Attention periodically samples competing stimuli during binocular rivalry

Matthew James Davidson[1,2]*, David Alais[3], Naotsugu Tsuchiya[1,2†], Jeroen J.A. van Boxtel[1,2†]

[1] School of Psychological Sciences, Faculty of Medicine, Nursing, and Health Sciences, Monash University.

[2]Monash Institute of Cognitive and Clinical Neurosciences, Monash University.

[3] School of Psychology, University of Sydney.

* corresponding author

† equal contribution

Abstract

The attentional sampling hypothesis suggests attention periodically enhances sensory processing. We investigated whether periodic sampling also acts as a gateway for attended stimuli competing for perceptual dominance. Here we report that during binocular rivalry crossmodal cues directed towards perceptually dominant stimuli resulted in an ~8 Hz attentional sampling rate, consistent with undivided attention. Conversely, crossmodal cues directed towards perceptually suppressed visual images resulted in an ~3.5 Hz attentional sampling rate, indicating that attentional sampling can be divided away from a conscious visual image. At these attentional sampling frequencies, the strength of inter-trial phase coherence over lateral-occipital and fronto-temporal cortex correlated with behavioral measures of changes in conscious perception, suggesting this mechanism acted to bind crossmodal stimuli into a single perceptual object. When crossmodal cues were not task-relevant, these effects disappeared, confirming that changes in conscious perception were dependent upon the allocation of attention, which flexibly samples stimuli in a task-dependent manner.

Main

In a difficult visual search task, perceptual performance does not monotonically increase over time, but fluctuates (Fiebelkorn, Saalmann, & Kastner, 2013; Landau & Fries, 2012). These fluctuations occur approximately 7-8 times per second (Dugué, McLelland, Lajous, & VanRullen, 2015; Fiebelkorn et al., 2013; VanRullen, Carlson, & Cavanagh, 2007), and have recently given rise to the attentional sampling hypothesis, wherein rhythmic enhancements to sensory processing are the result of covert shifts in the spatiotemporal allocation of attention (Dugué et al., 2015; Dugué, Xue, & Carrasco, 2017; Dugué & VanRullen, 2014; Tamber-Rosenau & Marois, 2016). The possibility of a periodic cortical mechanism acting to prioritize incoming sensory stimulation has received growing support (Herbst & Landau, 2016; Rufin VanRullen, 2016a, 2016b; Zoefel & VanRullen, 2017), particularly from psychophysical and electrophysiological investigations involving the manipulation of attention and detection of near-threshold stimuli (VanRullen, 2016b).

In their original investigation, Landau and Fries (2012; adapted since in Chen, Wang, Wang, Tang, & Zhang, 2017; Hogendoorn, 2016) observed that following a cue to reorient attention, accurate probe detection within left or right visual hemifields oscillated in a counterphase, 4 Hz rhythm. Equivalent fluctuations in perceptual performance have been observed when attention is divided between two objects or locations, or when visual cues to reorient attention are only partially informative of target location - requiring a subsequent shift in attention (Fiebelkorn et al., 2013; Huang, Chen, & Luo, 2015; Landau & Fries, 2012; Song, Meng, Chen, Zhou, & Luo, 2014). Similar periodicities also emerge when comparing temporal modulations in detection accuracy for difficult compared to easy visual search tasks (Chen et al., 2017; Dugué & VanRullen, 2014), implying that attention may covertly sample external stimuli in a task-dependant manner. In support of this notion, Landau, Schreyer, Van Pelt, and Fries (2015) investigated lateralized gamma-

band amplitude as a proxy for attention, and observed that pre-target 4 Hz phase modulations in gamma could predict subsequent hits and misses for targets presented in the two attended hemifields – suggesting that an intrinsic 8 Hz attentional rhythm had been sequentially distributed across two locations in their paradigm. Here, we investigated whether a task-dependent attentional sampling mechanism persists when monitoring separate sensory modalities, or two competing, mutually exclusive perceptual outcomes.

To investigate these issues, we adapted a binocular rivalry paradigm in which incompatible images are presented to each eye, resulting in stochastic perceptual alternations between each image (Alais, 2012; Alais & Blake, 2005; Maier, Panagiotaropoulos, Tsuchiya, & Keliris, 2012). We then induced or delayed perceptual transitions using crossmodal auditory and tactile stimulation that was either congruent or incongruent with the perceived visual image (Lunghi, Morrone, & Alais, 2014) and found distinct sampling signatures for each case in behaviour and the electroencephalogram (EEG). These distinct signatures for crossmodal stimulation that was either congruent or incongruent with the currently dominant visual stimulus show that the locus of attentional sampling can flexibly orient away from a consciously perceived image during competition for perceptual dominance.

Results

Attending to congruent low-frequency crossmodal stimulation promotes perceptual dominance and decreases suppression during binocular rivalry

We manipulated the conscious visibility of images across two sessions of 24 x 3-minute binocular rivalry trials. Subjects (*N*=34) reported the content of their visual consciousness via button press while neural activity was simultaneously recorded via 64-channel EEG (Figure 1a, Methods). Rivalry stimuli were orthogonal sinusoidal gratings which underwent continuous contrast modulation, one at 4.5 Hz and the other at 20 Hz. In each 3-minute trial, we intermittently presented 12 crossmodal cues (mean duration 2.6 s) which were amplitude-modulated signals presented in the auditory and/or tactile modality at a frequency congruent with one of the visual stimuli (4.5 or 20 Hz). For one of their two sessions (day 1 for *n*=18), subjects were instructed to focus attention on reporting their visual percept alone – ignoring any crossmodal cues. For their alternate session, subjects were instructed to report their visual percept as well as attend to crossmodal cues, by counting the number of times that the temporal frequency of crossmodal cues coincided with their conscious visual percept at crossmodal cue offset (see Methods). This paradigm allowed us to contrast the effects of attending to low- or high-frequency crossmodal cues, which could be matched or mismatched to the visually dominant image, with mismatched events indicating a match to the suppressed visual stimulus (Figure 1b).

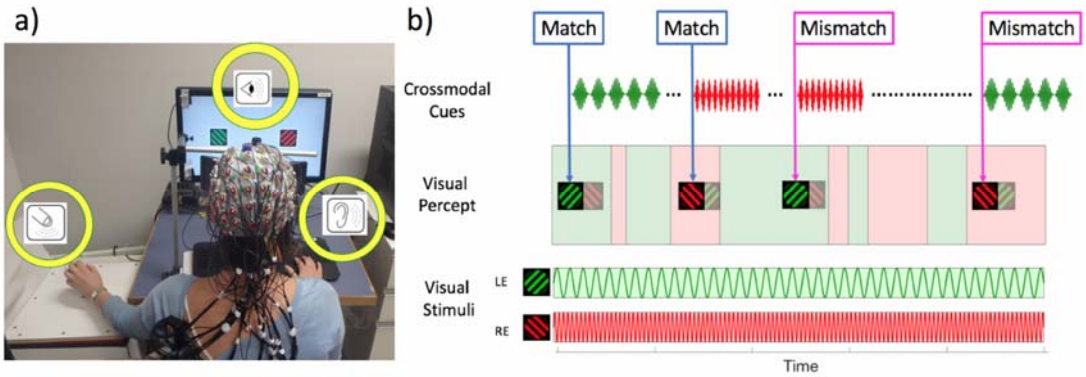


Figure 1. Experimental paradigm. a) The experimental setup. Subjects viewed binocular rivalry stimuli through a mirror stereoscope and reported their changing percept with their right hand. Occasionally, they received crossmodal cues that were auditory and/or tactile, delivered via earphones or vibrotactile device beneath their left hand. Neural activity during the task was recorded through 64-channel EEG. b) A schematic time course showing stimulus presentation, reported visual percept, and the two cue conditions, either matched or mismatched, defined with respect to the dominant visual percept at crossmodal cue onset. Each eye was presented with a continuous 4.5 or 20 Hz sinusoidal flicker (bottom), with crossmodal cues (also 4.5 or 20 Hz) presented at varying inter-stimulus intervals (top). The red/green shading (middle) indicates the changing dominant visual percepts during binocular rivalry (green indicates 4.5 Hz visual flicker, red 20 Hz visual flicker). Due to the alternation dynamics of binocular rivalry, the crossmodal cue either matched the current contents of visual consciousness, or mismatched (and thus matched the nonconscious visual representation). To manipulate attention, for one of their two experimental sessions subjects were tasked with counting how many times the frequency of crossmodal cues coincided with that of their perceived visual image at crossmodal cue offset (see Methods).

Following the onset of a mismatched crossmodal cue, the probability of perceiving the congruent visual image significantly increased when comparing attended low-frequency cues to all other cue types (Figure 2a). This cueing effect was also observed as an increased probability of maintaining the current visual percept when the dominant visual percept at cue onset was congruent with an attended low-frequency crossmodal cue (Figure 2b). As the overall crossmodal effects were strongest for the attended low-frequency stimuli (see Discussion), we focused our subsequent analyses on this cue condition.

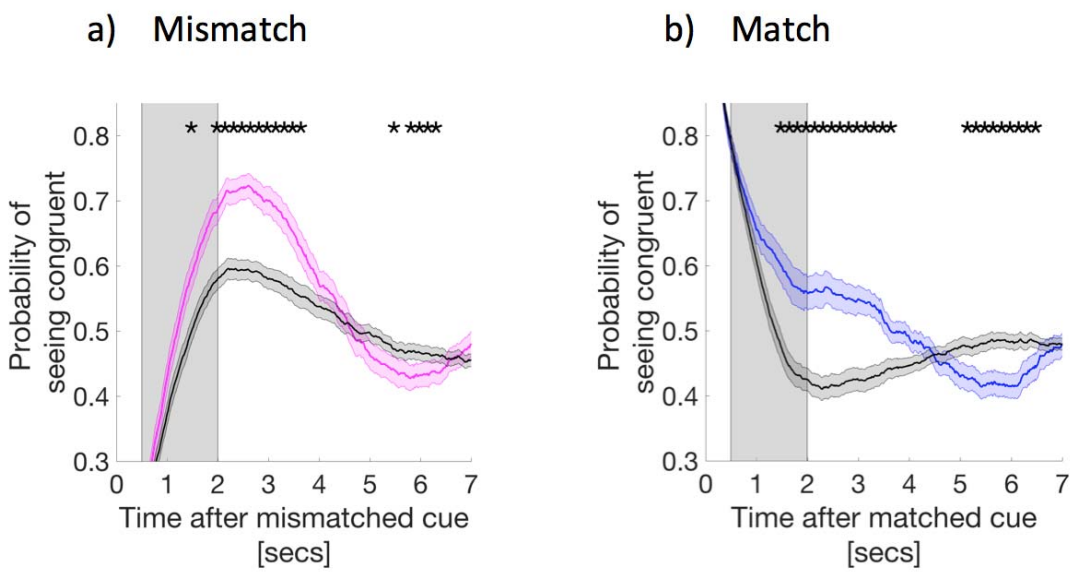


Figure 2. Time course of the influence of crossmodal cues on binocular rivalry. Button-press data were epoched after crossmodal cue onset, averaged per subject, then further averaged across subjects for each condition. Y-axis represents the proportion of button-presses reporting the congruent visual flicker. Colored lines and their shading show mean and ±1 standard error across 34 subjects for attended low-frequency cues, showing a) following mismatched cues and b) matched cues. Black lines with gray shading represent the combined response to all other cue types (see Supplementary Figure 1 for

*separated cue types). The time window following cue onset that was used to estimate the frequency of attentional sampling in subsequent analyses is indicated by the gray background. * indicates significant difference at each time point (repeated measures ANOVA followed by post hoc t-tests). We use FDR q = .05, for the statistical threshold unless noted otherwise.*

Attention flexibly samples multiple stimuli in the presence of crossmodal conflict

We checked the timing of reported perceptual switches for evidence of periodic sampling, following the onset of attended low-frequency crossmodal cues. Our study design allowed us to compare attentional sampling in two situations. Namely when the low-frequency crossmodal cues were congruent with either the dominant (matched) or suppressed (mismatched) visual image at cue onset. Previous investigations have shown approximately 8 Hz fluctuations in perceptual performance when participants attended a single visual location (Laura Dugué, McLelland, et al., 2015; Fiebelkorn et al., 2013) as well as ~4 Hz sampling when monitoring multiple visual features (Fiebelkorn et al., 2013) or locations (Landau & Fries, 2012; Landau et al., 2015). Therefore, we hypothesised that ~4 Hz perceptual sampling in our paradigm would indicate the distribution of attention away from the single dominant visual stimulus.

Figures 3a and 3b show the unfiltered time course of the number of perceptual switches after the onset of an attended low-frequency crossmodal cue, summed across 34 subjects. These time course show periodic fluctuations in the number of perceptual switches, with distinct frequencies based on cue type. When the frequency of the cue matched the conscious visual flicker at cue onset, subsequent changes in perception occurred at 7.3-8.1 Hz (Figure 3b and 3c blue, FDR q= .005, p <.001, permutation tests) consistent with findings that show attention samples from single locations at a rate of

approximately 7-8 Hz. When crossmodal cues were mismatched with the dominant visual image at cue onset, the power spectrum of perceptual switches peaked at 3.3 Hz (Figure 3a and 3c magenta). This slower rhythm of perceptual changes following mismatched cues suggests that attentional sampling was divided away from the conscious visual representation before triggering a visual perceptual switch to match multisensory stimuli. Given that the remaining three combinations (attended high-, ignored low- and ignored high-frequency cues) all failed to exhibit any sustained perceptual changes during binocular rivalry (Supplementary Figure 2), we did not pursue further spectral or neural analyses of these conditions.

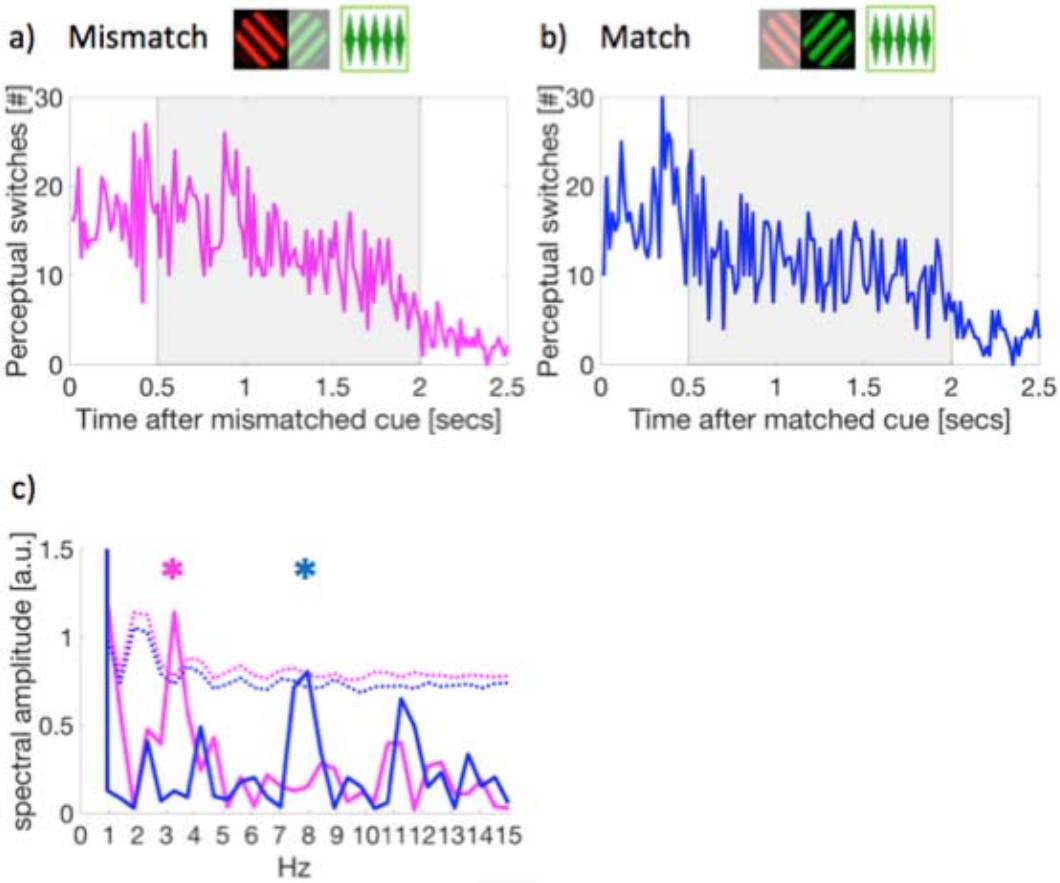


Figure 3. Attentional sampling during binocular rivalry following a crossmodal cue. Time course of the number of button presses indicating a perceptual switch when the low-frequency crossmodal cue (a) mismatched or (b) matched that of the perceived visual flicker at cue onset. c) The power spectra for the time course of perceptual switches. The magenta and blue lines are for the mismatched and matched cue conditions, respectively. Dashed lines indicate the statistical threshold for FDR q = .005 according to permutation tests (see Methods).

The neural correlates of divided and focused attentional sampling

As selective attention has been argued to reset the phase of ongoing oscillatory activity (Frey, Ruhnau, & Weisz, 2015; Lakatos et al., 2009; Van Atteveldt, Murray, Thut, & Schroeder, 2014), we investigated whether evidence of oscillatory phase reset would be present after cue onset, as a proxy for attentional sampling. Specifically, we calculated the inter-trial phase coherence (ITPC) at our observed behavioral sampling frequencies for candidate neural correlates of attentional sampling in the EEG signal. Compared with the two seconds preceding cue onset, both mismatched and matched attended low-frequency cues increased both 3.5 and 8 Hz ITPC over a range of central-parietal regions (see Supplementary Figure 3).

To isolate the neural correlates of divided and focused attentional sampling, we compared the evoked ITPC between mismatched and matched cue conditions, specifically at the attentional sampling frequencies (3.5 Hz and 8 Hz) that are behaviourally relevant in our attended low-frequency crossmodal condition. We found that the mismatched cues induced stronger ITPC than the matched cues, at 3.5 Hz over right fronto-central-temporal electrodes [FT8, C6] and at 8 Hz over right parietal-occipital electrodes [P6, PO8]. We

continued by investigating these two regions of interest (ROI) for our subsequent analyses on the neural correlates of behavioral attentional sampling.

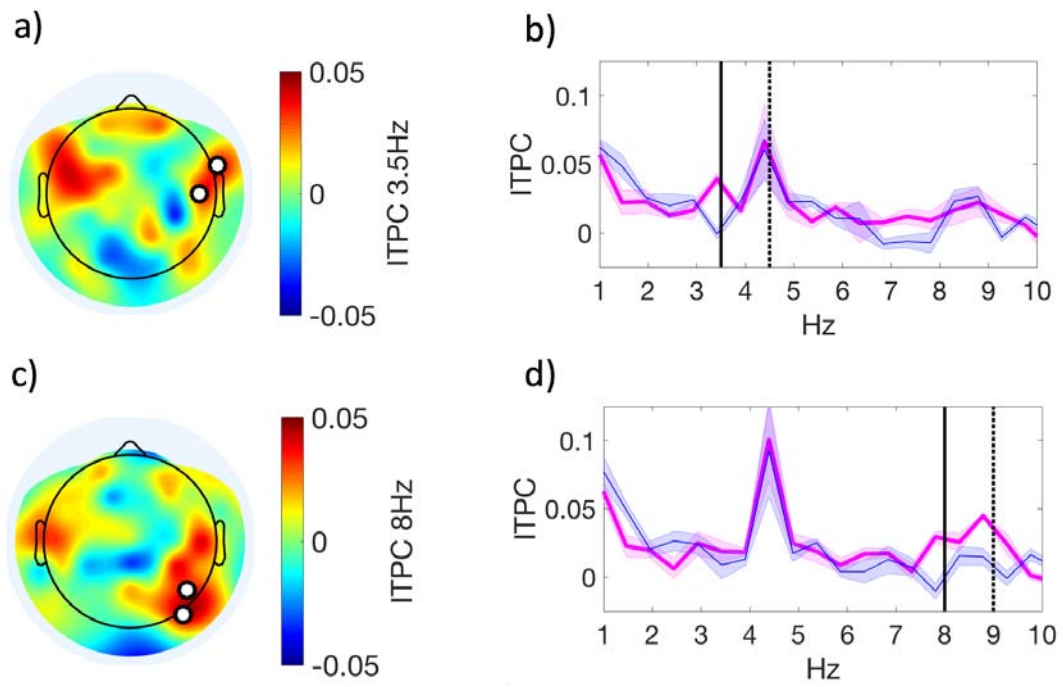


Figure 4. Neural correlates of divided and focused attentional sampling for the attended low-frequency condition. Topography of the difference between mismatched and matched cue conditions at a) 3.5 Hz ITPC and c) 8 Hz ITPC. White disks note the electrodes that showed significantly stronger ITPC for the mismatched than the matched condition (non-parametric cluster-based tests (FDR q= .005). b and d) ITPC across frequencies averaged over the white electrodes in a and c), respectively. The magenta and blue lines show the ITPC across frequencies for the mismatched and matched conditions. Dotted and solid vertical black lines mark in (b), the stimulus frequency at 4.5 Hz and the behaviorally-observed attentional sampling frequency at 3.5 Hz, and in (d) the stimulus harmonic at 9 Hz and the 8 Hz sampling frequency observed behaviorally. We used a 2-second time window, which gives us a half bandwidth of 0.5Hz, to sufficiently distinguish

these frequencies (see Methods). Shading shows ±1 standard error of the mean across 34 subjects.

Attentional-sampling ITPC strength within ROIs predicts perceptual outcome

Next, we investigated whether the evoked ITPC in these ROI predicted the magnitude of behavioral effects across subjects. Specifically, we compared the magnitude of evoked ITPC at the attentional sampling frequencies during the period 0-2s after cue onset (Figure 5; x-axes), to the subsequent behavioral probability of perceiving a congruent visual flicker 2-4 s after cue onset (y-axes). For each comparison, we subtracted the respective values (either ITPC or behavioral probability) averaged across all other conditions from those in the attended low-frequency conditions.

In the first ROI ([FT8, C6]), we found that the magnitude of 3.5 Hz ITPC and the behavioral probability of perceiving a congruent visual percept were positively correlated for mismatched cues ($r(32)= .39$, $p=.024$, Figure 5a), but negatively correlated for matched cue types ($r(32)= -.40$, $p=.018$, Figure 5b). As to our second ROI ([P6,PO8]), we found that 8 Hz ITPC was not correlated with behavioral probability for mismatched cues, yet positively correlated for the matched cue types ($r(32) = .38$, $p = .026$).

This pattern of results, in particular, the dissociation of correlations for the mismatched and matched cue types (Figure 5), is consistent with an interpretation that the distinct frequencies of 3.5 and 8 Hz observed behaviorally may reflect flexible phase-resetting of the neural mechanisms of attentional sampling in response to multisensory conflict.

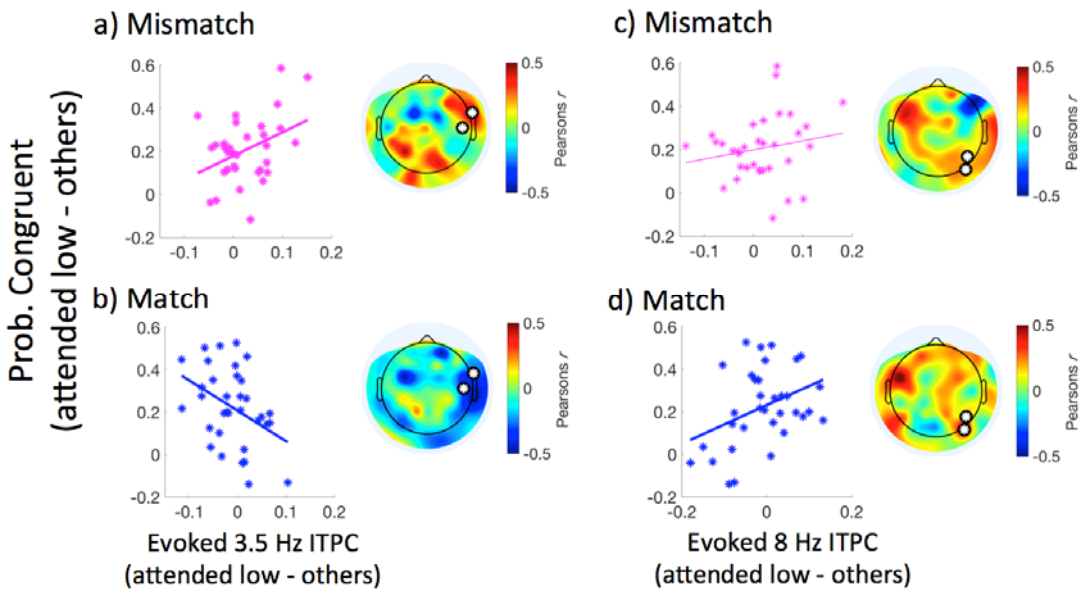


Figure 5. Correlations within ROIs between the behavioral effects and ITPC evoked by crossmodal cues, for 3.5 Hz during (a) mismatch and (b) match as well as 8Hz during (c) mismatch and (d) match. The x-axes represent the magnitude of evoked ITPC during the period 0-2s after cue onset. The y-axes represent the behavioral probability of perceiving a congruent visual flicker 2-4 s after cue onset. Each point in the scatter plots represent the magnitude of ITPC or behavioral probability for the difference between the attended low-frequency condition and the average across all other conditions for a single subject. Topoplots display the strength of these respective correlations encoded in color (Pearsons r).

General Discussion

Using a combination of binocular rivalry and crossmodal cues that either matched or mismatched the perceived visual percept, we observed evidence for both focused and divided attentional sampling that depended on cue congruency, with distinct

electrophysiological correlates in the EEG. Crossmodal cues that were congruent with the dominant visual stimulus resulted in undivided attention, sampling at a rate of ~8 Hz. However, divided and slower rates of attentional sampling at ~3.5 Hz emerged when crossmodal cues were incongruent with the dominant visual stimulus, with the neural correlates of these processes found over parieto-occipital and fronto-centro-temporal electrodes, respectively.

Implications for crossmodal and attentional modulation in binocular rivalry

Traditionally, top-down, voluntary attention has been thought to have limited control over perceptual dynamics during binocular rivalry (Paffen & Alais, 2011; for bottom-up control, including crossmodal stimulation, see Conrad, Bartels, Kleiner, & Noppeney, 2010; Deroy et al., 2016; Guzman-Martinez, Ortega, Grabowecky, Mossbridge, & Suzuki, 2012; Kang & Blake, 2005; Lunghi & Alais, 2013; Lunghi, Binda, & Morrone, 2010; Lunghi et al., 2014; van Ee, van Boxtel, Parker, & Alais, 2009). Our results clearly show additional dependence on the top-down deployment of attention, as without explicit instruction to attend to crossmodal signals, no facilitatory crossmodal effects emerged (Supplementary Figure 2; see also Jack & Hacker, 2014; Talsma, Senkowski, Soto-Faraco, & Woldorff, 2010; van Ee et al., 2009). This interaction between low-level stimulus features (temporal frequency) and the allocation of attention indicates the facilitative role of both crossmodal stimuli (Deroy et al., 2016; Deroy, Chen, & Spence, 2014) and attention for perceptual transitions during binocular rivalry (Dieter, Brascamp, Tadin, & Blake, 2016; Dieter & Tadin, 2011; Paffen & Alais, 2011; Zhang, Hong, Gao, Gao, & Röder, 2011).

Attention flexibly samples at 8 Hz when focused, and 3.5 Hz when divided during binocular rivalry

Past research has suggested that attention samples sensory input at approximately 7-8 Hz within a single spatial location (Laura Dugué, McLelland, et al., 2015; Fiebelkorn et al., 2013; R. VanRullen et al., 2007; Zoefel & VanRullen, 2017). In our binocular rivalry paradigm, we also observed an 8 Hz rhythm following matched cues, suggesting that attention was sampling from a single entity.

In contrast, perceptual sampling has also been observed at ~4 Hz, when cues have encouraged dividing attention between two objects or locations (Fiebelkorn et al., 2013; Huang et al., 2015; Landau & Fries, 2012; Landau et al., 2015; Song et al., 2014). As such, the ~3.5 Hz rhythm we observed when crossmodal cues mismatched with the conscious visual stimuli extends the evidence for divided attentional sampling to separate entities during binocular rivalry.

Neural correlates of attentional sampling

Distinct neural correlates of these attentional sampling rhythms were found in the EEG. We found significantly greater 3.5 Hz ITPC strength for mismatched compared to matched cue types over right-fronto-centro-temporal electrodes [FT8 and C6], suggesting this region may be a candidate neural correlate for divided periodic attentional sampling. Following mismatched cues, increased 3.5 Hz ITPC in this region also positively correlated with the likelihood of switching to the previously suppressed, yet crossmodally congruent visual image across subjects (Figure 5a). Following matched cues, by contrast, increased 3.5 Hz ITPC in this ROI was negatively correlated with the likelihood of maintaining a congruent visual percept (Figure 5b).

Following visual only stimulation, previous research has identified a pre-target ~4 Hz phase-dependency for peri-threshold perception when attention is divided across visual hemifields (Landau et al., 2015). We note that in our paradigm attention was not divided between visual hemifields, yet interestingly, our ROI corresponds to the right superior temporal gyrus (Koessler et al., 2009), which has been implicated in the reorientation of attention to previously unattended locations (Corbetta & Shulman, 2002; Downar, Crawley, Mikulis, & Davis, 2000). Taken together, our results show that periodic attentional sampling exists during binocular rivalry, when visual stimuli spatially overlap and compete for perceptual dominance.

We also found behavioral and neural correlates of focused attentional sampling during binocular rivalry when cues were consistent with the prevailing visual percept. Specifically, 8 Hz ITPC over parieto-occipital electrodes was positively correlated with the likelihood of maintaining a congruent perceptual outcome. Previously, phase-dependent peri-threshold perception has been reported for focused attention tasks in the visual domain (Busch & VanRullen, 2010; Busch, Dubois, & VanRullen, 2009; Hanslmayr, Volberg, Wimber, Dalal, & Greenlee, 2013; Mathewson, Gratton, Fabiani, Beck, & Ro, 2009), and has primarily implicated an approximately 7 Hz component located over fronto-central electrodes(Busch & VanRullen, 2010; Busch et al., 2009).

Given the differences between paradigms, we did not expect our ROIs for divided and focused attentional sampling to coincide with those reported in previous research regarding phase-dependent perception. Nonetheless, it is notable that we identified fronto-centro-temporal [FT8, C6] and parietal-occipital [P6, PO8] ROIs in the EEG, with cortical projections to the superior-temporal and post-central gyri, as well as middle-temporal and middle occipital gyri respectively (Koessler et al., 2009). Each of these regions has previously been implicated in the reorienting of visuo-spatial attention (Corbetta &

Shulman, 2002; Downar et al., 2000; Dugué, Merriam, Heeger, & Carrasco, 2017), and for the integration of multisensory stimuli into a coherent percept (Beauchamp, 2005; Bushara et al., 2003; Calvert & Thesen, 2004; Driver & Noesselt, 2008; Zhang et al., 2011).

Attentional sampling may divide between sensory modalities

Our analysis so far has revealed that when a crossmodal cue mismatched the visually dominant stimulus during binocular rivalry, that attentional sampling was divided between separate entities. But between which entities was attention divided? One possibility is that attentional sampling during mismatched cues was divided between two sensory modalities, as the brain tries to resolve a conflict between concurrent auditory/tactile and visual information. When the conflict is resolved at a time when all modalities are congruent, a single 'object' is perceived, and attention is undivided. As mentioned above, the location of our ROIs are consistent with this multisensory integration account. Figure 7a provides a schematic of this interpretation. If the neural activity in our ROIs is indeed representative of divided sampling between modalities, it constitutes the first evidence that an attentional sampling mechanism can flexibly orient between temporally co-modulating crossmodal stimuli. Although the facilitative role of attention in multisensory integration remains controversial (Hartcher-O'Brien et al., 2016; Talsma et al., 2010), we see it as a viable possibility that this mechanism resolved perceptual ambiguity through a visual perceptual switch to the competing image, rendering the multisensory stimuli congruent.

Crossmodal attentional sampling may also explain why we observed low- but not high-frequency behavioral effects in the present task, particularly in the context of previous investigations regarding the binding of multisensory stimulus attributes (Fujisaki

& Nishida, 2005, 2010; Lunghi et al., 2014; Vroomen & Keetels, 2010), and the limits of crossmodal temporal judgements (Fujisaki & Nishida, 2005, 2010; Holcombe, 2009; Vroomen & Keetels, 2010). For example, Fujisaki and Nishida (2005) have shown that judgements of temporal synchrony between rhythmic sensory streams degrade above ~4 Hz, arguing that salient features become difficult to discriminate in this context. This is consistent with the absence of a high-frequency crossmodal effect in our design, as explicit judgements of temporal synchrony between crossmodal and visual rhythmic sensory streams were required from our subjects during attend conditions – with the added ambiguity induced via binocular rivalry. When sensory input is ambiguous, crossmodal congruency has been shown to augment attentional control over perceptual selection (van Ee et al., 2009). Such perceptual difficulties may be specifically enhanced in our binocular rivalry design, because compared to an earlier behavioral paradigm (Lunghi et al., 2014), we required an increased size of visual stimuli to facilitate subsequent steady-state visually evoked potential analyses (Davidson et al., in prep). Larger stimuli have been shown to induce more piecemeal rivalry, with more ambiguous perceptual dominance than rivalry between smaller stimuli (Blake, O'Shea, & Mueller, 1992). Together, the dependence on perceived crossmodal congruency to resolve ambiguous perceptual selection, combined with decreased temporal discrimination at higher frequencies (Fujisaki & Nishida, 2005) may account for the frequency specificity of our behavioral results during divided crossmodal attentional sampling.

Attentional sampling may divide between conscious and nonconscious images

Another possibility, which is not mutually exclusive with the first, is that the 3.5 Hz rhythm in our paradigm reflects divided attentional sampling between dominant and suppressed visual images during binocular rivalry (Figure 7b). The frequency of divided

attentional sampling that we observed is consistent with those obtained when visual attention has been divided between two objects or locations (Fiebelkorn et al., 2013; Landau & Fries, 2012). As our binocular rivalry stimuli occupied the same location, attention in our paradigm was likely divided between either features or objects, instead of locations. Indeed, feature-based attention has already been shown to modulate neural processes when an attended target is suppressed during continuous flash suppression (Kanai, Tsuchiya, & Verstraten, 2006). During binocular rivalry, perceptual dominance has also been shown to be influenced by object-based attention (Mitchell, Stoner, & Reynolds, 2004), with unconscious selection mechanisms argued to mediate perceptual fluctuations (Lin & He, 2009). This interpretation is also consistent with the locations of our ROI, as the right inferior parietal(Lumer, Friston, & Rees, 1998; Roy, Jamison, He, Engel, & He, 2017), and ventral temporal cortices (Tong, Nakayama, Vaughan, & Kanwisher, 1998) have each been implicated in visual perceptual rivalry (Tong, Meng, & Blake, 2006).

This second alternative is also indirectly supported by the temporal limits of binocular rivalry when conflicting visual stimuli are presented asynchronously, without temporal overlap between the two eyes (O'Shea & Blake, 1986; van Boxtel, Knapen, Erkelens, & van Ee, 2008; van Boxtel, van Ee, & Erkelens, 2007). The maximum stimulus onset asynchrony that can sustain this type of rivalry is approximately 350±50 ms, beyond which alternating stimuli introduced to one eye are perceived immediately, without rivalry occurring(van Boxtel et al., 2008). This limit is consistent with a 7-8 Hz attentional sampling rhythm distributed between the two conflicting stimuli (each sampled at ~3-4 Hz). When stimuli are presented rapidly enough they are temporally-bound together and can engage in ongoing rivalry; when stimuli are presented at less than 3-4 Hz, they are temporally individuated by attention, and rivalry ceases.

The suggestion that attention can sample between conscious and nonconscious vision is consistent with a view that the underlying neuronal processes for attention and consciousness are supported by distinct neural mechanisms. Specifically, evidence for attentional sampling of a suppressed image suggests that attention is not sufficient for consciousness (Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006; Koch & Tsuchiya, 2007; Lamme, 2003; Tsuchiya & Boxtel, 2013). (However, this interpretation remains consistent with a view that attention may also be necessary for conscious perception (Chica & Bartolomeo, 2012; Cohen & Dennett, 2011; Merikle & Joordens, 1997; O'Regan & Noë, 2001; Posner, 1994; 2012)).

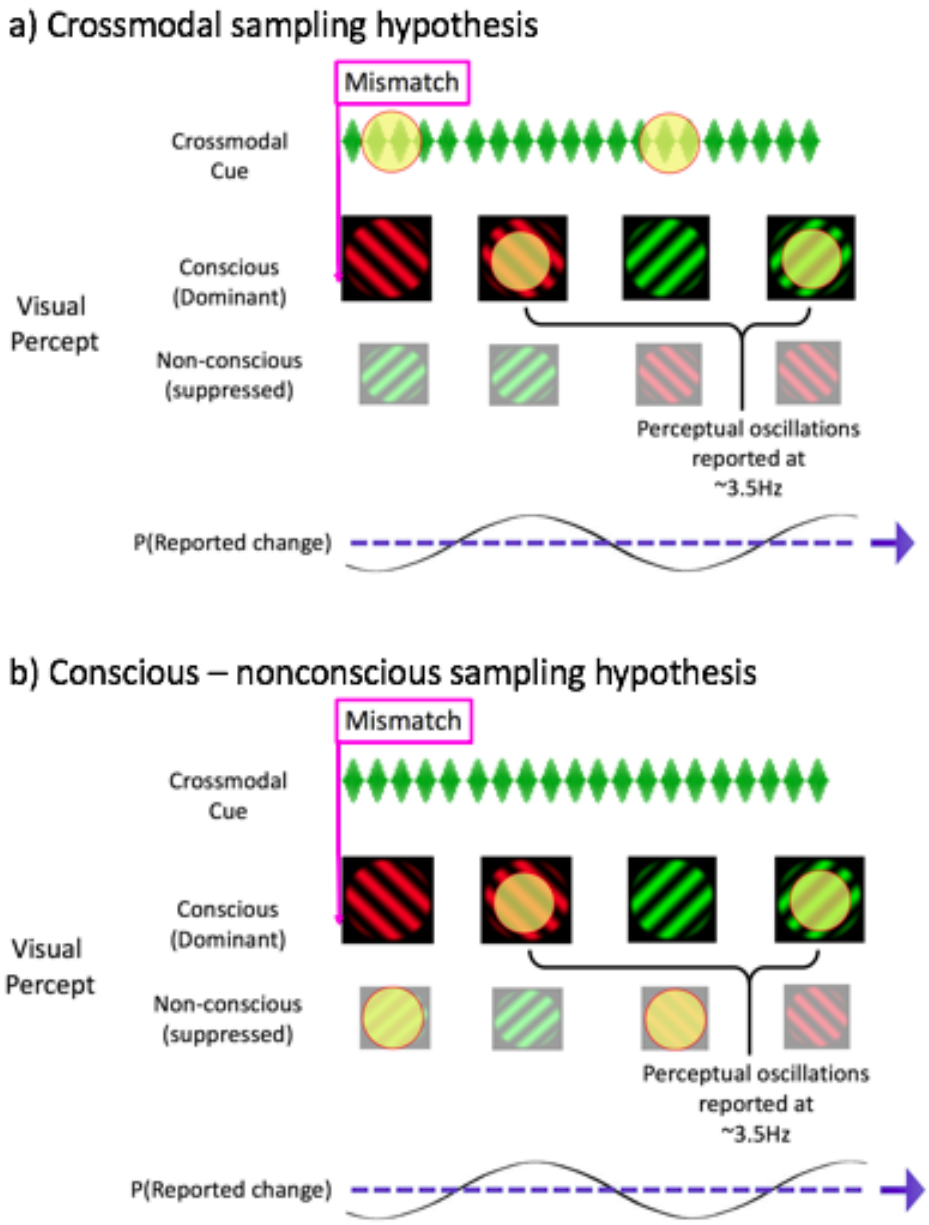


Figure 7. Two possible interpretations of attentional sampling during mismatched crossmodal cues. Schematic representation of attentional sampling and perceptual oscillations during binocular rivalry. a) Crossmodal sampling hypothesis: While perceiving the high-frequency visual flicker, an attended low-frequency crossmodal cue mobilises attention to sample between the dominant image and mismatched crossmodal cue at ~3.5 Hz. As a consequence, the likelihood of reporting changes in the conscious percept are modulated at ~3.5 Hz. b) Conscious-nonconscious sampling hypothesis: The

onset of a mismatched cue prompts attention to sample between separate visual features, which in our paradigm consists of dominant and suppressed visual images.

Conclusion

The attention sampling hypothesis continues to garner traction from various psychophysical and neuronal paradigms (VanRullen, 2016b). Here we have provided novel evidence in support of attentional sampling during binocular rivalry, through the use of crossmodal cues matched to either conscious or nonconscious visual stimuli. Future targeted experimentation can confirm whether attention can indeed sample across modalities (Figure 7a), as well as if attention can sample between conscious and nonconscious neural representations (Figure 7b). Investigating the interactions between crossmodal stimuli and conscious perception represent a fruitful avenue for experimentation (Faivre, Arzi, Lunghi, & Salomon, 2017), here uncovering the previously unknown dependence for attention and consciousness on rhythmic neural dynamics of the human brain.

Methods

Participants

Thirty-four healthy individuals (21 female, 1 left handed, average age 23 ± 4.7) were recruited via convenience sampling at Monash University, Melbourne, Australia. All had normal or corrected-to-normal vision and gave written informed consent prior to participation. Monash University Human Research and Ethics Committee approved this

study, and subjects were paid 15 AUD per hour of their time, over an approximate total time of 5 hours.

Apparatus and Stimuli

Stimuli were generated using Psychtoolbox (Brainard, 1997) and custom MATLAB scripts. Each visual stimulus was viewed through a mirror stereoscope placed at an approximate viewing distance of 55 cm from computer screen (29 x 51 cm, 1080 x 1920 pixels, 60 Hz refresh rate) with the subjects' heads stabilized via chin rest. Rivalry stimuli were red and green gratings displayed on a black background, with a white frame to aid binocular fusion, embedded within the wider gray background of the remaining portions of the screen (Figure 1a). Beside each white framed image, colored arrows indicated the direction for button press (Right = red. Left = Green). Gratings were sinusoidal with spatial frequency of 0.62 cycles per degree, oriented ± 45° from vertical, and subtended 6.5° visual angle (240 x 240 pixels on the display). Visual stimuli were contrast-modulated at either 4.5 or 20 Hz using a temporal sinusoidal envelope. The phase of each grating was static throughout each 3-minute binocular rivalry trial, yet shifted after each trial to reduce the effects of visual adaptation.

For crossmodal stimuli 50 Hz carrier tones were amplitude modulated by 4.5 or 20 Hz sine waves to create digital waveforms, which were either 2, 3.1 or 4 seconds in duration. For tactile stimulation, subjects clasped a wooden ball with their left hand attached to a Clark Synthesis Tactile Sound Transducer (TST429 platinum) housed in a custom sound insulated box (Lunghi et al., 2014). Auditory stimulation was delivered binaurally through Etymotic HD5 noise reduction headphones, with ACCU-Fit foam ear tips to reduce ambient noise.

Stimulus timing

Accurate stimulus timing of synchronous visual and crossmodal stimuli was ensured with a WDM-compatible hardware independent low latency ASIO driver (www.asio4all.com), which was necessary to minimize audio buffer duration to sub millisecond intervals and reduce latency compensation. The time-course of stimulus presentation was also physically recorded in the EEG for offline analysis. Photodiodes were used to record the flicker-envelope of visual stimuli and stored as separate channels in the ongoing EEG. The waveforms for crossmodal stimulation were simultaneously sent to both the presentation hardware and external electrode channels using a digital splitter (Redback A2630 4 Channel Headphone Distribution Amplifier). Stimulus presentation lag was assessed by computing the difference between the recorded frames of trigger-codes and actual physical trace within the EEG as part of data pre-processing. We adjusted the relative timing of behavioral and EEG data accordingly as part of this analysis. In most cases, no adjustment was necessary, requiring a maximum change of 3 frames in duration on <1% of trials across all participants.

Calibration of visual stimuli

A maximum of 10 one-minute binocular rivalry trials were performed prior to experimentation on the first day for all subjects. These blocks served to familiarize subjects with reporting their visual percepts during binocular rivalry, and to calibrate approximately equal dominance durations for the flickering stimuli in each eye. Contrast values for left/right eye, green/red color, and low/high frequency stimulus combinations

(in total 8 combinations) were adjusted on the logarithmic scale until approximately equivalent total dominance durations were reached (between 1:1 and 1:1.5), with the additional requirement that the average perceptual duration for each stimulus was longer than 1 second.

Calibration of auditory stimuli

Prior to experimentation, subjects were also tasked with equating the perceptual intensity of tactile and auditory stimulation for each of low and high frequency condition, to achieve approximately equal phenomenological intensity across subjects and stimulus conditions. For all subjects, the amplitude of tactile vibrations was set to the same comfortable, above threshold level (approximately equivalent to 65 dB SPL). In the absence of visual stimulation, simultaneous auditory and tactile stimuli were then presented in a staircase procedure, with subjects adjusting the amplitude of auditory tones to match the perceived intensity of simultaneous tactile vibrations. They performed the matching task separately within low-frequency auditory tones and tactile vibrations and within high-frequency auditory tones and tactile vibrations. This calibration procedure was performed on each day of testing, to account for differences in the insertion depth of inner-ear headphones across separate days.

Experimental Procedure and Behavioral Analysis

24 three-minute binocular rivalry blocks were presented on two separate days of testing. In each block, subjects reported their dominant visual percept during rivalry while receiving occasional crossmodal cues, which were either auditory, tactile, or simultaneous auditory and tactile. In a given three-minute block, we presented only one of the three

types of crossmodal cues, the order of which was randomized for each subject and each day of experimentation. In each block, 12 trials of crossmodal cues were presented. Each cue was either low (4.5 Hz) or high (20 Hz) frequency auditory and/or tactile stimulation. Six cues were presented for each frequency, composed of three x 2 s, two x 3.1 s, and one x 4 s cues. To increase uncertainty of the timing of the cues, we inserted three null cues without any crossmodal stimulation for a duration of 2.6 s (the average of crossmodal cue durations). We randomized the order of all cues, which were separated with uniform jittering by 7-10 s ISI within each block.

Across all sessions subjects were told to focus on accurately reporting their dominant visual percept at all times via button press. Over two sessions on separate days, subjects distributed attention between visual rivalry and crossmodal cues based on separate task instructions. On Day 1 for n= 18 or Day 2 for n=16, subjects were instructed to ignore the crossmodal cues and to focus on reporting visual rivalry only. For the alternate session, subjects were instructed to distribute attention across both visual rivalry and crossmodal cues. To ensure attention on task, these alternate days included task instructions for subjects to silently count whenever the temporal frequency of an attended crossmodal cue matched with that of their dominant visual percept at the time of crossmodal cue offset. Due to the varied duration of crossmodal cues, this task ensured that attention was allocated consistently throughout the presentation of crossmodal cues. To familiarize subjects with these task demands, an additional two practice trials were included during the calibration procedure on the relevant day of experimentation. Although 34 subjects were retained for final analysis, five others were recruited and began the experiment, yet failed to complete their second day of experimentation. One other subject was removed

for failing to understand task instructions and limit their movement during EEG recording.

Behavioral data analysis

We performed a number of preprocessing steps to accurately estimate the timing of changes in visual consciousness during binocular rivalry. First, we categorised each time-point according to the flicker frequency of the dominant visual stimulus reported. To analyze the time course of probability of button press (Figure 2a), we then categorized button-presses as either congruent or incongruent with the ongoing crossmodal stimulus frequency. Then, we obtained the probability of congruent button press as a function of time per subject, by averaging responses at each time point across all 144 trials per attention x frequency cue subtype. The averaged probability across subjects at each time point was compared among 4 conditions (low- vs high-frequency crossmodal cue x attended vs unattended) with one-way repeated-measures ANOVA. We defined significant differences among conditions at those time points that survived corrections for multiple comparisons with post hoc *t*-tests with least significant differences using FDR at q=0.05 (Benjamini & Yekutieli, 2001).

For our spectral analysis (Figure 3), we defined perceptual switches as the first time-point recording a change in button-state after cue onset. Matched and mismatched crossmodal cues were categorized based on the temporal frequency of the crossmodal cue and the dominant visual stimulus at cue onset.

For the spectral analysis of perceptual switches, we applied a single-taper Fourier transform to the period 0.5 - 2 seconds after cue onset. This window was selected to restrict the analysis so that all the analyzed trials were during an attended cueing period (as the minimum crossmodal cue duration was 2 seconds), and to remove transient button presses occurring early in the cue period, which were unlikely to be caused by crossmodal match or mismatch.

To assess statistical significance we used a non-parametric randomization procedure (Maris & Oostenveld, 2007). After obtaining the spectral amplitude for the observed data across subjects, we generated a null distribution of perceptual switch times during the same cue period by randomly shifting switch-times within each participant, while keeping the number of perceptual switches the same. This procedure creates surrogate trials to test the null hypothesis that there are no temporal effects on the timing of perceptual switches. We repeated the same spectral analysis performed on our observed data on 5000 shuffled data sets, which also controls for any bias introduced by our analysis procedure, to estimate the null distribution of Fourier-transformed power, at each frequency, under the null distribution. We then compared the observed amplitude with the null distribution. We regarded the spectral amplitude at a certain frequency to be significantly above chance at corrected $p < 0.005$ level if the observed amplitude is higher than the top 99.5% of the null distribution.

EEG recording and analysis

EEG was recorded at a sampling rate of 1000 Hz using three BrainAmp amplifiers and 64-channel ActiCap (BrainProducts), with the impedance of each electrode kept below 10 kΩ. Ground and reference electrodes were AFz and FCz at recording, respectively. After

re-referencing the data to the average of all channels, we performed linear detrending and bandpass filtering (0.1- 60 Hz with a Hamming-windowed finite impulse response filter) and down-sampled the data to 250 Hz before time-frequency analysis.

We performed all time-frequency analyses using the Chronux toolbox (http://chronux.org; Bokil, Andrews, Kulkarni, Mehta, & Mitra, 2010) and custom MATLAB scripts. To resolve our frequencies of interest (especially between 3.5 and 4.5 Hz), we used a single-taper Fourier transform with a time-window of 2 seconds, which resulted in a half bandwidth (*W*) of 0.5 Hz (*W*=1/2). This half band width is consequently capable of resolving differences between 3.5 and 4.5 Hz, as demonstrated in Figure 4c.

To assess the neural correlates of attentional sampling (Figures 4-5), we analyzed the intertrial phase coherence (ITPC) within electrodes, over multiple time-frequency points (Bastos & Schoffelen, 2016). ITPC is an amplitude-normalized measure of the degree to which EEG responses are phase locked to the onset of an exogenous cue, ranging between 0 (random phase over trials) and 1 (perfect phase coherence over trials). To compute ITPC, the consistency of phase angles is computed as the length of the average of unit phase vectors in the complex plane over trials. For a given time, t, and frequency, f,

$$\mathrm{ITPC}(t,f) = \left|\frac{1}{N} * \sum_{n=1}^{N} e^{i(\theta(t,f,n))}\right|$$

where N is the number of trials, and θ is the phase angle at time t, frequency f, and trial n.

Due to the dynamic nature of perceptual alternations during binocular rivalry, the number of available trials for analysis in each mismatched and matched cue type ranged from 12 to 36 trials across subjects (after averaging first across participants, the mean number of trials was 24 (±1.5) trials across matched / mismatched and attention conditions). Because the bias level of ITPC is strongly influenced by the number of trials,

we took additional measures to equate the number of mismatched and matched cue types for the analysis. To achieve this aim, the maximum number of trials recorded for a given cue combination was identified across participants. Following this, participants with fewer numbers of trials had their observed number of trials supplemented by resampling with replacement from their recorded trials, equating them to the predefined maximum for each condition. Upon this resampled dataset, the ITPC was computed, and this process repeated 100 times. As the difference in ITPC between auditory, tactile, and combined auditory and tactile cues was not significant in any statistical tests comparing the effects of frequency of the crossmodal cues, we proceeded by combining crossmodal cues.

False discovery rate corrections were applied to ITPC data based on non-parametric spatial cluster-based analyses (Maris & Oostenveld, 2007). Similar to the permutation and shuffling tests described above, here the null hypothesis to test is that there is no spatial-cluster of significant difference in ITPC values between the mismatched or matched cue conditions. For all analyses, the minimum number of neighbouring channels for a cluster was set to two (minimum inter-sensor electrode distance 3.5 cm), with the threshold for significance to be a candidate member of a cluster set to an observed test-statistic equivalent to $p < 0.05$.

To apply the non-parametric correction for multiple comparisons, candidate clusters were identified based on the observed data, and criteria outline above. Within each cluster, the sum of observed *t*-scores for each electrode when comparing mismatched to matched cue types was then retained for further analysis. To create the null-distribution, condition labels (mismatch and match) were randomly shuffled for each electrode within each participant to create an equivalent dataset testing the null hypothesis that no significant difference exists in ITPC values. The test-statistics were then re-calculated for our candidate cluster sites when comparing these surrogate data sets, and

repeated 2000 times to obtain a distribution of test-statistics at our cluster-based electrode locations under the null hypothesis. Against this distribution, the observed test-statistic for the candidate cluster was then compared to a critical value, and accepted when exceeding the top 99.95% of the null distribution (or corrected to $p < .005$).

Finally, when assessing the neural correlates of attentional sampling, we correlated an EEG-based ITPC value with a behavioral measure of attentional sampling. Specifically, we used the difference in the evoked ITPC value for each electrode at 3.5 or 8 Hz over 0-2 second after cue onset between attended low-frequency condition and the mean across the remaining three conditions (i.e., attended high-frequency, ignored low- and high- frequency conditions). We computed this difference in the evoked ITPC separately for mismatched and matched conditions, while equatiting the number of trials by resampling with replacement (see above). As our behavioral measure, we used the difference in the mean probability of perceiving a congruent visual percept 2-4 seconds after cue onset between attended low-frequency and other conditions, separately for mismatched and matched conditions. This region was selected to investigate whether the strength of early attentional sampling could account for the significant differences between cue conditions regarding the emergence of a congruent perceptual outcome.

Author Contributions

D.A., N.T., and J.J.A.V.B. conceived of the initial extension of Lunghi et al (2014) into an EEG paradigm. D.A. provided tactile-stimulation equipment and code for auditory stimulus construction. M.J.D designed and programmed the research experiment reported in this article, with feedback from N.T. and J.J.A.V.B. M.J.D collected the experimental data. MJD performed the initial behavioral and EEG analysis, with feedback from J.J.A.V.B. and N.T. M.J.D conceived of and performed the attentional sampling analysis

with feedback from N.T. and J.J.A.V.B. M.J.D. wrote the article, with feedback from D.A, N.T. and J.J.A.V.B.

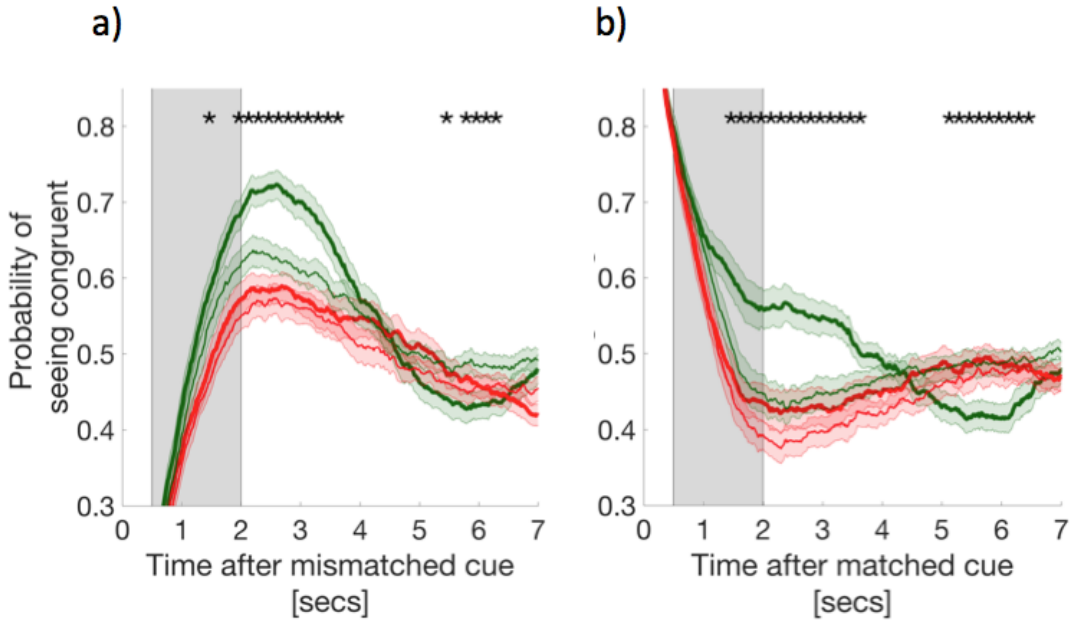
Acknowledgements

M.J.D. was supported by an Australian Government Research Training Program Scholarship. D.A. was funded by an Australian Research Council Discovery Project (DP150101731). N.T. was funded by an Australian Research Council Future Fellowship (FT120100619). The authors thank Bryan Paton and Claudia Lunghi for technical advice, and Brandon Lam for early piloting of the experimental paradigm.

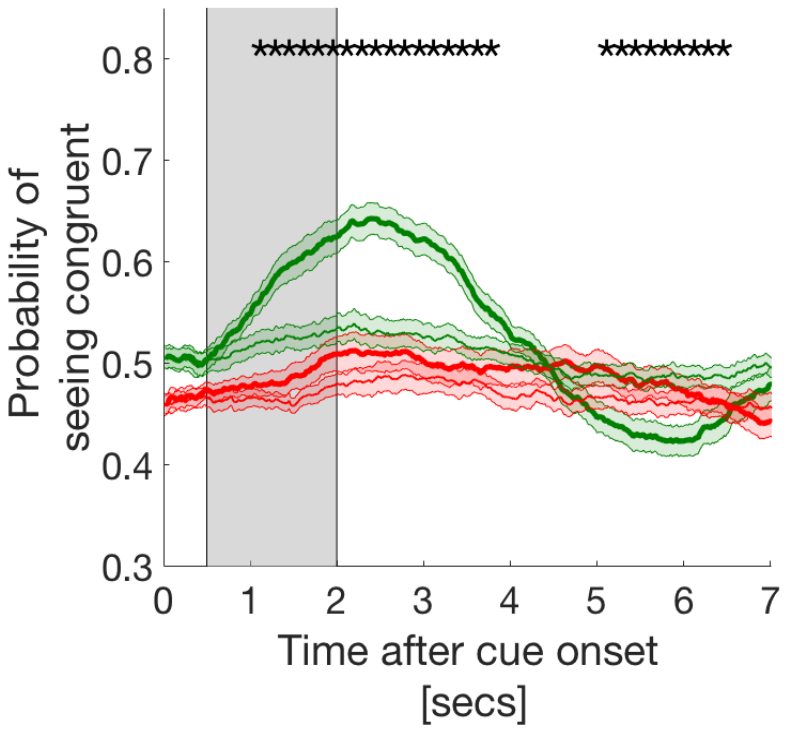
Competing Financial Interests

The authors declare they have no competing financial interests.

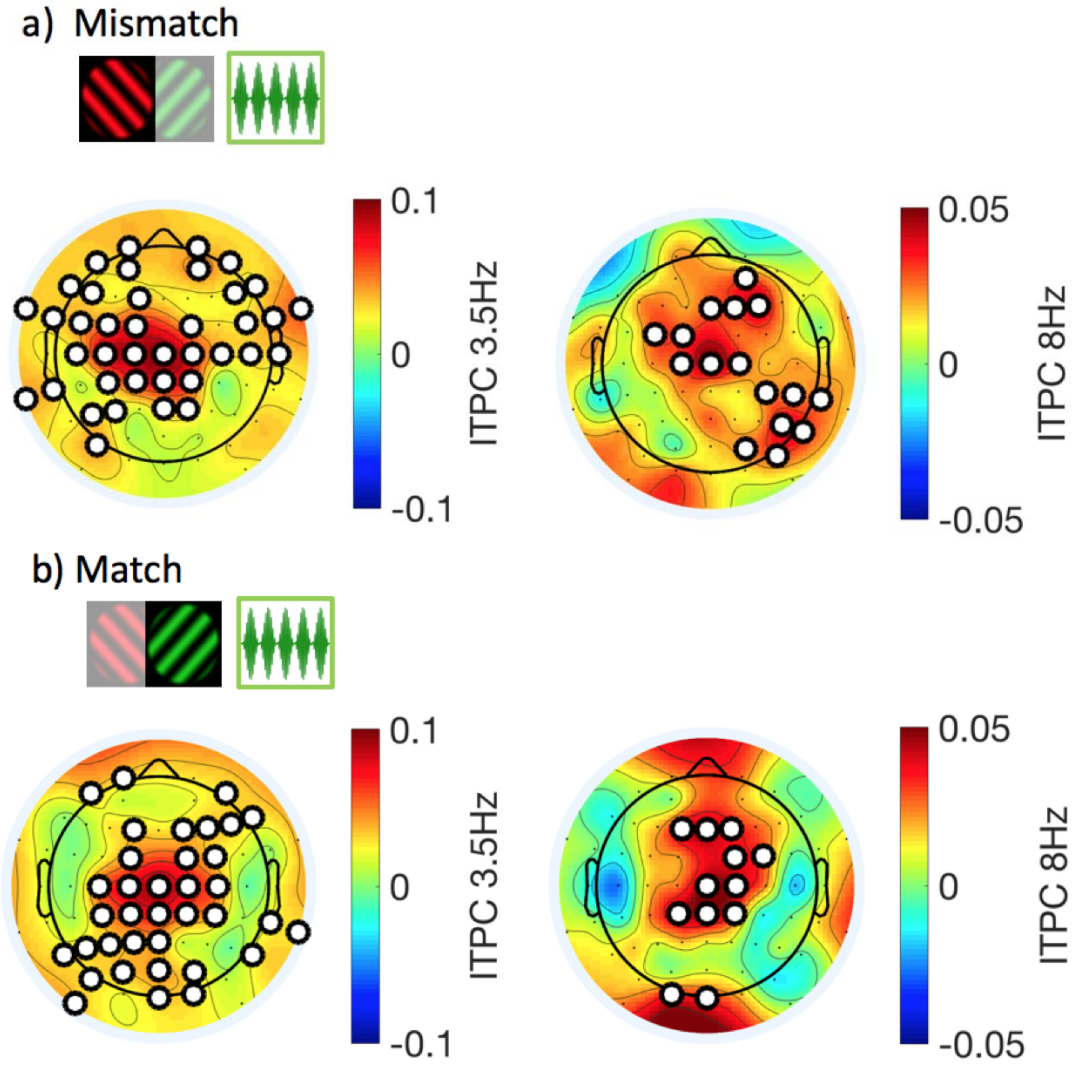
Supplementary Figures



Supplementary Figure 1: *Time course of the influence of crossmodal cues on binocular rivalry. Same format as the Figure 2, but separately for attended and ignored cues (thick and thin lines) and low and high frequency (green and red colors). for a) mismatched cues and b) matched cues. Shading represents ±1 standard error of the mean across N=34 subjects. * indicates significant difference at each time point (repeated measures ANOVA followed by post hoc t-tests).*



Supplementary Figure 2: *Time course of the influence of crossmodal cues on binocular rivalry for both mismatched and matched cue types. Same format as supplementary figure 2, colored lines and their shading show mean and ±1 standard error across 34 subjects separately for attended and ignored cues (thick and thin lines) and low and high frequency (green and red colors). * indicates significant difference at each time point (repeated measures ANOVA followed by post hoc t-tests). Only attended-low frequency cues facilitated a sustained change in conscious percept during rivalry shown above, and as such we did not pursue further spectral or neural analyses on the remaining conditions.*



Supplementary Figure 3: *ITPC measured in the first 2 seconds after attended low-frequency cue onset, compared to the baseline period 2 seconds prior to cue onset. White disks show the electrodes with significant increases in ITPC after non-parametric spatial-cluster based corrections for false discovery rate.*

References

Alais, D. (2012). Binocular rivalry: Competition and inhibition in visual perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, *3*(1), 87–103. http://doi.org/10.1002/wcs.151

Alais, D., & Blake, R. (2005). *Binocular rivalry*. MIT press.

Bastos, A. M., & Schoffelen, J.-M. (2016). A Tutorial Review of Functional Connectivity Analysis Methods and Their Interpretational Pitfalls. *Frontiers in Systems Neuroscience*, *9*(January), 1–23. http://doi.org/10.3389/fnsys.2015.00175

Beauchamp, M. S. (2005). See me, hear me, touch me: Multisensory integration in lateral occipital-temporal cortex. *Current Opinion in Neurobiology*, *15*(2), 145–153. http://doi.org/10.1016/j.conb.2005.03.011

Benjamini, Y., & Yekutieli, D. (2001). The Control of the False Discovery Rate in Multiple Testing under Dependency. *The Annals of Statistics*, *29*(4), 1165–1188.

Blake, R., O'Shea, R. P., & Mueller, T. J. (1992). Spatial zones of binocular rivalry in central and peripheral vision. *Visual Neuroscience*, *8*(5), 469–478.

Bokil, H., Andrews, P., Kulkarni, J. E., Mehta, S., & Mitra, P. P. (2010). Chronux: A platform for analyzing neural signals. *Journal of Neuroscience Methods*, *192*(1), 146–151. http://doi.org/10.1016/j.jneumeth.2010.06.020

Busch, N. A., & VanRullen, R. (2010). Spontaneous EEG oscillations reveal periodic sampling of visual attention. *Proceedings of the National Academy of Sciences*, *107*(37), 16048–16053. http://doi.org/10.1073/pnas.1004801107

Busch, N. A, Dubois, J., & VanRullen, R. (2009). The phase of ongoing EEG oscillations predicts visual perception. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(24), 7869–7876.

http://doi.org/10.1523/JNEUROSCI.0113-09.2009

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. http://doi.org/10.1038/nn993

Calvert, G. A., & Thesen, T. (2004). Multisensory integration: Methodological approaches and emerging principles in the human brain. *Journal of Physiology Paris*, *98*(1–3 SPEC. ISS.), 191–205. http://doi.org/10.1016/j.jphysparis.2004.03.018

Chen, A., Wang, A., Wang, T., Tang, X., & Zhang, M. (2017). Behavioral Oscillations in Visual Attention Modulated by Task Difficulty. *Frontiers in Psychology*, *8*(September), 1630. http://doi.org/10.3389/fpsyg.2017.01630

Chica, A. B., & Bartolomeo, P. (2012). Attentional routes to conscious perception. *Frontiers in Psychology*, *3*(JAN), 1–12. http://doi.org/10.3389/fpsyg.2012.00001

Cohen, M. A., & Dennett, D. C. (2011). Consciousness cannot be separated from function. *Trends in Cognitive Sciences*, *15*(8), 358–364. http://doi.org/10.1016/j.tics.2011.06.008

Conrad, V., Bartels, A., Kleiner, M., & Noppeney, U. (2010). Audiovisual interactions in binocular rivalry. *Journal of Vision*, *10*(10), 27. http://doi.org/10.1167/10.10.27

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), 201–215. http://doi.org/10.1038/nrn755

Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences*, *10*(5), 204–211. http://doi.org/10.1016/j.tics.2006.03.007

Deroy, O., Chen, Y., & Spence, C. (2014). Multisensory constraints on awareness. *Philosophical Transactions of the Royal Society of London. Series B, Biological*

Sciences, *369*(1641), 20130207. http://doi.org/10.1098/rstb.2013.0207

Deroy, O., Faivre, N., Lunghi, C., Spence, C., Aller, M., & Noppeney, U. (2016). The Complex Interplay Between Multisensory Integration and Perceptual Awareness. *Multisensory Research*, *29*(March), 1–22. http://doi.org/10.1163/22134808-00002529

Dieter, K. C., Brascamp, J., Tadin, D., & Blake, R. (2016). Does visual attention drive the dynamics of bistable perception? *Attention, Perception, & Psychophysics*. http://doi.org/10.3758/s13414-016-1143-2

Dieter, K. C., & Tadin, D. (2011). Understanding attentional modulation of binocular rivalry: a framework based on biased competition. *Frontiers in Human Neuroscience*, *5*(155).

Downar, J., Crawley, A. P., Mikulis, D. J., & Davis, K. D. (2000). A multimodal cortical network for the detection of changes in the sensory environment. *Nature Neuroscience*, *3*(3), 277–283. http://doi.org/10.1038/72991

Driver, J., & Noesselt, T. (2008). Multisensory Interplay Reveals Crossmodal Influences on “Sensory-Specific” Brain Regions, Neural Responses, and Judgments. *Neuron*, *57*(1), 11–23. http://doi.org/10.1016/j.neuron.2007.12.013

Dugué, L., Marque, P., & VanRullen, R. (2015). Theta Oscillations Modulate Attentional Search Performance Periodically. *Journal of Cognitive Neuroscience*, *27*(5), 945–958. http://doi.org/10.1162/jocn

Dugué, L., McLelland, D., Lajous, M., & VanRullen, R. (2015). Attention searches nonuniformly in space and in time. *Proceedings of the National Academy of Sciences*, *112*(49), 15214–15219. http://doi.org/10.1073/pnas.1511331112

Dugué, L., Merriam, E. P., Heeger, D. J., & Carrasco, M. (2017). Specific Visual Subregions of TPJ Mediate Reorienting of Spatial Attention. *Cerebral Cortex*, (Table

1), 1–16. http://doi.org/10.1093/cercor/bhx140

Dugué, L., & Vanrullen, R. (2014). The dynamics of attentional sampling during visual search revealed by Fourier analysis of periodic noise interference. *Journal of Vision*, *14*, 1–15. http://doi.org/10.1167/14.2.11.doi

Dugué, L., Xue, A. M., & Carrasco, M. (2017). Distinct perceptual rhythms for feature and conjunction searches. *Journal of Vision*, *17*(3), 22. http://doi.org/10.1167/17.3.22

Faivre, N., Arzi, A., Lunghi, C., & Salomon, R. (2017). Consciousness is more than meets the eye: a call for a multisensory study of subjective experience†. *Neuroscience of Consciousness*, *3*(1), 1–8. http://doi.org/10.1093/nc/nix003

Fiebelkorn, I. C., Saalmann, Y. B., & Kastner, S. (2013). Rhythmic sampling within and between objects despite sustained attention at a cued location. *Current Biology*, *23*(24), 2553–2558. http://doi.org/10.1016/j.cub.2013.10.063

Frey, J. N., Ruhnau, P., & Weisz, N. (2015). Not so different after all: The same oscillatory processes support different types of attention. *Brain Research*, *1626*(March), 183–197. http://doi.org/10.1016/j.brainres.2015.02.017

Fujisaki, W., & Nishida, S. (2005). Temporal frequency characteristics of synchrony-asynchrony discrimination of audio-visual signals. *Experimental Brain Research*, *166*(3–4), 455–464. http://doi.org/10.1007/s00221-005-2385-8

Fujisaki, W., & Nishida, S. (2010). A common perceptual temporal limit of binding synchronous inputs across different sensory attributes and modalities. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1692), 2281–90. http://doi.org/10.1098/rspb.2010.0243

Guzman-Martinez, E., Ortega, L., Grabowecky, M., Mossbridge, J., & Suzuki, S. (2012). Interactive coding of visual spatial frequency and auditory amplitude-modulation rate. *Current Biology*, *22*(5), 383–388. http://doi.org/10.1016/j.cub.2012.01.004

Hanslmayr, S., Volberg, G., Wimber, M., Dalal, S. S., & Greenlee, M. W. (2013). Prestimulus oscillatory phase at 7 Hz gates cortical information flow and visual perception. *Current Biology*, *23*(22), 2273–2278. http://doi.org/10.1016/j.cub.2013.09.020

Hartcher-O'Brien, J., Talsma, D., Adam, R., Vercillo, T., Macaluso, E., & Noppeney, U. (2016). The Curious Incident of Attention in Multisensory Integration: Bottom-up vs. Top-down. *Multisensory Research*, *29*(6–7), 557–583. http://doi.org/10.1163/22134808-00002528

Herbst, S. K., & Landau, A. N. (2016). Rhythms for cognition: The case of temporal processing. *Current Opinion in Behavioral Sciences*, *8*, 85–93. http://doi.org/10.1016/j.cobeha.2016.01.014

Hogendoorn, H. (2016). Voluntary Saccadic Eye Movements Ride the Attentional Rhythm. *Journal of Cognitive Neuroscience*, *28*(10), 1625–1635. http://doi.org/10.1162/jocn

Holcombe, A. O. (2009). Seeing slow and seeing fast: two limits on perception. *Trends in Cognitive Sciences*, *13*(5), 216–221. http://doi.org/10.1016/j.tics.2009.02.005

Huang, Y., Chen, L., & Luo, H. (2015). Behavioral Oscillation in Priming: Competing Perceptual Predictions Conveyed in Alternating Theta-Band Rhythms. *Journal of Neuroscience*, *35*(6), 2830–2837. http://doi.org/10.1523/JNEUROSCI.4294-14.2015

Jack, B. N., & Hacker, G. (2014). Predictive Coding Explains Auditory and Tactile Influences on Vision during Binocular Rivalry. *Journal of Neuroscience*, *34*(19), 6423–6424. http://doi.org/10.1523/JNEUROSCI.1040-14.2014

Kanai, R., Tsuchiya, N., & Verstraten, F. a J. (2006). The Scope and Limits of Top-Down Attention in Unconscious Visual Processing. *Current Biology*, *16*(23), 2332–2336. http://doi.org/10.1016/j.cub.2006.10.001

Kang, M.-S., & Blake, R. (2005). Perceptual synergy between seeing and hearing revealed during binocular rivalry. *Psichologija*, *32*, 7–15.

Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: two distinct brain processes. *Trends in Cognitive Sciences*, *11*(1), 16–22. http://doi.org/10.1016/j.tics.2006.10.012

Koessler, L., Maillard, L., Benhadid, A., Vignal, J. P., Felblinger, J., Vespignani, H., & Braun, M. (2009). Automated cortical projection of EEG sensors: Anatomical correlation via the international 10–10 system. *NeuroImage*, *46*(1), 64–72. http://doi.org/10.1016/j.neuroimage.2009.02.006

Lakatos, P., O'Connell, M. N., Barczak, A., Mills, A., Javitt, D. C., & Schroeder, C. E. (2009). The Leading Sense: Supramodal Control of Neurophysiological Context by Attention. *Neuron*, *64*(3), 419–430. http://doi.org/10.1016/j.neuron.2009.10.014

Lamme, V. A. F. (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences*, *7*(1), 12–18. http://doi.org/10.1016/S1364-6613(02)00013-X

Landau, A. N., & Fries, P. (2012). Attention samples stimuli rhythmically. *Current Biology*, *22*(11), 1000–1004. http://doi.org/10.1016/j.cub.2012.03.054

Landau, A. N., Schreyer, H. M., Van Pelt, S., & Fries, P. (2015). Distributed Attention Is Implemented through Theta-Rhythmic Gamma Modulation. *Current Biology*, *25*(17), 2332–2337. http://doi.org/10.1016/j.cub.2015.07.048

Lin, Z., & He, S. (2009). Seeing the invisible: The scope and limits of unconscious processing in binocular rivalry. *Progress in Neurobiology*, *87*(4), 195–211. http://doi.org/10.1016/j.pneurobio.2008.09.002

Lumer, E. D., Friston, K. J., & Rees, G. (1998). Neural correlates of perceptual rivalry in the human brain. *Science (New York, N.Y.)*, *280*(5371), 1930–1934. http://doi.org/10.1126/science.280.5371.1930

Lunghi, C., & Alais, D. (2013). Touch Interacts with Vision during Binocular Rivalry with a Tight Orientation Tuning. *PLoS ONE*, *8*(3), 1–8. http://doi.org/10.1371/journal.pone.0058754

Lunghi, C., Binda, P., & Morrone, M. C. (2010). Touch disambiguates rivalrous perception at early stages of visual analysis. *Current Biology : CB*, *20*(4), R143-4. http://doi.org/10.1016/j.cub.2009.12.015

Lunghi, C., Morrone, M. C., & Alais, D. (2014). Auditory and tactile signals combine to influence vision during binocular rivalry. *The Journal of Neuroscience*, *34*(3), 784–92. http://doi.org/10.1523/JNEUROSCI.2732-13.2014

Maier, A., Panagiotaropoulos, T. I., Tsuchiya, N., & Keliris, G. A. (2012). Introduction to research topic–binocular rivalry: a gateway t o studying consciousness. *Binocular Rivalry: A Gateway to Consciousness*, *6*(7), 5.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. http://doi.org/10.1016/j.jneumeth.2007.03.024

Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M., & Ro, T. (2009). To See or Not to See: Prestimulus Phase Predicts Visual Awareness. *Journal of Neuroscience*, *29*(9), 2725–2732. http://doi.org/10.1523/JNEUROSCI.3963-08.2009

Merikle, P. M., & Joordens, S. (1997). Parallels between Perception without Attention and Perception without Awareness. *Consciousness and Cognition*, *6*(January 2014), 219–236. http://doi.org/10.1006/ccog.1997.0310

Mitchell, J. F., Stoner, G. R., & Reynolds, J. H. (2004). Object-based attention determines dominance in binocular rivalry. *Nature*, *429*(6990), 410–413. http://doi.org/10.1038/nature02584

O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual

consciousness. *Behavioral and Brain Sciences*, *24*(5), 939–973. http://doi.org/10.1017/S0140525X01000115

O'Shea, R. P., & Blake, R. (1986). Dichoptic temporal frequency differences do not lead to binocular rivalry. *Perception & Psychophysics*, *39*(1), 59–63. http://doi.org/10.3758/BF03207584

Paffen, C. L. E., & Alais, D. (2011). Attentional Modulation of Binocular Rivalry. *Frontiers in Human Neuroscience*, *5*(September), 1–10. http://doi.org/10.3389/fnhum.2011.00105

Posner, M. I. (1994). Attention: the mechanisms of consciousness. *Proceedings of the National Academy of Sciences*, *91*(16), 7398–7403.

Posner, M. I. (2012). Attentional networks and consciousness. *Frontiers in Psychology*, *3*(MAR), 1–4. http://doi.org/10.3389/fpsyg.2012.00064

Roy, A. V., Jamison, K. W., He, S., Engel, S. A., & He, B. (2017). Deactivation in the posterior mid-cingulate cortex reflects perceptual transitions during binocular rivalry: Evidence from simultaneous EEG-fMRI. *NeuroImage*, *152*(October 2016), 1–11. http://doi.org/10.1016/j.neuroimage.2017.02.041

Song, K., Meng, M., Chen, L., Zhou, K., & Luo, H. (2014). Behavioral Oscillations in Attention: Rhythmic Pulses Mediated through Band. *Journal of Neuroscience*, *34*(14), 4837–4844. http://doi.org/10.1523/JNEUROSCI.4856-13.2014

Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences*, *14*(9), 400–410. http://doi.org/10.1016/j.tics.2010.06.008

Tamber-Rosenau, B. J., & Marois, R. (2016). Central attention is serial, but midlevel and peripheral attention are parallel—A hypothesis. *Attention, Perception, & Psychophysics*. http://doi.org/10.3758/s13414-016-1171-y

Tong, F., Meng, M., & Blake, R. (2006). Neural bases of binocular rivalry. *Trends in Cognitive Sciences*, *10*(11), 502–511. http://doi.org/10.1016/j.tics.2006.09.003

Tong, F., Nakayama, K., Vaughan, J. T., & Kanwisher, N. (1998). Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron*, *21*(4), 753–759. http://doi.org/10.1016/S0896-6273(00)80592-9

Tsuchiya, N., & Boxtel, J. J. Van. (2013). *Attention and consciousness in different senses*. http://doi.org/10.3389/978-2-88919-132-1

Van Atteveldt, N., Murray, M. M., Thut, G., & Schroeder, C. E. (2014). Multisensory integration: Flexible use of general operations. *Neuron*, *81*(6), 1240–1253. http://doi.org/10.1016/j.neuron.2014.02.044

van Boxtel, J. J. A., Knapen, T., Erkelens, C. J., & Van Ee, R. (2008). Removal of monocular interactions equates rivalry behavior for monocular, binocular, and stimulus rivalries. *Journal of Vision*, *813*(15), 1–17. http://doi.org/10.1167/8.15.13

van Boxtel, J. J., van Ee, R., & Erkelens, C. J. (2007). Dichoptic masking and binocular rivalry share common perceptual dynamics. *Journal of Vision*, *7*(14), 3.1-11. http://doi.org/10.1167/7.14.3

van Ee, R., van Boxtel, J. J. A., Parker, A. L., & Alais, D. (2009). Multisensory congruency as a mechanism for attentional control over perceptual selection. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(37), 11641–11649. http://doi.org/10.1523/JNEUROSCI.0873-09.2009

VanRullen, R. (2016a). Perceptual cycles. *Trends in Cognitive Sciences*, *20*(10), 723–735. http://doi.org/10.1167/15.12.1401

VanRullen, R. (2016b). Perceptual Rhythms. *Stevens' Handbook of Experimental Psychology*, 1–51.

VanRullen, R., Carlson, T., & Cavanagh, P. (2007). The blinking spotlight of attention.

Proceedings of the National Academy of Sciences, *104*(49), 19204–19209. http://doi.org/10.1073/pnas.0707316104

Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony. *Attention, Perception & Psychophysics*, *72*(4), 871–884. http://doi.org/10.3758/APP

Zhang, D., Hong, B., Gao, X., Gao, S., & Röder, B. (2011). Exploring steady-state visual evoked potentials as an index for intermodal and crossmodal spatial attention. *Psychophysiology*, *48*(5), 665–675. http://doi.org/10.1111/j.1469-8986.2010.01132.x

Zoefel, B., & VanRullen, R. (2017). Oscillatory mechanisms of stimulus processing and selection in the visual and auditory systems: State-of-the-art, speculations and suggestions. *Frontiers in Neuroscience*, *11*(MAY), 1–13. http://doi.org/10.3389/fnins.2017.00296