Distinct metacognitive and decisional process for fission and fusion illusion

Rao Xie*
Donders Centre for Neuroscience (DCN), Radboud University
Heyendaalseweg 135, Nijmegen, The Netherlands
rao.xie@donders.ru.nl
Steffen Buergers*
Royal HaskoningDHV
Centre for Computational Neuroscience and Cognitive Robotics (CNCR)
B15 2TT, Birmingham, UK
steffen.burgers@rhdhv.com
Uta Noppeney
Donders Centre for Neuroscience (DCN), Radboud University
Heyendaalseweg 135, Nijmegen, The Netherlands
Centre for Computational Neuroscience and Cognitive Robotics (CNCR)
B15 2TT, Birmingham, UK
uta.noppeney@donders.ru.nl

Abstract

Combining sensory inputs into a single unified perceptual decision poses an important cognitive challenge in everyday life. In the fission and fusion illusion, the number of perceived flashes is biased towards the number of concurrently presented beeps. While both illusions have been explained by Bayesian principles of reliability weighted integration, neuroimaging research has pointed towards divergent underlying processes. In this electroencephalography (EEG) study, observers were presented with 1 or 2 flashes together with 0, 1 or 2 beeps. On each trial, they performed a flash discrimination task followed by a confidence rating. Behaviourally, we observed that metacognitive efficiency in the 2-beep condition was significantly higher than in the 0-beep and 1-beep conditions. Multivariate EEG decoding locked to observers' decisional responses showed that the neural representations generalized between the 0 and 1-beep conditions, yet were distinct from the fission illusion. Our findings show that despite being governed by shared computational principles the fission and fusion illusions are distinct in terms of their neural decisional processes and their metacognitive outcomes.

Keywords: sound induced flash illusion; confidence; metacognitive efficiency; EEG decoding

Introduction

To form a coherent percept human observers integrate sensory signals weighted according to their relative reliabilities consistent with Bayesian principles. When sensory signals are brought into conflict, reliability weighted integration often results in crossmodal biases as exemplified in the sound induced flash illusion (SIFI) (Shams, Kamitani, & Shimojo, 2000). Because of the greater temporal precision of the auditory sense, observers tend to perceive two flashes when a single flash is paired with two sounds (fission illusion), but one flash when two flashes are paired with one sound (fusion illusion). Despite sharing the same computational principles, human neuroimaging research has revealed that they rely on partly distinct neural processes. Event related potential (ERP) analyses have demonstrated that the fission and fusion illusion differ in their early audiovisual interactions (Mishra, Martinez, Sejnowski, & Hillyard, 2007). fMRI studies revealed enhanced blood oxygen level dependent (BOLD) activity in retinotopically defined V1 and V2 for an illusory second flash similar to a real second flash, and attenuated activity for a fused single flash similar to a real single flash (Watkins, Shams, Josephs, & Rees, 2007). One key difference between the fission and fusion illusion is that only the fission illusion invokes a percept of an additional event that is not present. By contrast, the fusion illusion resembles observers' failure to detect a second flash in purely visual one versus two flash discrimination paradigms. Therefore, the fission and fusion illusion may differ not only in early interactions, but particularly in late neural activity leading up to observers' perceptual decision and associated metacognitive processes. To elucidate the decisional and metacognitive processes in the sound induced flash illusions, we presented observers with one or two flashes paired with 0, 1, or 2 beeps. Observers performed a flash discrimination task together with a confidence rating. Behaviourally, we assessed observers' metacognitive efficiency when monitoring their flash discrimination performance in different beep conditions. Using multivariate pattern classification and generalization across auditory conditions on observers' response-locked ERPs, we investigated whether late decisional processes rely on shared or distinct neural representations across the sound conditions. We hypothesized that the neural and metacognitive processes are shared across the fusion illusion and unisensory visual flash discrimination, but distinct from those in the fission illusion where observers perceive an additional illusory flash.

Methods

Experimental procedure

In a two-flash discrimination task (Fig.1A), 20 human observers were presented with 1 or 2 flashes together with 0, 1 or 2 sounds. On each trial, participants reported their perceived number of flashes and confidence level (4-point Likert scale). To obtain approximately equal probability of 'one flash' and 'two flash' responses to identical stimuli, we adjusted the asynchrony between the two flashes and/or sounds within a trial separately for the '2 flash & 0 sound' (2F0B), '2 flash & 1 sound' (2F1B) and '1 flash & 2 sound' (1F2B) conditions in adaptive staircases individually for each participant.

EEG recording and analysis

EEG signals were recorded from 64 channels using Ag/AgCl active electrodes arranged in an extended international 10–20 layout (ActiCap, Brain Products GmbH) at a sampling rate of 1,000 Hz, referenced at FCz with a high-pass filter of 0.1 Hz. EEG data is resampled to 64 Hz for the decoding analysis.

In MVPA, lasso regression classifiers (regularization = 1 or 200) including all electrodes were trained to discriminate "see 2 flashes" from "see 1 flash" perceptual outcomes for 2F0B, 2F1B and 1F2B conditions separately at each time point. To investigate the neural dynamics, MVPA was performed both on stimulus-locked (0ms-600ms) and response-locked (800ms time window leading up to the response) EEG data. Specifically, decoding within and cross condition allows a direct representation comparison among auditory conditions.

The current analysis is based on the same study published by (Buergers & Noppeney, 2022).

Results

Metacognitive efficiency in 2-beep condition is higher than 0- and 1- beep condition

As expected from the staircase procedure, in the critical conditions (2F0B, 2F1B and 1F2B), participants responded "see2" approximately 50% of the time and there was no significant d' condition effect ($F = 2.04, p = 0.14$), suggesting that titration

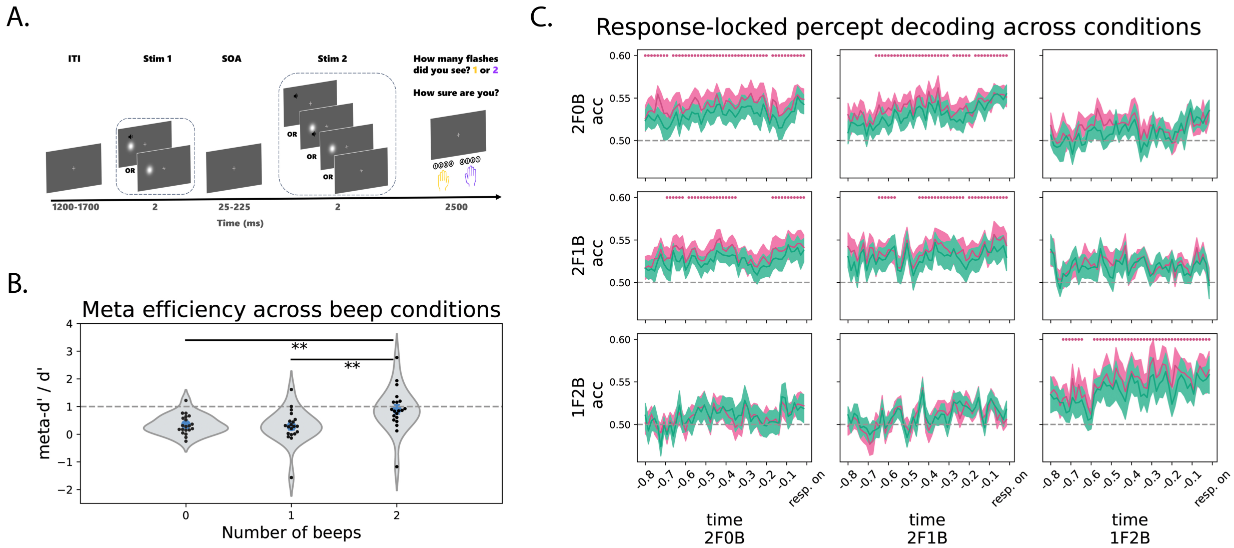

Figure 1: (A) Example trial. Observers were presented with either one or two flashes (together with 0ounds). They reported whether they perceived one or two flashes (by hand used) and how confident they are (by finger used). SOA is titrated for each subject individually through a staircase procedure. (B) Metacognitive efficiency $(meta - d'/d')$. Observers showed significantly higher (and close to optimal) metacognitive efficiency in the 2-beep condition compared to both 0-beep and 1-beep condition. Blue dot denotes mean in group level and error bars denote ± 1 SE, $p < 0.01(**)$ (C) Percept decoding ('see 1 flash vs. see 2 flashes) accuracy in response-locked data $(N = 20)$ $(mean + SE)$. Pink and green color represents regularization parameter 1 and 200 respectively. '2F0B': 2 flashes & 0 beeps condition, '2F1B': 2 flashes & 1 beep condition, '1F2B': 1 flash & 2 beeps condition; 'resp. on': response onset; 'acc': accuracy.

of performance accuracy lead to approximately equal sensitivity between auditory contexts.

Further, to estimate participants' metacognitive ability of assessing perceptual evidence when making confidence judgement, metacognitive efficiency was computed ($meta - d'/d'$ (Maniscalco & Lau, 2012)) separately for each auditory condition (i.e., 0-beep, 1-beep, 2-beep). A significant difference in metacognitive efficiency was observed ($F = 6.25, p = 0.003$). Specifically, the metacognitive efficiency in the 2-beep condition was higher than both the 0-beep and 1- beep conditions (Fig.1B; 0-beep – 1-beep: $t(19) = 0.39, p = 0.698$; 0-beep – 2-beep: $t(19) = -2.90, p = 0.006$; 1-beep – 2-beep: $t(19) = -2.78, p = 0.008$). Particularly, in the 2-beep condition, metacognitive efficiency approached an average value of 1, signifying near-perfect access to perceptual evidence.

Representations generalize between 0- and 1-beep but not 2-beep condition

Response-locked percept decoding (Fig.1C) showed sustained above-chance decoding accuracy throughout this window in all conditions (non-parametric permutation test, diagonals in Fig.1C; 0-beep: -800 - -690ms, -659 - -188ms, -157 - 0ms: $p < 0.05$; 1-beep: -643 - -565ms, -439 - -220ms, -188 - 0ms: $p < 0.05$; 2-beep: -737 - -643ms, -580 – 0ms: $p < 0.05$). Importantly, cross-condition analysis revealed significant generalization between 0- and 1-beep conditions but not to 2-beep condition or vice versa (off-diagonals in Fig.1C; 0-beep – 1-beep: -659 - -298ms, -267 - -188ms, -157 – 0ms: $p < 0.05$, 1-beep – 0-beep: -690 - -612ms, -580 - -345ms, -157 – 0ms). In addition, a very similar confidence decoding pattern among conditions was observed, where confidence level was significantly decoded within each auditory condition and generalized only between 0- and 1- beep but not to 2-beep context.

Discussion

This study unveils distinct decisional processes of fission and fusion illusions assessed by metacognitive efficiency and neural representation. In the decisional window, percept and confidence in the fusion condition exhibited shared EEG features with the unisensory condition. Such similarities were absent in the illusory flash. Furthermore, the fission condition exhibited markedly enhanced metacognitive efficiency, nearing optimal level when compared to other conditions. This shows that observers utilize the information available for making perceptual decision to make confidence judgement when encountering an invoked event. Recent investigation into perceptual reality monitoring emphasizes the important role that metacognition may play in distinguishing between perceived and imagined sources of information (Dijkstra, Kok, & Fleming, 2022). Schizophrenia and related psychosis have implicated aberrant metacognitive activities as a core of dysfunction (Griffin & Fletcher, 2017). Thus, developing a computational account of illusions may have important implications for studying hallucinatory experiences in psychotic disorders.

References

Buergers, S., & Noppeney, U. (2022). The role of alpha oscillations in temporal binding within and across the senses. *Nature human behaviour*, *6*(5), 732–742.

Dijkstra, N., Kok, P., & Fleming, S. M. (2022). Perceptual reality monitoring: Neural mechanisms dissociating imagination from reality. *Neuroscience & Biobehavioral Reviews*, *135*, 104557.

Griffin, J. D., & Fletcher, P. C. (2017). Predictive processing, source monitoring, and psychosis. *Annual review of clinical psychology*, *13*, 265–289.

Maniscalco, B., & Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and cognition*, *21*(1), 422–430.

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. *Journal of Neuroscience*, *27*(15), 4120–4131.

Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature*, *408*(6814), 788–788.

Watkins, S., Shams, L., Josephs, O., & Rees, G. (2007). Activity in human v1 follows multisensory perception. *Neuroimage*, *37*(2), 572–578.