The Figure is in the Brain of the Beholder: Neural Correlates of Individual Percepts in the Bistable Face-Vase Image

---

A Thesis

Presented to

The Division of Philosophy, Religion, Psychology, and Linguistics

Reed College

In Partial Fulfillment

of the Requirements for the Degree

Bachelor of Arts

Phoebe Bauer

May 2015

Approved for the Division

(Psychology)

Michael Pitts

Acknowledgments

I think some people experience a degree of unease when being taken care of, so they only let certain people do it, or they feel guilty when it happens. I don't really have that. I love being taken care of.
Here is a list of people who need to be explicitly thanked because they have done it so frequently and are so good at it:

Chris: thank you for being my support system across so many contexts, for spinning with me, for constantly reminding me what I'm capable of both in and out of the lab. Thank you for validating and often mirroring my emotions, and for never leaving a conflict unresolved.

Rennie: thank you for being totally different from me and yet somehow understanding the depths of my opinions and thought experiments. Thank you for being able to talk about magic. Thank you for being my biggest ego boost and accepting when I internalize it.

Ben: thank you for taking the most important classes with me so that I could get even more out of them by sharing. Thank you for keeping track of priorities (quality dining: yes, emotional explanations: yes, fretting about appearances: nu-uh).

#AshHatchtag & Stella & Master Tran: thank you for being a ceaseless source of cheer and laughter and color and love this year.

Orestis: thank you for having the time and the energy and the compassion and the knowhow because I don't think this experiment would have been done in time without you (or at least, my relationship with Michael wouldn't be as good as it is).

Isabel: thank you for always knowing the precise nuances of what I find beautiful. Thank you for being able to talk about every single thing I think is worth thinking about.

Eli: thank you for completing my soul. Thank you for listening when I monologue, for being my truest feelings equal, for doing whatever it is you've done that makes me comfortable telling you the things I've barely told myself. Thank you for being an anchor for my most adventurous self. ~~

Emilime: thank you for growing and changing and keeping track and holding on. Thank you for being the easiest person in my life to get along with.

Flip: thank you for being there when I'm ready and being ready when I'm there. Thank you for reminding me that love is the most important thing and that it's solid and simple and wholesome.

Bonnie: thank you for all your deep magical thought bubbles.

Amanda: thank you for eight years of sage wisdom and sass encouragement.

Simon: thank you for understanding the things I wouldn't think to explain. Thank you for forgiving and being worth forgiveness.

Mom: thank you for teaching me to love books and puzzles and flowers and flavors and fabric. Thank you for teaching me that my choices are the ones that matter.

Dad: thank you for teaching me to love bones and branches and birds and mountains and the sky. Thank you for letting me follow my bliss without worrying about the feasibility.

Michael: thank you for ignoring my self-deprecation and anxiety and for holding faith in the value of this project. Thank you for showing me how to see patterns and stories in overwhelming amounts of data. Thank you for your contagious excitement, for inspiring me to keep reading and keep thinking, and for constantly proving to me that neuroscience is *fun*.

Enriqueta: thank you for words of encouragement and for being available and attentive. Thank you for keeping every lab endeavor grounded in reality and for holding us all accountable for our claims.

Table of Contents

List of Tables

List of Figures

Abstract

The purpose of the present study was to examine the neural correlates of individual percepts in the bistable face-vase figure. Stimuli were presented intermittently (interleaved with blank interstimulus intervals) to allow time-locking of ERPs. A larger N170 was observed when participants reported seeing the face as the figure than when they reported seeing the vase. No N170 difference was found between the percepts in a spatially equivalent control stimulus with identifiable facial features removed. A later negativity was observed between percepts in the control stimulus. These results demonstrate that distinct percepts arising from a single common retinal input are associated with differential, category-specific neural activity.

Comparisons were also made between “reversal” and “stable” trials, and two components identified in prior research were replicated: a reversal negativity (RN), from ~200-350ms over posterior electrodes, and a late positive complex (LPC) from ~400-600ms over central-parietal electrodes. Both reversal components were modulated during an active control task in which participants were asked to volitionally control their percepts on each individual trial. Theories about the responsibility of top-down and bottom-up processing for visual bistability and, in particular, evidence in the current findings for the manifestation of each in the neural components are discussed. Methods for furthering the exploration of these contrasting (but likely coexisting) mechanisms are outlined.

for M & D

for making me bright

Chapter 1

Introduction

1.1 Bistable Images

A multistable figure is an ambiguous image that can be perceived in multiple mutually exclusive ways. While looking at the same unchanged physical stimulus, an observer can experience one of a variety of different percepts. A bistable image is one with two valid perceptual interpretations, each arising from a single constant retinal input. When someone looks at a bistable figure the two percepts will alternate spontaneously, such that each will be seen for part of the time but they will never be perceived simultaneously. Examples of multistable images include the Necker cube (which can be arrayed into a lattice composed of nine individual cubes), Shroeder's staircase, and Lemmo's ambiguous cheetahs (Figure 1.1).

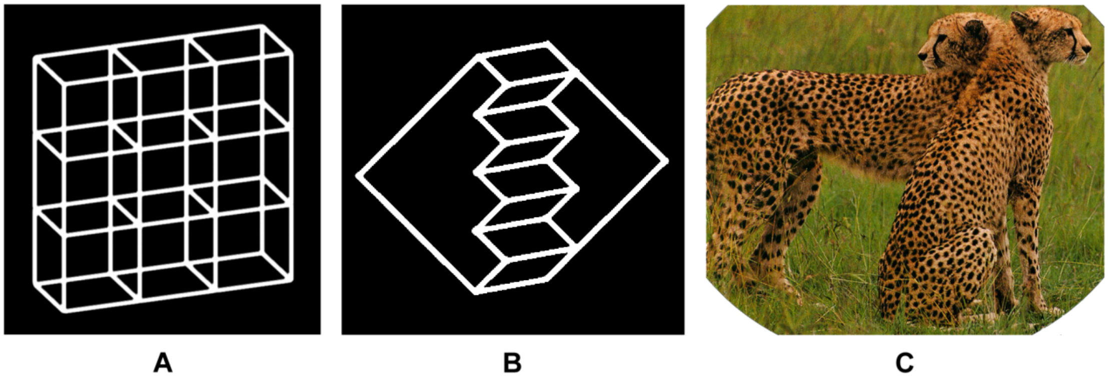


Figure 1.1: Examples of multistable images.

The Necker lattice (A) is composed of nine Necker cubes stacked on top of each other in a 3 x 3 array. Schroeder's staircase (B), modified to be tilted 45 degrees, can be seen as four stairsteps up to the left or three stairsteps up to the right. The sitting cheetah in Lemmo's ambiguous cheetahs (C) can be seen as looking over its shoulder to the left while the standing cheetah looks forward, or the sitting cheetah can be seen as looking forward to the right while the standing cheetah turns to look behind itself to the left.

Line drawings like the cube or the staircase have a pair of illusory three-dimensional percepts, but it may be technically inaccurate to refer to these images as bistable, because they can also be seen as a two-dimensional pattern of lines. However, if we consider only the three-dimensional interpretations, the perceptual experience when viewing these images is dependent on perspective: for the cube and the staircase, differences between percepts are due to which face or corner of the figure is oriented towards the front (in relation to the observer). For other bistable figures, percepts are dependent on edge assignment, wherein the contours between each portion can be grouped with either the figure or the background, and those seen as the figure are given priority in subsequent processing (Driver & Baylis, 1996). Rubin's face-vase drawing is an example of a bistable image dependent on figure-ground segregation. It can be seen either as a single central vase or as two facial profiles looking in at each other, and each individual percept causes the other figure to fade into the background. (Figure 1.2).

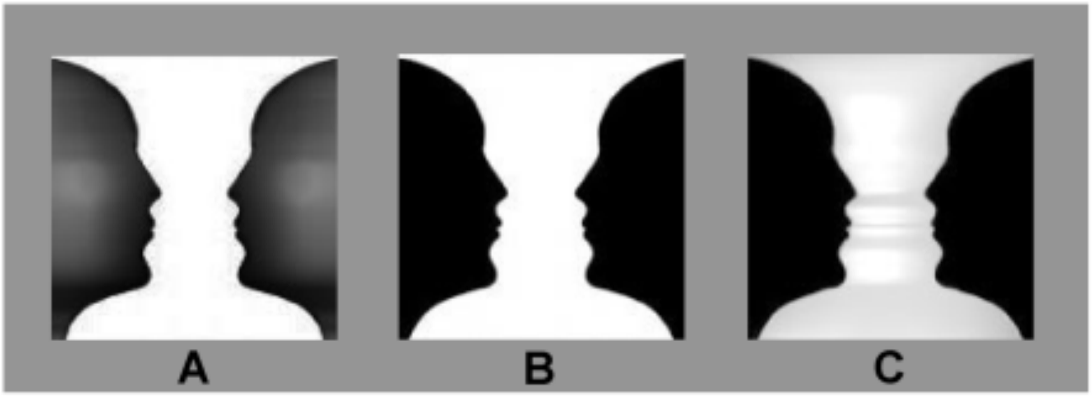


Figure 1.2: Three variants of Rubin's bistable face-vase drawing.
The normal face-vase (B), can be seen either as two peripheral face profiles looking inwards (in black), or as a single central vase (in white). Depth cues such as light and shading can bias figure-ground segregation towards perceiving the faces (A) or the vase (C) as figure more often.

1.2 Why look at bistability?

Ambiguous figures are interesting not only as a peculiarity, but because of what they can reveal about the visual system, in particular the higher levels of visual processing involved in interpreting sensory input. Multistability is a product of normal visual processes – edge assignment and figure-ground segregation are utilized when viewing most things in the world, but they usually result in a single optimal interpretation (Wagemans et al., 2012). When the object of perception leads to two or more valid and unique interpretations, the perceptual system does not pick one, but instead adopts each interpretation for a certain amount of time (typically a few seconds) before switching to another interpretation in an continuously alternating cycle.

Gestalt psychologists theorized that the interpretation that wins out and is perceived is that which is the 'best,' which Hochberg and McAlister (1953) quantified as that defined by the smallest amount of information. They measured the proportion of the time that various projections of a Necker cube were seen as a two-dimensional pattern, rather than a three-dimensional solid, and found that more simple projections (Figure 1.3 C) were seen as flat for the majority of the time. When alternate percepts are equally simple or 'good', the visual system will switch between them and the figure will be multistable, either between two- and three-dimensions (Figure 1.3 B) or between three-dimensional percepts from different angles (Figure 1.3 A). The Gestalt understanding also includes goodness based on experience, which is why the Necker lattice in Figure 1.1 may be slightly biased towards the left – because we are more used to viewing objects from above than below (Williams, McCoy, & Purves, 1998). It is also worth noting that if a person comprehends only one percept, they can hold the figure in their visual field for quite a long time without it alternating at all. In fact, if shown an unambiguous version of the figure before ever seeing the ambiguous version, one may be entirely biased towards that interpretation until the other aspect is pointed out (Leeper, 1935).

Ambiguous figures also serve as a useful tool for cognitive neuroscience. Because multiple distinct percepts arise from a single physical stimulus, the neural activity measured in response to such a stimulus avoids the common confound of differential sensory input. Rather, any differences measured due to changes in percept are interpretable as neural correlates of perceptual switching (not sensory switching), directed attention, or consciousness of a particular percept.

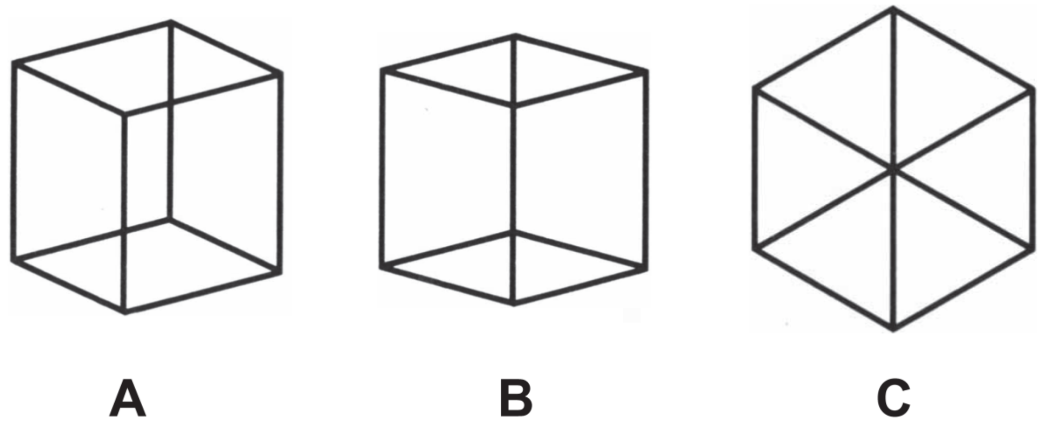


Figure 1.3: The Kopfermann cubes examined by Hochberg; image taken from Attneave, 1971.
A cube like (A), which is made up of 16 line segments and has 25 angles and 10 intersections, is usually perceived in one of its three-dimensional interpretations (only bidimensional 1.3% of the time), while (B) has 13 segments, 19 angles, and 8 intersections and is seen in two and three dimensions equally (bidimensional on 49% of responses), and (C), which has only 12 segments, 17 angles, and 7 intersections, is most often reported to be a regular hexagon divided into six equilateral triangles (60% of responses).

1.3 Theoretical Accounts of Multistability

Leopold and Logothetis (1999) postulate that perceptual reversals are a result of a high-level ‘exploratory’ mechanism that directs perceptual systems to periodically reorganize and refresh incoming information. This cognitive network would generally function via eye-movements, and process easily disambiguated objects in the visual field. When applied to a bistable stimulus in a scenario where the eyes must be held at a fixation point, this exploratory mechanism would direct covert attention such that the experienced percept alternates.

The exploratory theory may be reconcilable with a bottom-up or feedforward view of processing. In this model reversals are a result of cycles of satiation (due to adaptation or fatigue) and recovery by the neural circuits responsible for each percept

(Kohler 1940; Long & Toppino, 2004; Orbach, Ehrlich, & Heath, 1963). The bottom-up view would characterize bistable perception as a sensory-driven process in which information flows from low-level sensory processing in visual cortex through intermediate stages to cognitive awareness. This approach would account for the exploratory model as a subconscious, automatic process that influences incoming sensory information so that a passive viewer will end up experiencing perceptual switches despite having no intention to look for one percept or another.

Yet the exploratory mechanism also seems in line with the antithetical top-down/feedback view, wherein attentional control by an active observer causes reversals. Here information flow from high-level nonsensory systems to lower level perceptual processing is emphasized, taking into consideration the impact of memory, learning, and decision making on reversals (Horlitz & O'Leary, 1993; Long & Toppino, 2004; Rock, Hall, & Davis, 1994). The exploratory model can be seen as a more active process, not necessarily a product of conscious decision-making, but certainly a higher level network that responds to momentary levels of satiation or fatigue and directs the sensory system to take in particular pieces of information based on the current state of perceptual processing.

It may be that both the top-down and bottom-up theories are correct, but that the contribution from each is contingent on context. As Long and Toppino (2004) thoroughly demonstrate, there is a copious amount of research in support of each process, and they convincingly argue that this suggests that they are both involved in bistable perception. Sensory input must be relayed to complex cognitive networks to produce a conscious experience of either percept, but higher-level information often influences low-level sensory processing. The exploratory theory provides a more nuanced model, taking into account both feed-forward and feedback information and allowing them to interact during dynamic viewing conditions. It also eliminates the need to characterize selective attention as flowing entirely in either direction, which is consistent with the well known distinction between top-down (endogenous) and bottom-up (exogenous) attention (Katsuki & Constantinidis, 2013).

1.4 Neural Correlates of Perceptual Reversals

The neural activity observed when a person views a bistable stimulus can be used to address the theoretical queries about what causes perceptual switches. Historically, attempts at recording brain activity related to perceptual reversals involved continuous, static presentations of bistable stimuli and button-press reports from participants to indicate that a perceptual switch had just occurred (Başar-Eroglu, Strüber, Stadler, Kruse, & Başar, 1993; İşoğlu-Alkaç et al., 1998; Strüber, Basar-Eroglu, Miener, & Stadler, 2001). With continuous presentation, the only thing to time-lock event related potentials (ERPs) to is the button-press, so in these studies a technique of "backwards-averaging" from the time of the response was used. These studies found a slow frontocentral positive wave associated with perceptual reversals beginning about 500ms prior to the button-press, but due to the inter- and intra-individual variability of the latency of participants' responses, this component was quite broad. The duration between an experienced reversal and participants' reports of that reversal are variable from trial-to-trial (and participant-to-participant), and therefore the measurement of the perceptual event is temporally uncertain. There are likely early ERP effects that are not identifiable without a more precise time reference (Kornmeier and Bach, 2004).

A solution to this time reference problem was developed by O'Donnell, Hendler, and Squires in 1988: they presented a Necker cube discontinuously, with breaks between presentations during which participants were asked to report which orientation they had perceived before the stimulus disappeared. Discontinuous presentation allows for much tighter control over the measurement of brain activity associated with perceptual switches, because ERPs to the percept can be measured by timelocking to stimulus onset. Kornmeier and Bach (2004) adjusted the paradigm to include much shorter interstimulus intervals (ISIs), because the 3.3s delays used by O'Donnell and colleagues were so long that they greatly decreased the prevalence of reversals (Kornmeier & Bach, 2004; Leopold, Wilke, Maier, & Logothetis, 2002; O'Donnell et al., 1988). Presenting a bistable figure for an 800ms duration with a 400ms ISI allows for only one of the two percepts to be experienced during each presentation, and for both perceptual repetitions and reversals to occur across successive trials.

The intermittent presentation paradigm has been utilized to examine ERP components associated with a switch in perception from one stimulus presentation to another by contrasting “reversal trials” with “stable trials”. Reversal trials are identified by asking participants to press a button whenever their perception differs from the preceding trial; stable trials, in contrast, refer to consecutive stimulus presentations on which the same percept is maintained. Many studies have compared reversals to stables and found a late positive component (LPC) about 300 or 400ms after stimulus onset at frontal and parietal electrode locations for reversal trials (Britz, Landis, & Michel, 2009; Intaitė, Koivisto, Ruksenas, & Revonsuo, 2010; Kornmeier & Bach, 2004, 2005, 2006; Kornmeier, Ehm, Bigalke, & Bach, 2007; Pitts, Gavin, & Nerger 2008; Pitts, Martinez, Stalmaster, Nerger, & Hillyard, 2009). This positivity likely corresponds to the one observed using static presentation, since the latter is maximal 125-250ms prior to the response, which (given a 600ms reaction time) would correspond to between 300 and 400ms after the occurrence of the reversal (İşoğlu-Alkaç et al., 1998; Kormeier & Bach, 2006).

In addition to the “late positive complex” (LPC), the aforementioned studies as well as one other (Pitts, Nerger & Davis, 2007) have identified a negative-going component over occipital and parietal sites between 200 and 400 ms post-stimulus-onset on reversal relative to stable trials. This “reversal negativity” (RN) has been shown using a variety of different bistable figures as well as unambiguous versions of certain figures: both it and the LPC are also present following exogenous reversals, but are found to occur earlier in time as compared to endogenous ones (Intaite et al., 2010; Kornmeier & Bach, 2006). The RN likely is not found in experiments using static presentation because it is too early relative to stimulus onset (and therefore the time of the shift in perception). The RN is also absent from O’Donnell’s data, possibly because the duration of their ISIs was so long that the few reversals that did occur are likely less directly comparable to standard reversals (Kornmeier & Bach, 2004).

Some of these authors suggest that the LPC is an index of cognitive processing and therefore is evidence in support of top-down theories of perceptual reversals (Kornmeier & Bach, 2006). Others interpret the processing indexed by the LPC as post-perceptual, a reflection of context updating in visual short term memory (Intaitė et al., 2010; Pitts et al., 2008). The RN, in contrast, because it occurs so early in response to a reversal, is viewed as a shift in selective attention, evidence of a bottom-up mechanism (Kornmeier & Bach, 2006; Pitts et al., 2008). The fact that it looks similar in response to ambiguous and unambiguous figures, merely shifted a bit earlier for the unambiguous, suggests that it does indeed correspond to some degree to low-level sensory processing (Intaitė et al., 2010; Kornmeier & Bach, 2006). However these studies mostly conclude that their findings are not conclusive evidence for or against either theoretical model.

1.5 Manipulations of Top-Down Influences

To try to further tease out the role of top-down and bottom-up processes in bistable perception, conditions can be implemented in which participants are asked to voluntarily control their perceptual experience. Behavioral experiments have demonstrated that such control is possible, and that when asked to speed up or slow the rate of reversals for bistable stimuli, participants are indeed able to do so (Liebert & Burk, 1985; Meng & Tong 2004; Toppino, 2003; van Ee, van Dam, & Brouwer, 2005). It must be noted however, that complete control over reversals isn't exhibited. In a condition where participants' intention is to increase the number of reversals as much as possible, some stable trials still occur, and likewise, in a condition where participants' intention is to maintain a single percept for as long as possible, some reversal trials occur (Pitts et al., 2008; Slotnick & Yantis, 2005). Bottom-up processing still plays a part, and some low-level sensory process seems to get in the way of active, higher level goals.

Neuropsychological data shows that patients with frontal lobe lesions have difficulty shifting between percepts relative to patients with posterior lesions or healthy controls (Ricci & Blundo, 1990). In particular, patients with frontal lesions have difficulty increasing the rate of reversals when compared to control participants, despite exhibiting comparable switching rates in a passive condition and even being able to maintain stable percepts for as long as controls (Windmann, Wehrmann, Calabrese, & Gunturkun, 2006). This suggests a distinction between the neural mechanisms responsible for these different types of volitional control of percept, perhaps depending on whether the system is trying to sustain visual attention or induce an attentional shift.

To examine the neural activity underlying perceptual reversals more directly, in 2008, Pitts et al. examined voluntary reversals as compared to involuntary ones using EEG. They observed an increase (making the component more negative) in the amplitude of the reversal negativity (RN) for conditions in which participants were instructed to speed-up reversals, as compared to the RN measured during passive trials (Pitts et al., 2008). They interpret this RN enhancement as evidence for the role of selective attention in perceptual reversals; a result of a voluntary (top-down) attentional shift.

1.6 Neural Correlates of Individual Percepts

Despite the large amount of research investigating perceptual switching of bistable stimuli, differential activity in response to each individual percept has not yet been examined, at least not with EEG techniques. Visual cortex activity has been shown to vary with reported percepts using fMRI (Andrews, Schluppeck, Homfray, Matthews, & Blakemore, 2002; Hesselmann, Kell, Eger, & Kleinschmidt, 2008), but this technique does not exhibit the temporal resolution necessary for identifying the brain's immediate response to a briefly presented stimulus. Steady state frequency tagging with MEG has also shown differential activity to distinct percepts, such that tagging different regions of a bistable figure elicited discernable oscillatory signals over visual cortex (Parkkonen, Andersson, Hämäläinen, & Hari, 2008). Single-unit recordings in non-human primates have shown that individual neurons in the inferior temporal cortex increase or decrease their firing rates according to the currently dominant percept, while neurons in primary visual cortex respond to their preferred stimulus regardless of what the animal reports perceiving (Leopold & Logothetis, 1999; Logothetis & Schall, 1989; Sheinberg & Logothetis, 1997; Dodd et al., 2001). It therefore seems reasonable to expect there to be some measurable difference in the ERPs associated with each of the two possible percepts.

In prior studies using Kornmeier and Bach's intermittent presentation paradigm, participants were always asked to report when reversal and/or stable trials occurred, which cannot reveal which percept was experienced on a given trial. However, this information becomes readily available with a slight change to the task: ask participants to respond with respect to which percept they see on each trial, rather than what they perceive in relation to the previous. With this alternate reporting task it is still possible to separate reversal and stable trials, by looking backwards in time to the reported percept on the preceding trial to determine if perception changed or stayed the same. It is not necessary to ask participants directly whether they experienced a reversal, because we know one occurs if they report seeing two different percepts on two successive trials. One study used this type of reporting with a Necker cube stimulus, but they do not report any analysis of neural activity recorded in response to individual percepts (Pitts et al., 2009).

1.7 The Present Study

We endeavored to conduct an EEG experiment to investigate the neural correlates of individual percepts. Rubin's face-vase stimulus was deemed an ideal bistable figure to use for this study due to the marked differences between percepts. In some bistable figures like the Necker cube, both percepts are the same shape - they are simply oriented differently. But the face-vase gives rise to two percepts with entirely independent identities. It is a particularly useful stimulus because the neural response to one of the two percepts may exhibit face-specific activity. Faces, even in profile, elicit a negative ERP component with a latency of 170ms known as the N170 (Eimer, 2000). Combined EEG-fMRI studies have shown that the N170 is correlated with activity in the face-selective areas observable using fMRI (Dalrymple et al., 2011, Nguyen & Cunnington, 2014, Yovel et al., 2008), including the areas implicated in perceiving the faces in this particular image (Hesselmann et al., 2008).

To verify that any observed ERP differences during the N170 time-frame are indeed face-specific, we created a bistable control figure with the same figure-ground ambiguity and central/peripheral layout as the face-vase but with smoothed contours that removed all identifiable facial features. The current investigation also sought to expand on the previously demonstrated differences between active and passive reversals by including passive conditions and conditions in which participants were instructed to voluntarily control their percepts.

We expected to replicate the previously reported reversal negativity for both the face-vase and the control stimulus. We predicted that this reversal negativity would grow in amplitude (becoming more negative) during the active task wherein participants were asked to control their perception. Most importantly, we hypothesized that a larger N170 would be elicited by the face-vase stimulus when participants reported face percepts compared to vase percepts, and that this difference would be absent in the control stimulus when comparing the peripheral to central percepts.

Chapter 2

Methods

2.1 Participants

A total of 25 Reed College community members between the ages of 18 and 29 (9 male, 16 female; mean age = 20.8, SD = 3.04) participated in the study. Participants had normal or corrected-to-normal vision and were screened to ensure that they had no history of a neurological condition or recent brain injury that might interfere with electrophysiological activity. Participants were compensated with psychology lottery tickets (1 ticket per half hour of participation), to be entered into a pool with other psychology experiment participants, of which one received $50 at the end of the semester. All procedures adhered to federal regulations and were approved by the Reed College Institutional Review Board. Written informed consent was obtained from each participant at the start of the experimental session.

2.2 Stimuli

Two bistable images served as the stimuli for this study: a modified Rubin's face-vase and a control stimulus designed to be as physically similar to the face-vase image as possible without including the same facial features. The control stimulus was dubbed the "backpack-chalice", since it can be perceived as either a central chalice or as two backpacks worn by people facing away from one another (Figure 2.1).

Both stimuli subtended a visual angle of 2.6° x 2.6° and were presented in the center of the screen using Presentation software (Neurobehavioral Systems, San Fransisco CA) on a Dell SA2311W Monitor with a frame rate of 60Hz. The background screen was grey, so that both portions of the stimulus would be visible. Participants maintained their gaze on a small (0.08°) centrally located black fixation dot that remained on screen while stimuli were presented and during the otherwise blank screen between presentations.

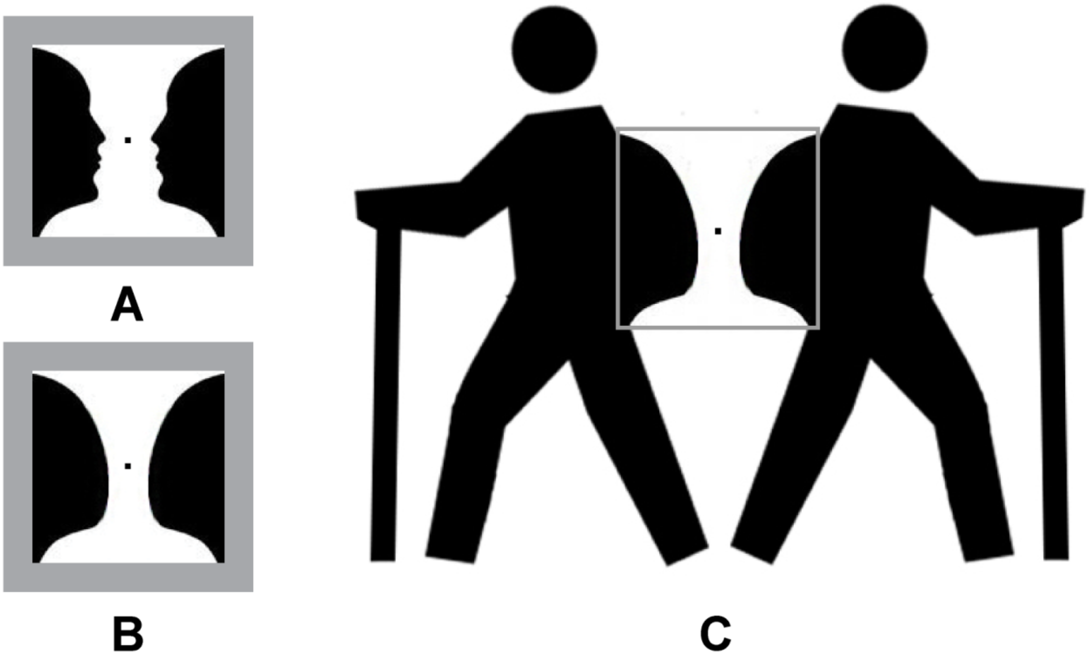


Figure 2.1: Experimental Stimuli.
The face-vase stimulus (A) and the backpack-chalice control stimulus (B) used in the present study, as well as the visual aid for the backpack-chalice (C) which was shown to participants prior to the start of experimental trials.

2.3 EEG Apparatus

EEG scalp voltages were recorded using a custom 96 channel Herrsching DE-8211 “Easycap”. Locations in the cap were modified from the standard system for equidistance. Electrode signals were amplified by BrainVision “Professional BrainAmp” amplifiers with a 500Hz digitization rate, and electrode impedances were kept below 5 kΩ. This was achieved with the use of a saline-based gel (ABRALYT HiCl) and gentle abrasion of the skin under each electrode. During task performance, eye position and eye movements were monitored by means of vertical and horizontal electrooculogram recordings. Participants’ responses were recorded with a Cedris RB-830 button pad.

2.4 Procedure

Participants were seated 75cm away from the computer monitor in order to maintain consistent retinal image size. Prior to any recordings, the stimuli were presented statically so that each percept could be identified. The visual aid for the backpack-chalice stimulus (consisting of two people positioned as though wearing the backpacks, shown in Figure 2.1) was also shown. Participants were allowed to alternate between viewing the stimulus alone and with the visual aid until they felt confident in their ability to interpret the backpacks as such.

During the experiment, stimuli were flashed on the screen for 800ms followed by a 400-600ms interstimulus interval (ISI) prior to onset of the subsequent stimulus. Participants were instructed to respond based on which percept they experienced on a given trial, pressing one button if they saw the faces or the backpacks, and another if they saw the vase or the chalice. The two stimuli were presented in separate blocks, so each button corresponded to only one percept at a given time. Participants were told to withhold any response if the image did not give rise to one clear percept. Which button corresponded to which percept was counterbalanced across participants, and the right index and middle fingers were used in all cases to press the two buttons. Figure 2.2 depicts the stimulus presentation paradigm.

During the passive task, participants were explicitly instructed not to voluntarily induce reversals or otherwise control their perceptual experience. They were told to refrain from trying to interpret the figure in a particular way and merely report which percept naturally popped out as the figure. In contrast, for the active task they were instructed to volitionally control their percepts. Specifically, they were asked to intentionally perceive a particular interpretation on each trial so that they would end up seeing about equal numbers of faces and vases (or of backpacks and chalices), as well as equal numbers of reversals and stables from one stimulus presentation to the next. Participants were cautioned that sometimes the stimulus would not give rise to the percept they intended, and to respond to what they actually saw on a given trial, rather than what they were intending to see.

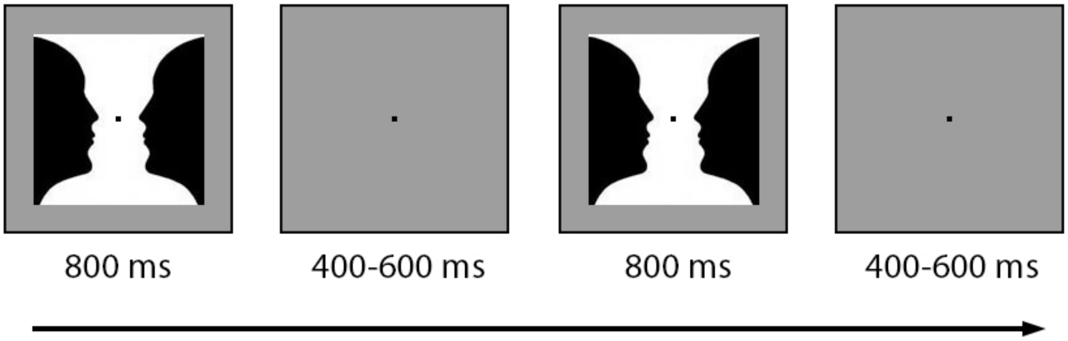


Figure 2.2: Trial schematic of the intermittent paradigm adopted from Kornmeier and Bach.
The stimulus was presented on and off for 800 ms with an ISI jittered between 400 and 600 ms. Participants were asked to respond prior to the onset of the subsequent stimulus with respect to which percept (the faces or the vase) they perceived on a given trial.

The goal of this instruction was to directly manipulate top-down influences on perception by creating an active task that would still allow for enough of each percept and each trial-type (stable or reversal) to make comparisons. Prior studies have implemented active tasks by simply asking participants to increase the rate of reversals. Because participants are not able to fully control their percepts, in these studies enough stable trials occurred to allow analysis of differences in brain activity between reversal and stable trials (Pitts et al., 2008). However, pilot tests in our lab with this particular face-vase stimulus revealed abnormally high levels of voluntary control, such that instructing participants to increase reversal rates resulted in a switch on almost every single trial. This does not leave enough stable trials to assess brain activity with a reasonable signal-to-noise ratio. The "equal numbers" instruction was adopted to maintain the general spirit of the active control condition while resulting in enough trials of each type for analysis.

To keep participants from falling into an automated motor response pattern, a warning tone (a 10ms auditory click) was presented whenever participants responded within 300ms of stimulus onset. Although complex visual processing can occur within 150ms (Thorpe, Fize, & Marlot, 1996) or even quicker (Thorpe, 2009), reaction times in such experiments are not below 300 ms. If in our study a button-press occurred earlier than 300ms, it would likely be in response to an expectation of what was going to be perceived, rather than a reaction to what actually was seen, and would therefore not be informative with regards to how to sort that particular trial. Any trial on which this auditory click was presented was excluded from analysis. Participants were informed that the sound would only occur on trials on which they responded too quickly, and that it was intended to bring their attention back to the task, to make sure they were truly responding to which percept they experienced. If more than 10% of trials across all the blocks involved a response within 300ms, the participant's data was excluded from further analysis ($n = 3$).

The experiment consisted of two long blocks (passive and active) made up of smaller blocks separated by stimulus (face-vase and backpack-chalice). Each short block consisted of 200 stimulus presentations, with breaks after every 50 presentations. The active condition consisted of 4 short blocks: 2 of each stimulus, in alternating order. For the passive condition, if at the end of 4 blocks there were not at least 150 trials in any condition, two more blocks were administered. The order of the long blocks (passive first or active first) and the short blocks (face-vase first or backpack-chalice first) was counterbalanced across participants. At least one practice block of 50 trials was given at the beginning of each long block to acclimate participants to the task and the speed of stimulus presentation. Each recording session lasted 120-180 minutes, including setup time and cap/electrode preparation.

2.5 Data Analysis

EEG data were processed using BrainVision Analyzer software (Brain Products, Germany). EEG was recorded using CPz as a reference, then re-referenced to the average of all channels. Trials were discarded semi-automatically from analysis if they contained an eye blink or eye movement artifact (EOG > 70 μV), or if any electrodes exceeded defined signal amplitudes. On average, 16.7% of trials were excluded due to these artifacts. Participants with fewer than 100 trials in any given condition after artifact rejection were excluded from analyses to ensure reasonable signal/noise ratio in the averaged ERP waveforms ($n = 2$). A total of 20 participants were included in the final dataset. ERPs were time-locked to stimulus onset, baseline corrected at -100 to 0 ms, and low-pass filtered at 25 Hz.

All data were analyzed using StatSoft STATISTICA 12.

Chapter 3

Results

3.1 Behavioral Results

Reversal rates during the passive condition for both the face-vase (43%) and the backpack-chalice (42%) stimuli were higher than has been reported in previous studies (Pitts et al., 2007, 2008; Kornmeier & Bach, 2004; Long & Toppino 2004). On average reversals occurred every 2.6 seconds (2 trials). During the active task participants successfully controlled their reversal rates, bringing them close to 50% (face-vase: 49%, backpack-chalice: 46%). Participants were slightly face-biased when viewing the face-vase stimulus (passive faces: 55%, vases: 45%) and slightly chalice-biased when viewing the backpack-chalice (passive backpacks: 43%, chalices: 57%). During the active task, participants were also successful at balancing their percept numbers (faces: 51%, vases: 49%, backpacks: 48%, chalices: 52%). Participants withheld responses indicating an intermediate, flat, or uncertain percept on 4% of trials.

Reaction times were similar across stimulus and active/passive conditions and on average fell within the 800ms interval during which the stimulus was still on screen (although this was certainly not the case for every single trial): Participants were slightly quicker to respond to the backpack-chalice percept (passive M = 607ms, SE = 13) than the face-vase (passive M = 614ms, SE = 15), and even faster in the active condition (active b-c M = 592ms, SE = 12, f-v M = 605ms, SE = 11).

3.2 Electrophysiological Results

3.2.1 Percept Differences

ERP differences due to which portion of the image was perceived as the figure (face vs. vase and backpack vs. chalice) were compared by averaging across trials on which participants reported seeing a particular percept. ERPs from representative electrodes for each percept in the face-vase image are shown for the passive task in Figure 3.1, for the active task in Figure 3.2, and for the backpack-chalice image in Figures 3.3 and 3.4.

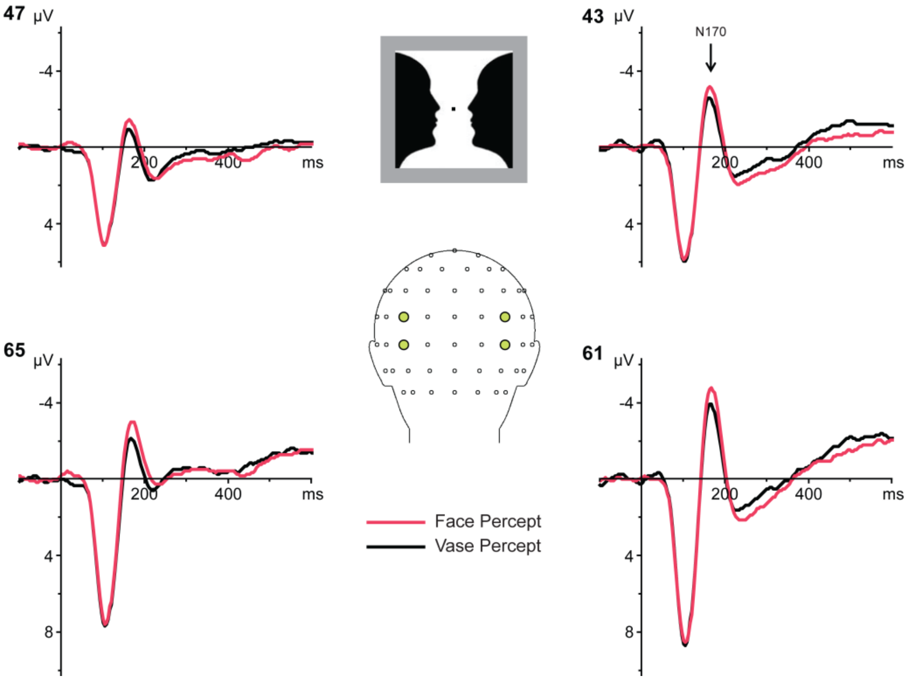


Figure 3.1: ERPs for the face and vase percepts during the passive task. Grand-averaged ERPs (*n* = 20) from four electrode sites for the face-vase stimulus, timelocked to stimulus onset. All graphs show amplitude (in microvolts), plotted by convention with negative up, as a function of time (in milliseconds post-stimulus onset). The backview of the scalp in the center shows the locations (blue circles) of the four electrodes plotted here. Note that statistical analyses were conducted on a larger set (24) of posterior electrodes. The N170 peak is indicated with an arrow at electrode 43.

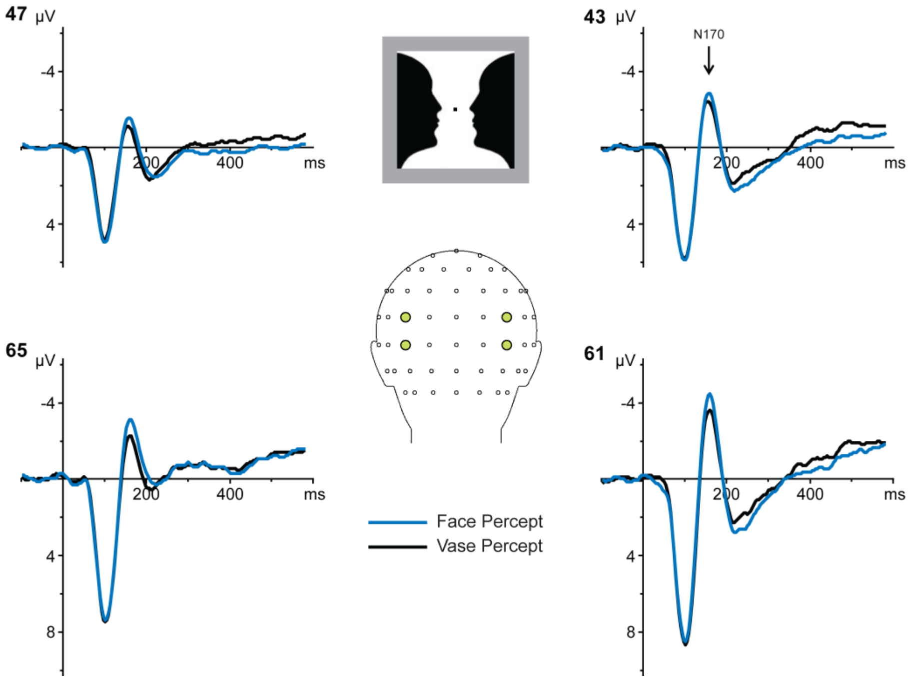


Figure 3.2: ERPs for the face and vase percepts during the active task. Plotting conventions are identical to Figure 3.1.

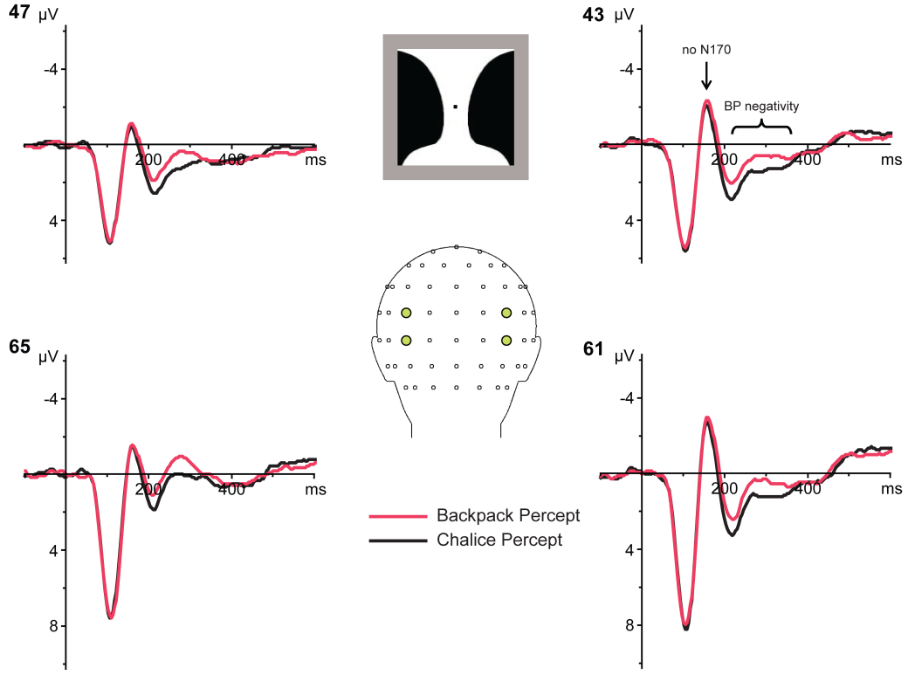


Figure 3.3: ERPs for the backpack and chalice percepts during the passive task. Plotting conventions are identical to Figure 3.1.The N170 effect was absent, but a subsequent percept difference was observed (180-280ms), labeled here for convenience as the backpack (BP) negativity.

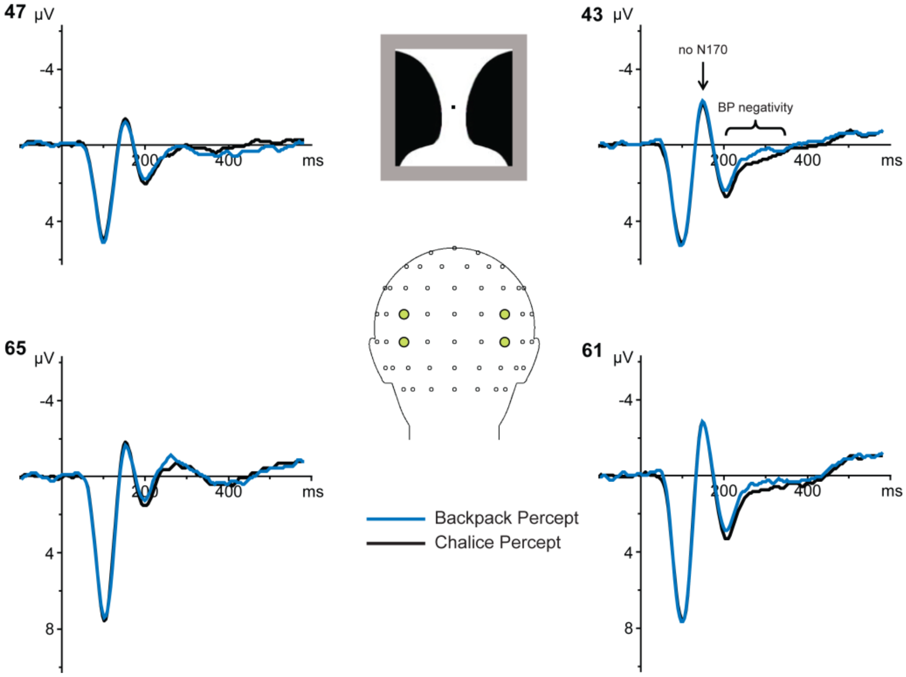


Figure 3.4: ERPs for the backpack and chalice percepts during the active task. Plotting conventions are identical to Figure 3.1.

3.2.1.1 N170 Component

An N170 effect was observed over occipital electrode sites; amplitudes were larger (more negative) for the face percept than the vase percept (Figures 3.1 and 3.2). This effect appeared to be absent for the backpack-chalice control stimulus. To assess this effect, peak amplitudes between 120 and 220ms at 24 lateral-occipital electrodes (42, 43, 44, 46, 47, 48, 60, 61, 62, 64, 65, 66, 75, 76, 77, 79, 80, 81, 87, 88, 89, 91, 92, and 93)[1] were extracted semi-automatically in each individual participant for each of the four percepts (face, vase, backpack, and chalice). A 2(condition: passive or active) x 2(stimulus: face-vase or backpack-chalice) x 2(percept: figure central or figures peripheral) x 2 (hemisphere) x 12(electrode: 12 in each hemisphere) repeated measures ANOVA was conducted, revealing main effects of stimulus, percept, hemisphere, and electrode, as well as an interaction of stimulus and percept. See Table 3.1 for *p*-values for these and the rest of the interactions. To explore the stimulus x percept interaction, paired

---

[1] For electrode locations overlaid with the international 10-20 system, see Appendix A.

t-tests were then performed on the means for each percept collapsed across condition, hemisphere, and electrode, revealing that the peak amplitudes for the face percept (*M* = -2.65 µV, *SE* = 0.48) were significantly more negative than those for the vase percept (*M* = -2.25 µV, *SE* = 0.49), *t*(19) = 3.74, *p* = 0.0014. No significant differences were seen between the peak amplitudes for the backpack (*M* = -1.64 µV, *SE* = 0.38) and chalice (*M* = -1.63 µV, *SE* = 0.37) percepts *t*(19) = 0.076, *p* = 0.94 (Figure 3.5).

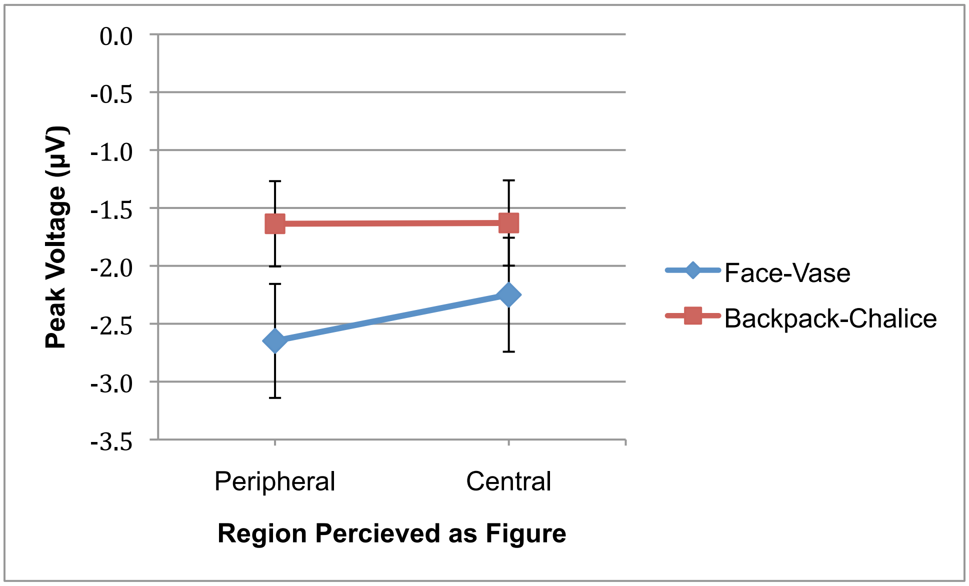


Figure 3.5: The interaction between percept and stimulus for peak N170 amplitudes. Peak amplitude of the N170 for each of the four percepts (peripheral: faces and backpacks; central: vase and chalice), averaged across participants, electrodes, hemisphere, and condition. Error bars represent ± 1 SEM. The N170 peak was significantly more negative for the faces than for the vase, but not for the backpacks compared to the chalice.

	DF	Error	*F*	*p*
Condition	1	19	0.06046	0.808415
Electrode	11	209	11.50358	0.000000
Hemisphere	1	19	5.21952	0.034006
Stimulus	1	19	16.58855	0.000649
Percept	1	19	6.47759	0.019763
Condition*Elect	11	209	1.61382	0.096569
Condition*Hemi	1	19	2.55841	0.126205
Elect*Hemi	11	209	4.57488	0.000003
Condition*Stim	1	19	1.16060	0.294831
Elect*Stim	11	209	6.69788	0.000000
Hemi*Stim	1	19	0.32794	0.573587
Condition*Percept	1	19	1.65321	0.213973
Elect*Percept	11	209	5.99096	0.000000
Hemi*Percept	1	19	0.04590	0.832641
Stim*Percept	1	19	11.45226	0.003114
Condition*Elect*Hemi	11	209	2.15879	0.017852
Condition*Elect*Stim	11	209	0.59076	0.835311
Condition*Hemi*Stim	1	19	0.00832	0.928284
Elect*Hemi*Stim	11	209	1.24876	0.256681
Condition*Elect*Percept	11	209	0.40972	0.950815
Condition*Hemi*Percept	1	19	0.35359	0.559104
Elect*Hemi*Percept	11	209	1.03309	0.418337
Condition*Stim*Percept	1	19	0.00019	0.989264
Elect*Stim*Percept	11	209	2.59861	0.004054
Hemi*Stim*Percept	1	19	0.66348	0.425425
Condition*Elect*Hemi*Stim	11	209	0.62580	0.805717
Condition*Elect*Hemi*Percept	11	209	0.79680	0.643183
Condition*Elect*Stim*Percept	11	209	2.80919	0.001946
Condition*Hemi*Stim*Percept	1	19	0.57509	0.457546
Elect*Hemi*Stim*Percept	11	209	1.13370	0.336406
C*E*H*S*P	11	209	0.51710	0.890635

Table 3.1: *F*- and *p*-values for the ANOVA examining N170 peak amplitude.

	DF	Error	*F*	*p*
Condition	1	19	0.185	0.672278
Electrode	11	209	35.747	0.000000
Hemisphere	1	19	2.689	0.117521
Stimulus	1	19	35.953	0.000009
Percept	1	19	4.249	0.053223
Condition*Elect	11	209	7.605	0.000000
Condition*Hemi	1	19	0.496	0.489762
Elect*Hemi	11	209	1.834	0.050124
Condition*Stim	1	19	0.032	0.858990
Elect*Stim	11	209	2.261	0.012755
Hemi*Stim	1	19	4.479	0.047743
Condition*Percept	1	19	4.182	0.054951
Elect*Percept	11	209	0.244	0.993919
Hemi*Percept	1	19	3.081	0.095340
Stim*Percept	1	19	0.001	0.973443
Condition*Elect*Hemi	11	209	1.362	0.192793
Condition*Elect*Stim	11	209	1.330	0.209270
Condition*Hemi*Stim	1	19	0.374	0.548095
Elect*Hemi*Stim	11	209	0.988	0.458660
Condition*Elect*Percept	11	209	0.501	0.901187
Condition*Hemi*Percept	1	19	0.704	0.411784
Elect*Hemi*Percept	11	209	0.750	0.689338
Condition*Stim*Percept	1	19	0.493	0.491081
Elect*Stim*Percept	11	209	1.685	0.078508
Hemi*Stim*Percept	1	19	0.457	0.507113
Condition*Elect*Hemi*Stim	11	209	0.928	0.514263
Condition*Elect*Hemi*Percept	11	209	1.192	0.294055
Condition*Elect*Stim*Percept	11	209	1.443	0.155835
Condition*Hemi*Stim*Percept	1	19	0.723	0.405888
Elect*Hemi*Stim*Percept	11	209	1.518	0.126645
C*E*H*S*P	11	209	0.391	0.958698

Table 3.2: *F*- and *p*-values for the ANOVA examining N170 peak latency.

The average latency of the N170 peak across all four percepts was 168.15 ms (*SE* = 1.01). Latencies were compared between percepts in the same 2 x 2 x 2 x 2 x 12 repeated measures ANOVA. A main effect of stimulus was seen, showing that the average latency for the face-vase stimulus (*M* = 170.92, *SE* = 1.52) was significantly different from that for the backpack-chalice stimulus (*M* = 165.39, *SE* = 1.18) ($F(1,19) = 35.95$, $p = 0.000009$, see Table 3.2 for other effects).

Scalp topographies for the difference waves between percepts (face-minus-vase and backpack-minus-chalice) show the distribution of the face-specific activity (Figure 3.6). The negativity for the backpack-chalice stimulus is later in time (as can be seen in the ERPs in Figure 3.3) and has a different distribution.

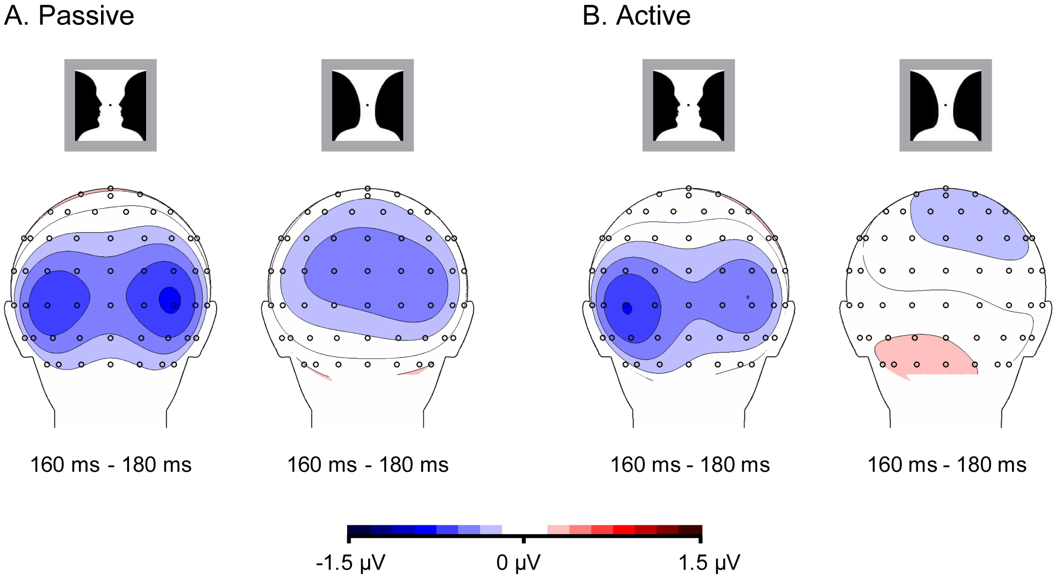


Figure 3.6: Difference maps for the N170.
Maps show the topography of the differences in peak amplitude for both the face-vase and the backpack-chalice stimuli (face minus vase; backpack minus chalice). The time-window shown is centered roughly around the average peak latency (168.15ms) and includes the average latency of each of the four individual percepts.

3.2.1.2 Backpack Negativity

A slightly later negativity was observed over occipital electrode locations for backpack percepts, as compared to chalice percepts (see Figure 3.3). This difference was analyzed by exporting mean amplitudes between 180 and 280ms at the same 24 electrode sites used for analyzing the N170. A 2 (stim) x 2 (condition) x 2 (percept) x 24 (electrode) repeated-measures ANOVA revealed a main effect of stimulus (see Table 3.3 for *p*-values and reports of the other effects), showing that the mean amplitude for the face-vase stimulus during this time-window (*M* = -0.23 µV, *SE* = 0.47) was significantly more negative than that of the backpack-chalice (*M* = 0.038 µV, *SE* = 0.42). To examine the visible difference between the ERPs for each percept in the backpack-chalice figure (see Figures 3.3 and 3.4), exploratory t-tests were conducted comparing one percept to the other for both stimuli. The mean amplitude for the backpack percept (*M* = -0.073 µV, *SE* = 0.41) was significantly more negative than that of the chalice (*M* = 0.15 µV, *SE* = 0.42), *t*(19) = 2.56, *p* = 0.019. The mean amplitudes for the face and vase percepts did not differ from one another (face *M* = -0.21 µV, *SE* = 0.47, vase *M* = -0.25 µV, *SE* = 0.48; *t*(19) = 0.44, *p* = 0.66).

	DF	Error	*F*	*p*
Stimulus	1	19	5.39442	0.031459
Condition	1	19	0.63531	0.435262
Electrode	23	437	12.20817	0.000000
Percept	1	19	2.11473	0.162208
Condition*Stim	1	19	1.66911	0.211872
Stim*Elect	23	437	2.21029	0.001161
Condition*Elect	23	437	1.95401	0.005569
Stim*Percept	1	19	4.01167	0.059663
Condition*Percept	1	19	2.89557	0.105131
Elect*Percept	23	437	3.68523	0.000000
Condition*Elect*Stim	23	437	0.42849	0.991409
Condition*Stim*Percept	1	19	0.10190	0.753048
Elect*Stim*Percept	23	437	7.59670	0.000000
Condition*Elect*Percept	23	437	1.13798	0.299703
Condition*Elect*Stim*Percept	23	437	2.70603	0.000044

Table 3.3: *F*- and *p*-values for the ANOVA examining mean amplitudes for the BPN.

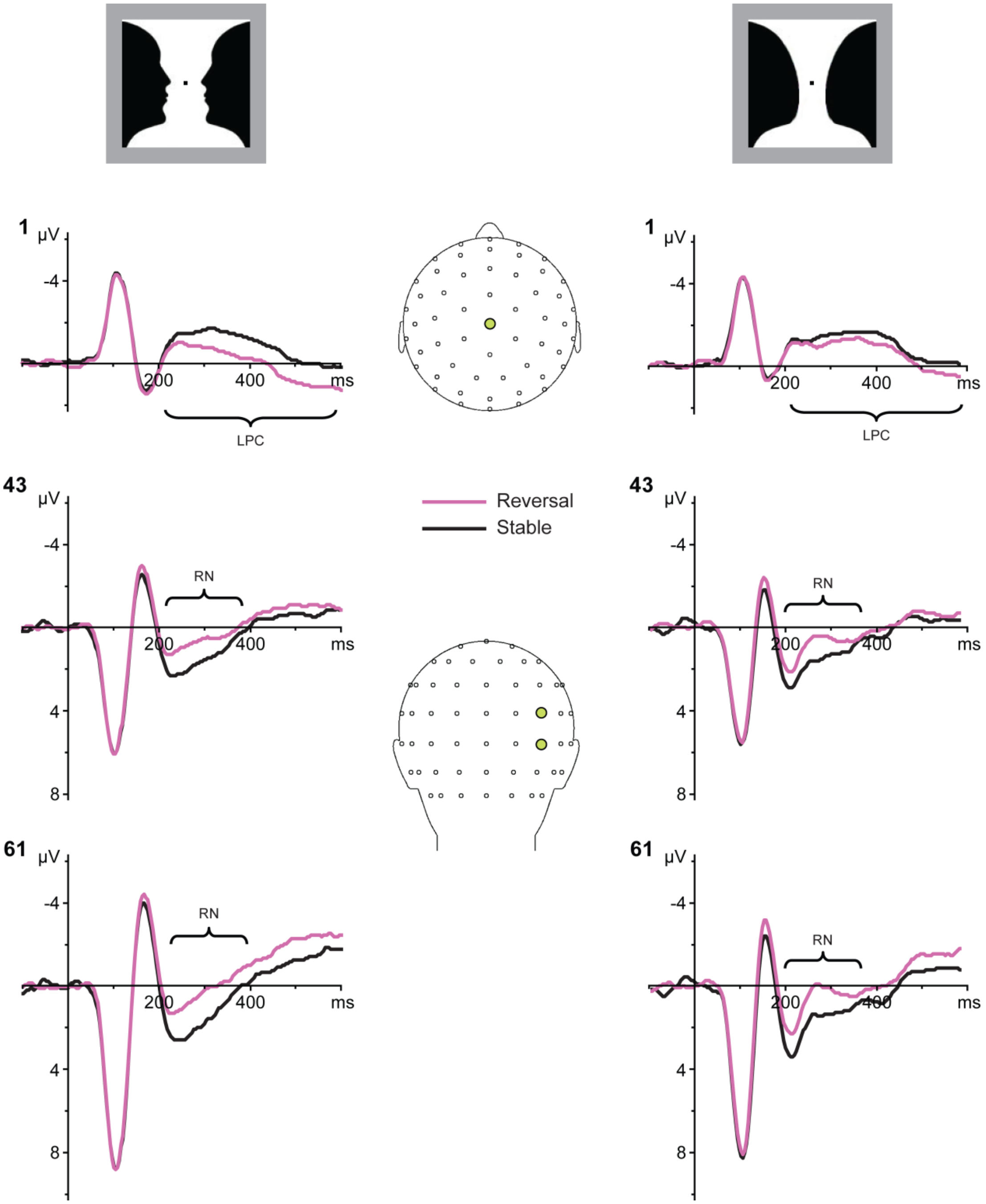


Figure 3.7: ERPs for reversal and stable trials during the passive task.

Grand-averaged ERPs (*n* = 20) from one central electrode site (1) and two right occipital sites (43 & 61) for both stimuli, time-locked to stimulus onset. Plotted electrodes are indicated on the scalp maps in the center with blue circles. All graphs show amplitude (in microvolts), plotted by convention with negative up, as a function of time (in milliseconds post-stimulus onset). The reversal negativity (RN) and the late positive complex (LPC) are indicated.

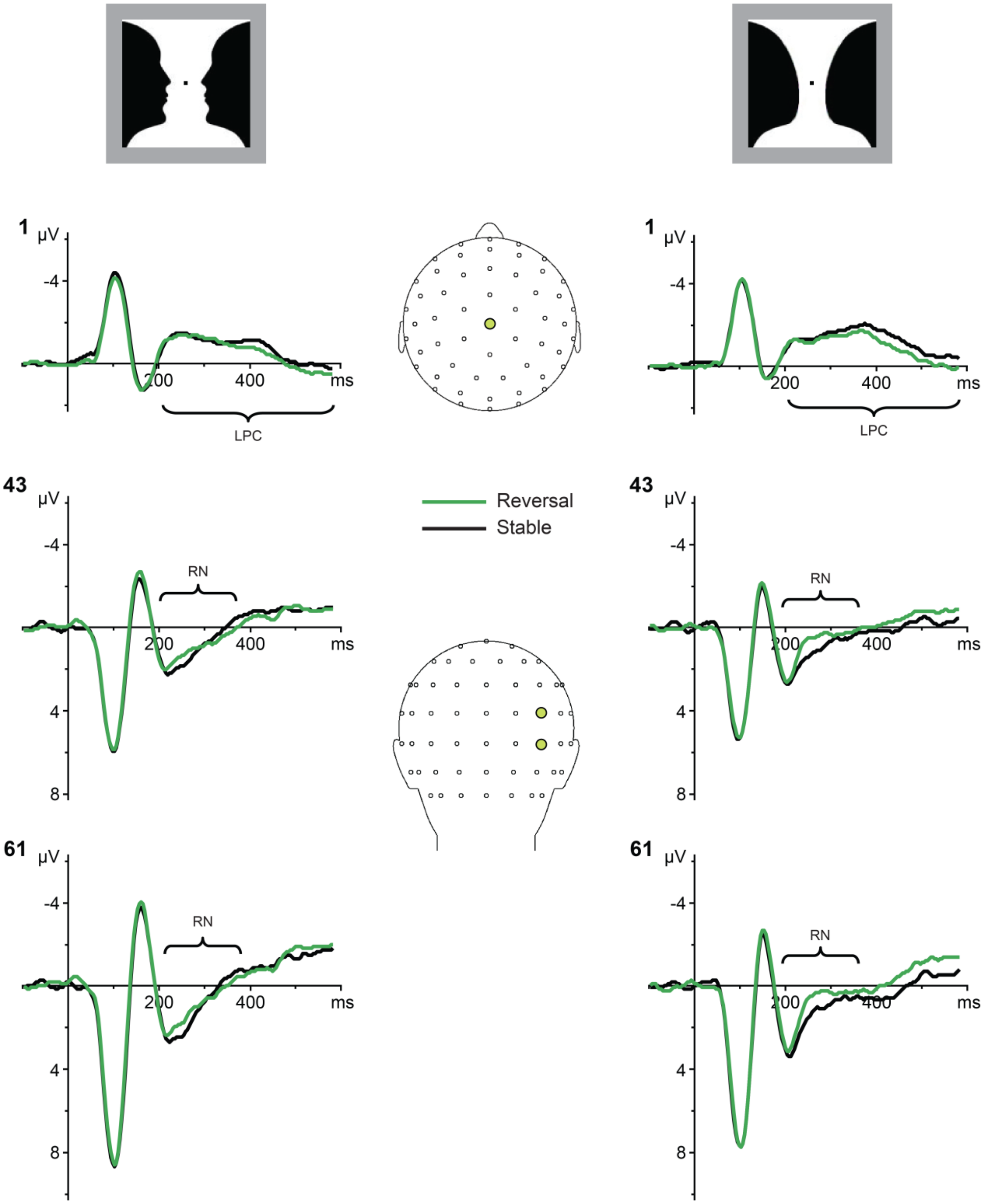


Figure 3.8: ERPs for reversal and stable trials during the active task.

Plotting conventions are identical to Figure 3.7.

3.2.2 Reversal vs. Stable Differences

Trials on which a “reversal” occurred (a change in percept relative to the previous trial) were compared to “stable” trials (those on which the reported percept was the same as it had been on the previous trial). Reversals and stables are compared for each stimulus individually, but are collapsed across percept.

3.2.2.1 Reversal Negativity (RN)

The reversal negativity was observed over occipital electrode locations for both stimuli during the passive task, and appeared much smaller for the active (Figures 3.7 and 3.8). Difference waves were created by subtracting ERPs for stable trials from reversal trials in each condition. Mean difference wave amplitudes across the timewindow of 200-300ms were analyzed at 18 lateral-occipital electrodes (42, 43, 44, 46, 47, 48, 60, 61, 62, 64, 65, 66, 75, 76, 77, 79, 80, and 81). A 2 (condition) x 2 (stimulus) x 2 (hemisphere) x 9 (electrode) repeated-measures ANOVA revealed main effects of condition, hemisphere, and electrode but not stimulus, and interactions of electrode and hemisphere, and electrode and stimulus (Table 3.4). The RN was right-lateralized in all conditions, such that mean difference amplitudes were larger in the right hemisphere ($M = -0.56$ μV, $SE = 0.121$) than in the left ($M = -0.38$ μV, $SE = 0.114$) when collapsed across condition and stimulus. Post-hoc, single sample t-tests confirmed that difference amplitudes for both passive ($M = -0.71$ μV, $SE = 0.146$; $t(19) = -4.85$, $p = 0.0001$) and active trials ($M = -0.23$ μV, $SE = 0.103$; $t(19) = -2.25$, $p = 0.036$) were significantly different from zero, although the difference was larger in the passive condition than in the active.

Topographic maps of the RN for both stimuli are shown in Figure 3.9. Scalp distributions are of the difference waves for reversal trials minus stable trials during the analyzed time-window.

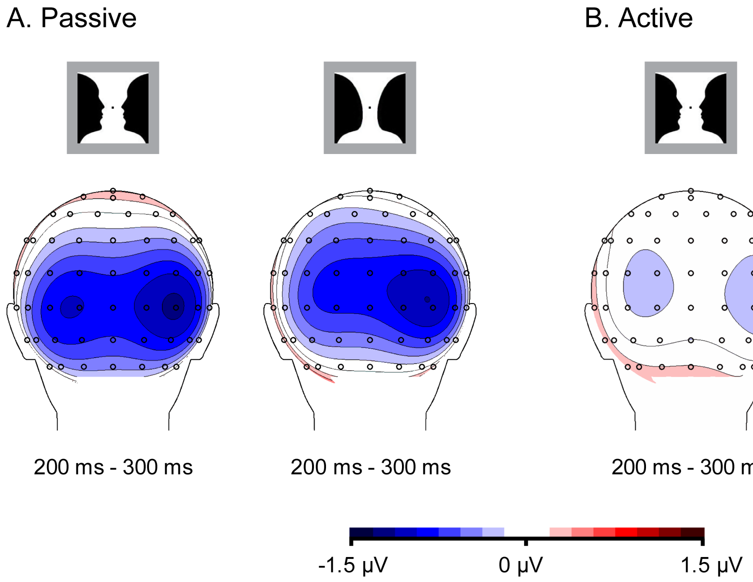


Figure 3.9: Difference maps for the RN.

Maps show the topography of the differences in mean amplitude between reversal and stable trials for both stimuli (reversal minus stable) during the analyzed time window.

	DF	Error	*F*	*p*
Condition	1	19	16.09397	0.000746
Electrode	11	209	8.42615	0.000000
Hemisphere	1	19	5.67860	0.027770
Stimulus	1	19	0.10942	0.744427
Condition*Elect	11	209	1.98138	0.052314
Condition*Hemi	1	19	2.63586	0.120950
Elect*Hemi	11	209	5.32633	0.000007
Condition*Stim	1	19	2.33968	0.142594
Elect*Stim	11	209	2.95214	0.004261
Hemi*Stim	1	19	0.18003	0.676112
Condition*Elect*Hemi	11	209	1.19375	0.306322
Condition*Elect*Stim	11	209	0.69931	0.691765
Condition*Hemi*Stim	1	19	0.59813	0.448808
Elect*Hemi*Stim	11	209	0.52610	0.835523
Condition*Elect*Hemi*Stim	11	209	1.61114	0.125881

Table 3.4: *F*- and *p*-values for the ANOVA examining RN difference amplitudes.

3.2.2.2 Late Positive Complex (LPC)

The LPC was observed across central electrode locations for both stimuli during the passive condition, but like the RN seemed to disappear during the active task (see Figures 3.7 and 3.8). Difference waves for the LPC were calculated just as for the RN, and mean amplitudes of these difference waves were analyzed from 300-600ms at central electrodes 1-6 and CPz. A 2 (condition) x 2 (stimulus) x 7 (electrode) ANOVA revealed a main effect of electrode but not of condition or stimulus, and a significant interaction of condition and stimulus, but none of the other interactions (See Table 3.5). Subsequent post-hoc paired t-tests to examine the condition x stimulus interaction showed that in the passive condition, the stimuli differed from one another ($t(19) = 3.20$, $p = 0.0047$), while in the active they did not ($t(19) = -0.94$, $p = 0.36$). However, unidirectional t-tests comparing the difference means for the individual stimuli in each condition showed that both the face-vase LPC ($M = 0.90$ µV, $SE = 0.11$; $t(19) = 8.29$, $p < 0.0000001$) and the backpack-chalice LPC ($M = 0.37$ µV, $SE = 0.15$; $t(19) = 2.57$, $p = 0.019$) differed from zero during the passive condition, but did not differ from zero during the active condition (f-v: $M = 0.23$ µV, $SE = 0.14$; $t(19) = 1.58$, $p = 0.13$; b-c: $M = 0.38$, $SE = 0.18$; $t(19) = 2.07$, $p = 0.052$) (See Figure 3.10).

	DF	Error	*F*	*p*
Intercept	1	19	29.18307	0.000033
Condition	1	19	3.68277	0.070125
Electrode	11	209	3.61473	0.002556
Stimulus	1	19	2.83519	0.108578
Condition*Elect	11	209	0.44244	0.848990
Condition*Stim	1	19	7.71443	0.011997
Elect*Stim	11	209	0.75079	0.610028
Condition*Elect*Stim	11	209	1.48594	0.189163

Table 3.5: *F*- and *p*-values for the ANOVA examining LPC difference amplitudes.

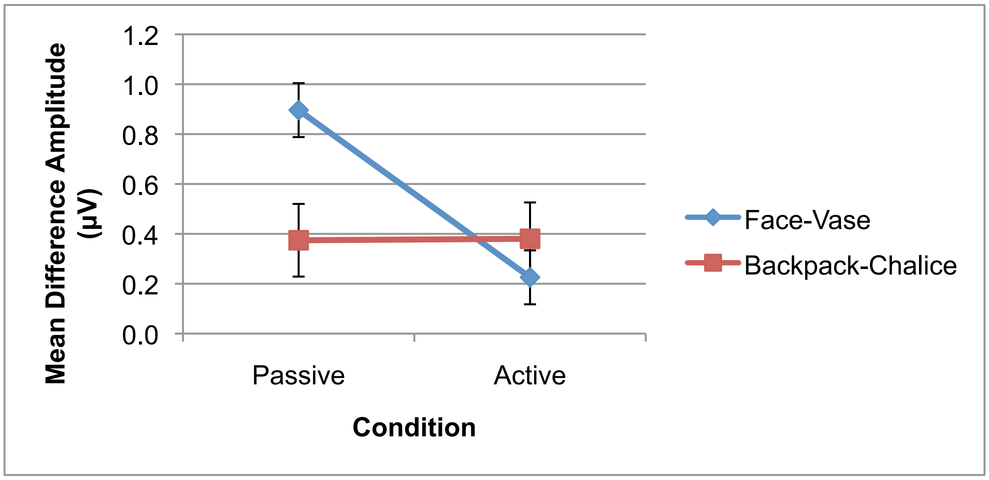


Figure 3.10: The interaction between stimulus and condition for the LPC.

Mean difference wave amplitudes (reversal-minus-stable) of the LPC for both stimuli in each condition, averaged across participants and electrode. Error bars represent ± 1 SEM. The LPC difference was significant for both stimuli in the passive condition, but not in the active.

Topographic maps of the LPC for both stimuli are shown in Figure 3.11. Scalp distributions are of the difference waves for reversal trials minus stable trials during the analyzed time-window.

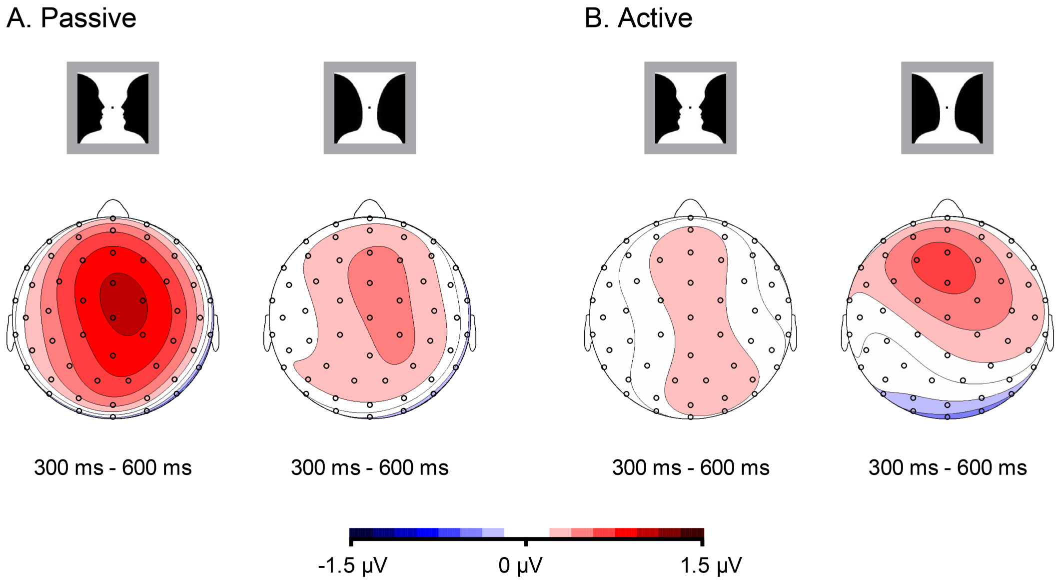


Figure 3.11: Difference maps for the LPC.

Maps show the topography of the differences in mean amplitude between reversal and stable trials for both stimuli (reversal minus stable) during the analyzed time window.

Chapter 4

Discussion

Previous EEG research on bistable images has compared perceptual reversals to perceptual stability, revealing two ERP components associated with reversals: the reversal negativity (RN) and the late positive complex (LPC). However it is unclear from previous studies what part of bistable perception either of these components actually index. To further elucidate the significance of these components as well as more thoroughly categorize the neural activity that occurs during bistable perception, we compared ERPs elicited by each individual percept. By examining the interplay between the neural processes underlying conscious perception of different interpretations of a bistable figure and those underlying switches and consistencies in perception, we can develop a more thorough understanding of the phenomenon of bistability and the neural basis of conscious visual perception.

4.1 Differences in Neural Activity Due to Percept

A larger N170 was seen for trials on which participants perceived the faces than for trials on which they perceived the vase, and no N170 difference was observed for the control stimulus. This negativity may be a neural component specific to conscious visual perception, since it differs according to subjective perception and not purely sensory input. While it is not clear from these results whether the N170 is a direct neural correlate of the awareness of faces, or an unconscious precursor necessary for conscious perception of a face, these findings demonstrate that distinct percepts arising from a single common retinal input are associated with differential, category-specific neural activity.

An amplitude difference was also observed in the backpack-chalice stimulus, wherein backpack percepts elicited more negative activity than did chalice percepts. This negativity was evident over occipital electrodes and showed a later onset and longer duration than the N170 effect seen in the face-vase comparison.

No difference was seen between the passive and active task for either of these effects, suggesting that the observed components are not due to expectation of a particular percept or otherwise a correlate of top-down attentional control. If the N170 difference or the BPN were moved earlier in time or were otherwise modulated during the active task, this might identify them as an index of the impact of top-down mechanisms. However, since both components were unaffected by the presence of volitional control, they seem to be associated with perceptual processing of the identity of an object, irrespective of expectation or intention.

4.1.1 Relevance of Percept-Based Neural Differences to Consciousness Research

Both of the percept-based ERP negativities constitute differential activity in response to a single physical stimulus (e.g. a greater N170 when *perceiving* the faces as figure than when they are still a part of the retinal input but are not being comprehended as such because the vase is perceived as figure). These results do not necessarily implicate these mid-latency ERP negativities as neural correlates of conscious perception. The N170 may correspond to the neural activity involved in attending to a face-shaped object, or the process of grouping sensory features together as a face (thereby creating that identity), rather than the conscious comprehension of what that identity is. Additionally, in this study participants consciously perceived *something* as a figure on every trial, and the differences observed in brain activity associated with each percept were relative differences. Thus, our results cannot definitively identify the N170 as a neural correlate of conscious face perception (an NCC for faces).

A more common method of trying to piece apart the brain activity that directly corresponds to consciousness is to compare “seen” stimuli, of which a participant is aware, to the same stimulus when it is shown but “unseen”. Neural components that appear in both cases must be due to low level sensory input or to something else unrelated to consciousness, perhaps indexing some other part of the task. Activity that corresponds to conscious perception must be present only in response to “seen” stimuli, making this a comparison that might be more useful for identifying true neural correlates of consciousness. However, even in seen/unseen comparisons, it is unclear whether activity specific to “seen” conditions is preconscious (but necessary for consciousness, since it is absent in the “unseen” condition) or post-perceptual (occurring as a downstream result of having consciously perceived something) (Aru et al., 2012). For example, the reason why a threshold stimulus is perceived on some trials and not perceived on other trials may be due to stochastic fluctuations in attention or to neural fatigue or satiation (Kitajo, Yamanaka, Ward, & Yamamoto, 2006). Differences in neural activity between seen and unseen conditions might therefore reflect these attentional fluctuations or low-level neural changes rather than mechanisms supporting conscious perception itself. In bistability, we face a similar problem: changes in perceptual experience may be due to fatigue or satiation, or some top-down exploratory mechanism (Long & Toppino, 2004; Leopold & Logothetis, 1999). What we can gain from both methods is an approximation of when and where in neural processing perceptual or attentional effects first manifest, as sensory input is held constant while perception changes (seen vs. unseen, or seen-as-A vs. seen-as-B).

4.1.2 Future Directions in Measuring Percept-Based Neural Differences

One potential area for further research is to investigate the component differences observed in the current study in relation to unambiguous figures. Specifically, it would be interesting to include unambiguous versions of the face-vase preceding trials on which the ambiguous version is presented, to subtly bias perception. When an observer is exposed to an unambiguous version of a bistable figure prior to viewing the ambiguous one, which percept is experienced in the ambiguous figure is contingent on how long the adaptation period lasted. If the unambiguous version was displayed for only a short period of time, a positive-bias towards the presented perceptual interpretation or “set” effect is seen for the ambiguous figure. If instead the unambiguous version is shown for a long enough duration, a reverse-bias effect is observed, such that the opposing percept is the one that is seen when the ambiguous figure appears (see citations within Long & Toppino, 2004).

EEG techniques could be used to examine these effects to further elucidate the impact of the aforementioned top-down and bottom-up mechanisms. Long and Toppino characterize the positive-bias effect as evidence of top-down influences on bistable perception, since it involves interpreting one identity (the single identity that the unambiguous figure gives rise to) and then cognitively maintaining it through presentation of a stimulus that lends itself to but does not necessitate or demand that identity (the unambiguous figure). In contrast, they group the reverse-bias effect with other evidence of bottom-up factors, explaining it as a direct result of neural fatigue or satiation. If the timing or the magnitude of the N170 or the BPN change following priming, this would indicate a direct impact of exogenous information on these specific components. If the neural response is altered (in comparison to that elicited by the type of continuously ambiguous presentation in the present study) following positive-bias priming, this would indicate the influence of top-down information on these specific components, showing that they are not actually so early as to be only sensory-driven. If instead some component differs after reverse-bias priming, this would identify it as a reflection of bottom-up activity. If both types of priming lead to differences, it would suggest that at the time of these components, both processes are already involved.

Another obvious follow-up to the present study is to simply try to replicate the findings using other bistable stimuli. This is the first study we know of to use EEG in an attempt to elucidate differential neural activity between percepts (although at least one attempt was made previously with the Necker cube and no ERP differences were found; personal correspondence with M.A. Pitts). The face-vase image was selected for the present study due to the easily identifiable neural component associated with one of its percepts (the N170) (Bentin, Allison, Puce, Perez, & McCarthy, 1996; Eimer, 2000). Prior researchers seem to have had the same idea, because it was also the stimulus selected for previous fMRI and MEG investigations of the neural activity corresponding to individual percepts (Andrews et al., 2002; Hesselmann et al., 2008; Parkkonen et al., 2008). With something like the Necker cube, the distinctions between percepts may be too subtle to elicit measurable differences in activity beyond the level of single-unit recordings. There are other bistable images that include facial features, although their percept pairs are less distinct than that of the face-vase (one has a full human body as the alternate percept to the face, and another has an animal; see Figure 4.1). The more distinct two percepts are, the more likely they are to activate different processing networks, and therefore the better chance there is of resolving neural differences at the scalp.

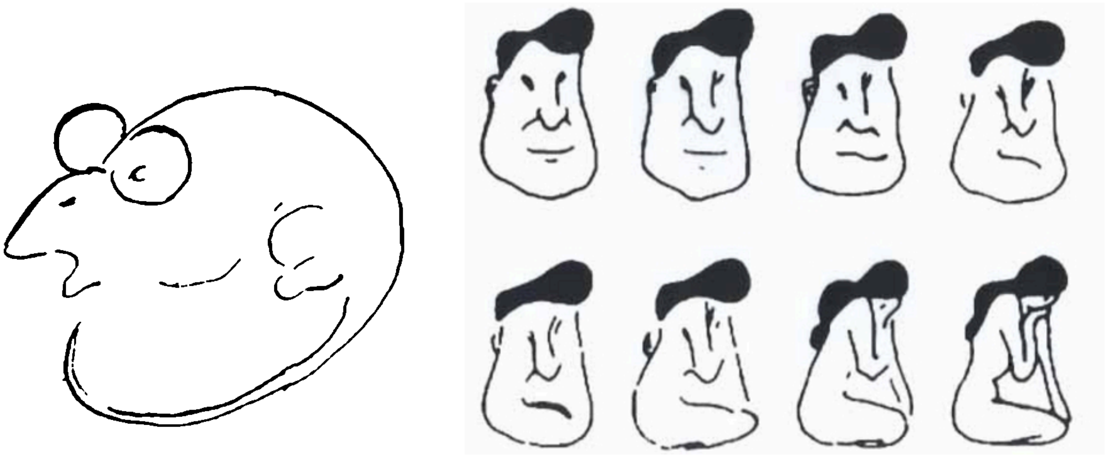


Figure 4.1: Other example bistable stimuli including faces.
The first can be seen either as a man's face, or as a rat, and is taken from Fisher (1967), while the second can be seen either as a man's face or as a woman's body, and is taken from Bugelski and Alampay (1961).

4.2 Differences in Neural Activity Due to Reversals

During passive viewing an occipital reversal negativity (RN) similar to that reported in prior research, in terms of topography and latency, was seen for both stimuli (Pitts et al., 2007, 2008, 2009; Kornmeier & Bach, 2004, 2005, 2006; Kornmeier et al., 2007; Britz et al., 2009; Intaite et al., 2010). A later positivity over central electrodes (LPC), also a replication of the aforementioned studies, was also observed for both stimuli, although it was significantly larger for the face-vase than for the backpack-chalice. Both components have previously been found for tasks requiring participants to report reversals or stables (instead of individual percepts), but as articulated by Pitts et al. (2009), such a task requires participants to hold the previous percept in memory and take it into consideration in making their response. This 2009 study was the first to show that both the RN and LPC are still elicited when the task is to respond based on individual percepts after every single trial, and our data successfully corroborates the evidence that the RN and LPC are not dependent on explicit inter-trial comparisons.

The dramatically higher passive reversal rates in the current experiment than has been seen in previous studies is worth noting. Our participants reported reversals on 43% of trials (on average every 2-3 seconds), while previous studies have resulted in passive reversals on 25-33% of trials (varying with differences in task or which bistable figure was used), or every 3-4 seconds (Pitts et al., 2007, 2008; Kornmeier & Bach 2007). The stimuli presented in the current study were much smaller than those used in previous experiments (Pitts et al. used bistable images subtending a visual angle of 3.3° x 3.3°, while Kornmeier and Bach showed Necker lattices at an angle of 7.5° x 7.5°; the stimuli in the current study were only 2.6° x 2.6°), and smaller image size has been shown to increase the rate of fluctuation (Washburn, Mallay, & Naylor, 1931). Reversal rates are also modulated by the duration of the ISI used in discontinuous stimulus presentation (Kornmeier & Bach, 2007; Orbach et al., 1963). For ISIs between zero ms (static, continuous presentation) and 400ms, reversal rates increase steadily, with a steeper slope for shorter stimulus durations. After 400ms, reversal rates decrease as ISIs grow steadily longer (Kornmeier & Bach, 2007; Leopold et al., 2002; O'Donnell et al., 1988). While the ISIs used in the current study are based on those outlined by Kornmeier and Bach as the optimal for mimicking the behavior of continuous stimulus presentation, it may be the case that for this particular bistable image a slightly altered temporal pattern would be more suitable. Baseline reversal rates should be assessed for static presentation with the face-vase and backpack-chalice images used in the present study, to see if the contrast with prior research is inherent to the stimuli, or if it can be modulated by altering ISI durations.

The rate of reversals during passive viewing is important because it may impact the observed ERP components. In addition to impacting reversal rate, ISI duration changes the latency and amplitude of the RN, with longer ISIs eliciting larger and earlier RNs (Kornmeier & Bach, 2007). Some investigators have divided participants into groups of "fast" and "slow" reversers, (Borsellino 1982; Strüber 2000; Strüber & Stadler 1999) but Kornmeier and Bach (2009) did not see specific patterns during active control conditions that warranted separating participants into two distinct subpopulations. When the data gathered during the active task in the current study are separated based on whether participants had to speed up or slow down their reversal rate (relative to passive), there are no differences between the ERPs from these two groups. However, interpretations of the RN may be dependent on the relationship it has with reversal rates.

One possible explanation of the RN component is that it is contingent on expectations, and that it appears in response to a change in perception because in some sense this is a violation of an expectation. This explanation would only seem parsimonious if, when reversal rates are high, the RN is smaller than when reversals are more rare (which data from the current study do not support – when RN amplitude was examined for individual participants with respect to reversal rates, no correlation was found). In contrast, the RN could be viewed as directed attention necessary for generating a percept, as Pitts et al. (2007) suggest by comparing it to Hillyard and Anllo-Vento's (1996) selective attention-dependent component (SN). The SN appears on trials on which an attended feature appears and is absent for trials showing an unattended feature. These authors speculate that the RN corresponds to feature selection requiring an attentional shift, which is why it differs from stable trials on which spatial and selective attention remain the same. Here there seems little reason for reversal rates to modulate the RN, since a shift in attention is required to process a different percept from the previous trial irrespective of how frequently reversals occur. This is also the case if the RN is viewed as a *response* to a change in perception – not an implementation of the change itself but a sort of flag for the fact that something is different now, even if that difference is not odd or unexpected.

The fact that the N170 precedes the RN may be evidence that the latter is related to post-perceptual processing of the switch, or at least that it is contingent on first interpreting whichever portion is seen as the figure as such. This would mean that the brain first processes that it is seeing a face and then does the work to identify that face as a switch from the previous trial. Alternatively, it may be the case that each of these processes are happening in concert and that identification of the percept neither precedes nor follows identification of the reversal. It could be that incoming sensory information is simultaneously relayed to multiple neural networks, such that the percept itself is interpreted at the same time as its relationship to the previous one. In this way, conscious awareness of the identity of the percept does not have to wait for checks with working memory or higher-order ideas about the meaning of perceptual changes. Each of these activities can be running concurrently, feeding information into each other such that decision making and motor activity can take place as quickly as possible. This latter possibility seems much more likely, given the complexity and efficiency of the brain.

The LPC is late enough that it is often described as a reflection of post-perceptual processing, possibly representing reconstruction of the reversed percept (Intaite et al., 2010; Kornmeier & Bach, 2007). Due to its similarity to the P300, it is interpreted as an index of awareness of (or attention to) the new percept, rather than a correlate of the reversal itself (Kornmeier & Bach, 2004, 2005; Pitts et al., 2008, 2009). But these interpretations are always in the context of the RN, wherein the RN is said to be a sensory component and the LPC, because it occurs later in time, is a secondary process. The current study adds a critical piece to these developing theories by implicating a perceptual difference that precedes both the RN and LPC.

4.2.1 Trying to Discriminate Top-Down From Bottom-Up

As articulated by Kornmeier and Bach (2007), it is unclear from the current neuroscience literature on bistable perception which of the observed neural components correspond to activity initiating a perceptual reversal, which correspond to sensory processing leading up to conscious comprehension of such a reversal, and which indicate a subsequent cognitive response to that reversal. The temporal precision of EEG seems useful for identifying neural activity in support of either the top-down or bottom-up theories, because it allows us to identify the order in which certain processes occur, but as has just been demonstrated, there are still a multitude of ways of interpreting these components when the only comparisons that can be made are of relative timing. This has lead researchers to try to incorporate other mechanisms of manipulating top-down or bottom-up information, by including unambiguous figures, or with active tasks like the one implemented in this study.

In our active task, both reversal components seem to disappear. This is in contrast to prior research, which has shown an increase in negative amplitude of the reversal negativity when participants were asked to control their perception (Pitts et al., 2008). Failure to replicate this finding may be due to differences in the nature of the active task: Pitts et al. asked their participants to "speed up the rate of reversals" which presumably prompted participants to try to switch as frequently as they possibly could. This would manifest as intentional shifts of attention as rapidly as can be accomplished. By contrast, for our task we asked participants to "keep everything balanced in equal numbers" as much as they possibly could. This task is accomplished by keeping track of what has been seen in the past and actively attempting to control what is seen on the current trial in accordance with what has been perceived the least frequently. Rather than an increase in attentional shifts, it is more of a modulation of perception at each point in time and an attempt to match a target percept held in memory to the incoming sensory information. This tightening of control may be culpable for the decrease in all of the observed component differences, because participants were actively trying to see a particular thing on any given trial, and therefore did not display the same neural response to passively occurring percepts or changes in perception.

These data can also be used in support of the idea that the RN is a response to a violation of expectations, in line with the predictive coding hypothesis (Friston, 2005; Friston & Kiebel, 2009). During the passive task, the RN appears as an error signal, because the brain's prediction based on immediately prior experience is that the same percept will be seen again. On reversal trials, this prediction is violated, and the brain generates an error signal to update predictions on future trials. In the active condition, participants were forming predictions on each trial and for the most part these were met, hence no error signal (no RN).

Most of these interpretations of the RN are of a bottom-up nature. However, the fact that both types of active task in some way impact the RN shows that top-down influences at least play a role in bistable perception. The effect suggests that introducing an element of volitional control over perception modulates selective attention or predictive coding. The work demonstrating that patients with frontal lobe lesions have difficulty increasing the rate of reversals when compared to controls (Ricci & Blundo, 1990), in addition to the fact that the ability to control reversals increases with meditation (Carter, Presti, Callistemon, Ungerer, Liu, & Pettigrew, 2005), show that cognitive control matters for bistable perception. But the impact of that control on individual neural components (like the RN) helps to identify the particular brain activity underlying that behavioral effect.

Other studies have attempted to manipulate the bottom-up processes involved in bistable perception by comparing the neural activity elicited by ambiguous images to that observed when viewing comparable, unambiguous versions of the stimuli that are biased towards one percept or the other. Presumably, any activity seen in response to both ambiguous and unambiguous images must be due to sensory driven figure comprehension or perception, or to subsequent conscious processing of a reversal or recognition of the percept, because in the unambiguous case, there is no top-down interpretation to be done. For all components seen for both ambiguous and unambiguous reversals (including the RN and the LPC), there is a delay of ~40ms when viewing the ambiguous images. This has lead researchers to suggest that 40ms is the amount of time necessary for disambiguation to take place (Kornmeier & Bach, 2006), but does not do much in the way of further identifying the role of either component. The intuition that the presence of the RN and the LPC during unambiguous perception shows they have nothing to do with top-down processing may be misguided. Gestalt theories of visual perception state that the visual system is constantly working to interpret ambiguous information, and while most of the time a single optimal interpretation easily wins out, there are still high-level processes involved.

4.2.2 Potentials for Further Research on Reversal-Based Neural Differences

4.2.2.1 Analyses on Current Datasets

There are multiple methods of analysis that have not yet been addressed that could very well be conducted on the current dataset. These analyses will likely be useful for further addressing theoretical queries about the neural underpinnings of bistable perception.

The first of these is to conduct source analyses so as to locate the neural generators of the observed ERP components. Although there are an infinite number of “inverse solutions” to electrical potentials measured at the scalp, there are a number of techniques that allow for estimating the sources of these signals. Previous research has identified a right-lateralized inferior occipital-temporal source for the RN, rendering it unlikely to be a reflection of an attentional shift, although it may be associated with the input of attention-biasing signals from fronto-parietal networks (Pitts, Martinez, Stalmaster, Nerger, & Hillyard, 2009). The same study suggests the inferior temporal and superior parietal regions as the generators of the LPC, although it is noted that these sources (like those of the P300 it is thought to resemble) are more difficult to localize. As these authors suggest, further examination of the sources of both components should be conducted, both for the sake of replication and in order to isolate contributing factors such as the active task and reversal rates.

The second analytical approach that should be taken is to look backwards in time, relative to stimulus onset, so as to examine the activity leading up to perceptual reversals. All the research outlined herein has been with respect to what occurs after a stimulus is presented. Many of the top-down and bottom-up arguments are about what causes a reversal to occur, and to address this question it seems necessary to know what is going on prior to stimulus onset. This is particularly important for analyzing the data from the active task, because it is worth questioning what happens when participants are actively attempting to make a switch occur, and this attempt is almost certainly taking place prior to stimulus onset. Some research has already been done: Britz, Landis, and Michel (2009) examined the momentary brain states preceding presentation of a bistable stimulus and found more activity in the right inferior parietal cortex before reversals than before stables. Similar investigations should be conducted using other datasets.

Yet a third method of analysis would be to sort reversal trials into a different bins based on how many stable trials they are preceded by. Predictions should be stronger with more immediate experience, so if the RN is an error signal, it should be larger when reversals occur after longer bouts of stable trials compared to shorter bouts. Alternatively, the knowledge that the figure is bistable and previous experience of both percepts may cause participants to believe that longer bouts of stable trials render a reversal more and more likely. If the RN is an index of expectation, it may be smaller for reversals after long bouts because these represent switches that have been anticipated for multiple trials. No correlation was found in our data using each participant's average RN amplitude and average reversal rate, but if the data were sorted by individual trials, interesting differences might be observed.

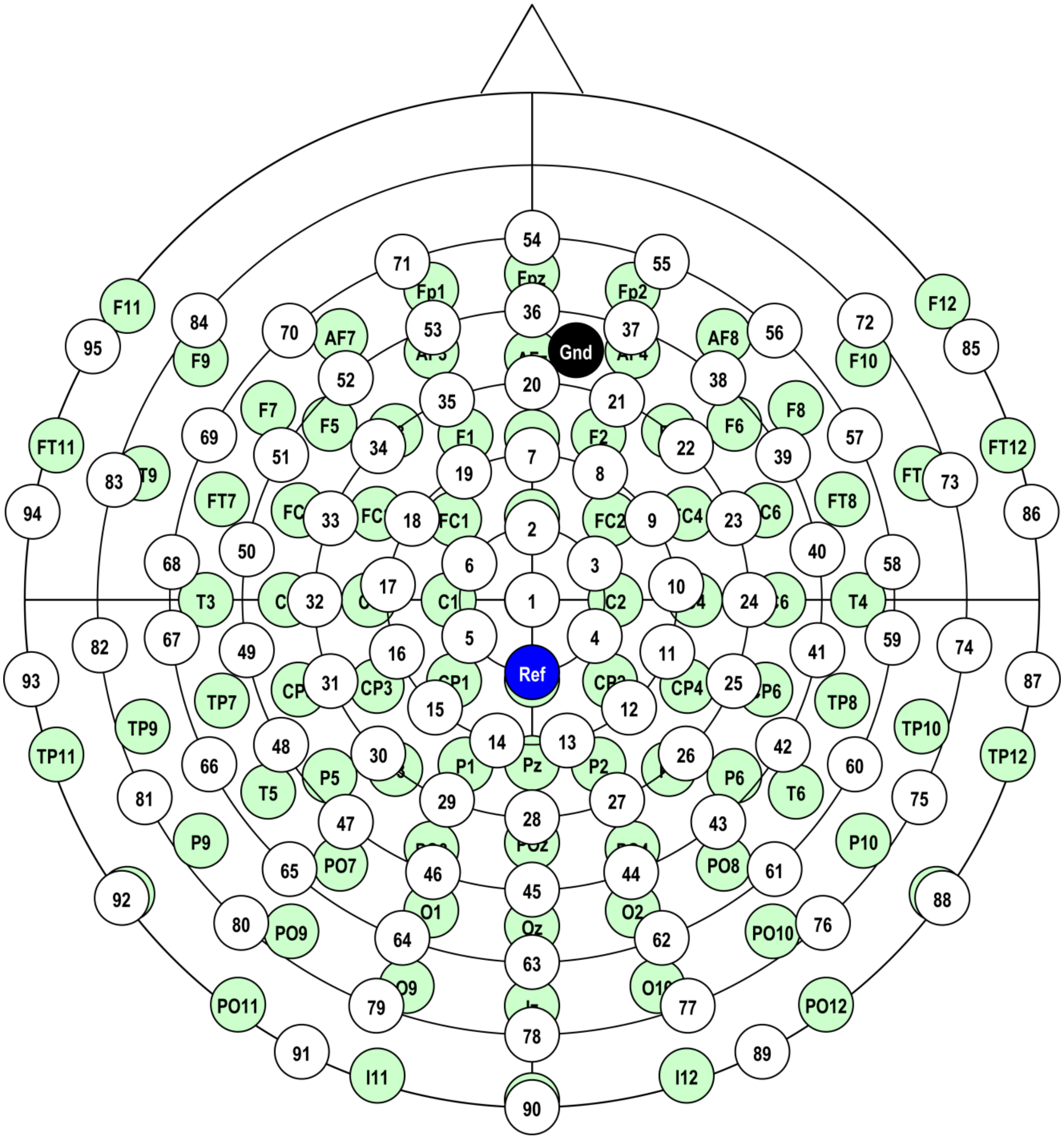
4.2.2.2 Further Studies

When viewing a bistable stimulus, the reversal rate increases systematically over time, which aligns with the theories of periodic neural satiation or fatigue and recovery as responsible for the occurrence of reversals (Long & Toppino, 2004). With each subsequent cycle, there is less and less complete recovery from fatigue, such that reversals happen more and more frequently until some threshold is reached. This pattern is somewhat irrelevant to designing EEG studies, which require so many stimulus repetitions that the reversal acceleration is basically unavoidable, but there are some interesting ways of manipulating this effect. If a bistable figure is rotated, moved to another place in the field of vision, or replaced with a larger or smaller version of itself, reversal rates return to baseline (Kohler, 1940; Toppino & Long, 1987). It would be interesting to examine how these effects manifest in ERPs – if it is indeed possible to lower reversal rates by slightly altering retinal input, even after many repetitions, then it seems relatively straightforward to pick a set of locations and/or sizes, present a bistable figure for a certain number of trials (enough to elicit the acceleration) in one before switching to another, and repeat enough blocks of each so as to acquire enough trials of the accelerated reversal rate and the newly lowered rate. We could then ask whether the RN changes in amplitude or latency as reversal rates increase, or as they revert back to baseline. If it does, it would suggest that the RN really is related to violations of expectations in a way that is modulated as those expectations change.

One piece of evidence in support of top-down influences on bistable perception is that given a secondary task, reversal rates decrease (Reisberg 1983; Reisberg & O'Shaughnessy, 1984). Raising mental load should have no impact on low level sensory processes. That the allocation of some cognitive resources to minimal secondary tasks decreases the rate of reversals implies that a certain amount of attention and active cognitive processing is required for perceptual reversals. How this allocation of attention impacts the ERPs is worth investigating. If, given a secondary task that must be accomplished in addition to reporting percepts, the RN and/or LPC are unchanged, this would suggest a certain degree of independence from top-down attention. If instead they are modulated by an increase in cognitive load, then they may be a reflection of active mechanisms of interpreting perceptual input.

Finally, it may be possible to measure the RN and the LPC in the absence of subjective reports. A recent study using binocular rivalry monitored nystagmus and pupil size as an indicator of perceptual reversals while recording fMRI data, and compared measured brain activity from trials on which participants also had to make an active report of perception to those on which objective measures alone were used (Frässle, Sommer, Jansen, Naber, & Einhäuser, 2014). The objective measures successfully matched active reports when both were collected. The authors found that without the active reporting task included, activation in frontal areas was absent. Occipital and parietal activity, however, remained, suggesting that the RN would likely still be observed in participants asked to view a bistable figure while eye movements and pupil size are used to infer when a switch occurs. This type of measure eliminates the need to rely on subjective self-report *and* eliminates the involvement of a separate motor task.

Appendix A: Complete List of Electrode Locations



Name	Theta	Phi	Name	Theta	Phi	Name	Theta	Phi
1	0	0	33	-60	-22	65	-100	48
2	20	90	34	-60	-45	66	-100	27
3	20	30	35	-60	-67	67	-100	6
4	20	-30	36	80	90	68	-100	-6
5	-20	30	37	80	70	69	-100	-27
6	-20	-30	38	80	50	70	-100	-48
7	40	90	39	80	30	71	-100	-69
8	40	62	40	80	10	72	120	40
9	40	34	41	80	-10	73	120	16
10	40	6	42	80	-30	74	120	-6
11	40	-21	43	80	-50	75	120	-27
12	40	-48	44	80	-70	76	120	-48
13	40	-76	45	80	-90	77	120	-69
14	-40	76	46	-80	70	78	120	-90
15	-40	48	47	-80	50	79	-120	69
16	-40	21	48	-80	30	80	-120	48
17	-40	-6	49	-80	10	81	-120	27
18	-40	-34	50	-80	-10	82	-120	6
19	-40	-62	51	-80	-30	83	-120	-16
20	60	90	52	-80	-50	84	-120	-40
21	60	67	53	-80	-70	85	140	30
22	60	45	54	100	90	86	140	10
23	60	22	55	100	69	87	140	-9
24	60	0	56	100	48	88	140	-36
25	60	-22	57	100	27	89	140	-63
26	60	-45	58	100	6	90	140	-90
27	60	-67	59	100	-6	91	-140	63
28	60	-90	60	100	-27	92	-140	36
29	-60	67	61	100	-48	93	-140	9
30	-60	45	62	100	-69	94	-140	-10
31	-60	22	63	100	-90	95	-140	-30
32	-60	0	64	-100	69	ECG	-180	55
Gnd	70	80	Ref	20	-90			

Bibliography

Andrews, T. J., Schluppeck, D., Homfray, D., Matthews, P., & Blakemore, C. (2002). Activity in the Fusiform Gyrus Predicts Conscious Perception of Rubin's Vase–Face Illusion. *NeuroImage*, *17*(2), 890–901.

Anllo-Vento, L., & Hillyard, S. A. (1996). Selective attention to the color and direction of moving stimuli: Electrophysiological correlates of hierarchical feature selection. *Perception & Psychophysics, 58*, 191–206.

Aru J., Bachmann T., Singer W., & Melloni L. (2012) Distilling the neural correlates of consciousness. *Neuroscience & Biobehavioral Reviews, 36(2)*, 737–746.

Attneave, F. (1971). Multistability in perception. *Scientific American*, 225(6), 63-71.

Başar-Eroglu, C., Strüber, D., Stadler, M., Kruse, P., & Başar, E. (1993). Multistable visual perception induces a slow positive EEG wave. *International Journal of Neuroscience*, *73*(1-2), 139–151.

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience, 8*(6), 551–565.

Borsellino, A., Carlini, F., Piani, M., Taccio, M. T., Marco, A. D., Penengo, P., & Trabucco, A. (1982). Effects of visual angle on perspective reversal for ambiguous patterns. *Perception*, *11*(3), 263 – 273.

Britz, J., Landis, T., & Michel, C. M. (2009). Right Parietal Brain Activity Precedes Perceptual Alternation of Bistable Stimuli. *Cerebral Cortex*, *19*(1), 55–65.

Bugelski, B. R., & Alampay, D. A., (1961). The role of frequency in developing perceptual sets. *Canadian Journal of Psychology*, 15, 205-211.

Carter, O. L., Presti, D. E., Callistemon, C., Ungerer, Y., Liu, G. B., & Pettigrew, J. D. (2005). Meditation alters perceptual rivalry in Tibetan Buddhist monks. *Current Biology, 15*(11), R412–R413.

Clark, V., & Hillyard, S. (1996). Spatial selective attention affects early extrastriate but not striate components of the visual evoked potential. *Journal of Cognitive Neuroscience, 8*, 387–402.

Dalrymple, K. A., Oruç, I., Duchaine, B., Pancaroglu, R., Fox, C. J., Iaria, G., … Barton, J. J. S. (2011). The anatomic basis of the right face-selective N170 IN acquired prosopagnosia: A combined ERP/fMRI study. *Neuropsychologia, 49(*9), 2553–2563.

Dodd, J. V., Krug, K., Cumming, B. G., & Parker, A. J. (2001). Perceptually bistable three-dimensional figures evoke high choice probabilities in cortical area MT. *The Journal of Neuroscience*, *21*(13), 4809–4821.

Driver, Jon, and Gordon C. Baylis. (1996). Edge-Assignment and Figure–ground Segmentation in Short-Term Visual Matching. *Cognitive Psychology, 31*(3), 248–306.

Eimer, M. (2000). The face-specific N170 component reflects late stages in the structural encoding of faces. *Neuroreport, 11*(10), 2319-2324.

Fisher, G. (1967). Measuring Ambiguity. *American Journal of Psychology, 80*(4), 541-557.

Frässle, S., Sommer, J., Jansen, A., Naber, M., & Einhäuser, W. (2014). Binocular Rivalry: Frontal Activity Relates to Introspection and Action But Not to Perception. *Journal of Neuroscience*, *34*(5), 1738–1747.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1456), 815–836.

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *364*(1521), 1211–1221.

Hesselmann, G., Kell, C. A., Eger, E., & Kleinschmidt, A. (2008). Spontaneous local variations in ongoing neural activity bias perceptual decisions. *Proceedings of the National Academy of Sciences, 105*(31), 10984–10989.

Hochberg, J. & McAlister, E. (1953). A Quantitative Approach to Figural "Goodness." *Journal of Experimental Psychology, 46*(5), 361-364.

Horlitz, K. L., & O'Leary, A. (1993). Satiation or availability? Effects of attention, memory, and imagery on the perception of ambiguous figures. *Perception & Psychophysics*, *53*(6), 668–681.

Intaitė, M., Koivisto, M., Rukšėnas, O., & Revonsuo, A. (2010). Reversal negativity and bistable stimuli: Attention, awareness, or something else? *Brain and Cognition*, *74*(1), 24–34.

İşoğlu-Alkaç, Ü., Başar-Eroğlu, C., Ademoğlu, A., Demiralp, T., Miener, M., & Stadler, M. (1998). Analysis of the electroencephalographic activity during the Necker cube reversals by means of the wavelet transform. *Biological Cybernetics, 79*(5), 437–442.

Katsuki, F., & Constantinidis, C. (2013). Bottom-Up and Top-Down Attention Different Processes and Overlapping Neural Systems. *The Neuroscientist*.

Kitajo, K., Yamanaka, K., Ward, L. M., & Yamamoto, Y. (2006). Stochastic resonance in attention control. *EPL (Europhysics Letters)*, *76*(6), 1029.

Kohler, W. (1940). *Dynamics in psychology*. New York: Liverlight.

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. (2005). The Necker cube—an ambiguous figure disambiguated in early visual processing. *Vision Research*, *45*(8), 955–960.

Kornmeier, J., & Bach, M. (2006). Bistable perception — along the processing chain from ambiguous visual input to a stable percept. *International Journal of Psychophysiology*, *62*(2), 345–349.

Kornmeier, J., Ehm, W., Bigalke, H., & Bach, M. (2007). Discontinuous presentation of ambiguous figures: How interstimulus-interval durations affect reversal dynamics and ERPs. *Psychophysiology*, *44*(4), 552–560.

Kornmeier, J., Hein, C. M., & Bach, M. (2009). Multistable perception: When bottom-up and top-down coincide. *Brain and Cognition*, *69*(1), 138–147.

Leeper, R. (1935). A Study of a Neglected Portion of the Field of Learning—the Development of Sensory Organization. *The Pedagogical Seminary and Journal of Genetic Psychology*, *46*(1), 41–75.

Leopold, D. A., & Logothetis, N. K. (1999). Multistable phenomena: Changing views in perception. *Trends in Cognitive Sciences, 3*(7), 254–264.

Leopold, D. A., Wilke, M., Maier, A., & Logothetis, N. K. (2002). Stable perception of visually ambiguous patterns. *Nature Neuroscience*, *5*(6), 605–609.

Liebert, R., & Burk, B. (1985). Voluntary control of reversible figures. *Perceptual and motor skills, 61*, 1307–1310.

Logothetis, N. K., & Schall, J. D. (1989). Neuronal correlates of subjective visual perception. *Science, 245*(4919), 761–763.

Long, G. M., & Toppino, T. C. (2004). Enduring interest in perceptual ambiguity: Alternating views of reversible figures. *Psychological Bulletin, 130*(5), 748–768.

Luck, S. J., Hillyard, S. A., Mouloua, M., Woldorff, M. G., Clark, V. P., & Hawkins, H. L. (1994). Effects of spatial cuing on luminance detectability: Psycho- physical and electrophysiological evidence for early selection. Journal of Experimental Psychology: Human Perception and Performance, 20, 887–904.

Meng, M., & Tong, F. (2004). Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *Journal of Vision, 4*, 539–551.

Nguyen, V. T., & Cunnington, R. (2014). The superior temporal sulcus and the N170 during face processing: Single trial analysis of concurrent EEG–fMRI. *NeuroImage, 86,* 492–502.

O'Donnell, B. f., Hendler, T., & Squires, N. k. (1988). Visual Evoked Potentials to Illusory Reversals of the Necker Cube. *Psychophysiology*, *25*(2), 137–143.

Orbach, J., Ehrlich, D., & Heath, H. A. (1963). Reversibility of the Necker cube: I. An examination of the concept of "satiation of orientation." *Perceptual and Motor Skills, 17,* 439–458.

Parkkonen, L., Andersson, J., Hämäläinen, M., & Hari, R. (2008). Early visual brain areas reflect the percept of an ambiguous scene. *Proceedings of the National Academy of Sciences, 105*(51), 20500–20504.

Pitts, M. A., Nerger, J. L., & Davis, T. J. R. (2007). Electrophysiological correlates of perceptual reversals for three different types of multistable images. *Journal of Vision, 7*(1), 6–6.

Pitts, M. A., Gavin, W. J., & Nerger, J. L. (2008). Early top-down influences on bistable perception revealed by event-related potentials. *Brain and Cognition, 67*(1), 11–24.

Pitts, M. A., Martínez, A., Stalmaster, C., Nerger, J. L., & Hillyard, S. A. (2009). Neural generators of ERPs linked with Necker cube reversals. *Psychophysiology*, *46*(4), 694–702.

Ricci, C., & Blundo, C. (1990). Perception of ambiguous figures after focal brain lesions. *Neuropsychologia, 28*(11), 1163–1173.

Reisberg, D. (1983). General mental resources and perceptual judgments. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 966–979.

Reisberg, D., & O’Shaughnessy, M. (1984). Diverting subjects’ concentration slows figural reversal. *Perception, 13,* 461–468.

Rock, I., Hall, S., & Davis, J. (1994). Why do ambiguous figures reverse? *Acta Psychologica*, *87*(1), 33–57.

Sheinberg, D. L., & Logothetis, N. K. (1997). The role of temporal cortical areas in perceptual organization. *Proceedings of the National Academy of Sciences, USA, 94,* 3408–3413.

Slotnick, S., & Yantis, S. (2005). Common neural substrates for the control and effects of visual attention and perceptual bistability. *Cognitive Brain Research, 24,* 97–108.

Strüber, D., Basar-Eroglu, C., Hoff, E., & Stadler, M. (2000). Reversal-rate dependent differences in the EEG gamma-band during multistable visual perception. I*nternational Journal of Psychophysiology, 38*(3), 243–252.

Strüber, D., Basar-Eroglu, C., Miener, M., & Stadler, M. (2001). EEG gamma-band response during the perception of Necker cube reversals. *Visual Cognition*, *8*(3-5), 609–621.

Strüber, D., & Stadler, M. (1999). Differences in top-down influences on the reversal rate of different categories of reversible figures. *Perception, 28*, 1185–1196.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of Processing in the Human Visual System. *Nature, 381*, 520-522.

Thorpe, S. (2009). The Speed of Categorization in the Human Visual System. *Neuron, 62*, 168-170.

Toppino, T. C., & Long, G. M. (1987). Selective adaptation with reversible figures: Don't change that channel. *Perception & Psychophysics, 42,* 37–48.

Toppino T (2003). Reversible-figure perception: Mechanisms of intentional control. *Perception & Psychophysics, 65*, 1285–1295.

van Ee, R., van Dam, L., & Brouwer, G. (2005). Voluntary control and the dynamics of perceptual bi-stability. *Vision Research, 45*, 41–55.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. *Psychological Bulletin*, *138*(6), 1172–1217.

Washburn, M. F., Mallay, H., & Naylor, A. (1931). The Influence of the Size of an Outline Cube on the Fluctuations of Its Perspective. *The American Journal of Psychology*, *43*(3), 484–489.

Williams, S. M., McCoy, A. N., & Purves, D. (1998). An empirical explanation of brightness. *Proceedings of the National Academy of Sciences*, *95*(22), 13301–13306.

Windmann, S., Wehrmann, M., Calabrese, P., & Gunturkun, O. (2006). Role of the prefrontal cortex in attentional control over bistable vision. *Journal of Cognitive Neuroscience, 18*(3), 456–471.

Yovel, G., Sadeh, B., Podlipsky, I., Hendler, T., & Zhdanov, A. (2008). The face-selective ERP component (N170) is correlated with the face-selective areas in the fusiform gyrus (FFA) and the superior temporal sulcus (fSTS), but not the occipital face area (OFA): a simultaneous fMRI-EEG study. *Journal of Vision, 8(*6):401.