



Audio–Visual Fission Illusion and Individual Alpha Frequency: Perspective on Buegers and Noppeney (2022)

Yasuki Noguchi

Abstract

■ Integrating visual and auditory information is an important ability in various cognitive processes, although its neural mechanisms remain unclear. Several studies indicated a close relationship between one's temporal binding window (TBW) for audio–visual interaction and their alpha rhythm in the brain (individual alpha frequency or IAF). A recent study by Buegers and Noppeney [Buegers, S., & Noppeney, U. The role of alpha oscillations in temporal binding within and across the senses. *Nature Human Behaviour*, 6, 732–742, 2022], however,

challenged this view using a new approach to analyze behavioral data. Conforming to the same procedures by Buegers and Noppeney, here, I analyzed the data of my previous study and examined a relationship between TBW and IAF. In contrast to Buegers and Noppeney, a significant correlation was found between occipital IAF and a new behavioral measure of TBW. Some possibilities that caused these opposing results, such as a variability of “alpha band” across studies and a large inter-individual difference in magnitude of the fission illusion, are discussed. ■

INTRODUCTION

Multisensory integration is a key function to perceive our environment efficiently (Hirst, McGovern, Setti, Shams, & Newell, 2020; Keil, 2020; Zhou, Cheung, & Chan, 2020). Of particular importance in everyday life is an integration between visual and auditory inputs, which plays a crucial role in speech perception (Conrey & Pisoni, 2006), person recognition (Robertson & Schweinberger, 2010) and emotional processing (Walker-Andrews, 1997), and so forth. It is known that temporal synchrony provides a critical cue for this audio–visual integration (Dixon & Spitz, 1980; Vroomen & Keetels, 2010). To be combined into a coherent percept, visual and auditory stimuli must co-occur within a limited interval that is called the temporal binding window (TBW).

Many studies have explored neural underpinnings of the TBW. In the research of visual perception, alpha oscillations (8–13 Hz) have been proposed to play a critical role in parsing visual inputs into discrete events (VanRullen, 2016). For example, two flashes successively presented were perceived as a single event when they fell into one alpha cycle, whereas those falling into separate cycles were perceived as two different events (Samaha & Postle, 2015). These results support a view that alpha rhythm serves as TBW through which a constant influx of visual signals is segmented into temporal units. Recent studies further provided a direct link between the speed of alpha

rhythms and performance efficiency in visual and visuo-spatial cognitive tasks (Bertaccini et al., 2022; Di Gregorio et al., 2022).

The close relationship between alpha cycle and TBW has been also reported in audio–visual interaction. Using the sound-induced double-flash (fission) illusion (Shams, Kamitani, & Shimojo, 2000), a typical example of an audio–visual integration, Cecere, Rees, and Romei (2015) reported that lower frequency of individual alpha rhythm (individual alpha frequency or IAF) was associated with a higher occurrence rate of the fission illusion (Cecere et al., 2015). Several studies have replicated this finding (Noguchi, 2022; Cooke, Poch, Gillmeister, Costantini, & Romei, 2019; Keil & Senkowski, 2017). These results indicate a link between one's duration of alpha cycle with audio–visual TBW; lower alpha rhythm provided a broader temporal unit for an audio–visual interaction, resulting in higher rate of an illusory flash induced by concurrent beeps. However, Buegers and Noppeney (2022) recently provided the data inconsistent with this view; they found no evidence for a relationship between IAFs and TBWs (Buegers & Noppeney, 2022). Neural correlates of the audio–visual TBW, therefore, is a matter of intense debate.

A hallmark of Buegers and Noppeney (2022) was that they separately investigated the effects of alpha frequency on two behavioral measures of the signal detection theory; sensitivity (d') and bias ($\text{Bias}_{\text{center}}$). Perceptual sensitivity is an index for a temporal resolution (precision) of a sensory system, whereas bias reflects a tendency of reporting one

Kobe University, Japan

particular percept (e.g., “two flash” response in the fission illusion). Buegers and Noppeney (2022) reported that neither the sensitivity nor bias for two-flash discrimination was associated with peak IAFs of 20 observers.

Inspired by their approach, here I re-analyzed the data in Noguchi (2022) and correlated the d' and bias for the fission illusion with IAF. Below are brief descriptions about stimuli, task, and data analyses (see Noguchi [2022] for details).

METHODS

Twenty-nine healthy participants (18–42 years) took part in the study. Data of two participants were discarded because of an excessive noise in EEG waveforms and thus replaced by those of two additional participants. This sample size (29) was determined by a power analysis (Type I error rate: 0.05, statistical power: 0.80, effect size: $r = .5$). An informed consent was obtained from each participant, and all experiments were carried out along with the guidelines and regulations of the ethics committee at Kobe University, Japan.

Structures of one trial was shown in Figure 1A. A combination of zero to two flashes (white circle, diameter: 4.69° , luminance: 55 cd/m^2) with zero to two beep sounds (pure tone, frequency: 4000 Hz, duration: 8 msec) produced eight types of trials; 0F1S (no-flash with one sound), 0F2S, 1F0S, 1F1S, 1F2S, 2F0S, 2F1S, and 2F2S. The flash was presented on a CRT monitor with the duration set at one frame of a refresh rate (60 Hz or 16.7 msec), although actual duration might be shorter because of phosphor decay rate (Ferri, Venskus, Fotia, Cooke, & Romei, 2018; Elze, 2010). A SOA between two flashes and that between two sounds were both set at 50 msec (Figure 1B). Participants reported the perceived number of flashes (zero, one, or two), ignoring all sounds. The eight types of trials were randomly intermixed in an experimental session of 126 trials. The flash was presented in a left visual field (eccentricity: 12.5°) for three experimental sessions but in a right field for the other three.

Behavioral data were analyzed conforming to procedures in Buegers and Noppeney (2022). First, I computed a perceptual sensitivity (d') and a response bias ($\text{Bias}_{\text{center}}$) when participants discriminated one-flash from two-flash trials under no sound (shown in green in Figure 1C). Using the data of 1F0S and 2F0S trials, they were obtained through the equations below:

$$d' = z(\text{HR}) - z(\text{FAR})$$

$$\text{Bias}_{\text{center}} = -0.5(z(\text{HR}) + z(\text{FAR}))$$

where the hit rate (HR) was a proportion of “two flash” responses in 2F0S trials and the false-alarm rate (FAR) was a proportion of “two flash” responses in 1F0S trials.

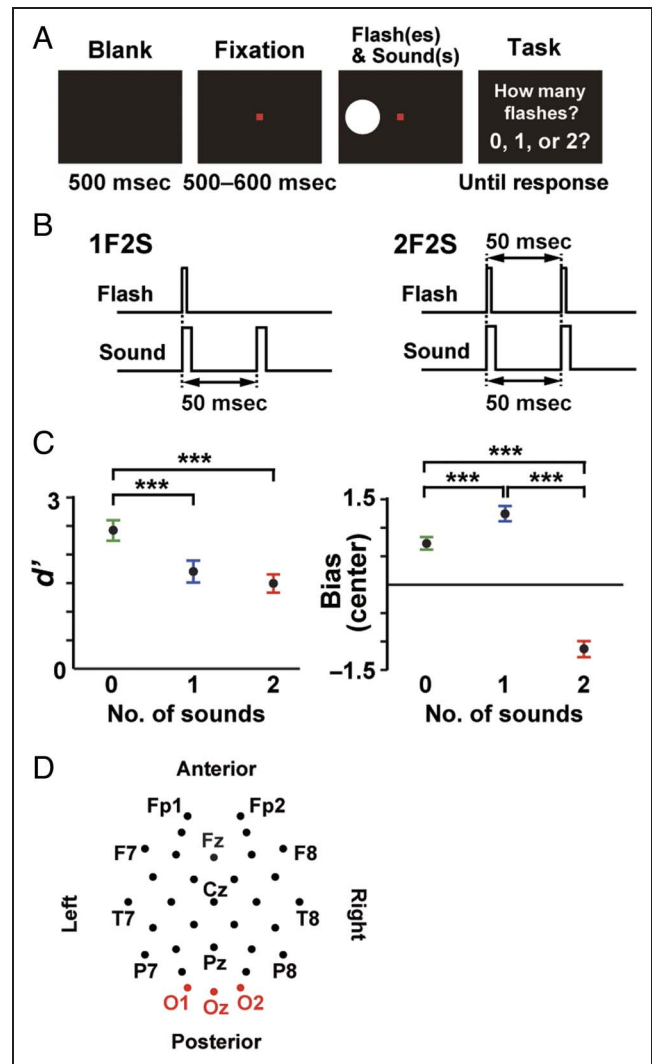


Figure 1. Experimental procedures and behavioral results. (A) Structures of one trial, reprinted from Noguchi (2022) with permission. Participants reported the number of flashes, neglecting all sounds. (B) Sequences of visual and auditory stimuli in the one-flash and two-sounds (1F2S) trial (left) and the 2F2S trial (right). (C) Behavioral results. In each of three contexts (zero sound/one sound/two sounds), perceptual sensitivity (d') and response bias ($\text{Bias}_{\text{center}}$) were computed and shown in left and right panels, respectively. Larger d' indicate a higher temporal resolution of visual system to discriminate one flash versus two flashes. Larger $\text{Bias}_{\text{center}}$ indicates that participants were more likely to report a one-flash than two-flash percept. (D) Two-dimensional layout of 32 EEG sensors. *** $p < .001$.

Likewise, another set of d' and $\text{Bias}_{\text{center}}$ to discriminate one flash versus two flashes under one sound was computed using the HR in 2F1S trials and FAR in 1F1S trials (blue in Figure 1C). A final set of d' and $\text{Bias}_{\text{center}}$ in the two-sound context was obtained from the HR in 2F2S and FAR in 1F2S trials (red in Figure 1C). The audio–visual fission illusion would be observed as a reduced d' in the two-sound context. Furthermore, if the fission illusion also involved a change in bias, this would be seen as a reduced $\text{Bias}_{\text{center}}$ in the two-sound context,

because $\text{Bias}_{\text{center}} < 0$ indicates that participants were more likely to report a two-flash than one-flash percept regardless of their perceptual sensitivities.

EEG signals were measured from 32 points over the scalp (Biosemi ActiveTwo system, sampling rate: 2048 Hz, analog low-pass filter: 417 Hz). Sensors of interest were set at O1, O2, and Oz (shown in red in Figure 1D) in light of a previous literature on alpha frequency and the fission illusion (Buegers & Noppeney, 2022; Cooke et al., 2019; Keil & Senkowski, 2017; Cecere et al., 2015). Using the Brainstorm toolbox for MATLAB (Tadel, Baillet, Mosher, Pantazis, & Leahy, 2011), EEG data of each trial (from -1000 to 0 msec relative to the first flash or sound) were converted into power spectral density (PSD). I confirmed by visual inspection that EEG waveforms in this period did not contain large noise after an application of digital band-pass filter (0.5 – 200 Hz, Butterworth). The conversion to PSD was performed by the Welch's method

(Hamming window, length: 1000 msec). No zero-padding technique was used. A peak frequency in alpha band (8 – 13 Hz) was then identified by searching a component frequency with a maximum power on the PSD. A trait alpha frequency for each participant was computed as an average of the peak frequencies across all 756 trials. For comparisons, peak frequencies of other bands (delta: 1 – 4 Hz, theta: 4 – 8 Hz, beta: 13 – 30 Hz, gamma: 31 – 58 Hz, and high gamma: 62 – 100 Hz) were also identified and correlated with d' (Figure 2A) and $\text{Bias}_{\text{center}}$ (Figure 2B).

In addition to the main analysis described above, I performed two analyses to test the robustness of the present data. First, IAF was computed using the data only in the two-sound trials (2F2S and 1F2S), not all 756 trials. This would match the data set for computing behavioral measure (d' or $\text{Bias}_{\text{center}}$) with that for EEG measure (IAF), enabling a fair comparison of those two (Figure 3A). Second, a time window of EEG analysis was set at -400 to 0 msec (Figure 3B), not -1000 to 0 msec. As shown in Figure 1A, a prestimulus period consisted of a blank screen of 500 msec followed by a fixation screen of 500–600 msec. Setting a time window at -400 to 0 msec might provide a better EEG measure that was not contaminated by visual response to an onset of the fixation point. A use of shorter time window, however, reduces frequency resolution of PSD. I resolved this issue by pooling EEG waveforms of all trials into a single (long) waveform. The Welch's method was applied to this long (concatenated) waveform to get a PSD for each participant.

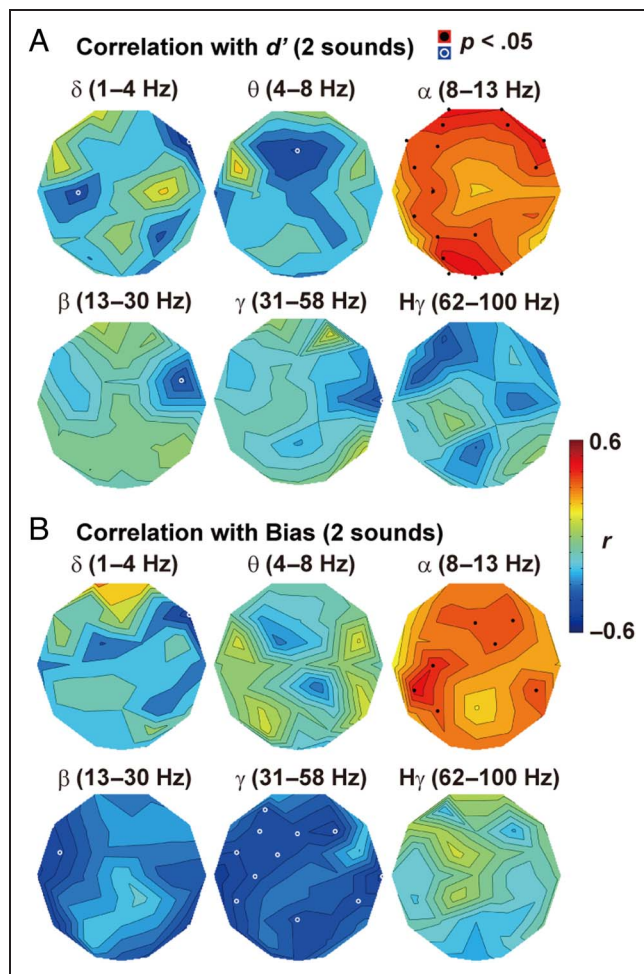


Figure 2. Correlation of behavioral and EEG measures. (A) r -maps between individual peak frequency and d' in the two-sound context (fission illusion). EEG sensors with significantly positive correlations ($p < .05$) are shown by black dots, whereas those with negative correlation are shown by white dots. (B) r -maps between individual peak frequency and $\text{Bias}_{\text{center}}$ in the two-sound context. Peak alpha frequencies at occipital sensors were significantly correlated with d' (A) but not with $\text{Bias}_{\text{center}}$ (B).

RESULTS

Perceptual sensitivity to discriminate one flash versus two flashes (d') is displayed in the left in Figure 1C. A one-way ANOVA over the three auditory contexts (zero sound/one sound/two sounds) yielded a significant main effect, $F(2, 56) = 13.94, p < .001, \eta^2 = .33$. Post hoc tests with the Bonferroni correction indicated significant differences between zero sound versus one sound ($p < .001$) and between zero sound versus two sounds ($p < .001$). The one-way ANOVA with the Greenhouse–Geisser adjustment also yielded a significant main effect for the $\text{Bias}_{\text{center}}$ (right panel of Figure 1C, $F(1.54, 43.21) = 145.56, p < .001, \eta^2 = .84$) with post hoc differences of all three pairs significant ($p < .001$). These data replicated Buegers and Noppeney (2022) at least for the fission illusion (two sounds).

Figure 2A shows r -maps between behavioral and EEG data. A correlation coefficient between the d' (two sounds) and peak frequency was calculated for each sensor position and color-coded over a 2-D layout of the 32 sensors (Figure 1D). Positive correlations with d' (shown in red) were selectively observed in peak alpha frequency (upper right). Correlations at occipital sensors (Oz: $r = .49$, O1: $r = .38$, and O2: $r = .40$) were all significant ($p < .05$, shown by black dots), indicating that individuals with higher alpha frequency showed larger d' (less fission

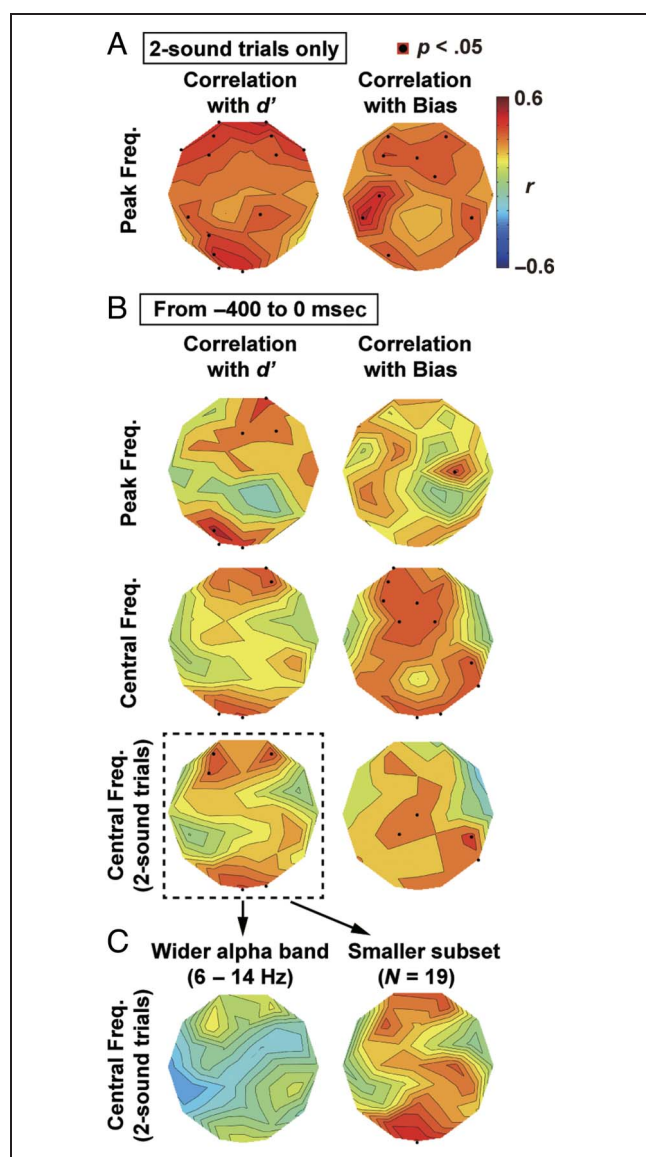


Figure 3. The r -maps of additional analyses. (A) Correlations between IAF and behavioral measures (d' and $\text{Bias}_{\text{center}}$). Although IAFs in Figure 2 were computed using the data of all (756) trials, those in this panel were obtained from 216 trials in the two-sound context (2F2S and 1F2S). (B) r -maps between behavioral measures and IAF at -400 to 0 msec (not at -1000 to 0 msec). Data of individual peak frequency, central frequency, and the central frequency of the two-sound trials are shown in top, middle, and bottom panels, respectively. (C) An effect of bandwidth and a screening of participants. When the central frequency was computed with a broader alpha band of 6 – 14 Hz, as in Buegers and Noppeney (2022), this caused a decrease in correlation coefficient (left). In contrast, correlations over the occipital cortex were kept significant when the data of participants who showed low task accuracy or no fission/fusion illusion were excluded from the analysis (right).

illusion). In contrast, the alpha frequencies over the occipital cortex showed no significant correlation with $\text{Bias}_{\text{center}}$ (Figure 2B, Oz: $r = .34$, O1: $r = .32$, and O2: $r = .34$), although strong correlations were seen over the non-occipital (temporal and frontal) regions (e.g., CP5: $r = .49$).

Results of the two additional analyses (see Methods section) are shown in Figure 3. I displayed in Figure 3A

the r -maps in which IAF was computed from the data of the two-sound trials (2F2S and 1F2S). Similar to Figure 2, significant correlations with d' , but not $\text{Bias}_{\text{center}}$, were observed at occipital sensors. Figure 3B shows the r -maps in which IAF was identified with a prestimulus data of -400 to 0 msec. The r -maps in the top indicate correlations between peak alpha frequency (pooled across all 756 trials) and behavioral measures, whereas those in middle indicate the correlation between central alpha frequency and behavioral measures. The central frequency was defined as a mean of component frequencies at alpha range (8 – 13 Hz) weighted by their powers. The r -maps of the central frequency on the two-sound trials are shown in the bottom of Figure 3B. The significant correlations between IAF with d' over the occipital cortex were consistently seen (Oz: $r = .41 \sim 0.50$) in all analyses.

DISCUSSION

In contrast to Buegers and Noppeney (2022), the present data support a close relationship between occipital IAF and one's temporal resolution of audio-visual integration (d'). How can we resolve this inconsistency? I provide below three possibilities that might reconcile the opposing results.

First, whereas Buegers and Noppeney (2022) searched for an IAF within a band of 6 – 14 Hz, I identified it using a narrower band of 8 – 13 Hz. Figure 2A suggest that a correlation between d' and peak frequency is highly selective to alpha band; no positive correlation with d' was observed in the r -map of individual theta (4 – 8 Hz) or beta (13 – 30 Hz) frequency. Although a definition of alpha band has been variable across studies, a use of a narrower band might be suitable to detect a relationship of IAF with the audiovisual TBW. Indeed, I observed a reduction in correlation coefficient at Oz (from $.42$ to $.15$) when an IAF was searched within the broader band of 6 – 14 Hz (Figure 3C, left).

Second, a detection of significant correlation might be prevented by the bias or a large intersubject variability in magnitudes of the fission illusion. As Buegers and Noppeney (2022) recognized, their data of individual TBWs measured by a classical method (psychometric thresholds) were susceptible to the bias and an intersubject variability. As a result, a scatterplot in their Yes–No SOA task (Figure 4b in Buegers & Noppeney, 2022) contained at least 6 outliers (out of 19) whose psychometric thresholds were unmeasurable. Excluding these outliers might reveal a negative correlation between IAFs and behavioral TBWs. Indeed, Keil and Senkowski (2017) showed a relationship between IAFs and rates of the fission illusion by excluding the data of 14 participants (out of 40) whose illusion rates were too high ($> 90\%$) or too low ($< 10\%$). The present data also suggest an effectiveness of the data screening. In the right of Figure 3C, I discarded the data of 10 participants who showed low task accuracy or no fission/fusion illusion (see Noguchi [2022], for precise criteria). The r -map of the remaining

19 participants indicated a positive correlation between d' and central IAF over a wide region in the occipital cortex (Oz: $r = .51$).

Finally, it is preferable to correlate behavioral and EEG data simultaneously recorded. In Buegers and Noppeney (2022), their two-interval forced-choice task provided an excellent behavioral measure that can reduce an effect of the bias. This measure, however, was correlated with EEG data in a resting (eyes-closed) session or different task sessions. Keil and Senkowski (2017) reported that, although the illusion rates were significantly correlated with IAFs simultaneously obtained, no correlation was seen when those rates were compared with resting-state IAFs. A combination of bias-free behavioral measure with simultaneous EEG data would benefit future studies to examine a relationship between the fission illusion and IAF.

On the other hand, there are some points to be improved in the present data. A major problem would be that I fixed the SOA between two flashes/sounds at 50 msec. Whereas Buegers and Noppeney (2022) tested a variety of SOAs ranging from 25 to 225 msec, only one SOA (50 msec) was implemented in the present study. Such an approach can seriously suffer from the problem of inter-individual variability described above. Setting appropriate SOAs that can cover the individual differences, hopefully through a preliminary thresholding procedure, would resolve this issue and provide better behavioral measures for the fission illusion.

Acknowledgments

I thank Nahomi Sato and Taeko Kaneda for their technical support.

Corresponding author: Yasuki Noguchi, Department of Psychology, Kobe University, 1-1 Rokkodai-cho, Nada, Kobe, 657-8501, Japan, or via e-mail: ynoguchi@lit.kobe-u.ac.jp.

Data Availability Statement

All data supporting the findings of this study are available from Yasuki Noguchi upon reasonable request.

Author Contributions

Yasuki Noguchi: Conceptualization; Funding acquisition; Writing—Original draft; Writing—Review & editing.

Funding Information

Yasuki Noguchi, Japan Society for the Promotion of Science (<https://dx.doi.org/10.13039/501100001691>), grant number: 19H04430.

Diversity in Citation Practices

Retrospective analysis of the citations in every article published in this journal from 2010 to 2021 reveals a persistent

pattern of gender imbalance: Although the proportions of authorship teams (categorized by estimated gender identification of first author/last author) publishing in the *Journal of Cognitive Neuroscience (JoCN)* during this period were M(an)/M = .407, W(oman)/M = .32, M/W = .115, and W/W = .159, the comparable proportions for the articles that these authorship teams cited were M/M = .549, W/M = .257, M/W = .109, and W/W = .085 (Postle and Fulvio, *JoCN*, 34:1, pp. 1–3). Consequently, *JoCN* encourages all authors to consider gender balance explicitly when selecting which articles to cite and gives them the opportunity to report their article's gender citation balance.

REFERENCES

- Bertaccini, R., Ellena, G., Macedo-Pascual, J., Carusi, F., Trajkovic, J., Poch, C., et al. (2022). Parietal alpha oscillatory peak frequency mediates the effect of practice on visuospatial working memory performance. *Vision*, 6, 30. <https://doi.org/10.3390/vision6020030>, PubMed: 35737417
- Buegers, S., & Noppeney, U. (2022). The role of alpha oscillations in temporal binding within and across the senses. *Nature Human Behaviour*, 6, 732–742. <https://doi.org/10.1038/s41562-022-01294-x>, PubMed: 35210592
- Cecere, R., Rees, G., & Romei, V. (2015). Individual differences in alpha frequency drive crossmodal illusory perception. *Current Biology*, 25, 231–235. <https://doi.org/10.1016/j.cub.2014.11.034>, PubMed: 25544613
- Conrey, B., & Pisoni, D. B. (2006). Auditory-visual speech perception and synchrony detection for speech and nonspeech signals. *Journal of the Acoustical Society of America*, 119, 4065–4073. <https://doi.org/10.1121/1.2195091>, PubMed: 16838548
- Cooke, J., Poch, C., Gillmeister, H., Costantini, M., & Romei, V. (2019). Oscillatory properties of functional connections between sensory areas mediate cross-modal illusory perception. *Journal of Neuroscience*, 39, 5711–5718. <https://doi.org/10.1523/JNEUROSCI.3184-18.2019>, PubMed: 31109964
- Di Gregorio, F., Trajkovic, J., Roperti, C., Marcantoni, E., Di Luzio, P., Avenanti, A., et al. (2022). Tuning alpha rhythms to shape conscious visual perception. *Current Biology*, 32, 988–998. <https://doi.org/10.1016/j.cub.2022.01.003>, PubMed: 35090592
- Dixon, N. F., & Spitz, L. (1980). The detection of auditory visual desynchrony. *Perception*, 9, 719–721. <https://doi.org/10.1068/p090719>, PubMed: 7220244
- Elze, T. (2010). Achieving precise display timing in visual neuroscience experiments. *Journal of Neuroscience Methods*, 191, 171–179. <https://doi.org/10.1016/j.jneumeth.2010.06.018>, PubMed: 20600318
- Ferri, F., Venskus, A., Fotia, F., Cooke, J., & Romei, V. (2018). Higher proneness to multisensory illusions is driven by reduced temporal sensitivity in people with high schizotypal traits. *Consciousness and Cognition*, 65, 263–270. <https://doi.org/10.1016/j.concog.2018.09.006>, PubMed: 30243043
- Hirst, R. J., McGovern, D. P., Setti, A., Shams, L., & Newell, F. N. (2020). What you see is what you hear: Twenty years of research using the sound-induced flash illusion. *Neuroscience & Biobehavioral Reviews*, 118, 759–774. <https://doi.org/10.1016/j.neubiorev.2020.09.006>, PubMed: 32937116
- Keil, J. (2020). Double flash illusions: Current findings and future directions. *Frontiers in Neuroscience*, 14, 298. <https://doi.org/10.3389/fnins.2020.00298>, PubMed: 32317920
- Keil, J., & Senkowski, D. (2017). Individual alpha frequency relates to the sound-induced flash illusion. *Multisensory*

- Research*, 30, 565–578. <https://doi.org/10.1163/22134808-00002572>, PubMed: 31287091
- Noguchi, Y. (2022). Individual differences in beta frequency correlate with the audio–visual fusion illusion. *Psychophysiology*, 59, e14041. <https://doi.org/10.1111/psyp.14041>, PubMed: 35274314
- Robertson, D. M. C., & Schweinberger, S. R. (2010). The role of audiovisual asynchrony in person recognition. *Quarterly Journal of Experimental Psychology*, 63, 23–30. <https://doi.org/10.1080/17470210903144376>, PubMed: 19672796
- Samaha, J., & Postle, B. R. (2015). The speed of alpha-band oscillations predicts the temporal resolution of visual perception. *Current Biology*, 25, 2985–2990. <https://doi.org/10.1016/j.cub.2015.10.007>, PubMed: 26526370
- Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions: What you see is what you hear. *Nature*, 408, 788. <https://doi.org/10.1038/35048669>, PubMed: 11130706
- Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., & Leahy, R. M. (2011). Brainstorm: A user-friendly application for MEG/EEG analysis. *Computational Intelligence and Neuroscience*, 2011, 879716. <https://doi.org/10.1155/2011/879716>, PubMed: 21584256
- VanRullen, R. (2016). Perceptual cycles. *Trends in Cognitive Sciences*, 20, 723–735. <https://doi.org/10.1016/j.tics.2016.07.006>, PubMed: 27567317
- Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony: A tutorial review. *Attention, Perception & Psychophysics*, 72, 871–884. <https://doi.org/10.3758/APP.72.4.871>, PubMed: 20436185
- Walker-Andrews, A. S. (1997). Infants' perception of expressive behaviors: Differentiation of multimodal information. *Psychological Bulletin*, 121, 437–456. <https://doi.org/10.1037/0033-2909.121.3.437>, PubMed: 9136644
- Zhou, H.-Y., Cheung, E. F. C., & Chan, R. C. K. (2020). Audiovisual temporal integration: Cognitive processing, neural mechanisms, developmental trajectory and potential interventions. *Neuropsychologia*, 140, 107396. <https://doi.org/10.1016/j.neuropsychologia.2020.107396>, PubMed: 32087206