

Involvement of the visual change detection process in facilitating perceptual alternation in the bistable image

Tomokazu Urakawa¹ · Mao Bunya¹ · Osamu Araki¹

Received: 8 June 2016/Revised: 5 January 2017/Accepted: 6 March 2017/Published online: 24 March 2017
© Springer Science+Business Media Dordrecht 2017

Abstract A bistable image induces one of two perceptual alternatives. When the bistable visual image is continuously viewed, the percept of the image alternates from one possible percept to the other. Perceptual alternation was previously reported to be induced by an exogenous perturbation in the bistable image, and this perturbation was theoretically interpreted to cause neural noise, prompting a transition between two stable perceptual states. However, little is known experimentally about the visual processing of exogenously driven perceptual alternation. Based on the findings of a previous behavioral study (Urakawa et al. in *Perception* 45:474–482, 2016), the present study hypothesized that the automatic visual change detection process, which is relevant to the detection of a visual change in a sequence of visual events, has an enhancing effect on the induction of perceptual alternation, similar to neural noise. In order to clarify this issue, we developed a novel experimental paradigm in which visual mismatch negativity (vMMN), an electroencephalographic brain response that reflects visual change detection, was evoked while participants continuously viewed the bistable image. In terms of inter-individual differences in neural and behavioral data, we found that enhancements in the peak amplitude of vMMN1, early vMMN at a latency of approximately 150 ms, correlated with increases in the proportion of perceptual alternation across participants. Our results indicate the involvement of automatic visual change detection in the induction of perceptual alternation,

similar to neural noise, thereby providing a deeper insight into the neural mechanisms underlying exogenously driven perceptual alternation in the bistable image.

Keywords Bistable · Perceptual alternation · Visual mismatch negativity · vMMN · EEG

Introduction

When an ambiguous visual image consistent with multiple conflicting percepts appears, its temporally stable percept is generated among all possible percepts for a short period. When a bistable visual stimulus, such as a bistable image or binocular rivalry stimulus, each of which induces one of two mutually exclusive percepts, is continuously viewed, the percept of the bistable stimulus intermittently alternates between two exclusive percepts (from percept A to conflicting percept B). Previous studies examined the neural mechanisms underlying perceptual alternation using functional magnetic resonance imaging (fMRI) (e.g., Lumer et al. 1998; Kleinschmidt et al. 1998; Leopold and Logothetis 1999; Tong et al. 2006; Sterzer and Kleinschmidt 2007; Sterzer et al. 2009). Perceptual alternation was found to involve visual areas and fronto-parietal cortical areas. The fronto-parietal areas were suggested to affect visual activities relevant to mutually exclusive percepts via top-down processes, which reorganized visual activities during perceptual alternation (e.g., Leopold and Logothetis 1999; Sterzer et al. 2009). In recent studies (Kanai et al. 2010; Megumi et al. 2015), the anatomical structure of the parietal cortex and its activity were reported to predict inter-individual differences in the behavioral characteristics of perceptual alternation.

Previous studies have mainly investigated the neural mechanisms responsible for endogenous perceptual

✉ Tomokazu Urakawa
urakawatom3@gmail.com

¹ Department of Applied Physics, Faculty of Science, Tokyo University of Science, 6-3-1 Niiijyuku, Katsushika-ku, Tokyo 125-8585, Japan

alternation. However, perceptual alternation is also induced by exogenous perturbation (e.g., a flash presented on the bistable image) (Kanai et al. 2005). In the visual processing of the bistable image, exogenous perturbation has been conceptualized as a generator of neural noise (e.g., Moreno-Bote et al. 2007). Neural noise has been ascribed to random fluctuations of neural activity and was emphasized to play an indispensable role in destabilizing the perceptual state of the bistable image and in initiating perceptual alternation (e.g., Brascamp et al. 2006; Kim et al. 2006; Moreno-Bote et al. 2007; Shpiro et al. 2009). However, limited information is available on the visual processes involved in exogenously driven perceptual alternation at the macroscopic level. Sudden increases in luminance (e.g., the flash onset) or in its contrast, typical external perturbations that generate neural noise, are a common stimulus property across various presentation schemes of visual images. Thus, these external perturbations are expected to pervasively activate cortical areas including those not necessarily crucial to externally driven perceptual alternation. In order to experimentally investigate neural activity relevant to externally driven perceptual alternation, a stimulation method that effectively taps confined and definable visual processing needs to be employed.

Electroencephalographic (EEG) studies have examined visual change detection processes using an oddball paradigm or its modified version (e.g., Alho et al. 1992; Czigler et al. 2002; Pazo-Alvarez et al. 2004; Maekawa et al. 2005; Astikainen et al. 2008; Kimura et al. 2009). In these studies, brain responses to an infrequently presented visual stimulus (deviant), embedded in a repetitively presented stimulus (or sequentially ruled successive stimuli) (standard), were shown to be larger in amplitude at a latency of approximately 130–250 ms than those to the standard. The augmented response to the deviant has been interpreted to reflect the detection of a change (i.e., a deviant) in the sequence of visual events and has been called visual mismatch negativity (vMMN). vMMN appeared even when participants did not pay attention to the deviant, and the response enhancement was interpreted to reflect automatic visual change detection. Magnetoencephalographic (MEG) and electroencephalographic (EEG) studies (Urakawa et al. 2010; Kimura et al. 2010) indicated that the enhanced brain response to the deviant over that to the standard mainly originated in the middle occipital or occipital area, indicating that these areas are relevant to the generation of vMMN.

The relationship between automatic visual change detection and behavior, in addition to perception, has been qualitatively unclear (Stefanics et al. 2014). Nevertheless, automatic visual change detection underlying the generation of vMMN appears to be involved in shaping a percept of the bistable image, similar to neural noise. A recent psychological study employed presentations of deviant and

standard stimuli around the bistable image (Urakawa et al. 2016) in order to examine the effects of the deviant on a percept of the bistable image. The findings obtained showed that an adaptation-induced bias for an initial percept of the bistable image, called the reverse-bias effect (e.g., Long and Toppino 2004; Long and Moran 2007), decreased in strength when the bistable image was simultaneously presented with the deviant, which symmetrically appeared around the bistable image and was synchronized with the onset of the bistable image. In this stimulation paradigm, the deviant surrounding the bistable image was expected to invoke vMMN. Based on the predictive coding framework (Mumford 1992; Rao and Ballard 1999; Friston 2005), an increasing number of studies recently proposed or noted that the generation of vMMN corresponded to the emergence of a prediction error, and the prediction error automatically emerges when the incoming sensory input is inconsistent with the preceding sequential rule of visual events (e.g., Winkler and Czigler 2012; Kimura 2012; Stefanics et al. 2014; O'Shea 2015; Stefanics et al. 2016). The amplitude of vMMN is expected to mirror the magnitude of the prediction error that is minimized by subsequent iterative and exploratory neural processes through cortical areas until the most likely causes of sensory inputs are identified. Based on this framework, the prediction error caused by the deviant adjacent to the bistable image was proposed to contribute to invoking the subsequent exploratory visual process that shapes the initial percept of the bistable image, resulting in a decrease in the strength of the reverse-bias effect (Urakawa et al. 2016).

By focusing on the effects of the automatic visual change detection process on the subsequent perceptual alternation of a continuously presented bistable image, it is possible to clarify one facet of neural processing underlying exogenously driven perceptual alternation. When the deviant and standard were added around the continuously presented bistable image, the deviant was expected to produce vMMN, an emergence of the prediction error. As proposed in the aforementioned psychological study regarding the biased initial percept of the bistable image (Urakawa et al. 2016), this emergence of the prediction error may operate on perceptual alternation in a manner by which the prediction error contributes to inducing the exploratory visual process in order to shape the upcoming percept of the bistable image, thereby prompting the induction of its perceptual alternation. Therefore, the present EEG study hypothesized that the amplitude of vMMN mirroring the emergence of the prediction error may reflect an exploratory process that shapes the percept of the bistable image and also that the automatic visual change detection process, which is considered to be responsible for the generation of vMMN, may be involved not only in detecting a visual change, but also in exogenously triggering perceptual alternation.

In order to clarify whether the automatic visual change detection process has an enhancing effect on subsequent perceptual alternation, we developed a novel stimulation paradigm in which a deviant interspersed among the standards was presented around the bistable image while participants continuously viewed the image, which was based on a previous study (Urakawa et al. 2016). Under this paradigm, we simultaneously recorded vMMN and the behavioral response (perceptual alternation in the bistable image), and then attempted to establish whether vMMN is relevant to the subsequent facilitation of perceptual alternation. As reported previously (Vogel and Awh 2008; Kanai and Rees 2011), evaluating the relationship between neural and behavioral data in terms of inter-individual differences is one of the powerful analytical approaches used to deduce the neural mechanisms underlying behavioral data. According to this concept, previous studies on perceptual alternation revealed inter-individual variability in behavioral data and the brain's anatomical structure or neural activity (Kanai et al. 2010; Kanai and Rees 2011; Megumi et al. 2015). In order to test our hypothesis, we focused on inter-individual variability and attempted to clarify whether an enhancement in the vMMN amplitude (the magnitude of the prediction error) is relevant to increases in perceptual alternation among participants.

Methods

Participants

Ten healthy volunteers (Ten males, age 22–32 years, mean \pm SD, 23.6 \pm 3.13 years), all of whom were right-

handed and had normal visual acuity, participated in this study. Informed consent was obtained from all participants.

Stimulus and tasks

Visual images were presented on a cathode ray tube (CRT) display (DELL P1137) using the MATLAB Psychophysics Toolbox (Brainard 1997; Pelli 1997). Figure 1 shows the experimental procedure for one trial of each of the two experimental conditions: DEV and STD conditions. Based on a previous study (Urakawa et al. 2016), we used eight symmetrically-located bars with a fixation point (this stimulus image is hereafter referred to as the bar image). The orientations of the bars in one bar image were identical and either horizontal or vertical. The size and luminance of a bar were $1.9^\circ \times 0.1^\circ$ and 27.6 cd/m^2 , respectively. The central positions of the bars were located 2.26° above and below, 3.03° diagonally, and 2.23° horizontally from the position of the fixation point. The luminance of the background was 0.44 cd/m^2 . In every condition, each trial was divided into 2 consecutive parts: first and second phases. In the first phase, a blank image containing a fixation point only appeared for 250 ms and the bar image was then repetitively presented 20 times, keeping the orientations of the bars constant. The duration and inter-stimulus interval (ISI) of the bar image were 250 and 250 ms, respectively (the first phase lasted 10 s from the initial blank image). The fixation point was continuously presented from the initial blank image, but was randomly changed to a cross in 10% of all image frames including blank frames. The time interval between two consecutive fixation changes was set at more than 1 s. During the first phase, participants were asked to continuously look at the fixation point. When a

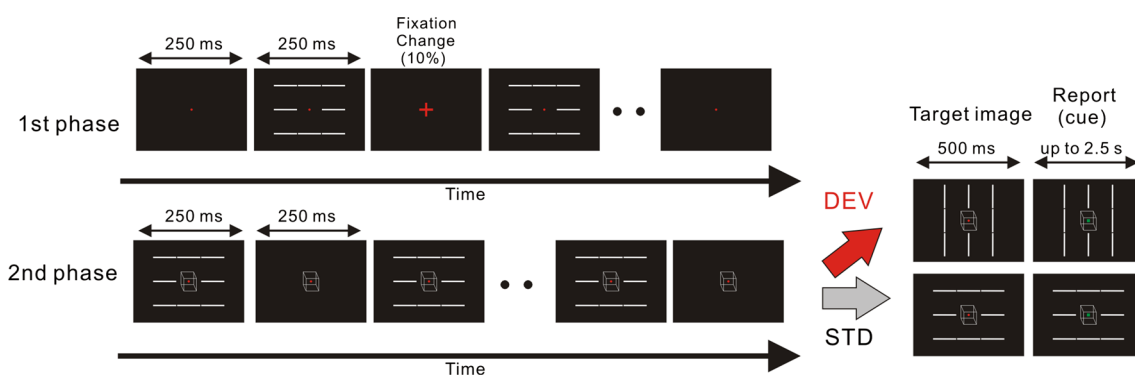


Fig. 1 Time course of stimulus presentation in one trial. Each trial consisted of two stimulation phases: the first and second phases. The first phase preceded the second phase, and the *bar image* was intermittently presented throughout these phases, keeping the orientations of the *bars constant*. In the first phase, the fixation point was immediately changed to a cross at a probability of 10%, and the participants were asked to respond to the change. In the second phase, the Necker cube was continuously presented at the center of the screen. Participants were asked to view the Necker cube and report its

initial percept in facing-orientation. Following the second phase, the target image was presented and the fixation point changed to a green square (a cue). The orientations of the *bars* in the target image changed by 90° in the DEV condition and did not change in the STD condition (the change in the orientations of the *bars* for the DEV condition was the deviant stimulus, and the non-changed orientations for the STD condition was the standard stimulus). In both conditions, participants were asked to report the facing orientation of the Necker cube after the cue

change occurred in the fixation point, participants were required to respond to the change as quickly as possible by pressing the middle arrow key of the keyboard in front of them. This task was set to prevent participants from paying attention to the bars, as in previous EEG studies (e.g., Czigler et al. 2002; Pazo-Alvarez et al. 2004; Maekawa et al. 2005; Astikainen et al. 2008; Kimura et al. 2009). At the end of the first phase, a blank image with a fixation point appeared for 250 ms, and the second phase was then started by the presentation of the bar image in which the Necker cube was centrally located. The presentation of the bar image and that of the Necker cube were independently manipulated from the beginning of the second phase. During the second phase, the Necker cube was continuously presented at the center of the display for 4.5 s while the bar image continued to be presented intermittently as in the first phase (for this period of 4.5 s, the bar image was intermittently presented 9 times and the last frame image in this period was a frame of the blank image). The orientation of the bars was kept constant during the 4.5-s time period. The size and mean luminance of the Necker cube were 1.7° (height) \times 1.6° (width) and 3.17 cd/m^2 , respectively. When the Necker cube appeared, participants were asked to report the current percept of the Necker cube by pressing a key on the keyboard. They were required to press the left arrow key when they perceived the Necker cube as being upper-left facing and the right arrow key when they perceived it as being lower-right facing. Once participants reported the facing orientation of the Necker cube, they were also instructed to pay attention to the initial percept of the Necker cube. This task was expected to reduce spontaneous and frequent perceptual alternation (Pelton and Solley 1968; Kanai et al. 2005). Following the 4.5-s exposure of the Necker cube, a bar image with the Necker cube immediately appeared (this image was hereafter referred to as a target image). The target image was presented for 500 ms, and the fixation point of this image then changed to a green square (a cue). When the cue appeared, participants were asked to report the current facing orientation of the Necker cube by pressing the left or right arrow key as described above. The target image with the cue was continuously presented for up to 2.5 s, this image immediately disappeared after participants responded, and the next first phase then started.

The present study employed a stimulation paradigm based on an oddball paradigm in which a standard stimulus was repetitively presented and a deviant was rarely presented (e.g., Alho et al. 1992; Czigler et al. 2002; Pazo-Alvarez et al. 2004; Maekawa et al. 2005; Astikainen et al. 2008; Kimura et al. 2009). The deviant breaks the sequential regularity of the visual events formed by the repetitively presented standard stimulus. As in previous studies (Astikainen et al. 2008; Kimura and Takeda 2013;

Urakawa et al. 2016), our stimulation method manipulated the orientations of the bars in order to break sequential regularity. A deviant was set for the DEV condition, but not for the STD condition. Based on a psychological study (Urakawa et al. 2016), we modified the oddball paradigm such that there was one deviant in one trial for the DEV condition, but not for the STD condition. In the DEV condition, the orientations of the bars in the target image changed by 90° relative to those of the preceding bars. This deviation in the orientations of the bars was the deviant (the target image with the deviant is hereafter referred to as the DEV target). In the STD condition, the orientations of the bars in the target image were identical to those of the preceding one, and, thus, there was no deviant (the target image without the deviant is hereafter referred to as the STD target). Each condition contained 120 trials. In the STD condition, the orientations of the bars in the first and second phases were set to be horizontal for 60 trials and vertical for the other 60 trials. In the DEV condition, the orientations of the bars changed from horizontal to vertical for 60 trials and the direction of the change was reversed for the other 60 trials. In the sequence of the stimulus presentation, trials of the DEV condition and those of the STD condition were randomized regardless of the orientations of the bars. The present study set 6 sessions, and one session contained 40 trials. Participants were given a rest between sessions if needed.

Analysis of behavioral data

Under the current stimulation paradigm, we obtained the number of trials in which the percept of the Necker cube (facing orientation) changed from before to after the DEV target and also the number of trials in which the percept changed from before to after the STD target. Trials in which perceptual alternation occurred before the presentation of the DEV/STD target were discarded from the analysis, and we recalculated the total number of trials and number of perceptual alternations for the DEV and STD target, respectively. The proportion of perceptual alternation was then obtained for the DEV and STD targets [the number of trials with perceptual alternation/the total number of trials (120)]. The proportion of perceptual alternation in the STD condition was subtracted from that in the DEV condition in each participant for the EEG analysis.

EEG recording

Neural activity in the DEV and STD conditions was recorded by an electroencephalography (EEG) processor with 57 electrodes (EEG-1200, Nihon Kohden, Tokyo, Japan; EasyCap GmbH, Herrsching, Germany). The

electrode layout was based on a modified version of the international 10–20 system. Impedance at each electrode was kept at less than 10 k Ω . EEG signals were digitized at 500 Hz and recorded with a 0.5–50 Hz band-pass filter online. In data acquisition, EEG signals were referenced to the right earlobe, and eye movement was monitored using horizontal and vertical bipolar electrooculograms (EOGs).

Analysis of EEG data

EEG signals were transformed offline to the average reference, and were low-pass filtered at 30 Hz. EEG epochs from 200 ms before to 600 ms after the onset of the DEV and STD targets were then collected. We calculated the mean of the EEG epochs across trials to obtain VEPs for the DEV and STD targets. In this calculation of VEPs, we excluded EEG epochs containing a deflection of greater than ± 100 μ V in at least one electrode from averages across trials in order to remove EEG signals containing artifacts. In each VEP, the mean amplitude for a period of -200 to 0 ms relative to the stimulus onset was used as the baseline. At least 88 artifact-free EEG signals were averaged in every condition for each participant. Based on previous studies (e.g., Guthrie and Buchwald 1991; Doniger et al. 2001; Sehatpour et al. 2006; Urakawa et al. 2015), the difference in the VEP amplitude between the DEV and STD conditions was evaluated using a series of two-tailed t tests through successive time points. When the t tests exceeded the 0.05 criterion for at least 11 subsequent time points (corresponding to 20 ms time interval in the present study), the amplitude difference between conditions was considered to be significant. As in previous studies (e.g., Alho et al. 1992; Czigler et al. 2002; Pazo-Alvarez et al. 2004; Maekawa et al. 2005; Astikainen et al. 2008; Kimura et al. 2009), VEP to the STD target was subtracted from that to the DEV target in order to obtain vMMN for each participant. We calculated Pearson's correlation coefficient between the peak latency/amplitude of vMMN and the differential proportion of perceptual alternation (the DEV condition–the STD condition) across participants.

Results

Behavioral data

In the first phase, participants were instructed to respond to a change in the fixation point. The mean detection rate of the change across participants and conditions was nearly 1 [0.96 ± 0.024 (SE)]. Figure 2 shows the proportion of perceptual alternation from before to after the DEV target and STD target for all participants. No value exceeded the range of the mean ± 2 SD for the DEV and STD targets.

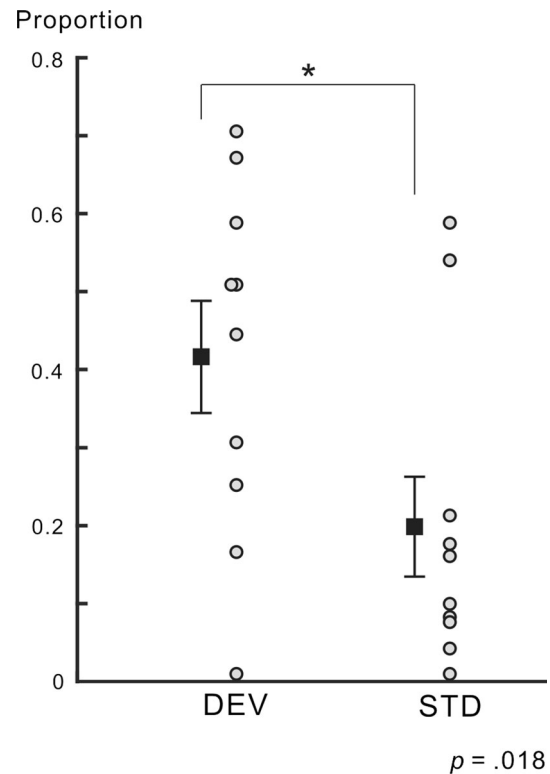


Fig. 2 Proportion of perceptual alternation. The proportions of perceptual alternation for all participants are shown for DEV and STD targets. The mean proportion is indicated by a square with \pm SE for each condition. The mean proportion of perceptual alternation in the DEV condition was significantly higher than that in the STD condition. Some data points are vertically shifted for display purposes

Thus, we did not regard any proportion value as an outlier value. Paired t tests revealed that the mean proportion of the perceptual alternation for the DEV target was significantly higher than that for the STD target ($t = 2.881$, $p = 0.018$, Cohen's $d = 1.012$), indicating that perceptual alternation was facilitated by the DEV target. As for the mean reaction time (RT) to the cue, no significant difference was observed between the conditions (the DEV condition, 944 ± 76.7 ms; the STD condition, 878 ± 72.0 ms; paired t test, $t = 1.543$, $p = 0.158$, Cohen's $d = 0.283$). Therefore, among participants reporting the facing orientation of the Necker cube, it was unlikely that the DEV target was effective enough to distract participants from the report itself.

The DEV target had two bar orientations (horizontal or vertical bars, see the “Methods” section). We attempted to clarify whether these two bar orientations differentially bias the perceived Necker facing following the onset of the DEV target when perceptual alternation occurred. In the horizontal orientation of the DEV target, the mean number of trials with perceptual alternation to the left facing cube was 14 ± 3.43 (SE), while that to the right facing cube was 10 ± 3.56 (SE). In the vertical orientation of the DEV target, the mean number of trials with perceptual

alternation to the left facing cube was 17.2 ± 4.30 (SE), while that to the right-facing cube was 8.3 ± 3.17 (SE). A two-way repeated-measure analysis of variance (ANOVA) with factors of the bar orientation of the DEV target (horizontal and vertical) and perceived Necker facing after perceptual alternation (perceptual alternation to the right facing cube and that to the left facing cube) revealed that there was no interaction between these factors ($F(1, 9) = 2.221$, $p = 0.170$, partial $\eta^2 = 0.198$), indicating that the two bar orientations of the DEV target did not differentially bias the following percept of the Necker cube to a particular face in perceptual alternation. In addition, the bar orientation of the DEV target did not significantly affect the number of trials with perceptual alternation ($F(1, 9) = 0.727$, $p = 0.419$, partial $\eta^2 = 0.075$).

EEG data

Figure 3a shows grand-averaged VEP waveforms at Oz and POz for the DEV and STD targets. The VEP amplitude appeared to be more negatively shifted for the DEV target than for the STD target at Oz, but not at POz, particularly at latencies of approximately 150–200 and 300–400 ms. Two-tailed t tests, successively performed for each consecutive time point (see Analysis of EEG data in the “Methods” section), revealed that the negative shift was significant at latencies of 126–158 and 280–406 ms at Oz. Figure 3b shows isocontour maps at latencies of 150 and 300 ms. VEP amplitudes recorded around posterior electrodes were more negatively displaced in the DEV target than in the STD target at the center of Oz. These results were consistent with previous findings (e.g., Alho et al. 1992; Czigler et al. 2002; Pazo-Alvarez et al. 2004; Maekawa et al. 2005; Astikainen et al. 2008; Kimura et al. 2009), and our stimulation paradigm effectively evoked vMMN.

With a focus on inter-individual variability in vMMN, we calculated vMMN at Oz for every participant. Figure 4a shows superimposed vMMN across participants with the mean. Inter-individual variability was observed in the peak latency and peak amplitude for vMMN. In accordance with previous findings (Maekawa et al. 2005), two consecutive peaks for posterior negativities (we hereafter refer to the first negativity as vMMN1 and second negativity as vMMN2) were observed across participants. vMMN1 and vMMN2 peaked at approximately 150 and 300 ms, respectively. Figure 4b shows isocontour maps at 150 and 300 ms. In both latencies, posterior negativity clearly appeared at Oz.

Correlation between vMMNs and perceptual alternation

The present study then attempted to clarify whether the differential proportion of perceptual alternation (the DEV

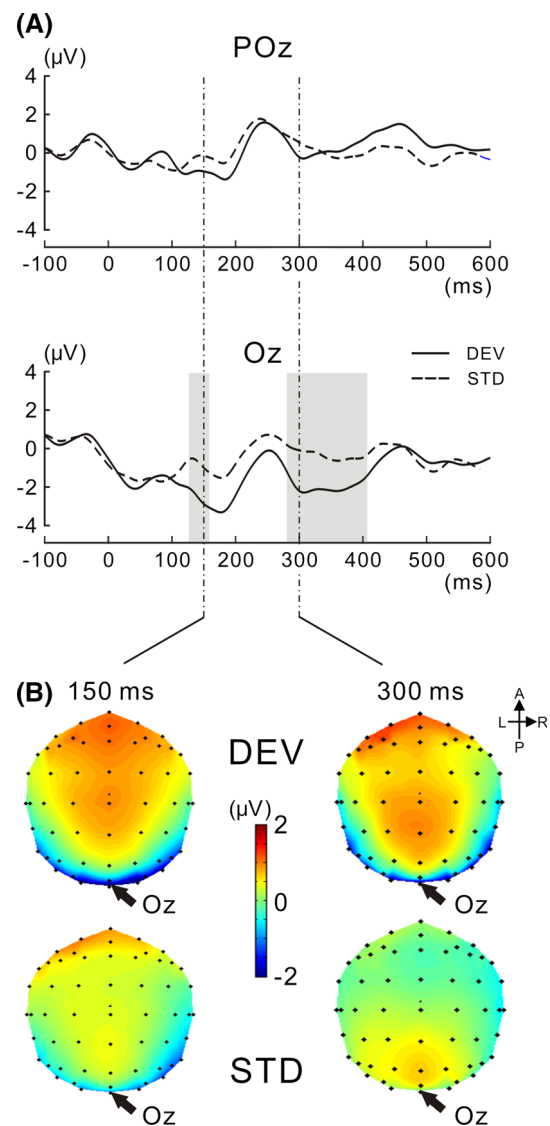


Fig. 3 VEPs to the target image for DEV and STD targets. **a** VEPs at POz and Oz are shown for the DEV and STD targets, respectively. The VEP amplitude for the DEV target was significantly higher than that for the STD target. The time interval in which there was a significant difference is shaded in gray (for the procedure of the statistical analysis, see the “Methods” section). **b** Isocontour maps at latencies of 150 and 300 ms are shown for DEV and STD targets. In both latencies, VEP was negatively enhanced in amplitude at the posterior electrodes, particularly at Oz

target—the STD target), a behavioral index, correlates with the peak latency and/or peak amplitude of vMMN across participants. A correlation analysis was performed for two different vMMNs (vMMN1 and vMMN2). No value of the behavioral index (the differential proportion) exceeded the range of the mean ± 2 SD. Thus, differential proportions were not regarded as outlier values.

Figure 5a shows the results of the correlation analysis for vMMN1. While no vMMN1 amplitude value was defined as an outlier (all vMMN1 amplitude values were in

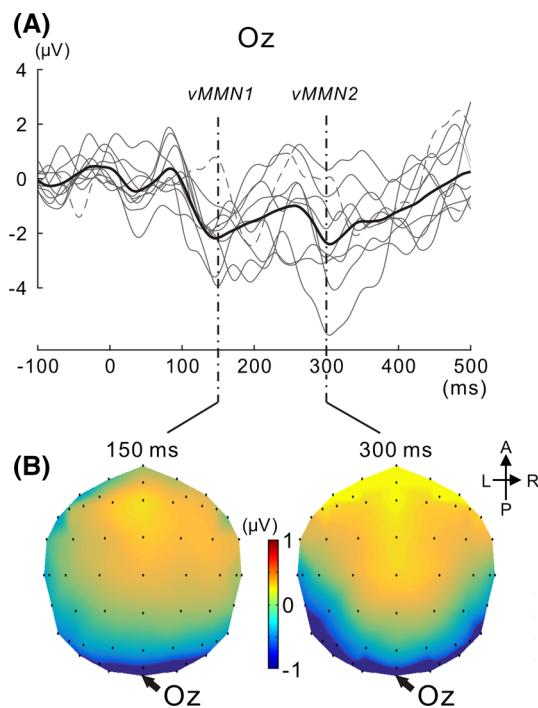


Fig. 4 vMMNs for all participants. **a** vMMNs for all participants are superimposed with their mean. The mean is illustrated in *bold black*. vMMNs including outlier data are depicted as a *dotted line* (outlier data were defined for each of the vMMN1 and vMMN2 latency ranges, see the “Results” section). Two successive vMMNs (vMMN1 and subsequent vMMN2) appeared. **b** Isocontour map of the vMMN mean across participants is shown at 150 and 300 ms, respectively. The most prominent negativity emerged at Oz for both vMMN1 and vMMN2

the range of its mean ± 2 SD), the vMMN1 latency value of one participant was not in the range of its mean ± 2 SD (this latency outlier was greater than the mean vMMN1 latency $+ 2$ SD). Therefore, data obtained for this participant were excluded from the correlation analysis of vMMN1. Using the remaining data ($N = 9$), we performed a normality test, the results of which indicated no significance for differential proportion, latency, or amplitude (the Shapiro–Wilk test, $p = 0.320$ for differential proportion; $p = 0.472$ for the latency of vMMN1; $p = 0.077$ for the amplitude of vMMN1). The correlation analysis revealed that the relationship between differential proportion and latency was not significant ($r = 0.153$, $p = 0.694$), whereas differential proportion negatively correlated with amplitude ($r = -0.689$, $p = 0.040$). This result is consistent with our hypothesis that enhancements in the amplitude of vMMN contribute to facilitating perceptual alternation.

Figure 5b shows the results of the correlation analysis for vMMN2. No vMMN2 amplitude value was defined as an outlier (all vMMN2 amplitude values were in the range of its mean ± 2 SD), whereas the vMMN2 latency value of

one participant was not in the range of the mean ± 2 SD (this latency data was greater than the mean vMMN2 latency $+ 2$ SD, and the participant with the latency outlier was not the same participant whose data were excluded from the correlation analysis for vMMN1). Therefore, data obtained for this participant were excluded from the correlation analysis for vMMN2. Using the remaining data ($N = 9$), we performed a normality test, the results of which indicated no significance for latency or amplitude (the Shapiro–Wilk test, $p = 0.472$ for the latency of vMMN2; $p = 0.977$ for the amplitude of vMMN2). The correlation analysis revealed the absence of a significant relationship between differential proportion and latency ($r = 0.500$, $p = 0.170$) and between differential proportion and amplitude ($r = -0.552$, $p = 0.124$). By focusing on inter-individual variability in each of the two different vMMNs, these results indicate that an enhancement in the amplitude of early vMMN (vMMN1), but not late vMMN (vMMN2) is associated with the facilitation of perceptual alternation across participants.

Discussion

Based on the predictive coding framework, the present study attempted to elucidate whether the automatic visual change detection process reflected in vMMN is relevant to the exogenously driven neural mechanisms triggering perceptual alternation in the bistable image. We made a novel stimulation paradigm in which both the Necker cube and a visual change evoking vMMN were presented. Our results showed that two successive vMMNs, vMMN1 at a latency of approximately 150 ms and subsequent vMMN2 at a latency of approximately 300 ms, were evoked under our stimulation paradigm. Enhancements in the peak amplitude of vMMN1 correlated with an increase in the proportion of perceptual alternation across participants; however, this was not the case for vMMN2. The correlation for vMMN1 supports the enhancement in the amplitude of vMMN1 corresponding to the magnitude in the prediction error that invokes perceptual alternation, similar to externally driven neural noise. In terms of inter-individual differences in neural and behavioral data, we suggest that the early automatic visual change detection process as early as 150 ms following the visual change (i.e., visual process underlying vMMN1) is involved in the exogenously driven neural mechanisms inducing perceptual alternation, which has not yet been reported in detail.

Previous studies on perceptual alternation proposed that early spatial attention (attending to a certain location within the bistable image) allowed a preferentially processed feature at the attended location, thereby contributing

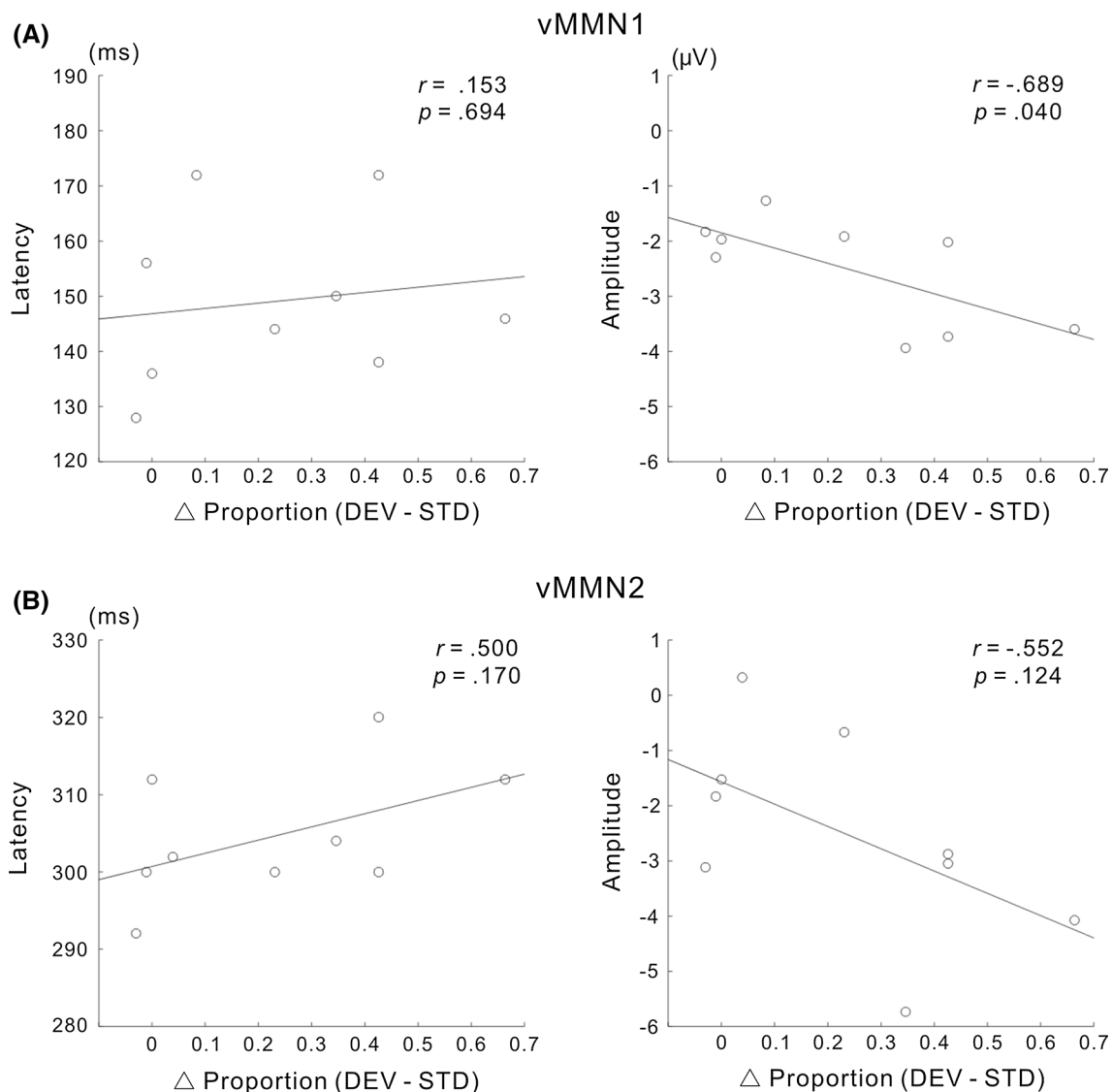


Fig. 5 Relationship between vMMNs and perceptual alternation across participants. The correlations between the differential proportion of perceptual alternation (DEV–STD) and the latency or amplitude of vMMN are shown for **a** vMMN1 and **b** vMMN2. The amplitudes of vMMN1 and vMMN2 appeared to decrease with an increase in the differential proportion (i.e., the amplitudes of both

vMMN1 and vMMN2 were augmented as the proportion of perceptual alternation increased). A negative correlation was observed between the differential proportion and amplitude of vMMN for vMMN1, but not for vMMN2. Data defined as outliers were excluded from these analyses (see the “Results” section)

to shaping an interpretation of the feature as “nearest” or “in the foreground” (Pitts et al. 2007, 2008). In our stimulation paradigm, the DEV target, for which the bar orientations abruptly changed, appeared to exogenously prompt participants to pay more attention to a certain location within the Necker cube than the STD target. Therefore, the DEV target may facilitate perceptual alternation more than the STD target, providing the present results (the proportion of perceptual alternation was higher for the DEV target than for the STD target). Nevertheless, this is unlikely for the following reasons. The bars were arranged symmetrically around the Necker cube and the

change in bar orientation simultaneously occurred for all bar locations in the present study. As discussed in our previous study (Urakawa et al. 2016), this stimulation method was not effective for inducing the stimulus-driven capture of attention toward a certain location on a visual image. Furthermore, our results support a percept of Necker facing following the DEV target in perceptual alternation not being significantly affected by the bar orientation of the DEV target. Taken together, our results do not appear to be fully accounted for by spatial attention in that the bars of the DEV target were unlikely to shift spatial attention and to bias a certain perceived Necker facing.

When perceptual alternation occurred under the condition that the bistable image was intermittently presented, VEPs time-locked to the image at the posterior electrodes were shown to be more negatively shifted in a time window of 150–350 ms; this negative shift started at approximately 160 ms and initially peaked at 250 ms (Kornmeier and Bach 2004). This negative enhancement in the VEP amplitude was quantified by calculating differential VEP (perceptual reversal–no perceptual reversal) and the negatively-going differential potential was called reversal negativity (RN). In our study, the DEV target facilitated subsequent perceptual alternation (see Fig. 2), and vMMN1 was similar to the RN in the early latency range (peak latency of vMMN 1: ca. 150 ms, see Figs. 3a, 4a). Thus, we cannot exclude the possibility that both vMMN1 and early RN at least partly reflect a common visual process relevant to perceptual alternation free of the external perturbation. Nevertheless, a recent study reported a positive correlation between the amplitude of RN peaking at approximately 150 ms (corresponding to the early RN) and number of perceptual reversals across participants (Russo and Pascalis 2016), which was in contrast to our results showing a negative correlation between the amplitude of vMMN1 and proportion of perceptual alternation (roughly corresponding to the number of perceptual alternations). Therefore, it is highly likely that vMMN1 did not simply correspond to the early phase of RN.

The vMMN2 peak at a latency of approximately 300 ms (see Fig. 4a) may have been contaminated, to a greater or lesser extent, by other posterior negativities irrelevant to vMMN itself. One of these posterior negativities is an attention-related component called selection negativity (SN) (e.g., Harter and Previc 1978; Eimer 1997). SN begins to appear at a latency of approximately 200 ms and persists for several hundred milliseconds when participants pay attention to a visual feature. Previous EEG studies on perceptual alternation suggested that SN emerges when early spatial attention allows a preferentially processed feature at the attended location (Pitts et al. 2007, 2008). As described above, our stimulation paradigm did not appear to effectively tap spatial attention. Thus, SN may not have been strongly invoked. Another possible posterior negativity emerging at the latency range of vMMN2 was late RN peaking at 250–300 ms. This late RN was reported to emerge for exogenously and endogenously driven perceptual alternation and was suggested to reflect visual processing leading to the “Gestalt reconstruction” in shaping a bistable image percept after disambiguation of the bistable image was completed (e.g., Kornmeier and Bach 2006). In our study, the proportion of perceptual alternation was higher for the DEV target than for the STD target, and the Gestalt reconstruction was expected to have occurred more frequently for the DEV target than for the STD target.

Therefore, we cannot exclude that late RN may have been more prominent for the DEV target than for the STD target, resulting in posterior negativity at the latency range of vMMN2. vMMN2 recorded under the present experimental paradigm may be confounded, to a greater or lesser extent, by SN and late RN. This potential contamination may weaken the correlation between the behavioral index and amplitude of vMMN2 free of other posterior negativities (i.e., SN and late RN) if present.

In the predictive coding framework (e.g., Friston 2005; Garrido et al. 2009), the enhancement of brain responses to the deviant over that to the standard under the oddball paradigm is interpreted to reflect the emergence of the prediction error in relation to a preceding sequential rule of sensory events. An increasing number of EEG studies have also stressed that vMMN reflects the prediction error (e.g., Winkler and Czigler 2012; Kimura 2012; Stefanics et al. 2014). The prediction error is expected to be minimized in the brain by iterative and exploratory neural processes through cortical areas in order to deduce the most likely causes of sensory inputs (Friston 2005). As previously discussed in a behavioral study (Urakawa et al. 2016), the enhancement of the prediction error for a visual scene, which is generated by the deviant added around the bistable image on the visual scene, may contribute to the subsequent exploratory processes shaping an initial percept of the bistable image. In a similar vein, this exploratory process may also invoke the induction of perceptual alternation in shaping the percept of a bistable image under the current experimental paradigm. In support of this concept, the present EEG study found a correlation between the enhanced amplitude of vMMN1 and increases in perceptual alternation, and this result is consistent with our hypothesis that increases in the strength of the prediction error, reflected in the enhanced amplitude of vMMN, correlates with the subsequent facilitation of perceptual alternation. Apart from the exogenously driven perceptual alternation, endogenous perceptual alternation was accounted for by the predictive coding framework (Hohwy et al. 2008), in which the prediction error from the perceptually suppressed image was posited to destabilize the dominant percept, triggering perceptual alternation. Although the prediction error in the previous study was not presumed to originate from external perturbation, it appears to be common to that driven by external perturbation in that its emergence may contribute to the induction of perceptual alternation.

In our stimulation paradigm, the orientations of the bars of the STD target were identical to those in preceding repetitively presented images, whereas the orientations of the bars of the DEV target were different to those in preceding images. Under this stimulation paradigm, the emergence of vMMN (VEP to the DEV target—that to the

STD target) may be partly ascribed to neural adaptation/refractoriness for the orientations of the bars in the STD target, which was expected to have been controlled or mitigated in order to record “genuine” vMMN (e.g., Czigler et al. 2002; Astikainen et al. 2008; Stefanics et al. 2014). Nevertheless, the predictive coding framework did not intrinsically separate neural adaptation from the detection of a violation in the sequential rule when dealing with neural processes reflected in vMMN. The diminishing neural response with a repetitively presented stimulus (i.e., neural adaptation) is interpreted to correspond to a reduction in the prediction error due to sensory inputs consistent with the preceding stimulus repetition/regularity.

The neural mechanisms underlying the induction of inter-individual variability for the amplitude of vMMN1 remains elusive. In previous studies, ongoing activity was expected to affect VEPs (e.g., Jansen and Brandt 1991; Fellingner et al. 2011) and concomitantly reflect subsequent behavioral responses (Ergenoglu et al. 2004; Mathewson et al. 2009). In relation to the perception of the bistable image, ongoing activity preceding endogenous perceptual alternation, called an EEG microstate, was reported to emerge at the right parietal area (Britz et al. 2009; Britz and Michel 2011). In the present study, ongoing activity preceding the DEV target at the parietal area may have somehow interacted more with the visual process underlying vMMN1 for participants with a potentially higher proportion of endogenous perceptual alternation. In this scenario, the inter-individual difference in ongoing activity may contribute to magnifying the prediction error invoked by the exogenous perturbation (enhancing the amplitude of vMMN1 to the DEV target) for participants with a potentially higher proportion of endogenous perceptual alternation, resulting in inter-individual variability in the vMMN1 amplitude.

In an MEG study using an oddball paradigm (Urakawa et al. 2010), the enhanced brain response to the visual deviant, reflecting the automatic visual change detection process, was reported to begin to appear at the middle occipital gyrus (MOG). The peak latency of MOG activity is similar to that of vMMN1 (both are approximately 150 ms), and the current orientation at the MOG was consistent with a negative shift in VEP over posterior electrodes (Urakawa et al. 2010). Therefore, it is plausible that the main source of vMMN1 lies in the MOG and also that the MOG is involved in the genesis of the prediction error in the subsequent visual process of the bistable image, thereby inducing perceptual alternation, similar to neural noise. In the case of endogenous perceptual alternation, the fronto-parietal areas, non-sensory areas higher than visual areas, have been suggested to send a top-down signal to visual areas relevant to two mutually-exclusive percepts (Leopold and Logothetis 1999; Sterzer et al. 2009). This top-down signal was interpreted to

be relevant to the reorganization of visual activities during perceptual alternation. It was further suggested that the front-parietal areas are activated in a feed-forward manner by the destabilization of the “balance of power” between visual representations coding for different percepts; this destabilization is putatively due to neural adaptation to the dominant percept or mutual inhibition between visual representations (Sterzer et al. 2009). In terms of the exogenous visual process, our present results appear to expand on this endogenous neural process in that the automatic visual change detection process at the MOG was additionally recruited when an external perturbation (a visual change) occurred around the bistable image. In a previous MEG study on automatic visual change detection (Urakawa et al. 2010), MOG activity was shown to be followed by activity at the right inferior frontal area. In addition, an fMRI study of endogenous perceptual alternation reported that activity at the right frontal area was reported to precede visual activities during alternation (Sterzer and Kleinschmidt 2007). Although the temporal resolution of MEG and that of fMRI are different (the former is greater than the latter), we suggest a tentative framework in which the automatic visual change detection process at the MOG, which monitors a change around the continuously presented bistable image, sends a signal of “a visual change in the bistable image” or a signal of “the emergence of the prediction error” to the right frontal area. The top-down process from the frontal area to visual areas may then be initiated, contributing to the reorganization of visual activities during perceptual alternation (from predictive coding, these processes correspond to neural processes relevant to minimization of the prediction error through the cortical hierarchy). In theoretical studies on the perceptual decision (e.g., Deco and Rolls 2009; Rolls and Deco 2010), neural noise in the bistable perceptual system, ascribed to random fluctuations in neural activity, was suggested to have a role in influencing the formation of the possible perception of the bistable image in an exploratory manner, and the role of this noise was interpreted to contribute to avoiding *perceptual deadlock* (i.e., perceptual locking of an ambiguous sensory scene, which results in no perceptual change to the alternative possibility). Since VEP used to obtain vMMN is time-locked in phase to the onset of the visual image, random fluctuations in neural activity that was not time-locked activity were reflected less in our results. Despite this difference in neural activity, we propose that automatic visual change detection in our proposed framework has a functional role not only in detecting a visual change around the bistable image, but also in avoiding perceptual deadlock.

In conclusion, our results showed that the automatic visual change detection process is relevant to exogenously driven perceptual alternation, similar to neural noise presumed previously (e.g., Moreno-Bote et al. 2007). The

present results suggest that automatic visual change detection is not merely related to change detection, but also to the exploratory shaping of a percept of the bistable image. This neural process may be automatically implemented in the early stage of visual processing following the onset of the visual change, thereby contributing to a new forthcoming perceptual perspective and exploratory adaptation of the ever-changing visual environment, which is full of ambiguous visual scenes.

Acknowledgements We thank Yasuhisa Kitai for his help in data acquisition and Emi Hasuo for comments on this manuscript. We are grateful to Hiroko Ichikawa for her helpful comments on the statistical analyses performed in this study. This work was supported by grants from the Japan Society for the Promotion of Science for Young Scientists (B) No. 26870426 to T. U.

References

- Alho K, Woods DL, Algazi A, Näätänen R (1992) Intermodal selective attention. II. Effects of attentional load on processing of auditory and visual stimuli in central space. *Electroenceph Clin Neurophysiol* 82:356–368
- Astikainen P, Lillstrang E, Ruusuvirta T (2008) Visual mismatch negativity for changes in orientation—a sensory memory-dependent response. *Eur J Neurosci* 28:2319–2324
- Brainard DH (1997) The psychophysics toolbox. *Spat Vis* 10:433–436
- Brascamp JW, Van Ee R, Noest AJ, Jacobs RH, Van Den Berg AV (2006) The time course of binocular rivalry reveals a fundamental role of noise. *J Vis* 6:1244–1256
- Britz J, Michel CM (2011) State-dependent visual processing. *Front Hum Neurosci* 2:370
- Britz J, Landis T, Michel CM (2009) Right parietal brain activity precedes perceptual alternation of bistable stimulus. *Cereb Cortex* 19:55–65
- Czigler I, Balazs L, Winkler I (2002) Memory-based detection of task-irrelevant visual changes. *Psychophysiology* 39:869–873
- Deco G, Rolls ET (2009) Stochastic dynamics as a principle of brain function. *Prog Neurobiol* 88:1–16
- Doniger GM, Foxe JJ, Schroeder CE, Murray MM, Higgins BA, Javitt DC (2001) Visual perceptual learning in human object recognition areas: a repetition priming study using high-density electrical mapping. *NeuroImage* 13:305–313
- Eimer M (1997) An event-related potential (ERP) study of transient and sustained visual attention to color and form. *Biol Psychol* 44:143–160
- Ergenoglu T, Demiralp T, Bayraktaroglu Z, Ergen M, Beydagi H, Uresin Y (2004) Alpha rhythm of the EEG modulates visual detection performance in humans. *Cogn Brain Res* 20:376–383
- Fellinger R, Klimesch W, Gruber W, Freunberger R, Doppelmayr M (2011) Pre-stimulus alpha phase-alignment predicts P1-amplitude. *Brain Res Bull* 85:417–423
- Friston K (2005) A theory of cortical responses. *Philos Trans R Soc B* 260:815–836
- Garrido MI, Kilner JM, Stephan KE, Friston KJ (2009) The mismatch negativity: a review of underlying mechanisms. *Clin Neurophysiol* 120:453–463
- Guthrie D, Buchwald JS (1991) Significance testing of difference potentials. *Psychophysiology* 28:240–244
- Harter MR, Previc FH (1978) Size specific information channels and selective attention: visual evoked potential and behavioral measures. *Electroencephalogr Clin Neurophysiol* 45:628–640
- Hohwy J, Roepstorff A, Friston K (2008) Predictive coding explains binocular rivalry: an epistemological review. *Cognition* 108:687–701
- Jansen BH, Brandt ME (1991) The effect of the phase of prestimulus alpha activity on the averaged visual evoked response. *Electroencephalogr Clin Neurophysiol* 80:241–250
- Kanai R, Rees G (2011) The structural basis of inter-individual differences in human behavior and cognition. *Nat Rev Neurosci* 12:231–242
- Kanai R, Moradi F, Shimojo S, Verstraten FAJ (2005) Perceptual alternation induced by visual transients. *Perception* 34:803–822
- Kanai R, Bahrami B, Rees G (2010) Human parietal cortex structure predicts individual differences in perceptual rivalry. *Curr Biol* 20:1626–1630
- Kim YJ, Grabowecy M, Suzuki S (2006) Stochastic resonance in binocular rivalry. *Vis Res* 46:392–406
- Kimura M (2012) Visual mismatch negativity and unintentional temporal-context-based prediction in vision. *Int J Psychophysiol* 83:144–155
- Kimura M, Takeda Y (2013) Task difficulty affects the predictive process indexed by visual mismatch negativity. *Front Hum Neurosci* 7:267
- Kimura M, Katayama J, Ohira H, Schröger E (2009) Visual mismatch negativity: new evidence from the equiprobable paradigm. *Psychophysiology* 46:402–409
- Kimura M, Ohira H, Schröger E (2010) Localizing sensory and cognitive systems for pre-attentive visual deviance detection: an sLORETA analysis of the data of Kimura et al. (2009). *Neurosci Lett* 485:198–203
- Kleinschmidt A, Büchel C, Zeki S, Frackowiak RS (1998) Human brain activity during spontaneously reversing perception of ambiguous figures. *Proc Biol Sci* 265:2427–2433
- Kornmeier J, 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, Bach M (2006) Bistable perception—along the processing chain from ambiguous visual input to a stable percept. *Int J Psychophysiol* 62:345–349
- Leopold DA, Logothetis NK (1999) Multistable phenomena: changing views in perception. *Trends Cogn Sci* 3:254–264
- Long GM, Moran CJ (2007) How to keep a reversible figure from reversing: teasing out top-down and bottom-up processes. *Perception* 36:431–445
- Long GM, Toppino TC (2004) Enduring interest in perceptual ambiguity views of reversible figures. *Psychol Bull* 130:748–768
- Lumer ED, Friston KJ, Rees G (1998) Neural correlates of perceptual rivalry in the human brain. *Science* 280:1930–1934
- Maekawa T, Goto Y, Kinukawa N, Taniwaki T, Hanbu S, Tobimatsu S (2005) Functional characterization of mismatch negativity to visual stimulus. *Clin Neurophysiol* 116:2392–2402
- Mathewson KE, Gratton G, Fabiani M, Beck DM, Ro T (2009) To see or not to see: prestimulus α phase predicts visual awareness. *J Neurosci* 29(9):2725–2732
- Megumi F, Bahrami B, Kanai R, Rees G (2015) Brain activity dynamics in human parietal regions during spontaneous switches in bistable perception. *NeuroImage* 107:190–197
- Moreno-Bote R, Rinzel J, Rubin N (2007) Noise-induced alternations in an attractor network model of perceptual bistability. *J Neurophysiol* 98:1125–1139
- Mumford D (1992) On the computational architecture of the neocortex. II. The role of cortico-cortical loops. *Biol Cybern* 66:241–251
- O’Shea RP (2015) Refractoriness about adaptation. *Front Hum Neurosci*. doi:10.3389/fnhum.2015.00038
- Pazo-Alvarez P, Amanedo E, Cadaveria F (2004) Automatic detection of motion direction changes in the human brain. *Eur J Neurosci* 19:1978–1986

- Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis* 10:437–442
- Pelton LH, Solley CM (1968) Acceleration of reversals of a Necker cube. *Am J Psychol* 81:585–588
- Pitts MA, Neger JL, Davis TJR (2007) Electrophysiological correlates of perceptual reversals for three different types of multistable images. *J Vis* 7:6
- Pitts MA, Gavin WJ, Neger JL (2008) Early top-down influences on bistable perception revealed by event-related potentials. *Brain Cogn* 67:11–24
- Rao RP, Ballard DH (1999) Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci* 2:79–87
- Rolls ET, Deco G (2010) *The noisy brain*. Oxford University Press, New York
- Russo E, Pascalis VD (2016) Individual variability in perceptual switching behavior is associated with reversal-related EEG modulations. *Clin Neurophysiol* 127:479–489
- Sehatpour P, Molholm S, Javitt DC, Foxe JJ (2006) Spatiotemporal dynamics of human object recognition processing: an integrated high-density electrical mapping and functional imaging study of “closure” processes. *NeuroImage* 29:605–618
- Shapiro A, Moreno-Bote R, Rubin N, Rinzal J (2009) Balance between noise and adaptation in competition models of perceptual bistability. *J Comput Neurosci* 27:37–54
- Stefanics G, Kremláček J, Czigler I (2014) Visual mismatch negativity: a predictive coding view. *Front Hum Neurosci*. doi:10.3389/fnhum.2014.00666
- Stefanics G, Kremláček J, Czigler I (2016) Mismatch negativity and neural adaptation: two sides of the same coin. Response: commentary: visual mismatch negativity: a predictive coding view. *Front Hum Neurosci*. doi:10.3389/fnhum.2016.00013
- Sterzer P, Kleinschmidt A (2007) A neural basis for inference in perceptual ambiguity. *Proc Natl Acad Sci USA* 104:323–328
- Sterzer P, Kleinschmidt A, Rees G (2009) The neural bases of multistable perception. *Trends Cogn Sci* 13:310–318
- Tong F, Meng M, Blake R (2006) Neural bases of binocular rivalry. *Trends Cogn Sci* 10:502–511
- Urakawa T, Inui K, Yamashiro K, Kakigi R (2010) Cortical dynamics of the visual change detection process. *Psychophysiology* 47:905–912
- Urakawa T, Ogata K, Kimura T, Kume Y, Tobimatsu S (2015) Temporal dynamics of the knowledge-mediated visual disambiguation process in humans: a magnetoencephalography study. *Eur J Neurosci* 41:234–242
- Urakawa T, Hirose N, Mori S (2016) Reduction in the reverse-bias effect by an abrupt break in the sequential regularity of visual events. *Perception* 45:474–482
- Vogel EK, Awh E (2008) How to exploit diversity for scientific gain. *Curr Dir Psychol Sci* 17:171–176
- Winkler I, Czigler I (2012) Evidence from auditory and visual event-related potential (ERP) studies of deviance detection (MMN and vMMN) linking predictive coding theories and perceptual object representations. *Int J Psychophysiol* 83:132–143