



The advantage of ambiguity? Enhanced neural responses to multi-stable percepts correlate with the degree of perceived instability

Benjamin J. Dyson*

Department of Psychology, Ryerson University, Toronto, ON, Canada

Edited by:

Idan Segev, The Hebrew University of Jerusalem, Israel

Reviewed by:

Lutz Jäncke, University of Zurich, Switzerland

Burkhard Pleger, Max Planck Institute for Human Cognitive and Brain Sciences, Germany

***Correspondence:**

Benjamin J. Dyson, Department of Psychology, Ryerson University, 350 Victoria Street, Toronto, ON M5B 2K3, Canada.
e-mail: ben.dyson@psych.ryerson.ca

Artwork can often pique the interest of the viewer or listener as a result of the ambiguity or instability contained within it. Our engagement with uncertain sensory experiences might have its origins in early cortical responses, in that perceptually unstable stimuli might preclude neural habituation and maintain activity in early sensory areas. To assess this idea, participants engaged with an ambiguous visual stimulus wherein two squares alternated with one another, in terms of simultaneously opposing vertical and horizontal locations relative to fixation (i.e., stroboscopic alternating motion; von Schiller, 1933). At each trial, participants were invited to interpret the movement of the squares in one of five ways: traditional vertical or horizontal motion, novel clockwise or counter-clockwise motion, and, a free-view condition in which participants were encouraged to switch the direction of motion as often as possible. Behavioral reports of perceptual stability showed clockwise and counter-clockwise motion to possess an intermediate level of stability compared to relatively stable vertical and horizontal motion, and, relatively unstable motion perceived during free-view conditions. Early visual evoked components recorded at parietal–occipital sites such as C1, P1, and N1 modulated as a function of visual intention. Both at a group and individual level, increased perceptual instability was related to increased negativity in all three of these early visual neural responses. Engagement with increasingly ambiguous input may partly result from the underlying exaggerated neural response to it. The study underscores the utility of combining neuroelectric recording with the presentation of perceptually multi-stable yet physically identical stimuli, in revealing brain activity associated with the purely internal process of interpreting and appreciating the sensory world that surrounds us.

Keywords: ambiguous stimuli, perceptual stability, perceptual intention, visual evoked responses, auditory evoked responses

INTRODUCTION

There is much that can be learnt about art by applying what we know about our perceptual and cognitive systems (e.g., Arnheim, 1974; Gombrich, 1977; Gregory, 1997; Solso, 2003). Although the phenomenological gap between the eventual experience of art and the initial sensory processes that underlie it can be significant (Dyson, 2009), common ground has been established between the gallery and the laboratory. For example, the later works of Piet Mondrian are generally ascribed artistic value, in addition to being composed of visual elements basic enough to allow for systematic experimentation. By manipulating the orientation (Latto et al., 2000), line spacing (McManus et al., 1993; Wolach and McHale, 2005), and color distribution (Locher et al., 2005) of the originals, it is possible to establish the extent to which Mondrian had established the “correct” organization of these features in accordance with a universal esthetic. Another way researchers have attempted to bridge the divide between art and science is to take advantage of the observation that differential perceptual and esthetic experiences may be derived from the same input: one only need consider proponents of optical art such as Bridget Riley and Victor Vasarely (Martinez-Conde and Macknik,

2010) to appreciate how the interpretation of ambiguous sensory information relies on an interaction between bottom-up and top-down processes (e.g., Dyson and Cohen, 2010). In this regard, examining the neurological indices associated with ambiguous experience has been an important step in understanding how purely internal processes such as perceptual interpretation might be expressed in the brain. One of the major advantages of this kind of approach is the way in which identical sensory stimulation can give rise to radically different perceptions, thereby revealing the neural correlates peculiar to conscious interpretation whilst at the same time controlling for differences in low-level stimulus processing (Jackson et al., 2008). One particularly intriguing possibility outlined by Zeki (2006) is that sensory information imbued with ambiguity might facilitate an esthetic experience by promoting brain activity and the active process of environmental interpretation (see Berlyne, 1971, for a historical antecedent of this idea). The present study will examine the extent to which variations in perceptual interpretation for the same sensory information map onto early evoked responses, and, how this potential advantage of ambiguity might be expressed in neural terms.

With respect to the kinds of ambiguous stimuli deployed in the laboratory, images as diverse as the Necker cube, the Rubin face/vase, the old/young woman, ambiguous cheetahs, dot lattices, overlaid gratings, spinning wheels, rotating spheres, and stereograms (e.g., Portas et al., 2000; Leopold et al., 2002; Sterzer et al., 2002; Gepshtein and Kubovy, 2005; see Windmann et al., 2006, for further examples) have all been used to inform how any one particular conscious percept arises from the numerous interpretations available to us. Moreover, the application of fMRI (e.g., Kleinschmidt et al., 1998; Portas et al., 2000), EEG/ERP (e.g., Kornmeier and Bach, 2004, 2005; Pitts et al., 2007), and MEG (e.g., Struber and Herrmann, 2002; Kaneoke et al., 2009) technologies have all contributed to our understanding of both the spatial and temporal positioning of perceptual interpretation relative to early sensory activity as well as later cognitive and response-based processes. For example, in terms of brain regions involved in the production of conscious interpretation, Rees et al. (2002) argues that whereas V1 activity does not modulate according to the content of subjective experience, activation of the ventral visual cortex seems to be a necessary (though not sufficient) condition for the advent of consciousness.

The application of ERP and MEG has been particularly useful in terms of revealing the latencies at which changes in perceptual interpretation are manifest. Both relatively early and relatively late time windows have been identified that distinguish between the experiences of perceptual change and perceptual maintenance in the absence of any low-level stimulus alteration. For example, with respect to an ambiguous Necker cube, perceptual change relative to perceptual maintenance has been marked by an increase in both positive-going and negative-going deflections 120 and 240 ms after stimulus onset, respectively, maximal at parietal and occipital scalp electrode sites (Kornmeier and Bach, 2004, 2005). Similar modulations have also been reported for other ambiguous figures such as the Rubin face/vase and Schroder's staircase (but not in an ambiguous Cheetah picture; Pitts et al., 2007). Further demarcations between perceptual stability and change have also been observed in later positive-going deflections around 300–400 ms after stimulus onset, with perceptual reversals again related to increased amplitude (Basar-Eroglu et al., 1993; Struber and Herrmann, 2002; Kornmeier and Bach, 2009). Therefore, one potential advantage of ambiguity at a neural level might be that the sensory system fails to habituate to input that has multiple interpretations. Enhanced neural responding may then provide a level of cortical engagement that facilitates the eventual development of an esthetic response (Zeki, 2006).

The current study considers this possibility using one particular type of ambiguous stimulus presentation known as stroboscopic alternating motion (after von Schiller, 1933; also known as “two-frame bi-stable apparent motion” or “bi-stable motion quartets”; Kohler et al., 2008). This particular stimulus consists of two alternating frames (see **Figure 1**), wherein two shapes (traditionally either squares or circles) are presented in both opposing vertical and horizontal locations from one another relative to fixation. In one frame (hereafter, V1), one shape is presented bottom-left whilst a second shape is presented top-right. In a second frame (hereafter, V2), one shape is presented top-left whilst a second shape is presented bottom-right. With repeated alternations of these two frames, two apparent motions have previously been

reported: vertical and horizontal, with the former motion characteristically being the easier of the two percepts to maintain when vertical and horizontal variation is of an identical physical magnitude (Chaudrun and Glaser, 1991; Kaneoke et al., 2009).

The aims of the present investigation are three fold. First, I will aim to increase the range of perceptual experience that can be derived from stroboscopic alternating motion by comparing the behavioral report of perceptual stability associated with the intentional production of traditional vertical and horizontal perceived motion (Kohler et al., 2008) with two hitherto unreported interpretations of this multi-stable percept: clockwise and counter-clockwise motion. In a previous study comparing a number of static ambiguous stimuli (i.e., Rubin face/vase, Necker cube) with a dynamic ambiguous “rotating” circle stimulus in which local elements alternated between gray and black, Windmann et al. (2006) reported participants had significant difficulty in perceiving changes between clockwise and counter-clockwise motion relative to the other types of ambiguous figure, although the reasons for this difficulty were not clear. In contrast, other research comparing the speed with which participants judged the motion of an ambiguously rotating cylinder to be clockwise or counter-clockwise with the interpretation of the Rubin face/vase image found little difference between the stimuli (Takei and Nishida, 2010). The clear benefit of comparing the difficulty with which rotational motion can be generated relative to vertical and horizontal motion using the current stimuli is that differences cannot be attributed to the use of physically different stimuli, nor to the differences between static and dynamic ambiguous images. In an additional attempt to further extend the range of perceptual experience in the current experiment, participants will also take part in fifth condition in which they intend to make the percept as unstable as possible by switching the direction of motion as often as they can (hereafter, “free-view”). Significant variations in the report of perceptual stability across these five conditions will support the view that stroboscopic alternating motion should be understood as a *multi-stable* rather than *bi-stable* stimulus, and help to evaluate the difficulty of intended rotational motion alongside potentially stable vertical and horizontal motion and potentially unstable free-view conditions with the same stimulus and, therefore, in the absence of low-level sensory differences.

Second I will consider the impact the degree of visual perceptual instability has upon the processing of auditory stimuli. Various accounts have been put forward for the independence or interdependence in the allocation of processing resource across different modalities (see Larsen et al., 2003). The empirical data continue to be equivocal (e.g., Dyson et al., 2005; Yucel et al., 2005; see Haroush et al., 2009, for a review), with recent evidence suggesting that the relationship between visual and auditory processing relies on a number of factors such as whether variation in visual task load is perceptual or memorial in nature (Muller-Gass and Schröger, 2007), the specific auditory event-related potentials interrogated (Harmony et al., 2000), as well as moment-to-moment fluctuations in visual task load (Haroush et al., 2009). In the current design, participants will attempt visual apparent motion both in the presence and absence of simultaneously alternating yet ignored auditory stimulation (A1, A2). By subtracting the neural response to trials in which participants receive just visual stimulation from trials in which participants receive both visual and auditory stimulation

[e.g., $(V1 + A1) - V1$; $(V2 + A2) - V2$], the resultant difference wave should provide some indication of the evoked responses generated by ignored auditory stimuli (although see Teder-Sälejärvi et al., 2002, to an objection to these kind of formulae in the context of cross-modal integration). If the relationship between auditory and visual perceptual processing is antagonist (e.g., Yücel et al., 2005), then those visual conditions related to high levels of perceptual instability should yield smaller auditory evoked responses.

Third I will examine the relationship between individual variation in the subjective report of perceptual stability and underlying visual neural activity. As previous electrophysiological data suggests (Basar-Eroglu et al., 1993; Struber and Herrmann, 2002; Kornmeier and Bach, 2004, 2005; Kornmeier and Bach, 2009), change in the interpretation of an ambiguous stimulus is reflected in increased amplitude at a number of points during visual processing. However, the literature is currently silent regarding the relationship between electrophysiological recording and variation in the interpretation of stroboscopic alternating motion. As such it is not clear whether stroboscopic alternating motion belongs to that sub-set of ambiguous images that modulate neural amplitude as a function of perceptual change (Pitts et al., 2007). Encouragingly, a recent study examining stroboscopic alternating motion under conditions of MEG recording reported differences between vertical and horizontal motion approximately 160 ms following frame onset (Kaneoke et al., 2009), placing one index of perceptual interpretation for this stimulus relatively early on in visual processing. In the current paradigm, perceptual switches will be indexed by the subjective rating of stability, with increased instability relating to an increased rate of perceptual switching. Therefore, if early visual evoked responses are predictive of perceptual performance then enhanced visual evoked responses should be associated with trials within which participants experience a high degree of perceptual instability.

MATERIALS AND METHODS

PARTICIPANTS

Twenty participants were initially run in the experiment. Four individuals had to be rejected on the basis of problems in EEG recording, and a further two individuals were rejected on the basis of their behavioral data (one due to their inability to perceive horizontal motion, and, one due to only being able to perceive intended motion on 42% of trials relative to a remaining group average of 93%). Consequently, the data from 14 participants were included in the final analyses (10 females; mean age = 25.86). All participants responded using their right hand, and all participants self-reported as being exclusively right-handed apart from one participant who self-reported as being able to use both left and right hand equally well. All participants received \$10 per hour or course credit for their involvement and provided informed consent prior to investigation according to the ethical guidelines established at Ryerson University. All participants reported normal hearing and normal or corrected-to-normal vision.

STIMULI AND APPARATUS

A white square of on-screen size of 18 mm × 18 mm and a gray fixation cross subtending 8 mm × 8 mm were generated. Visual stimuli were then differentially lateralized to generate one presentation (V1) in which one square was presented top-right and a second square was presented bottom-left relative to central fixation.

A second presentation (V2) constituted one square presented top-left and a second square presented bottom-right relative to central fixation. All horizontal and vertical variations were 44 mm away from central fixation, relative to square center. Further visual stimuli were generated constituting a clockwise rotated arrow (*clock*), a counter-clockwise rotated arrow (*counter*), an up-down pointing arrow (*vertical*), a left-right pointing arrow (*horizontal*), and, an up-down + left-right pointing arrow (*free*), all within a central on-screen area of 35 mm × 35 mm (see **Figure 1** for a schematic overview). Two 500 ms tones (low: 400 Hz, and, high: 800 Hz) were also created, each having 10 ms linear onset and offsets. Auditory stimuli were then differentially lateralized to generate one presentation (A1) in which a high tone was presented to the right ear concurrently with a low tone presented to the left ear. A second combined presentation was also generated (A2) in which a low tone was presented to the right ear concurrently with a high tone presented to the left ear (after Deutsch, 1974). Combined auditory presentation was calibrated monaurally at approximately 70 dB SPL(C) using Sennheiser HD202 headphones, and a Scosche SPL1000 sound level meter. Stimulus presentation and response collection was controlled by Presentation (Neurobehavioral Systems) and the experiment was completed in a quiet room.

DESIGN

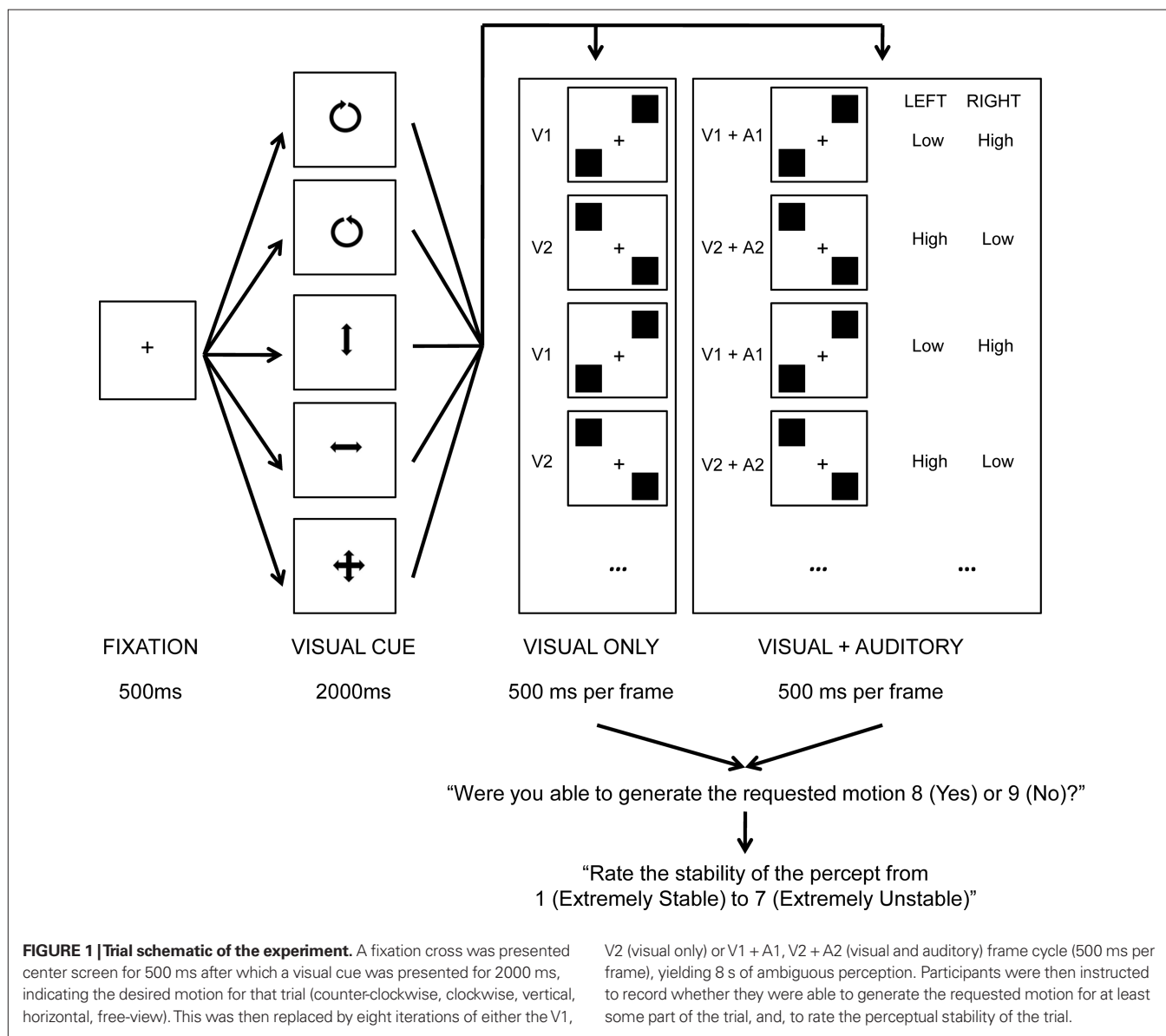
Ten experimental conditions were established involving the orthogonal combination of visual intention cue (clockwise, counter-clockwise, vertical, horizontal, free-view), with the presentation of auditory stimulation (sound, silence) concurrently with visual stimulation. Participants initially completed 10 practice trials (1 presentation × 10 conditions), followed by 3 blocks of 100 experimental trials (10 presentations × 10 conditions). Trial presentation was randomized across both condition and participant. Participants were encouraged to take breaks to avoid fatigue in between blocks, and to withhold responses to avoid fatigue within blocks.

PROCEDURE

Each trial began with the presentation of a black screen for 500 ms, subsequently replaced by a visual cue center screen indicating the desired interpretation of motion for 2000 ms. Participants were asked to prepare to interpret the forthcoming visual stimulation according to the cue, and to ignore auditory input if it was present. This was followed by an 8 s period in which visual, or, visual and auditory stimuli were presented, alternating across two frames presented for 500 ms each, thereby resulting in eight presentation of frame one (V1, or, V1 + A1) and eight presentations of frame two (V2, or, V2 + A2; see **Figure 1**). Participants were encouraged to focus on a central fixation throughout stimulus presentation. Following stimulus presentation, participants were asked whether or not they were able to generate the requested visual motion for at least part of the trial, and then, how stable the experienced percept was on a 7-point scale (1 = extremely stable, 7 = extremely unstable). Both post-trial prompts were on-screen until the participant responded.

ERP RECORDING

Electrical brain activity was continuously digitized using ActiView (Bio-Semi; Wilmington, NC, USA), with a band-pass filter of 0.16–100 Hz and a 1024 Hz sampling rate. Recordings made from



FPz, AFz, FC1, FCz, FC2, F1, Fz, F2, Cz, CPz, Pz, PO3, POz, PO4, Pz, Oz, M1, and M2 were stored for off-line analysis. Horizontal and vertical eye movements were recorded using channels placed at the outer canthi and at inferior orbits, respectively. Common mode sense (CMS) was taken from an independent electrode situated between P1 and PO3, while driven right leg (DRL) was situated between P2 and PO4. Data pre-processing was conducted using BESA 5.3 Research (MEGIS; Gräfelfing, Germany) using a 0.5 (6 db/oct; forward) to 30 (24 db/oct; zero phase) Hz band-pass filter. The contributions of both vertical and horizontal eye movements were reduced from the EEG record using the automated VEOG and HEOG artifact option in BESA. Only trials in which participants reported success in perceiving the intended motion for at least part of the trial were included in analyses, and the evoked response from first and last frames within each set of 16 were not included thereby avoiding anticipatory responses. This led to the

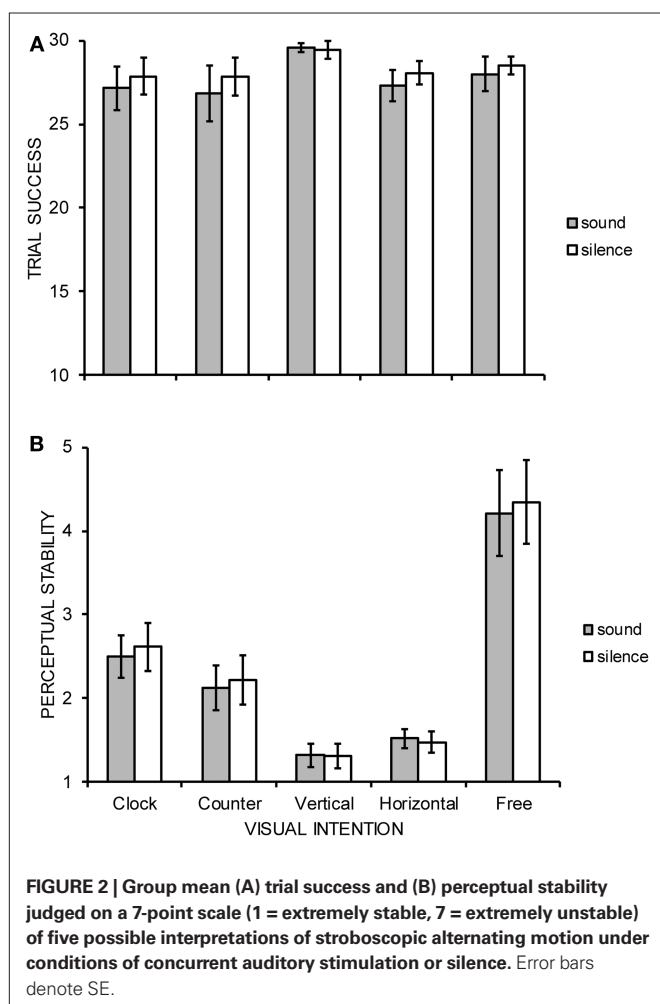
availability of a potential 210 epochs per frame per experimental condition. Individual epochs defined as 100 ms pre-stimulus baseline and 500 ms post stimulus activity were rejected on the basis of amplitude difference exceeding $75 \mu\text{V}$, gradient between consecutive time points exceeding $75 \mu\text{V}$, or, signal lower than $0.01 \mu\text{V}$, within any channel. For the purposes of the current analyses, data were collapsed across frame and serial position, averaged and finally baselined according to the pre-stimulus interval. Visual stimuli were analyzed according to an average reference, auditory stimuli were analyzed according to an average mastoid reference $[(M1 + M2)/2]$.

RESULTS

BEHAVIORAL DATA

Both the number of trials in which participants reported perceiving the requested motion for at least some part of the trial as indexed by the first behavioral measure, and, the subjective stability of

those successful trials as indexed by the second behavioral measure were subjected to separate, two-way within-participant ANOVAs constituting the factors of visual intention (clockwise, counter-clockwise, vertical, horizontal, free-view) and audition (sound, silence). All data were analyzed with Greenhouse–Geisser degrees of freedom correction and all significant interactions between factors were further explored using Tukey's HSD test ($p = 0.05$). The ANOVA data are summarized in **Figure 2**. The number of successful trials did not significantly differ across factors with participants indicating successful perceived motion on approximately 28 of 30 trials per condition: intention main effect [$F_{(1.7, 21.7)} = 1.32$, $MSE = 17.23$, $p = 0.284$], audition main effect [$F_{(1, 13)} = 2.01$, $MSE = 4.60$, $p = 0.179$], and, intention \times audition interaction [$F_{(3.0, 39.1)} = 1.02$, $MSE = 1.28$, $p = 0.394$]. A significant main effect of intention [$F_{(1.8, 23.3)} = 18.11$, $MSE = 2.16$, $p < 0.001$] upon perceptual stability was revealed, indicating free-view (4.28) to be more significantly unstable than all other conditions, and, clockwise (2.56) to be significantly more unstable than both vertical (1.32) and horizontal (1.50) perceived motion. Neither the main effect of audition [$F_{(1, 13)} = 2.39$, $MSE = 0.05$, $p = 0.146$] nor the interaction between intention and audition reached statistical significance [$F_{(2.6, 33.2)} = 0.81$, $MSE = 0.08$, $p = 0.478$].



AUDITORY ERP RESPONSE

Auditory evoked responses were generated by calculating the difference wave between silence and sound trials (see **Figure 3A**), revealing maximal activity at fronto-central sites (Fz, FC1, FCz, FC2, Cz). P1 (maximal positivity between 60 and 160 ms post stimulus onset), N1 (maximal negativity between 110 and 210 ms post stimulus onset), and, P2 (maximal positivity between 190 and 290 ms post stimulus onset) peak latencies were sought and mean amplitudes for auditory components were quantified in terms of a 30 ms time window (15 ms either side of peak latency) and analyzed with respect to visual intention only. With respect to P1, N1, and P2, neither peak latencies [$F_{(2.7, 35.9)} = 0.34$, $MSE = 428$, $p = 0.885$; $F_{(3.1, 39.9)} = 0.40$, $MSE = 232$, $p = 0.760$; $F_{(2.8, 36.8)} = 0.65$, $MSE = 809$, $p = 0.581$, respectively] nor mean amplitudes [$F_{(3.2, 41.00)} = 1.04$, $MSE = 0.65$, $p = 0.388$; $F_{(3.1, 32.58)} = 0.14$, $MSE = 0.80$, $p = 0.942$; $F_{(3.4, 44.8)} = 1.91$, $MSE = 0.62$, $p = 0.134$, respectively] differed significantly as a function of visual intention.

VISUAL ERP RESPONSE

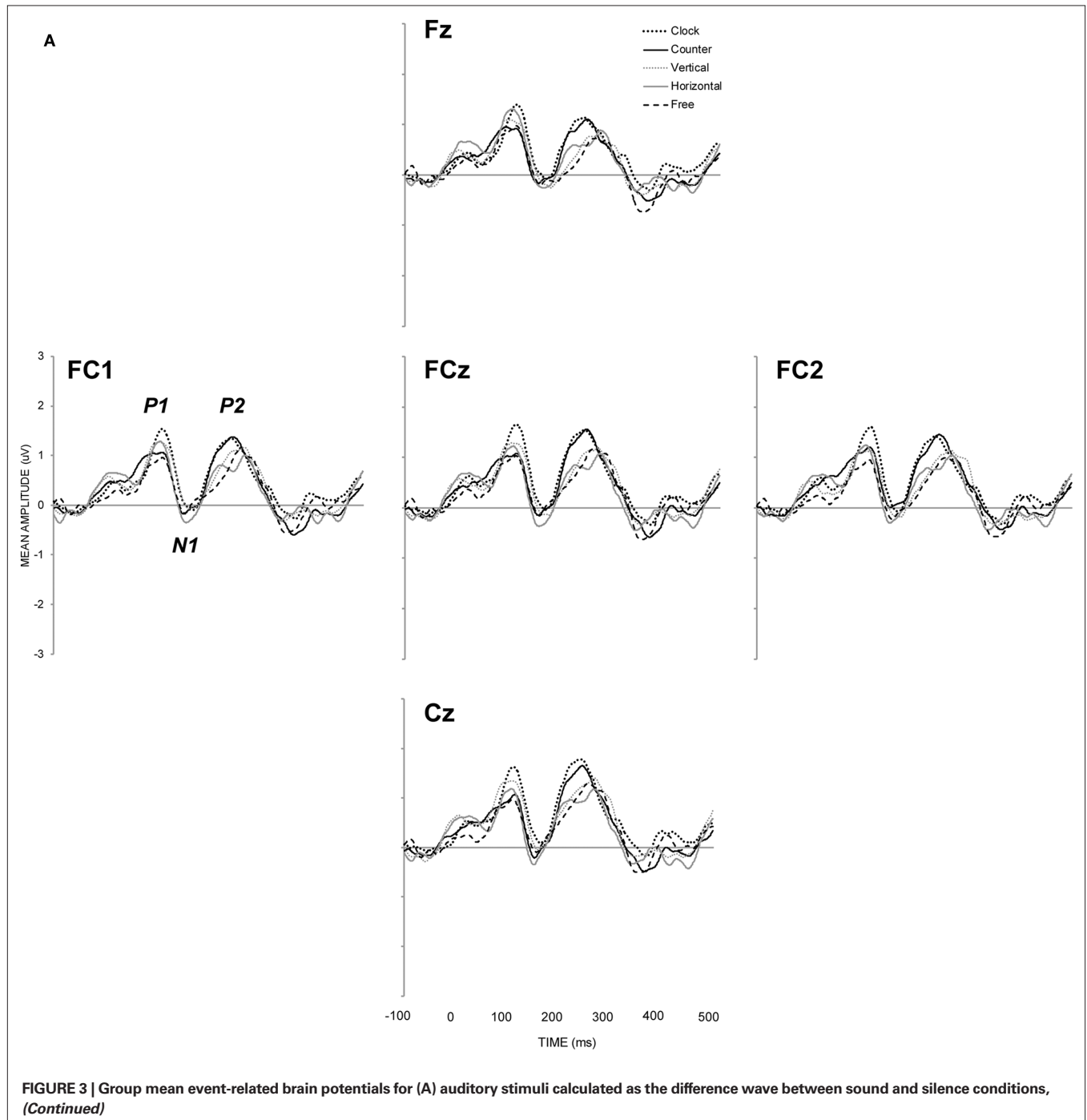
Figure 3 shows neural responses to visual stimulation at parietal–occipital electrode sites with (**Figure 3B**) and without (**Figure 3C**) the concurrent presentation of auditory stimulation. C1 (maximal negativity between 25 and 125 ms post stimulus onset), P1 (maximal positivity between 70 and 170 ms post stimulus onset), and, N1 (maximal negativity between 130 and 270 ms post stimulus onset) peak latencies were sought across sites showing maximal activity (PO3, POz, PO4, and Oz). Mean amplitudes for C1, P1, and N1 were again quantified according to a 30 ms time window. Visual ERP components were subjected to the same two-way, repeated measures ANOVA as in behavioral analyses, the results of which are shown in **Table 1**. In **Figure 4**, the left hand panels show mean amplitudes across the 10 conditions of interest and the right hand panels show the group correlation between mean amplitude and perceptual stability.

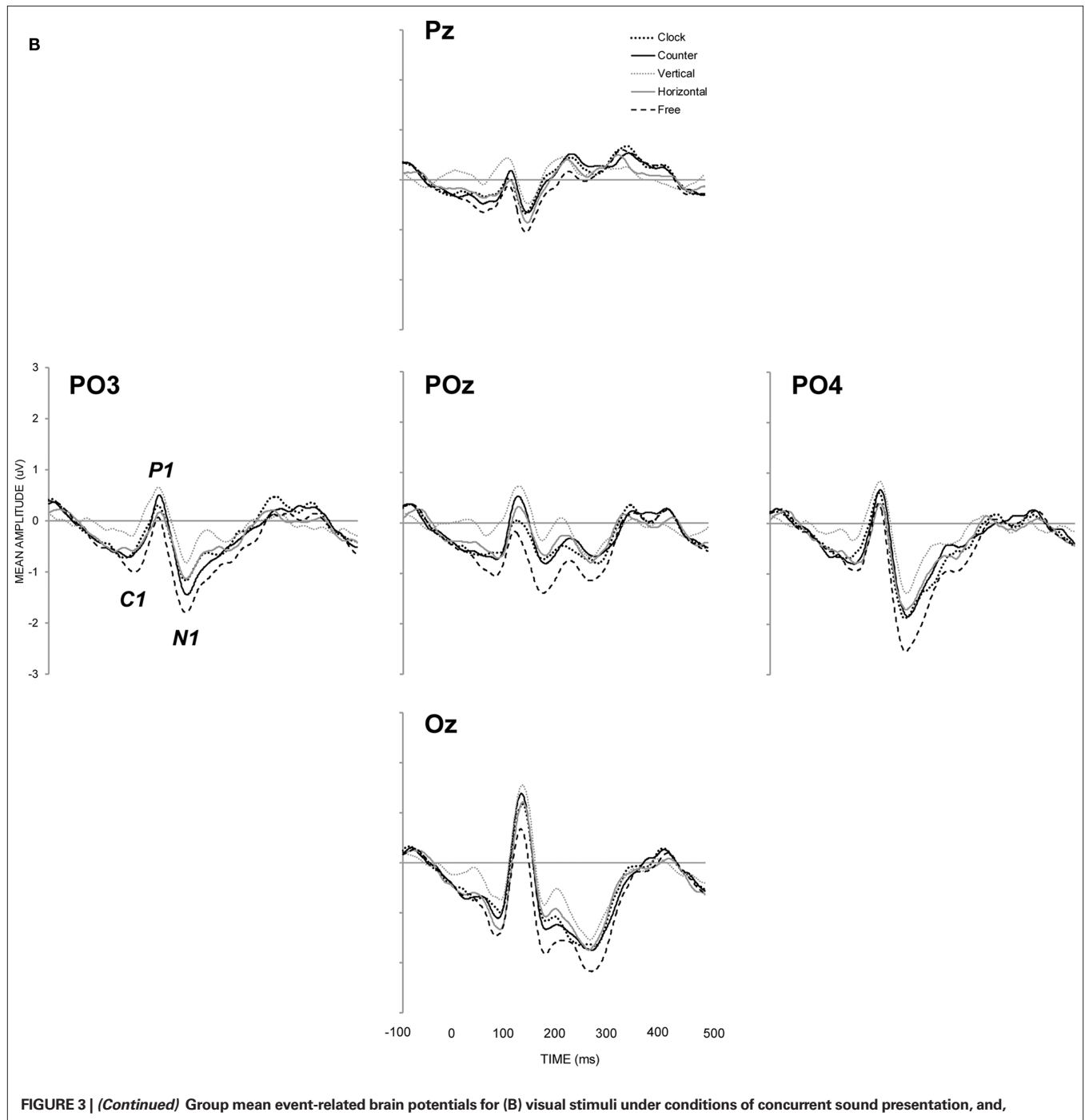
C1 (71 ms) mean amplitude revealed a significant main effect of intention ($p = 0.001$), with the free-view condition generating larger negativity ($-0.95 \mu\text{V}$) than vertical ($-0.31 \mu\text{V}$) and horizontal ($-0.58 \mu\text{V}$) motion, and, clockwise ($-0.70 \mu\text{V}$) and counter-clockwise ($-0.66 \mu\text{V}$) motion generating larger negativity than vertical motion. Individual correlations between mean amplitude and perceptual stability were calculated for each participant, and the average correlation was significantly different from zero [$r = -0.45$; $t(13) = 7.17$, $p < 0.001$], in that larger negativity was associated with increase perceptual instability. P1 (118 ms) mean amplitude revealed no significant group differences with respect to visual intention or audition, but the average correlation between mean amplitude and perceptual stability was revealed to be significant [$r = -0.28$; $t(13) = 2.36$, $p = 0.034$], indicating that smaller positivity was associated with increased perceptual instability. N1 (190 ms) mean amplitude revealed significant main effects of both intention ($p = 0.001$) and audition ($p = 0.004$). In terms of visual intention, free-view ($-1.81 \mu\text{V}$) trials generated significantly larger negativity than clockwise ($-1.31 \mu\text{V}$), counter-clockwise ($-1.32 \mu\text{V}$), horizontal ($-1.06 \mu\text{V}$), and, vertical ($-0.82 \mu\text{V}$) intended motion trials. The pairwise comparison between counter-clockwise motion

and vertical motion was also significant. Individual correlations between mean amplitude and perceptual stability once again revealed a significant negative correlation [$r = -0.50$; $t(13) = 5.73$, $p < 0.001$], indicating that mean amplitude increased as perceptual instability increased. Visual N1 was also larger under conditions of silence ($-1.47 \mu\text{V}$) relative to sound presentation ($-1.06 \mu\text{V}$), possibly due to the spatio-temporal interactions between parietal and occipital visual N1 and fronto-central auditory P2 components (Vidal et al., 2008).

DISCUSSION

Electrophysiological and behavioral measures of performance were collected as participants engaged with stroboscopic alternating motion (von Schiller, 1933) under a variety of different intended visual motions, and, under conditions of concurrent auditory stimulus presentation or silence. The data supported the idea that stroboscopic alternating motion is a multi-stable rather than bi-stable stimulus, in the revelation of two hitherto unreported interpretations: clockwise and counter-clockwise

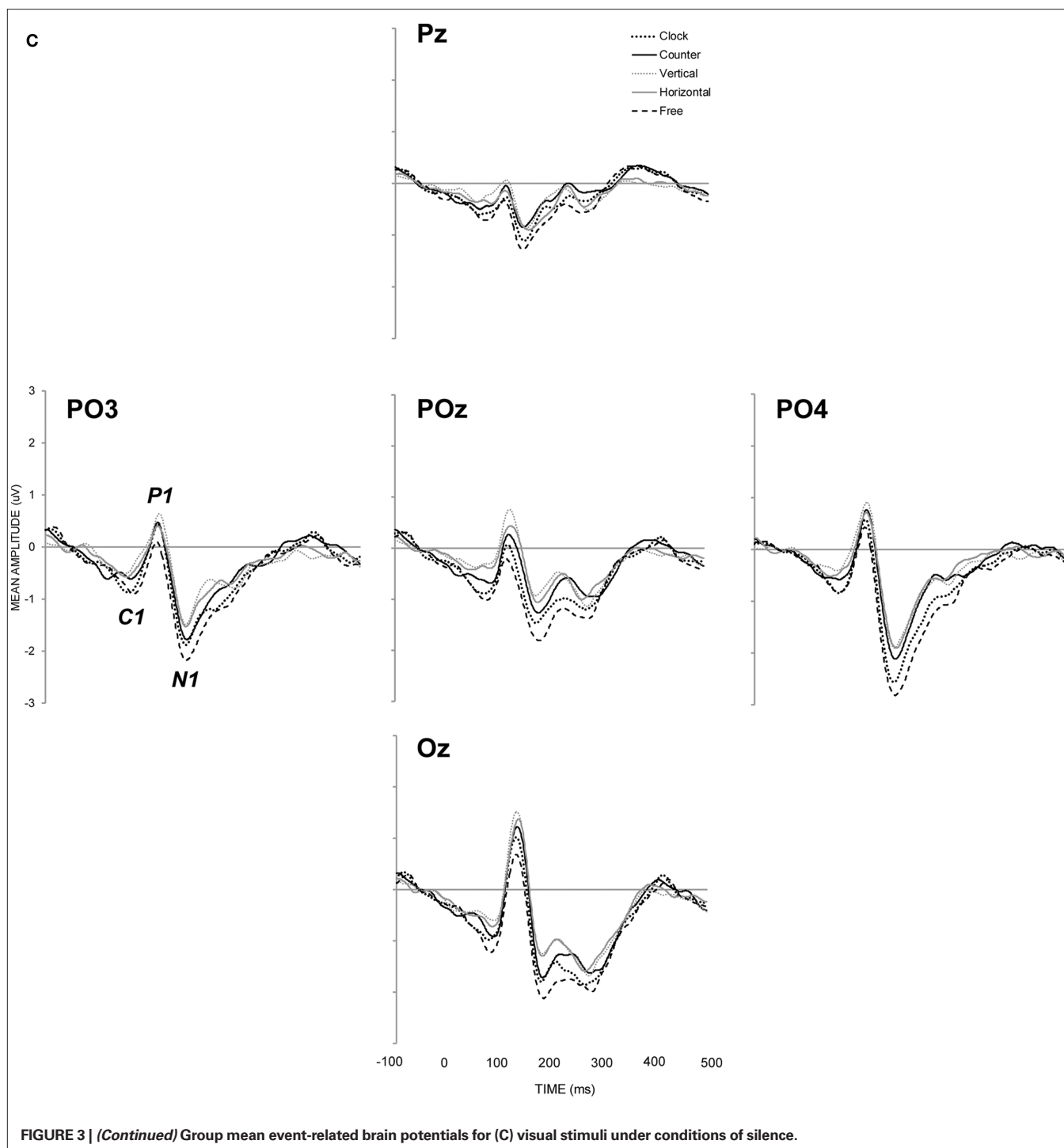




rotation. Behavioral reports of perceptual stability showed rotational motion to yield less perceptual stability than intended vertical or horizontal motion, but greater perceptual stability than intentionally unstable motion (free-view condition). Early cortical responses to ignored auditory stimuli observed at fronto-central sites did not seem to modulate as a function of visual intention. However, behavioral reports of perceptual stability were correlated with early visual exogenous responses recorded

from parietal–occipital sites (C1, P1, and N1) both at a group and individual level, in that increased negativity reflected increased perceptual instability.

In terms of the relative ease with which different motions could be imposed on the stroboscopic display, the current data echoed the previous observation that when vertical and horizontal variation is of an identical physical magnitude, vertical motion is perceived more readily than horizontal motion (Chaudrun and Glaser, 1991;



Kaneoke et al., 2009). This was represented by significant pairwise comparisons in the electrophysiological data, in that C1 and N1 amplitudes were reduced for vertical (but not horizontal) trials relative to counter-clockwise trials. While there are a number of possible explanations for this effect (see Chaudrun and Glaser, 1991, for a discussion), it would perhaps be important to rule out the quotidian explanation that since most contemporary experiments into stroboscopic motion take place on landscape oriented computer

monitors, vertical motion brings the shapes closer to the boundary of the monitor than does horizontal motion, thereby potentially creating increased perceived distance in the former case. This could easily be achieved by turning the monitor on its side in future investigations.

The current experiment also presents evidence to suggest that two new interpretations of the previously *bi-stable* stroboscopic stimulus are available: namely, clockwise and counter-clockwise motion. The subjective report of perceptual stability suggests that

Table 1 | Summary of separate ANOVAs for peak latency and mean amplitude for visual C1, P1, and N1 components across parietal-occipital (PO3, POz, PO4, Oz) electrodes.

ERP component								
Factor	Peak latency				Mean amplitude			
	df	F	MSE	p	df	F	MSE	p
C1								
I	3.1, 39.8	0.92	365	0.443	2.8, 37.1	7.07	0.30	0.001
A	1, 13	<0.01	296	0.984	1, 13	0.79	0.08	0.390
I × A	3.4, 44.2	2.20	166	0.094	2.2, 29.6	1.57	0.27	0.223
P1								
I	2.8, 36.4	2.18	170	0.111	2.5, 33.1	2.94	0.95	0.055
A	1, 13	3.51	91	0.084	1, 13	0.56	0.33	0.468
I × A	2.7, 36.2	1.21	45	0.320	3.4, 44.3	0.85	0.17	0.488
N1								
I	1.5, 20.1	0.96	280	0.380	2.1, 28.4	8.86	0.78	0.001
A	1, 13	1.26	261	0.282	1, 13	12.10	0.50	0.004
I × A	2.5, 33.1	0.56	121	0.619	2.6, 34.0	0.99	0.23	0.401

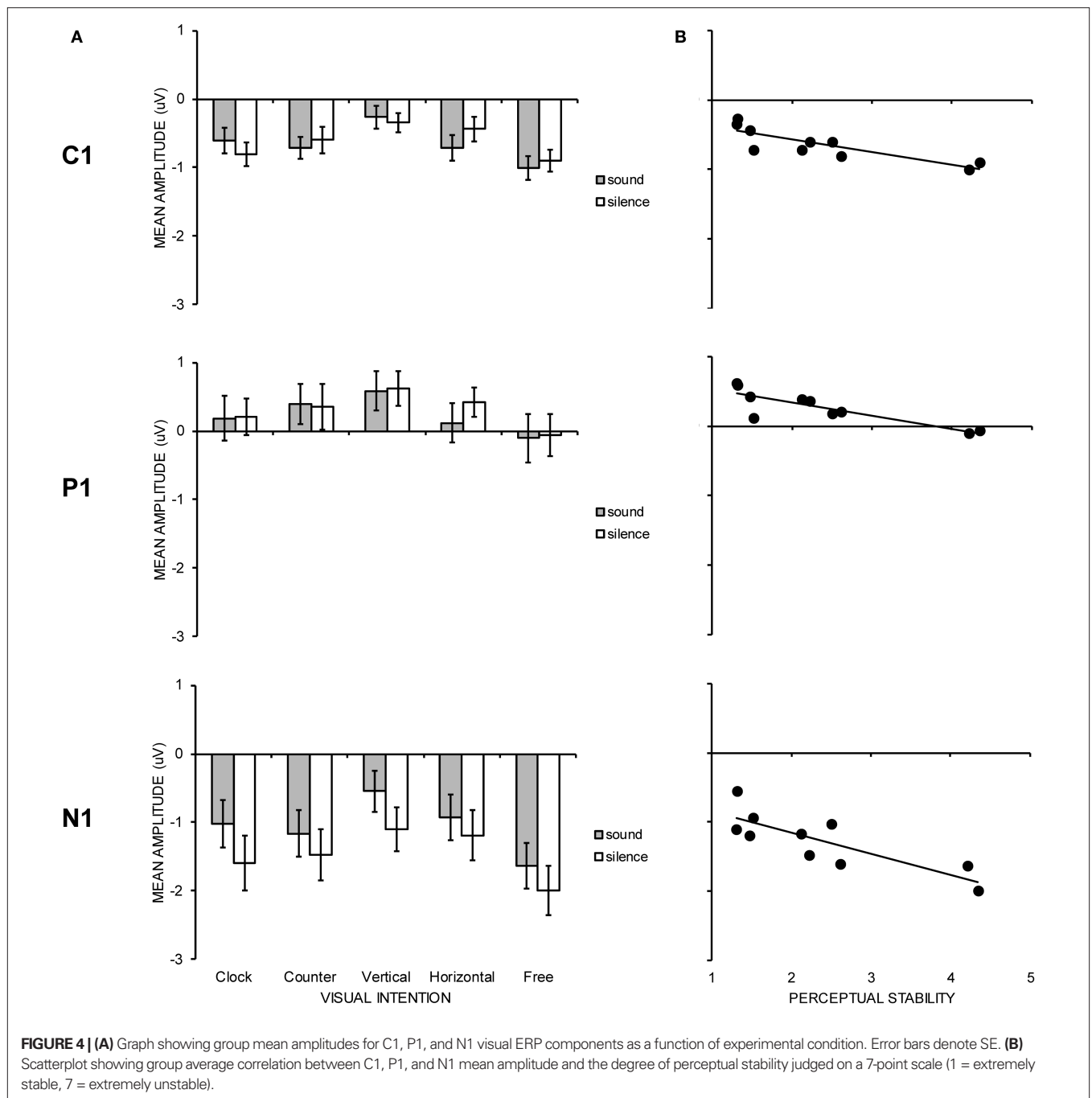
I, intention; A, audition. Statistically significant effects in bold.

these rotational motions achieved intermediate stability: significantly less stable than either vertical or horizontal motion but significantly more stable than the perception experienced under free-view conditions. Moreover, the difficulties in maintaining rotational motion cannot be attributed to the use of different stimuli or the use of static versus dynamic stimuli, as in previous studies (Windmann et al., 2006; Takei and Nishida, 2010). Further work would help to elucidate the mechanisms involved with rotational motion (cf., Jackson et al., 2008; Liesefeld and Zimmer, 2011) and at least one interpretation of rotational motion is available using a frame-by-frame analysis of the current data. Specifically, it would be possible to evaluate the idea that part of the reason for increased perceptual instability during intended rotational motion is that participants were simply mimicking clockwise and counter-clockwise rotation by recombining vertical and horizontal movements across different frames, thereby instigating a predictable perceptual switch on every trial. With respect to **Figure 1**, clockwise rotation could be mimicked by intending a vertical movement at V1 followed by a horizontal movement at V2, whereas a counter-clockwise rotation could be mimicked by intending a horizontal movement at V1 followed by a vertical movement at V2. If participants generated this pseudo-circular motion by combining vertical and horizontal movement, then a double dissociation should be predicted: the clockwise condition should be more similar to the vertical condition and the counter-clockwise condition should be more similar to horizontal condition in V1, and these associations should be reversed in V2.

With respect to the relationship between resource allocation across visual and auditory processing (Haroush et al., 2009), the data must remain silent as no significant differences in auditory evoked response were observed as a function of visual intention. **Figure 3A** hints at differential auditory P2 activity as a function of rotational versus non-rotational motion, but this failed to reach

statistical significance. Support for the independence between visual and auditory resource cannot be based on support of the null hypothesis as in the current experiment, as it is always possible that an effect in audition could be observed by further increasing the difficulty of the visual task (Valtonen et al., 2003). As above, it is perhaps worth considering what a more detailed analysis might reveal in terms of the interactions between auditory and visual stimulation. The auditory stimuli selected in this experiment had weak synesthetic (Martino and Marks, 2001) relations with the visual stimuli on a frame-by-frame basis, in that the vertical positions of the squares were assigned a high or low pitch value, and the horizontal positions of the squares were assigned to left or right headphone presentation (cf., Mudd, 1963). In this regard, concomitant auditory stimulation for V1 (high right square with low left square) was a high pitched tone in the right ear and a low pitched tone in the left ear (A1), whilst for V2 (low right square with high left square) the auditory stimulation was a low pitched tone in the right ear and a high pitched tone in the left ear (A2). One of the reasons for pursuing these frame-by-frame interactions further is because the consistency between audio and visual information is potentially violated on half the trials as a result of an auditory *Octave Illusion* (Deutsch, 1974). In short, when participants are asked to report what they hear on the basis of the acoustic presentation described above, a common (although by no means sovereign) perception is of hearing a high pitched tone in the right ear for A1 but a low pitched tone in the left ear for A2. While conflicting account of the illusion have been put forward in terms of either suppression or fusion between ears (Chambers et al., 2004; Deutsch, 2004), it is clear that the illusion is only apparent on half the trials (A2). Therefore, in terms of congruency between auditory and visual presentation, V1 + A1 should provide a stronger sense of perceptual unity than V2 + A2. Consequently, an interaction between frame and visual evoked responses in the presence or absence of sound might provide evidence for the octave illusion even when participants are instructed to ignore the acoustic information.

The most critical finding however was the provision of clear evidence for the neural expression of perceptual instability. Using perceptual instability as an index of the frequency of perceptual switching in the current design, the data echo the previous findings of amplitude modulation as a function of perceptual switching in a number of other ambiguous figures (Basar-Eroglu et al., 1993; Struber and Herrmann, 2002; Kornmeier and Bach, 2009). Specifically, the data replicate the observation of a relatively late larger negativity for perceptual reversals, but not the observation of a relatively early larger positivity for perceptual reversals (Kornmeier and Bach, 2004, 2005; Pitts et al., 2007). Rather, the degree of perceptual instability across the 10 experimental conditions was characterized by increased negativity throughout the exogenous visual response, beginning around the time of C1 (approximately 70 ms) and returning to baseline around 400 ms after frame onset (see **Figures 3B,C**). Perhaps one reason why the current ERP data show an early and consistent increase in negativity during perceptually unstable trials, rather than both increases in negativity and positivity (Kornmeier and Bach, 2004, 2005), is a result of the averaging protocol. In contrast to previous studies where electrophysiological analyses were compared between times at which participants report an



intentional (or spontaneous) switch in perception compared to when the percept is stable (e.g., Kornmeier and Bach, 2009), the current analysis utilized an aggregated measure of performance whereby all frames within the trial (apart from the first V1 or V1 + A1 frame and the last V2 or V2 + A2 frame) were considered to increase signal-to-noise ratio. Therefore, it is possible that the significant modulations in neural responses reported above are the result of both sustained and transient evoked responses, although a purely sustained response account would struggle to argue for why significant neural modulation was observed at some time points and not others. Nevertheless, the recent

work of Liesefeld and Zimmer (2011) provides one example of a sustained evoked response that behaves in a manner analogous to the transient evoked responses observed here. In the context of mental rotation, these researchers showed that a negative-going, parietal slow-wave component increased as the amount of rotation increased, and also increased for counter-clockwise relative to clockwise rotational requests. The association between increased negativity and increased task complexity in their data provides a useful link with the association between increased negativity and increased perceptual instability in the current study. In order to tease apart the contributions of sustained

and transient evoked response in future work, the comparison between longer and shorter epochs within the same paradigm would be of particular benefit.

A final set of issues revolve around the observation that perceptual changes can be initiated under a number of conditions such as the induction of change by presenting disambiguated versions of the image (e.g., Sterzer and Kleinschmidt, 2006; Takei and Nishida, 2010), the intention to change as a function of task demands (e.g., Kohler et al., 2008), and, the spontaneous experience of change as a result of neural fatigue, attentional diversion, rate of stimulus delivery, and, eye movements (e.g., Baker and Graf, 2010). Consequently, there may be subtle differences between conditions that produce perceptual instability as a result of the *failure of stable intention* (i.e., vertical, horizontal, clockwise, counter-clockwise) and conditions that produce perceptual instability as a result of the *success of unstable intention* (i.e., free-view). Although the current data suggest that neural responses during free-view trials were only quantitatively rather than qualitatively different from all other trials, these subtly different styles of processing may reveal different patterns of neural performance, perhaps in pre-frontal areas implicated in holding particular perceptual interpretations in mind (Windmann et al., 2006). Moreover, it is also important to acknowledge variability in the ratings of perceptual stability between participants and consider how certain interpretations of the external world become more or less likely as a function of long-term experience. Distinct neural differences between visual artists and non-artists (e.g., Bhattacharya and Petsche, 2005; Kottlow et al., 2011), and, between musicians and non-musicians (e.g., Ohnishi et al., 2001; Schlaug, 2001) have been reported and it is a matter for future work as to how such sensory expertise might modulate the experience of ambiguous or multi-stable stimuli.

In sum, it would appear that the more ambiguous the sensory information, the more accentuated the brain's response to it. This current observation made with stroboscopic alternating motion

(von Schiller, 1933) appears to be a characteristic of some (but not all; Pitts et al., 2007) bi- or multi-stable stimuli tested in the laboratory (Windmann et al., 2006). It remains a question for future research whether the relationship between perceptual report and neural activity also apply to the sensory stimuli to which esthetic value is ascribed. For example, Ishai et al. (2007; see also Pepperell, 2011) discuss the consequences of examining indeterminate art and note that, relative to paintings containing recognizable objects, indeterminate paintings yielded longer response latencies (see also Dyson and Cohen, 2010). Therefore, ambiguity within artwork may have the net result of facilitating viewer or listener engagement, and it is interesting to consider how these effects might have their origins in neural activity. Further studies that integrate both brain and behavior metrics of esthetic reaction will be ideally positioned to evaluate how early neural responses might predict one's eventual reaction to artwork. Interacting with ambiguity then may provide a natural, neurological route to promoting deeper and richer reactions to the world (Zeki, 2006). As always, it will be the combination of bottom-up and top-down factors that generate the final perceptual experience: a carefully crafted ambiguous percept on the part of the artist combined with a willing to invoke ambiguity on the part of the perceiver may serve to engage the brain to the fullest extent. The likelihood that the viewer or listener has an esthetic response to their sensory world may then begin with an exaggerated cortical response to it.

AUTHOR'S NOTE

Benjamin J. Dyson is supported by an Early Researcher Award granted by Ontario Ministry of Research and Innovation. Data was presented at the 18th Annual Meeting of the Cognitive Neuroscience Society, 2nd–5th April, San Francisco, USA. The author would like to thank Idan Segev, Lutz Jäncke and Burkhard Pleger for comments on earlier versions of the manuscript.

REFERENCES

- Arnheim, R. (1974). *Art and Visual Perception: A Psychology of the Creative Eye*. Berkeley: University of California Press.
- Baker, D. H., and Graf, E. W. (2010). Extrinsic factors in the perception of bistable motion stimuli. *Vision Res.* 50, 1257–1265.
- Basar-Eroglu, C., Struber, D., Stadler, M., and Kruse, P. (1993). Multistable visual perception induces a slow positive EEG wave. *Int. J. Neurosci.* 73, 139–151.
- Berlyne, D. E. (1971). *Aesthetics and Psychobiology*. New York: Appleton-Century-Crofts.
- Bhattacharya, J., and Petsche, H. (2005). Drawing on mind's canvas: differences in cortical integration patterns between artists and non-artists. *Hum. Brain Mapp.* 26, 1–14.
- Chambers, C. D., Mattingley, J. B., and Moss, S. A. (2004). Reconsidering evidence for the suppression model of the octave illusion. *Psychon. Bull. Rev.* 11, 642–666.
- Chaudrun, A., and Glaser, D. A. (1991). Metastable motion anisotropy. *Vis. Neurosci.* 7, 397–407.
- Deutsch, D. (1974). An auditory illusion. *Nature* 251, 307–309.
- Deutsch, D. (2004). The octave illusion revisited again. *J. Exp. Psychol. Hum. Percept. Perform.* 30, 355–364.
- Dyson, B. J. (2009). "Aesthetic appreciation of pictures," in *Encyclopedia of Perception*, ed. B. Goldstein (New York, NY: Sage), 11–13.
- Dyson, B. J., Alain, C., and He, Y. (2005). Effect of visual attentional load on auditory scene analysis. *Cogn. Affect. Behav. Neurosci.* 5, 319–338.
- Dyson, B. J., and Cohen, R. (2010). Translations: effects of viewpoint, feature and naming on identifying repeatedly copied drawings. *Perception* 39, 157–172.
- Gepshtein, S., and Kubovy, M. (2005). Stability and change in perception: spatial organization in temporal context. *Exp. Brain Res.* 160, 487–495.
- Gombrich, E. H. (1977). *Art and Illusion: A Study in the Psychology of Pictorial Representation*, 5th Edn. Oxford: Phaidon Press.
- Gregory, R. L. (1997). *Eye and Brain: The Psychology of Seeing*, 5th Edn. Oxford: Oxford University Press.
- Harmony, T., Bernal, J., Fernandez, T., Silva-Pereyra, J., Fernandez-Bouzas, A., Marosi, E., and Reyes, A. (2000). Primary task demands modulate P3a amplitude. *Brain Res. Cogn. Brain Res.* 9, 53–60.
- Haroush, K., Hochstein, S., and Deouell, L. Y. (2009). Momentary fluctuations in allocation of attention: cross-modal effects of visual task load on auditory discrimination. *J. Cogn. Neurosci.* 22, 1440–1451.
- Ishai, A., Fairhall, S. L., and Pepperell, R. (2007). Perception, memory and aesthetics of indeterminate art. *Brain Res. Bull.* 73, 319–324.
- Jackson, S., Cummins, F., and Brady, N. (2008). Rapid perceptual switching of a reversible biological figure. *PLoS ONE* 3, e3982. doi: 10.1371/journal.pone.0003982
- Kaneoke, Y., Urakawa, T., Hirai, M., Kakigi, R., and Murakami, I. (2009). Neural basis of stable perception of an ambiguous apparent motion stimulus. *Neuroscience* 159, 150–160.
- Kleinschmidt, A., Büchel, C., Zeki, S., and Frackowiak, R. S. (1998). Human brain activity during spontaneously reversing perception of ambiguous figures. *Proc. R. Soc. B Biol. Sci.* 265, 2427–2433.
- Kohler, A., Haddad, L., Singer, W., and Muckli, L. (2008). Deciding what to see: the role of intention and attention in the perception of apparent motion. *Vision Res.* 48, 1096–1106.
- Kornmeier, J., and Bach, M. (2004). Early neural activity in Necker-cube reversal: evidence for low-level processing of a gestalt phenomenon. *Psychophysiology* 41, 1–8.
- Kornmeier, J., and Bach, M. (2005). The Necker cube- an ambiguous figure

- disambiguated in early visual processing. *Vision Res.* 45, 955–960.
- Kornmeier, J., and Bach, M. (2009). Object perception: when our brain is impressed but we do not notice it. *J. Vis.* 9, 1–10.
- Kottlow, M., Praeg, E., Luethy, C., and Jancke, L. (2011). Artists' advance: decreased upper alpha power while drawing in artists compared with non-artists. *Brain Topogr.* 23, 392–402.
- Larsen, A., McIlhagga, W., Baert, J., and Bunsden, C. (2003). Seeing or hearing? Perceptual independence, modality confusions, and crossmodal congruity effects with focused and divided attention. *Percept. Psychophys.* 65, 568–574.
- Latto, R., Brain, D., and Kelly, B. (2000). An oblique effect in aesthetics: homage to Mondrian (1872–1994). *Perception* 29, 981–987.
- Leopold, D. A., Wilke, M., Maier, A., and Logothetis, N. (2002). Stable perception of visually ambiguous patterns. *Nat. Neurosci.* 5, 605–609.
- Liesefeld, H. R., and Zimmer, H. D. (2011). The advantage of mentally rotating clockwise. *Brain Cogn.* 75, 101–110.
- Locher, P., Overbeeke, K., and Jan Stappers, P. (2005). Spatial balance of color triads in the abstract art of Piet Mondrian. *Perception* 34, 169–189.
- Martinez-Conde, S., and Macknik, S. L. (2010). Art as visual research: kinetic illusions in Op art. *Sci. Am.* 20, 48–55.
- Martino, G., and Marks, L. E. (2001). Synesthesia: strong and weak. *Curr. Dir. Psychol. Sci.* 10, 61.
- McManus, I., Cheema, B., and Stoker, J. (1993). The aesthetics of composition: a study of Mondrian. *Empirical Stud. Arts* 11, 83–94.
- Mudd, S. A. (1963). Spatial stereotypes of four dimensions of pure tone. *J. Exp. Psychol.* 66, 347–352.
- Muller-Gass, A., and Schröger, E. (2007). Perceptual and cognitive task difficulty has differential effects on auditory distraction. *Brain Res.* 1136, 169–177.
- Ohnishi, T., Matsuda, H., Asada, T., Argua, M., Hirakata, M., Nishikawa, M., Katoh, A., and Imabayashi, E. (2001). Functional anatomy of musical perception in musicians. *Cereb. Cortex* 11, 754–760.
- Pepperell, R. (2011). Connecting art and the brain: an artist's perspective on visual indeterminacy. *Front. Hum. Neurosci.* 5:84. doi: 10.3389/fnhum.2011.00084
- Pitts, M. A., Nerger, J. L., and Davis, T. J. R. (2007). Electrophysiological correlates of perceptual reversals for three different types of multistable images. *J. Vis.* 7, 1–14.
- Portas, C. M., Strange, B. A., Friston, K. J., Dolan, R. J., and Frith, C. D. (2000). How does the brain sustain a visual percept? *Proc. Biol. Sci.* 267, 845–850.
- Rees, G., Kreiman, G., and Koch, C. (2002). Neural correlates of consciousness in humans. *Nat. Rev. Neurosci.* 3, 261–270.
- Schlaug, G. (2001). The brain of musicians: a model for functional and structural adaptation. *Ann. N. Y. Acad. Sci.* 930, 281–299.
- Solso, R. L. (2003). *The Psychology of Art and the Evolution of the Conscious Brain*. Cambridge, MA: MIT Press.
- Sterzer, P., and Kleinschmidt, A. (2006). A neural basis for inference in perceptual ambiguity. *Proc. Natl. Acad. Sci. U.S.A.* 104, 323–328.
- Sterzer, P., Russ, M. O., Preibisch, C., and Kleinschmidt, A. (2002). Neural correlates of spontaneous direction reversals in ambiguous apparent visual motion. *Neuroimage* 15, 908–916.
- Struber, D., and Herrmann, C. S. (2002). MEG alpha activity decrease reflects destabilization of multistable percepts. *Brain Res. Cogn. Brain Res.* 14, 370–382.
- Takei, S., and Nishida, S. (2010). Perceptual ambiguity of bistable visual stimuli causes no or little increase in perceptual latency. *J. Vis.* 10, 1–15.
- Teder-Sälejärvi, W. A., McDonald, J. J., Di Russo, F., and Hillyard, S. A. (2002). An analysis of audio-visual crossmodal integration by means of event-related potential (ERP) recordings. *Brain Res. Cogn. Brain Res.* 14, 106–114.
- Valtonen, J., May, P., Mäkinen, V., and Tiitinen, H. (2003). Visual short-term memory load affects sensory processing of irrelevant sounds in human auditory cortex. *Brain Res. Cogn. Brain Res.* 17, 358–367.
- Vidal, J., Giard, M. H., Roux, S., Barthelemy, C., and Bruneau, N. (2008). Cross-modal processing of auditory-visual stimuli in a no-task paradigm. A topographic event-related potential study. *Clin. Neurophysiol.* 119, 763–771.
- von Schiller, P. (1933). Stroboskopische Alternativversuche [Stroboscopic alternative motion]. *Psychol. Forsch.* 17, 179–214.
- Windmann, S., Wehrmann, M., Calabrese, P., and Güntürkün, O. (2006). Role of the prefrontal cortex in attentional control over bistable vision. *J. Cogn. Neurosci.* 18, 456–471.
- Wolach, A. H., and McHale, M. A. (2005). Line spacing in Mondrian paintings and computer-generated modifications. *J. Gen. Psychol.* 132, 281–291.
- Yucel, G., Petty, C., McCarthy, G., and Belger, A. (2005). Graded visual attention modulates brain responses evoked by task-irrelevant auditory pitch changes. *J. Cogn. Neurosci.* 17, 1819–1828.
- Zeki, S. (2006). "The neurology of ambiguity," in *The Artful Mind*, ed. M. Turner (Oxford: Oxford University Press), 243–270.

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 28 February 2011; paper pending published: 07 June 2011; accepted: 18 July 2011; published online: 23 August 2011.

Citation: Dyson BJ (2011) The advantage of ambiguity? Enhanced neural responses to multi-stable percepts correlate with the degree of perceived instability. *Front. Hum. Neurosci.* 5:73. doi: 10.3389/fnhum.2011.00073

Copyright © 2011 Dyson. This is an open-access article subject to a non-exclusive license between the authors and Frontiers Media SA, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and other Frontiers conditions are complied with.