inference can be further supported by a previous behavioral finding that the temporal window (i.e., the temporal distance between the transient onset and the coincidence of two disks) for the promotion of bouncing percept was narrower for the transient visual flash than for the transient sound (Shimojo et al., 2001).

Following the postcoincidence P2 component, we also found that the amplitude of the centroparietal P3 component was significantly greater on streaming than bouncing trials in V condition. In contrast, the opposite was true in VF condition. That is, the P3 amplitude was significantly larger on bouncing than streaming trials when the coincident flash was presented. The P3 wave has been thought to be an endogenous component reflecting processes involved in post-perceptual evaluation or categorization of task-defined events (Johnson, 1984, 1986; Polich, 2007; Rik et al., 2014). Therefore, the quite different patterns of P3 results between VF and V condition suggests that the additional visual flash presented at the moment of two disks' coincidence influence not only the relatively early perceptual stage of processing (indexed by postcoincidence P2) but also the postperceptual phase (indexed by postcoincidence P3) when facing the streaming/bouncing motion display. It is worth mentioning that although several behavioral studies have taken advantage of signal detection paradigm to characterize the perceptual (sensitivity, indexed by d') and decisional (response criterion, indexed by c) processes in terms of the *sound*-induced bouncing effect (Grassi & Casco, 2012; Grove et al., 2012), the current ERP study was the first, to our knowledge, to investigate the respective contributions of perceptual and decisional processes to the occurrence of *flash*-induced bouncing effect.

As to the specific psychological processes reflected by the present P3 results, in addition to the well-known characteristic that target stimuli with lower probability elicited larger P3 amplitude (for review, see Polich, 2007), previous investigations also reported that the P3 amplitude would be increased if subjects were more certain about their responses (Johnson, 1984, 1986). Consistent with the findings of Johnson (1984, 1986), several ERP studies conducting response-based analysis have found that the amplitude of P3 was much larger when subjects discriminated the targets correctly than when the targets were discriminated incorrectly (e.g., Feng, Störmer, Martinez, McDonald, & Hillyard, 2014; Sergent et al., 2005). Moreover, P3 amplitude elicited by stimuli in the easy task was found to be also greater than that elicited by stimuli in the difficult task (Qin, Xiao, Li, Sha, & Li, 2009). Based on the evidence above, it is reasonable to infer that the amplitude of P3 component would be larger when participants make their decision more confidently. If that is the case, the current findings of larger postcoincidence P3 on streaming than bouncing trials in V condition but the opposite was true in VF condition might indicate that subjects were more certain about their streaming than bouncing responses when no coincident flash was presented but were more certain about their bouncing than streaming responses when the flash was presented. This interpretation fits well with our behavioral results and previous findings that streaming percept was dominant in the baseline streaming/bouncing display but bouncing percept became dominant when the additional flash was presented (Adams & Grove, 2018; Burns & Zanker, 2000; Kawabe & Miura, 2006; Watanabe & Shimojo, 1998, 2001a).

5 | CONCLUSION

The present study used ERP recordings to explore the temporal dynamics of the flash-induced bouncing effect. In summary, we found that the amplitude of postcoincidence parietooccipital P2 component (190–230 ms after coincidence) elicited by the visual motion was

significantly attenuated on bouncing trials (compared with the streaming trials) when the flash was presented at the moment of two disks' coincidence, but this effect did not occur when the flash was absent. Besides, the postcoincidence parietal P3 component (330–430 ms) was found to be larger on bouncing than streaming trials when the flash was presented, but the opposite was true when no flash was presented. These electrophysiological findings indicate that the flash-induced bouncing effect occurs at both perceptual and postperceptual stages of processing, thus providing evidence for both the perceptual hypothesis (Berger & Ehrsson, 2017; Dufour et al., 2008; Remijn et al., 2004; Watanabe & Shimojo, 2001b) and the inference hypothesis (Adams & Grove, 2018; Grove et al., 2012, 2016; Grove & Sakurai, 2009; Zeljko & Grove, 2016) in terms of the flash-induced bouncing effect. Furthermore, the current findings also suggest that the bouncing effect induced by a visual flash and that induced by an auditory tone might rely on distinct neural pathways at the perceptual stage of processing. Therefore, it may be interesting for future studies to investigate these two kinds of transient-induced bouncing effects systematically and characterize their respective psychophysiological mechanisms.

CONFLICT OF INTEREST

The authors declare no potential conflict of interest.

AUTHOR CONTRIBUTIONS

Song Zhao, Hui Zhong, Xinyin Huang, and Wenfeng Feng designed the research; Hui Zhong, Song Zhao, and Wenfeng Feng performed the research; Hui Zhong, Song Zhao, Tingji Chen, and Wenfeng Feng analyzed the data; Song Zhao, Hui Zhong, Xinyin Huang, and Wenfeng Feng wrote the paper.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

ORCID

Song Zhao https://orcid.org/0000-0001-6453-9214
Wenfeng Feng https://orcid.org/0000-0002-7664-5863

REFERENCES

Adams, K. L., & Grove, P. M. (2018). The effect of transient location on the resolution of bistable visual and audiovisual motion sequences. *Perception*, 47(9), 927–942.

Anllo-Vento, L., & Hillyard, S. A. (1996). Selective attention to the color and direction of moving stimuli: Electrophysiological correlates of hierarchical feature selection. *Perception & Psychophysics*, 58(2), 191–206.

Ao, X. Y., Fan, S. L., He, X., & Chen, L. (2000). The asymmetry of the upper and the lower visual-field in spatial selective attention: An high resolution ERP study. *NeuroImage*, 11(5), S40. https://doi.org/10.1016/s1053-8119(00)90974-1

Berger, C. C., & Ehrsson, H. H. (2017). The content of imagined sounds changes visual motion perception in the cross-bounce illusion. *Scientific Reports*, 7, 40123. https://doi.org/10.1038/srep40123

Bertenthal, B. I., Banton, T., & Bradbury, A. (1993). Directional bias in the perception of translating patterns. *Perception*, 22(2), 193–207.

Boylan, C., Clement, R. A., & Harding, G. F. (1984). Lateralization of the flash visual-evoked cortical potential in human albinos. *Investigative Ophthalmology & Visual Science*, 25(12), 1448–1450.

Britz, J., Landis, T., & Michel, C. M. (2009). Right parietal brain activity precedes perceptual alternation of bistable stimuli. *Cerebral Cortex*, 19(1), 55–65.

Burns, N. R., & Zanker, J. M. (2000). Streaming and bouncing: Observations on motion defined objects. *Clinical & Experimental Ophthalmology*, 28(3), 220–222.

Bushara, K. O., Hanakawa, T., Immisch, I., Toma, K., Kansaku, K., & Hallett, M. (2003). Neural correlates of cross-modal binding. *Nature Neuroscience*, 6(2), 190–195.

Di Russo, F., Aprile, T., Spitoni, G., & Spinelli, D. (2008). Impaired visual processing of contralesional stimuli in neglect patients: A visual-evoked potential study. *Brain*, 131(3), 842–854.

Dufour, A., Touzalin, P., Moessinger, M., Brochard, R., & Després, O. (2008). Visual motion disambiguation by a subliminal sound. *Consciousness and Cognition*, 17(3), 790–797.

Dunn, B. R., Dunn, D. A., Languis, M., & Andrews, D. (1998). The relation of ERP components to complex memory processing. *Brain and Cognition*, 36(3), 355–376.

Federmeier, K. D., & Kutas, M. (2002). Picture the difference: Electrophysiological investigations of picture processing in the two cerebral hemispheres. *Neuropsychologia*, 40(7), 730–747.

Feng, W., Störmer, V. S., Martinez, A., McDonald, J. J., & Hillyard, S. A. (2014). Sounds activate visual cortex and improve visual discrimination. *Journal of Neuroscience*, 34(29), 9817–9824.

Freunberger, R., Klimesch, W., Doppelmayr, M., & Höller, Y. (2007). Visual P2 component is related to theta phase-locking. *Neuroscience Letters*, 426(3), 181–186.

Grassi, M., & Casco, C. (2010). Audiovisual bounce-inducing effect: When sound congruence affects grouping in vision. *Attention Perception and Psychophysics*, 72(2), 378–386.

Grassi, M., & Casco, C. (2012). Revealing the origin of the audiovisual bounce-inducing effect. *Seeing and Perceiving*, 25(2), 223–233.

Grove, P. M., Ashton, J., Kawachi, Y., & Sakurai, K. (2012). Auditory transients do not affect visual sensitivity in discriminating between objective streaming and bouncing events. *Journal of Vision*, 12(8), 1–11.

Grove, P. M., Robertson, C., & Harris, L. R. (2016). Disambiguating the stream/bounce illusion with inference. *Multisensory Research*, 29, 453–464.

Grove, P. M., & Sakurai, K. (2009). Auditory induced bounce perception persists as the probability of a motion reversal is reduced. *Perception*, 38(7), 951–965.

Hesselmann, G., Kell, C. A., Eger, E., & Kleinschmidt, A. (2008). Spontaneous local variations in ongoing neural activity bias perceptual decisions. *Proceedings of the National Academy of Sciences of the United States of America*, 105(31), 10984–10989.

Hesselmann, G., Kell, C. A., & Kleinschmidt, A. (2008). Ongoing activity fluctuations in hMT+ bias the perception of coherent visual motion. *Journal of Neuroscience*, 28(53), 14481–14485.

Hipp, J. F., Engel, A. K., & Siegel, M. (2011). Oscillatory synchronization in large-scale cortical networks predicts perception. *Neuron*, 69(2), 387–396.

Johnson, R., Jr. (1984). P300: A model of the variables controlling its amplitude. *Annals of the New York Academy of Sciences*, 425(1), 223–229.

Johnson, R., Jr. (1986). A triarchic model of P300 amplitude. *Psychophysiology*, 23(4), 367–384.

Kawabe, T., & Miura, K. (2006). Effects of the orientation of moving objects on the perception of streaming/bouncing motion displays. *Perception & Psychophysics*, 68(5), 750–758.

Leblanc, E., Prime, D. J., & Jolicoeur, P. (2008). Tracking the location of visuospatial attention in a contingent capture pardigm. *Journal of Cognitive Neuroscience*, *20*(4), 657–671.

Lee, P., Hsieh, J., Wu, C., Shyu, K., Chen, S., Yeh, T., & Wu, Y. (2006). The brain computer interface using flash visual evoked potential and independent component analysis. *Annals of Biomedical Engineering*, *34*(10), 1641–1654.

Liu, Q., Wu, Y., Yang, Q., Campos, J. L., Zhang, Q., & Sun, H. (2009). Neural correlates of size illusions: An event-related potential study. *Neuroreport*, *20*(8), 809–814.

Luck, S. J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology*, *54*(1), 146–157.

Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, *31*(3), 291–308.

Maniglia, M., Grassi, M., Casco, C., & Campana, G. (2012). The origin of the audiovisual bounce-inducing effect: A TMS study. *Neuropsychologia*, *50*(7), 1478–1482.

Martens, S., Munneke, J., Smid, H., & Johnson, A. (2007). Quick minds don't blink: Electrophysiological correlates of individual differences in attentional selction. *Journal of Cognitive Neuroscience*, *18*(9), 1423–1438.

McDonald, J. J., Teder-Salejarvi, W. A., Russo, F. D., & Hillyard, S. A. (2003). Neural substrates of perceptual enhancement by cross-modal spatial attention. *Journal of Cognitive Neuroscience*, *15*(1), 10–19.

Metzger, W. (1934). Beobachtungen über phänomenale identität. *Psychologische Forschung*, *19*(1), 1–60.

Omoto, S., Kuroiwa, Y., Otsuka, S., Baba, Y., Wang, C., Li, M., … Suzuki, Y. (2010). P1 and P2 components of human visual evoked potentials are modulated by depth perception of 3-dimensional images. *Clinical Neurophysiology*, *121*(3), 386–391.

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148.

Qin, J., Xiao, F., Li, F., Sha, W., & Li, H. (2009). The characteristic of extrapolation in numerical inductive inference: An ERP study. *Brain Research*, *1295*, 142–148.

Remijn, G. B., Ito, H., & Nakajima, Y. (2004). Audiovisual integration: An investigation of the 'streaming-bouncing' phenomenon. *Journal of Physiological Anthropology and Applied Human Science*, *23*(6), 243–247.

Rik, V. D., Martijn, A., Jongsma, M. L. A., Kessels, R. P. C., & Francesco, D. R. (2014). P300 development across the lifespan: A systematic review and meta-analysis. *PLoS One*, *9*(2), e87347. https://doi.org/10.1371/journal.pone.0087347

Sekuler, A. B., & Sekuler, R. (1999). Collisions between moving visual targets: What controls alternative ways of seeing an ambiguous display? *Perception*, *28*(4), 415–432.

Sekuler, R., Sekuler, A. B., & Lau, R. (1997). Sound alters visual motion perception. *Nature*, *385*(6614), 308.

Sergent, C., Baillet, S., & Dehaene, S. (2005). Timing of the brain events underlying access to consciousness during the attentional blink. *Nature Neuroscience*, *8*(10), 1391–1400.

Shimojo, S., Scheier, C., Nijhawan, R., Shams, L., Kamitani, Y., & Watanabe, K. (2001). Beyond perceptual modality: Auditory effects on visual perception. *Acoustical Science and Technology*, *22*(2), 61–67.

Van der Burg, E., Talsma, D., Olivers, C. N. L., Hickey, C., & Theeuwes, J. (2011). Early multisensory interactions affect the competition among multiple visual objects. *NeuroImage*, *55*(3), 1208–1218.

Watanabe, K., & Shimojo, S. (1998). Attentional modulation in perception of visual motion events. *Perception*, *27*(9), 1041–1054.

Watanabe, K., & Shimojo, S. (2001a). When sound affects vision: Effects of auditory grouping on visual motion perception. *Psychological Science*, *12*(2), 109–116.

Watanabe, K., & Shimojo, S. (2001b). Postcoincidence trajectory duration affects motion event perception. *Perception & Psychophysics*, *63*(1), 16–28.

Williams, Z. M., Elfar, J. C., Eskandar, E. N., Toth, L. J., & Assad, J. A. (2003). Parietal activity and the perceived direction of ambiguous apparent motion. *Nature Neuroscience*, *6*(6), 616–623.

Wlotko, E. W., & Federmeier, K. D. (2007). Finding the right word: Hemispheric asymmetries in the use of sentence context information. *Neuropsychologia*, *45*(13), 3001–3014.

Zeljko, M., & Grove, P. M. (2016). Sensitivity and bias in the resolution of stream-bounce stimuli. *Perception*, *46*(2), 1–27.

Zhao, S., Wang, Y., Feng, C., & Feng, W. (2020). Multiple phases of cross-sensory interactions associated with the audiovisual bounce-inducing effect. *Biological Psychology*, *149*, 107805. https://doi.org/10.1016/j.biopsycho.2019.107805

Zhao, S., Wang, Y., Jia, L., Feng, C., Liao, Y., & Feng, W. (2017). Precoincidence brain activity predicts the perceptual outcome of the streaming/bouncing motion display. *Scientific Reports*, *7*, 8832. https://doi.org/10.1038/s41598-017-08801-5

Zhao, S., Wang, Y., Xu, H., Feng, C., & Feng, W. (2018). Early cross-modal interactions underlie the audiovisual bounce-inducing effect. *NeuroImage*, *174*, 208–218.

Zvyagintsev, M., Nikolaev, A. R., Sachs, O., & Mathiak, K. (2011). Early attention modulates perceptual interpretation of multisensory stimuli. *Neuroreport*, *22*(12), 586–591.