

**From the holdings of the Department of Special Collections , Princeton University Library,
Princeton, NJ**

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyright law. We do not charge any permission or usage fees for the publication of images of material in our collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff. For more information regarding copyright, credit, and citation please review the information on the Copyright, Credit and Citation Guidelines page of our website.

<https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines>. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. <https://library.princeton.edu/special-collections/ask-us>

Doctoral Dissertations, 2011-2026

Psychology

2018

Brain Mechanisms of Subjective Visual Awareness and Visual Attention

Webb, Taylor

Princeton, NJ : Princeton University

<http://arks.princeton.edu/ark:/88435/dsp017d278w72h>

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. <https://www.copyright.gov/title17/title17.pdf>

BRAIN MECHANISMS OF SUBJECTIVE VISUAL AWARENESS AND VISUAL ATTENTION

TAYLOR WHITTINGTON WEBB

A DISSERTATION PRESENTED TO THE FACULTY OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE BY THE DEPARTMENT OF PSYCHOLOGY

ADVISOR: MICHAEL S. A. GRAZIANO

JUNE 2018

© Copyright by Taylor Whittington Webb, 2018. All rights reserved.

Abstract

Our psychological processes are sometimes accompanied by a sense of awareness, a property which is still poorly understood. The experiments presented in this dissertation contribute to our emerging understanding of the behavioral and brain basis of visual awareness, with a special focus on its relationship with visual attention. First, in a series of a behavioral experiments, the functional relationship between attention and awareness is considered. The results show that attention is less well controlled in the absence of awareness, contradicting the notion that awareness is an ‘epiphenomenon’ - a property that accompanies mental processes but does not interact with them. Then, a series of brain imaging and brain stimulation studies investigate the role of the temporoparietal junction (TPJ) in awareness. These studies show that temporary disruption of the TPJ with transcranial magnetic stimulation (TMS) interferes with visual awareness, and that the involvement of the TPJ in visual awareness cannot be accounted for as an artifact of visual attention. These studies also investigate how the involvement of the TPJ in awareness relates to its involvement in a range of other cognitive processes, and how these processes are supported by a network of other brain regions throughout the parietal and frontal cortex.

Contents

Abstract	iii
Contents	iv
Chapter 1. Introduction	1
Chapter 2. Behavioral experiments investigating the relationship between visual attention and visual awareness	9
Chapter 3. Parcellation of TPJ involvement in multiple cognitive domains	38
Chapter 4. Transcranial magnetic stimulation to the TPJ produces neglect-like effects	56
Chapter 5. Cortical networks involved in visual awareness independent of visual attention	67
Chapter 6. Conclusion	93
References	96

Chapter 1

Introduction

When we look out onto the world, our mental processes - perception, attention, memory, decision making, and so on - are sometimes accompanied by a peculiar property, often referred to as 'subjective awareness'. Our brains do not merely carry these processes out 'in the dark'. In some instances, we also become aware of these processes themselves. How this happens remains poorly understood. In this dissertation, a range of behavioral, brain imaging, and brain stimulation experiments are reported which make a significant contribution to our understanding of two fundamental questions regarding subjective awareness.

The first question concerns whether awareness has a specific functional role, or is merely an 'epiphenomenon' - accompanying mental processes without actually contributing to them. This may seem like a puzzling question, given the subjective centrality of awareness in our mental lives. A surprising conclusion of much of the contemporary research on awareness, however, is that much of our mental processes can indeed be carried out without awareness (Lau & Rosenthal, 2008). In Chapter 2, behavioral experiments are reported which suggest that at least one functional role of awareness may be to help with the control of attention (Webb et al., 2016a).

The second question concerns the specific brain regions that underlie subjective awareness. A nearly three decade long search for the 'neural correlates of consciousness' (Crick & Koch, 1990) has resulted in a wide array of proposals, ranging from primary

sensory cortices (Tong, 2003), to the prefrontal cortex (Dehaene & Naccache, 2001), to subcortical regions (Crick & Koch, 2005). Though subjective awareness is most likely to be the product of a network of brain regions, multiple lines of evidence, including both neuroimaging experiments (Lumer et al., 1998; Dehaene et al., 2001; Rees et al., 2002; Webb et al. 2016b), and studies on the clinical condition of hemispatial neglect (Vallar & Perani, 1986; Corbetta & Shulman, 2011), suggest that the temporoparietal junction (TPJ) may play an especially important role.

One concern, however, is the ubiquitous involvement of the TPJ in a range of other tasks, including attention (Corbetta & Shulman, 2000; Corbetta & Shulman, 2002) and social cognition (Saxe & Kanwisher, 2003; Saxe & Powell, 2006). In Chapter 3, multiple functional MRI (fMRI) experiments are reported which map out the involvement of distinct subcomponents of the TPJ in these and other cognitive domains (Igelström et al., 2016). Building on this, in Chapter 4, it is shown that transcranial magnetic stimulation (TMS) to a specific, dorsal subcomponent of the TPJ (TPJd), disrupts visual awareness (Kelly et al., 2014). Finally, in Chapter 5, using fMRI, it is shown that this same subcomponent, in addition to an anterior component (TPJa), is involved in subjective awareness independently of its involvement in attention (Webb et al., 2016b). Together, these experiments suggest that the TPJ is an important node in the cortical network that supports subjective awareness.

1.1 Defining Awareness

An important first step is to clarify what is, and isn't, meant by the term 'awareness'. A distinction is commonly made between 'objective awareness' and 'subjective awareness'. As an example, if we wanted to determine whether a research participant was

‘objectively aware’ of an oriented line stimulus, we might present them with many trials of this stimulus and ask them to discriminate its orientation. If their performance at this task was better than what would be expected by chance alone, this would be an objective indication that the stimulus had, to some extent, been processed by the participant’s brain. It wouldn’t matter whether the participant vehemently denied seeing any stimulus whatsoever. The fact that the stimulus reliably influenced their behavior in this manner would, by definition, mean that they were ‘objectively aware’ of it.

‘Subjective awareness’, on the other hand, is defined precisely as whether or not a participant, in his or her own opinion, has perceived a stimulus. Importantly, ‘objective’ and ‘subjective’ awareness can dissociate. Strong demonstrations of dissociation come from clinical conditions such as blindsight and hemispatial neglect. Blindsight typically emerges as a result of damage to early cortical visual areas, after which patients report a subjective blind spot in their visual field, despite some degree of preserved (‘objective’) visual discrimination capacities within that region of visual space (Weiskrantz, 1986). In hemispatial neglect, a similar phenomenon has been described (Marshall and Halligan, 1988), usually on the left half of visual space following damage to frontal and parietal regions in the right hemisphere (Corbetta & Shulman, 2011).

Beyond the striking dissociations observed in the clinical literature, subtler dissociations can be evoked in the laboratory with ordinary participants. Using various visual masking techniques (Breitmeyer & Öğmen, 2006), stimulus conditions can be constructed under which participants report no visual awareness for a stimulus, despite some degree of preserved ability to make discriminations about that stimulus. It has even been shown that ‘objective awareness’ for a stimulus can be balanced while manipulating ‘subjective awareness’ (Lau & Passingham, 2006). Thus, there is ample

reason to suspect that ‘objective’ and ‘subjective’ measures of awareness do not measure the same underlying brain processes, and should be treated differently. This dissertation is concerned primarily with subjective awareness. For simplicity, this property will often be referred to simply as ‘awareness’.

1.2 Is Awareness an Epiphenomenon?

The notion of epiphenomenalism was articulated clearly by Thomas Huxley (1874), when he compared consciousness to the “steam whistle which accompanies the work of a locomotive engine ... without influence upon its machinery”. The idea that consciousness emerges from our mental processes, but is not a necessary or functional part of those processes, remains popular today (Lau & Rosenthal, 2008). This view has been bolstered by a growing list of mental processes that appear to be possible without consciousness - attention (Kentridge et al., 2008), cognitive control (Lau & Passingham, 2007), learning (Seitz et al., 2009) and many more.

One problem for the epiphenomenalist view is the question of why, if awareness plays no functional role, the brain would expend precious metabolic resources on the computations that underlie it. Thus, a number of hypotheses have been advanced concerning what specific functional role awareness might play. These include proposals that awareness is important for integrating information (Tononi, 2008), for broadcasting information between disparate brain regions (Baars, 1997), or for resolving conflicts between competing brain processes (Morsella, 2005).

One proposal (Graziano & Kastner, 2011; Graziano, 2013; Graziano & Webb, 2014; Webb & Graziano, 2015), referred to as the ‘attention schema theory’, suggests that awareness serves as a schematic model of one’s own mental processes. In this way,

the attention schema theory is related to other proposals which view awareness as a form of metacognition, such as 'higher order thought theory' (Rosenthal, 2004), but differs from those proposals by suggesting a specific functional role for awareness in helping to control attention.

Despite earlier views which equated attention and awareness (Posner, 1994; De Brigard & Prinz, 2009), it is now well understood that attention is possible in the absence of awareness (McCormick, 1997; Lamme, 2004; Tsushima et al., 2006; Jiang et al., 2006; Koch & Tsuchiya, 2007; Kentridge et al., 2008; Hsieh et al., 2011; Norman et al., 2013; Lin & Murray, 2015; Webb et al., 2016a). It is possible to visually process a stimulus, at the expense of other visual stimuli, despite being unaware of that stimulus. Nevertheless, the existence of interactions between attention and awareness (Tsushima et al., 2006; McCormick, 1997; Lin & Murray, 2015) suggest an important link between these processes. The attention schema theory proposes that awareness is to the control of attention as the body schema (Graziano & Botvinick, 2002) is to motor control - a simplified, schematic model of the system which allows it to be controlled more effectively (Graziano & Webb, 2014; Webb & Graziano, 2015).

This idea is tested in a series of behavioral experiments that are presented in Chapter 2 (Webb et al., 2016a). In those experiments, it is shown that, in the absence of awareness for a visual stimulus, attention is more susceptible to bottom-up influences, and less well-controlled in line with task demands. These results suggest that awareness plays a role in the control of attention (perhaps in addition to other functions), and is therefore unlikely to be an epiphenomenon.

1.2 Brain Structures Involved in Awareness

Francis Crick and Christof Koch, in an influential paper (1990), advocated the search for the 'neural correlates of consciousness'. A number of proposals have since been made regarding the essential brain structures that underlie awareness. These proposals have been quite diverse. Awareness has been attributed to primary sensory cortices (Tong, 2003), parietal cortex and prefrontal cortex (Lumer et al., 1998; Dehaene et al., 2001; Rees et al., 2002; Carmel et al., 2006; Lau & Passingham 2006), temporal cortex (Karnath et al., 2001), and even subcortical areas (Crick & Koch, 2005). Others have argued that awareness is not a function of a particular brain region, but depends instead on features such as feedback vs. feedforward projections (Lamme, 2001), or the extent of information integration within a neural circuit (Tononi, 2008).

Though no consensus has yet emerged, a number of studies suggest an important role for a network of frontal and parietal regions, especially when controlling for lower level perceptual processes (Lumer et al., 1998; Dehaene et al., 2001; Rees et al., 2002; Carmel et al., 2006; Lau & Passingham 2006). This view is also supported by evidence from clinical impairments of awareness. Hemispatial neglect, in which patients suffer a devastating lack of awareness for one half of space, most commonly emerges following damage to the parietal and frontal cortices (Brain, 1941; Critchley, 1953; Heilman & Valenstein, 1972; Vallar & Perani, 1986; Mort et al., 2003; Corbetta & Shulman, 2011). A less obvious source of evidence comes from the case of blindsight. Though blindsight usually results from damage to the visual cortex, it has been shown that this causes impaired function in parietal and frontal cortices which may account for the observed deficit in awareness (Persaud et al., 2011).

An area of special interest is the temporoparietal junction (TPJ), both because it consistently emerges in neuroimaging studies of awareness (Lumer et al., 1998; Dehaene et al., 2001; Rees et al., 2002), but also because it is the most common lesion site for hemispatial neglect (Vallar & Perani, 1986; Corbetta & Shulman, 2011). The TPJ, however, is also known to participate in a number of other cognitive domains. It is thought to play a prominent role in processes as diverse as attention (Corbetta & Shulman, 2000; Corbetta & Shulman, 2002), social cognition (Saxe & Kanwisher, 2003; Saxe & Powell, 2006), episodic memory (Konishi et al., 2000), and target detection (Stevens et al., 2000). What, therefore, is the relationship between these processes and the apparent involvement of the TPJ in awareness?

Before considering this question, in Chapter 3, a series of experiments are reported (Igelström et al., 2016) which clarify the specific subcomponents of the TPJ that are involved in attention, social cognition, episodic memory, and target detection. These experiments reveal that different subcomponents of the TPJ play a domain-specific role in different processes, but that a dorsal region of the TPJ (TPJd) plays a domain-general role in all of these processes.

In Chapter 4, it is shown that TMS to this region (TPJd) causes a neglect-like impairment in visual awareness (Kelly et al., 2014). Finally, bringing this line of work together with the behavioral methods introduced in Chapter 2, Chapter 5 demonstrates that both TPJd and an anterior region of the TPJ (TPJa) are involved in subjective awareness independently of any involvement in attention (Webb et al., 2016b). Importantly, however, this work also shows that TPJd and TPJa, though especially important for awareness, participate as nodes in a network of interconnected regions including the temporal, insular, and prefrontal cortices. This work refines our

understanding of the specific brain regions that support awareness, and suggests that TPJd may serve as an important domain-general hub for awareness and other processes.

Chapter 2

Behavioral experiments investigating the relationship between visual attention and visual awareness

It is now well established that attention and awareness are separable. People can attend to a stimulus in the absence of awareness of that stimulus (McCormick, 1997; Lamme, 2004; Tsushima et al., 2006; Jiang et al., 2006; Koch & Tsuchiya, 2007; Kentridge et al., 2008; Hsieh et al., 2011; Norman et al., 2013; Lin & Murray, 2015; Webb et al., 2016a). Yet attention and awareness are not entirely independent, given that interactions between them have been reported (Tsushima et al., 2006; McCormick, 1997; Lin & Murray, 2015; Webb et al., 2016a). The specific relationship between attention and awareness, however, has remained uncertain.

In this chapter, results from a series of behavioral experiments are reported which clarify this relationship. In Behavioral Experiment 1 (Webb et al., 2016a), the time course of attention to a task-relevant stimulus is compared in the presence versus absence of awareness for that stimulus. The results show that attention is less stable over time when participants are unaware of the stimulus. In Behavioral Experiments 2 (Webb et al., 2016a) and 3, the inhibition of a task-irrelevant stimulus is compared in the presence versus absence of awareness for that stimulus. The results show that attention to the task-irrelevant stimulus is less effectively inhibited when participants are unaware of that stimulus. In behavioral experiment 4 (Webb et al., 2016a), the effect of bottom-up stimulus strength on attention is compared in the presence versus absence

of awareness for the stimulus. The results show that attention is driven by bottom-up stimulus strength more when participants are unaware of the stimulus.

Taken together, these results suggest that awareness plays a special role in the top-down control of attention, whereas attention in the absence of awareness is more likely the result of bottom-up influences. These results are discussed in the context of the attention schema theory (Graziano & Kastner, 2011; Graziano, 2013; Graziano & Webb, 2014; Webb & Graziano, 2015). Furthermore, by precisely quantifying the psychophysical relationship between attention and awareness, these results lay the foundation for experiments investigating the brain networks that separately underlie these processes, discussed in subsequent chapters.

2.1 Behavioral Experiment 1

To date, most research on the relationship between attention and awareness has focused on establishing whether attention is possible without awareness. By contrast, relatively few studies have focused on directly comparing attention with and without awareness. In this first behavioral experiment, the time course of attention to a visual stimulus is compared in the presence versus absence of visual awareness for that stimulus.

A classic method for measuring the time course of attention to a specific stimulus was introduced by Posner (1980). In this paradigm, a ‘cue’ stimulus is presented, biasing attention to either the left or the right side of visual space. Participants then make a response to a subsequent ‘target’ stimulus, presented on either the left or the right. To the extent that the cue biases attention, responses to the subsequent target stimulus should be faster and more accurate when the target appears on the same vs. opposite

side as the cue. By varying the time between presentation of the cue and target, the time course of the cue's effect on attention can be measured.

In the present experiment, this classic paradigm is extended by incorporating visual masking, to control awareness of the cue stimulus. This allows the direct comparison of the time course of attention to a visual stimulus at different levels of awareness for that stimulus. The results show that attention is initially drawn to the stimulus regardless of whether participants are aware of it, but that attention is less stable over time in the absence of awareness.

2.1.1 Subjects

Seventy-five participants were tested (44 females, 18–48 years old, normal or corrected-to-normal vision). All participants provided informed consent, and all procedures were approved by the Princeton Institutional Review Board.

2.1.2 Behavioral Paradigm

Participants sat 30 cm from the monitor and used a chin rest to stabilize the head. Stimuli were presented with the MATLAB Psychophysics Toolbox (Brainard, 1997). Figure 2.1 shows the behavioral paradigm. Each trial began with a white central fixation point on a black background. Participants were instructed to fixate during the trial. After 1 second, the cue period began and lasted 4 refresh cycles (~50 ms). Throughout the cue period, the cue (a white spot 1.1° in diameter) was presented 6° to the left of fixation (1/3 of trials), to the right of fixation (1/3 of trials), or not presented (1/3 of trials). The cue was followed by a mask. In the long cue/mask interval condition (1/2 of trials), a cue/mask interval of 4 refresh cycles (~50 ms) was inserted between the cue and the mask. Thus, the time from cue onset to mask onset was ~100 ms, intended to allow the

cue to be subjectively visible to the participants. In the short cue/mask interval condition (1/2 of trials), the mask immediately followed the cue. Thus, in this trial type, the time from cue onset to mask onset was ~50 ms, intended to render the cue subjectively invisible to the participants. During the mask period, on all trials, two metacontrast masks (Breitmeyer & Öğmen, 2006) were presented, 6° to the left and right of fixation. The masks themselves therefore did not preferentially attract bottom-up attention to one side. Each mask consisted of a white ring with an outer diameter of 1.5° and an inner diameter of 1.1°. The duration of the mask period varied between 115 and 540 ms, depending on the total duration of the trial as described next.

After the mask period, a target was added to the already-present white circles that composed the mask. The target was either on the left (1/2 of trials) or right (1/2 of trials) of fixation. The location of the cue did not predict the location of the target. The target was presented at one of five times after cue onset: 12 refresh cycles (~165 ms), 20 refresh cycles (~270 ms), 28 refresh cycles (~380 ms), 36 refresh cycles (~485 ms), and 44 refresh cycles (~590 ms). Because of these varied target presentation times, the mask period between the cue and the target also varied accordingly. The target consisted of one white line segment extending from the top of the mask ring and one white line segment extending from the bottom of the mask ring. The segments were collinear, forming an implied line through the ring. The line was tilted toward the left (1/2 of trials) or toward the right (1/2 of trials) by 3°. Participants were required to discriminate the orientation by pressing the F key on a keyboard if the line was tilted to the left and the J key if it was tilted toward the right. The target period lasted 1 second, and participants were instructed to respond as quickly as possible during that period. The short target response period encouraged a speeded response. Participants performed

the discrimination task with a mean accuracy of 90% (SD = 8%) and a mean latency of 673 ms (SD = 127 ms). The 10% of trials with incorrect responses or no responses during the 1-second response window were excluded from analysis. Trials with a latency < 300 ms were also excluded because the RT was too short to be a plausible response to the discrimination task, and therefore, those trials probably represented mistaken key presses.

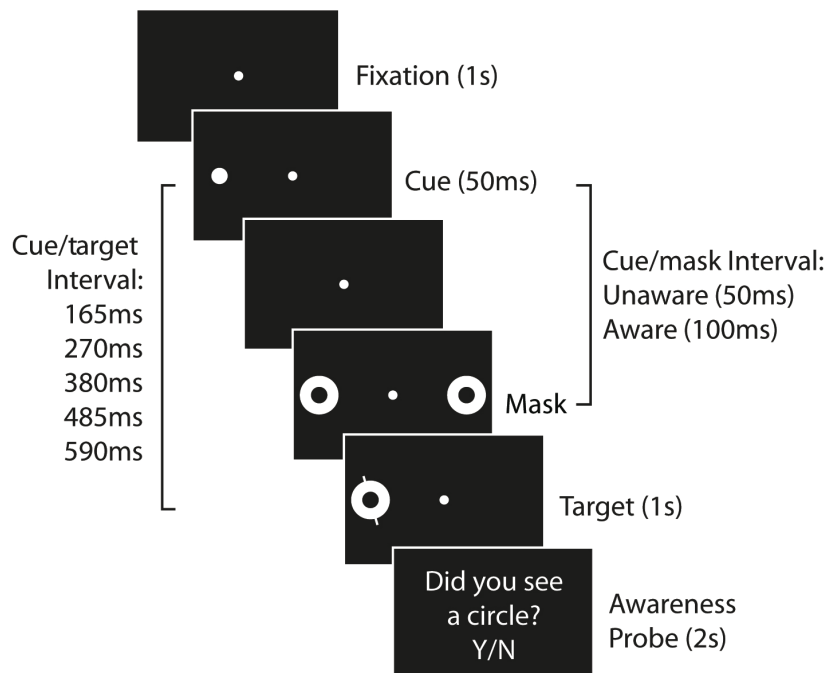


Figure 2.1: Behavioral paradigm for measuring time course of attention to a visual stimulus (the ‘cue’ stimulus) with and without awareness of that stimulus.

On each trial, after the 1-second target period, the fixation point and all other stimuli disappeared, and a question was presented on the screen: “Did you see a circle? Y/N.” Participants were instructed that this question referred to the cue stimulus presented at the beginning of some trials. Participants pressed the F key to indicate that they had seen the cue on that trial and the J key to indicate that they had not. The awareness probe remained on the screen for 2 seconds, and participants were instructed

to respond during that time. On <1% of trials, participants failed to respond during that time window, and those trials were excluded from analysis. The mean awareness RT was 580 ms (SD = 322 ms). Choosing a cutoff for the minimum RT is more difficult for the awareness response. Note that the cue was presented early in the trial, and participants had as much as 1.5 seconds (depending on trial type) before the awareness probe to decide if they had seen the cue. Moreover, the awareness probe was presented at a predictable time 1 second after the target onset. Therefore, participants could, in principle, anticipate the timing of the question and respond with extremely short latency. This may account for the large spread in latency to the awareness probe. On the basis of the distribution of RTs, we chose to exclude trials with a cue awareness RT of < 50 ms. After accounting for overlap with trials on which participants either didn't respond to the target or responded incorrectly (which were already excluded from analysis), this resulted in the exclusion of an additional 7% of trials; however, the same pattern of statistical results was produced without excluding these trials (as shown in *Results*).

Each participant performed practice trials followed by 9 runs of 24 trials each (216 trials total). Target orientation (tilted left or right), target location (left or right), cue condition (left, right, or no cue), and cue/mask interval (long or short) were randomized and counter-balanced within each run, and the cue did not predict the location of the target. The design was a mixed, within/ between-subject design. The aware versus not aware manipulation (long or short cue/mask interval) was within participants. However, different participants were tested for each of the five time intervals between the cue and the target. The reason was that to test each participant at all 5 time points would have required an estimated 7.5 hr of testing per participant,

which was prohibitive on practical grounds. Thus, we tested each participant with only one cue/target time interval. For each of the five cue/target time intervals, 15 participants were tested, for a total of 75 participants.

2.1.3 Analysis

To determine whether the division of trials into “aware” and “unaware” was justified, the participants’ responses to the awareness probe were analyzed. Responses to the awareness question from the 1/3 of trials on which no cue was presented were used to assess participants’ false positive rate, the rate at which participants indicated that a cue was present when none was. The rate was low (mean = 11%, SD = 14%), indicating that participants were not guessing about the presence of the cue. Responses to the awareness question from the 2/3 of trials on which a cue was presented were used to compute cue detection rates. To assess whether the metacontrast manipulation had the intended effect on awareness, these cue detection rates were analyzed in a 5×2 mixed factorial ANOVA: 5 time points (between-subject variable) $\times$ 2 awareness conditions (within-subject variable). A highly significant main effect of cue/mask interval confirmed that participants tended to report awareness of the cue on long cue/mask interval trials (awareness reported on 80% of trials) and tended to report no awareness of the cue on short cue/mask interval trials (awareness reported on 30% of the trials; $F_{1,70} = 371, p = 2 \times 10^{-16}$). No other significant effects were found for the cue detection rate, showing that performance was stable across the 5 time points. Thus, not only did the manipulation separate most trials into “aware” and “unaware”, but also that separation was stable across all five cue/target intervals. This confirmation allowed for the main analysis described next.

In a Posner task with high-accuracy performance, accuracy is typically an insensitive measure of attention, whereas latency is the preferred measure. The logic of the Posner task is that if the cue draws attention to itself and the subsequent target appears at the same location as the cue, then the response to the target will be faster. Likewise, if the cue draws attention to itself and the subsequent target appears at the opposite location to the cue, then the response to the target will be slower as attention is shifted from the cue to the target. To quantify this effect of the cue on attention, we used a difference of response latencies: $\Delta t = [\text{average target RT on cue/target mismatch trials}] - [\text{average target RT on cue/target match trials}]$. A larger Δt indicates that the cue drew more attention to itself. For each participant, two Δt scores were computed, one for long cue/mask interval trials (aware trials) and one for short cue/mask interval trials (unaware trials). The data were then analyzed in a 5×2 mixed factorial ANOVA: 5 time points (between-subject variable) $\times$ 2 awareness conditions (within-subject variable) to determine whether the time course of attention was the same or different depending on the presence or absence of awareness.

2.1.4 Results

Behavioral Experiment 1 tested the time course of attention to a brief visual cue over the 600 ms period following cue presentation. The time course was tested with and without awareness of the cue. Figure 2.2A shows the results. The red line shows the results for the aware condition, and the blue line shows the results for the unaware condition. Although attention was present in the unaware condition, it did not behave in the same way. Moreover, the change in attention was more complex than a simple increase of attention in the aware condition. The change therefore cannot be attributed to the stimulus simply having greater perceptual signal strength in the aware condition.

Indeed, at 1 time point (270 ms), attention was actually greater without awareness than with it. The most striking change was that attention was less stable over time in the unaware condition than in the aware condition.

The data were analyzed with a 5×2 mixed factorial ANOVA. The main effect of awareness was not significant ($F_{1,70} = 1.11, p = .30$), showing that awareness did not cause an overall increase or decrease in attention. The main effect of time was significant ($F_{4,70} = 4.11, p = .005$), showing that attention changed as a function of time. The interaction was also significant ($F_{4,70} = 3.38, p = .02$), confirming that the two curves had significantly different shapes. A 5×1 ANOVA on aware trials (Figure 2.2A, red line) revealed no significant effect of time ($F_{4,70} = 0.67, p = .61$), indicating that attention was relatively stable during the trial in the presence of awareness. In contrast, a 5×1 ANOVA on unaware trials (Figure 2.2A, blue line) revealed significant variation over time ($F_{4,70} = 9.1, p = 5 \times 10^{-6}$). The primary difference between the aware and unaware conditions in this task, therefore, was that, in the presence of awareness, attention to the cue was more stable throughout the tested time period.

Figure 2.2B shows the same result as in Figure 2.2A, but with a selective subset of the data. Here, the “confirmed aware” trials meet two constraints: They are the 80% of trials with a long cue/mask interval on which the participant also explicitly indicated that the cue had been seen. Likewise, the “confirmed unaware” trials are the 70% of trials with a short cue/mask interval on which the participant also indicated that the cue had not been seen. This more selective data set shows a nearly identical pattern to the one in Figure 2.2A. The same statistical effects were found (5×2 ANOVA, no significant main effect of awareness, $F_{1,70} = 0.82, p = .37$; significant main effect of time, $F_{4,70} = 5.8, p = .0004$; significant interaction, $F_{4,70} = 2.82, p = .03$; 5×1 ANOVA, no significant

effect of time in aware trials, $F_{4,70} = 1.14, p = .35$; significant effect of time in unaware trials, $F_{4,70} = 10.1, p = 1 \times 10^{-6}$).

We also investigated whether the observed pattern of results was affected by our decision to exclude trials with a cue awareness RT < 50ms. This was not the case. The same pattern of statistical results was produced after including these trials (5×2 ANOVA, no significant main effect of awareness, $F_{1,70} = 1.11, p = .30$; significant main effect of time, $F_{4,70} = 4.11, p = .005$; significant interaction, $F_{4,70} = 3.34, p = .015$; 5×1 ANOVA, no significant effect of time in aware trials, $F_{4,70} = 0.67, p = .61$; significant effect of time in unaware trials, $F_{4,70} = 9.12, p = 6 \times 10^{-6}$).

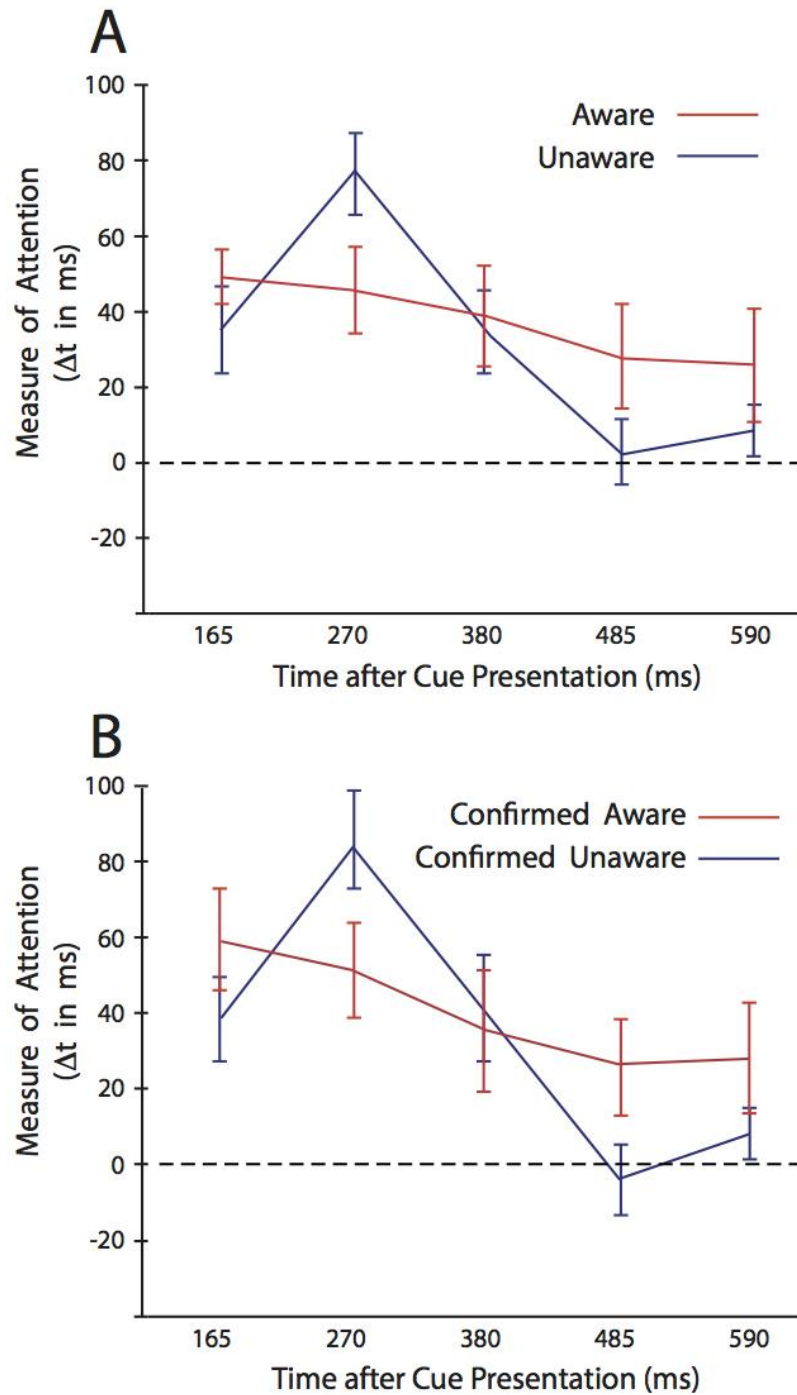


Figure 2.2: Results of Behavioral Experiment 1, showing time course of attention with and without awareness. A. All data from ‘aware’ (long cue/mask interval) and ‘unaware’ (short cue/mask interval) conditions. B. A more selective subset of the data from trials on which participants confirmed they were either aware or unaware of the cue.

2.1.5 Discussion

The primary result of Behavioral Experiment 1 was that attention to a visual stimulus was more stable and sustained over time when participants were aware of the stimulus. One might ask why attention should be sustained in this case. In the Posner paradigm, sustained attention is ordinarily observed when the cue stimulus is task-relevant, usually because the location of the cue predicts the location of the target. In the present paradigm, however, the cue location did not predict the location of the target.

In this case, sustained attention to the cue is best explained by the fact that participants were required to report on the presence of the cue on every trial, making the cue task-relevant. This feature of the task required participants to sustain some degree of attention to the cue stimulus, and to maintain its representation in working memory. The results suggest that the process of sustaining attention to a stimulus in a stable manner, at least over the tested timescale of 600ms, may in some way depend on awareness of that stimulus. One interesting direction for future experiments may be to investigate whether this relationship also holds for longer timescales, providing a possible link to the literature on vigilance (Haider et al., 1964).

2.2 Behavioral Experiment 2

Behavioral Experiment 1 investigated how awareness impacts the time course of attention to a task-irrelevant stimulus. In Behavioral Experiments 2 and 3, the effect of awareness on attention to a task-irrelevant stimulus is explored.

In Behavioral Experiment 1, the cue was task-relevant because participants had to report on its presence at the end of the trial. Thus, to modify the paradigm and make the cue entirely task-irrelevant requires eliminating this report. This, of course, poses a

challenge, because it eliminates any direct measure of awareness. The results of Behavioral Experiment 1, however, show that, on most trials, the mask succeeded in rendering the cue seen on long cue/mask interval trials and unseen on short cue/ mask interval trials. Moreover, using the participants' own reports on subjective awareness to further specify the trials did not change the pattern of results (compare Figure 2.2A and 2.2B). These results suggest that the end-of-trial awareness report can be eliminated without fundamentally altering the paradigm. In Behavioral Experiment 2, the cue stimulus is rendered task-irrelevant by eliminating this report.

2.2.1 Subjects

In Behavioral Experiment 2, we tested 75 new participants not tested in Behavioral Experiment 1 (56 females, 18–42 years old, normal or corrected-to-normal vision). All participants provided informed consent, and all procedures were approved by the Princeton Institutional Review Board.

2.2.2 Behavioral Paradigm

The methods were identical to those of Behavioral Experiment 1, with one exception. After the target period, participants were not asked whether they had seen the cue. The cue was therefore rendered task-irrelevant.

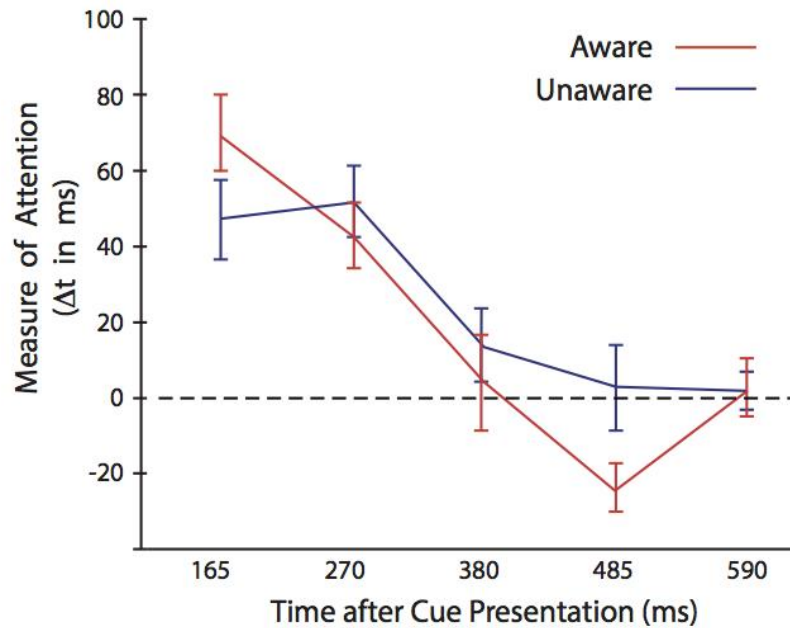


Figure 2.3: Results of Behavioral Experiment 2, showing time course of attention to a task-irrelevant stimulus with and without awareness.

2.2.3 Results

Figure 2.3 shows the results of Behavioral Experiment 2. In the aware condition (red line), the cue drew attention initially. At 165 ms, the measure of attention was significantly above zero. However, unlike in Behavioral Experiment 1 when the stimulus was behaviorally relevant and attention was relatively sustained, here with a task-irrelevant stimulus attention was attenuated over time. By 485 ms, attention showed evidence of active suppression, often referred to as ‘inhibition of return’. This pattern of early facilitation, followed by active suppression, is a commonly observed pattern for task-irrelevant stimuli in the Posner paradigm (Posner, 1980). In contrast, in the unaware condition (blue line), there was no evidence of active suppression. Attention was initially drawn to the cue and then fell back to zero, just as in the unaware condition of Behavioral Experiment 1, but did not dip below zero.

A 5×2 mixed factorial ANOVA was performed. The main effect of awareness was not significant ($F_{1,70} = 0.99, p = .32$), showing that awareness did not cause an overall increase or decrease in attention. The main effect of time was significant ($F_{4,70} = 12.71, p = 8 \times 10^{-8}$), showing that attention changed as a function of time. The interaction was also significant ($F_{4,70} = 3.2, p = .02$), confirming that the two curves had significantly different shapes. The negative attention effect in the aware condition at 485 ms was significantly below zero (one-sample t-test, $t_{14} = 2.9, p = .01$) and significantly below the unaware condition (paired t-test, $t_{14} = 2.36, p = .03$).

2.3 Comparison of Behavioral Experiments 1 and 2

Figure 2.4 shows the data from Behavioral Experiments 1 and 2 replotted for comparison. Figure 2.4A shows all results from aware trials. The thick line shows trials when the cue was task-relevant (participants were asked about their awareness of the cue after each trial), and the thin line shows trials when the cue was task-irrelevant (participants were not asked about the cue). These two curves differ markedly. Two characteristics stand out. First, when the cue was task-relevant, attention was relatively sustained over the tested time period. Second, when the cue was task-irrelevant, the attention initially drawn to the cue was then actively suppressed. With awareness of the cue, attention to the cue was controlled in a task-relevant manner, whether the task required sustained attention or suppressed attention.

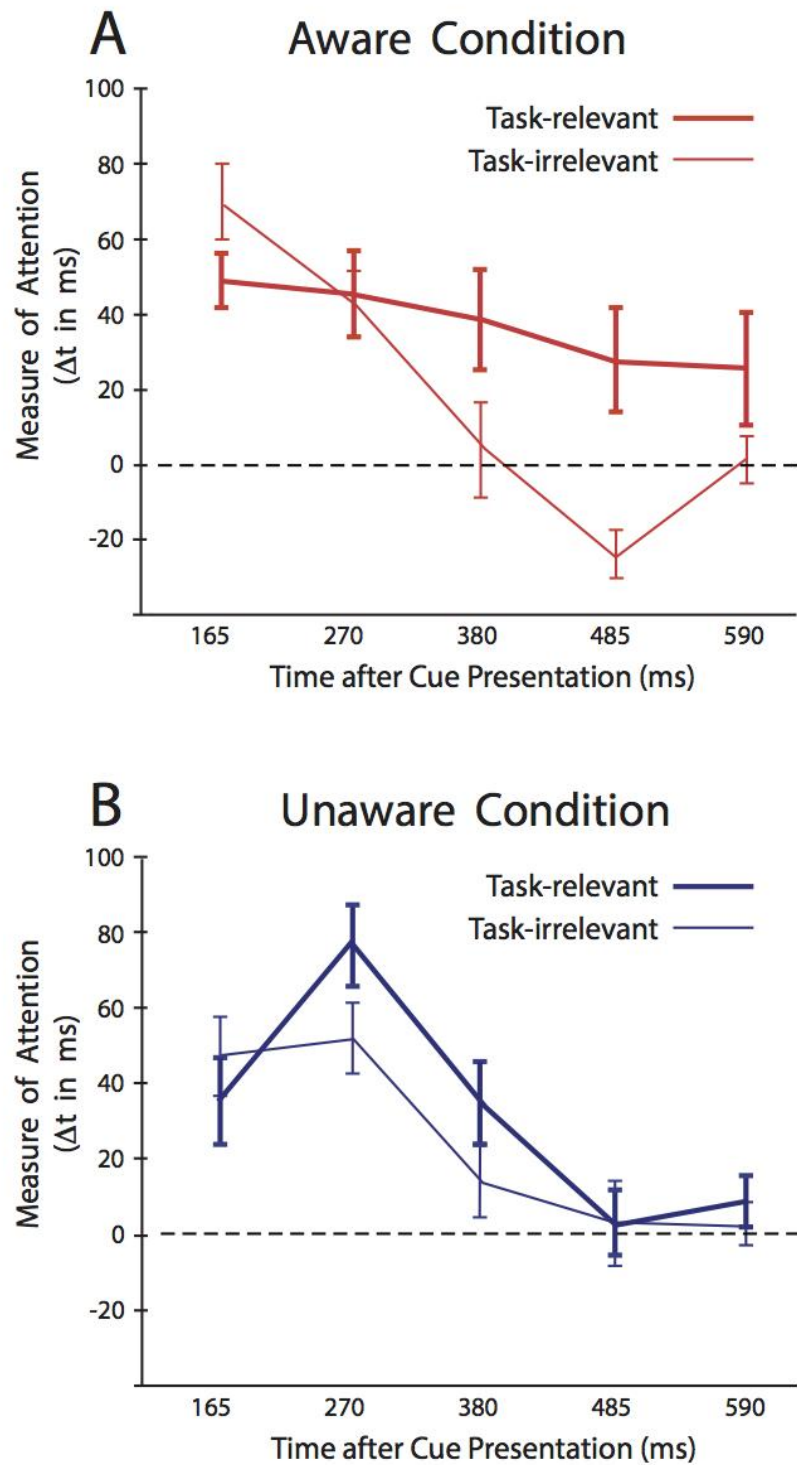


Figure 2.4: Comparison of results from Behavioral Experiments 1 and 2. A. Aware condition in Experiment 1 ('Task-relevant') vs. Experiment 2 ('Task-irrelevant'). B. Unaware condition in Experiment 1 ('Task-relevant') vs. Experiment 2 ('Task-irrelevant').

Figure 2.4B shows the results from unaware trials. The thick line shows trials when the cue was task-relevant, and the thin line shows trials when the cue was task-irrelevant. The two curves are not significantly different at any time point. Most importantly, attention was not sustained when the cue was task-relevant, nor was attention actively suppressed when the cue was task-irrelevant.

An omnibus analysis ($5 \times 2 \times 2$ mixed factorial ANOVA: 5 time points $\times$ 2 awareness conditions $\times$ 2 task-relevance conditions) revealed no significant main effect of awareness ($F_{1,140} = 0.03, p = .87$), a significant main effect of time ($F_{4,140} = 13.93, p = 1 \times 10^{-9}$), and a significant three-way interaction ($F_{4,140} = 3.27, p = .01$).

2.4 Behavioral Experiment 3

The results of Behavioral Experiment 2 suggested that active suppression of a task-irrelevant stimulus may depend on awareness of that stimulus. This conclusion is consistent with the results of another recently published experiment (Lin & Murray, 2015). In this experiment, active suppression of attention to a task-irrelevant stimulus was found on a different timescale, over the course of many trials as opposed to within a single trial. As in Behavioral Experiment 2, this effect was found only when participants were aware of the task-irrelevant stimulus.

For Behavioral Experiment 3, we carried out a replication of the experiment reported by Lin & Murray (2015), and carried out an additional analysis of the results which was important in confirming the interpretation in terms of active suppression of attention.

2.4.1 Subjects

In Behavioral Experiment 3, we tested 55 participants (34 females, 18–51 years old, normal or corrected-to-normal vision). All participants provided informed consent, and all procedures were approved by the Princeton Institutional Review Board.

2.4.2 Behavioral Paradigm

The behavioral paradigm was an exact replication of the paradigm reported by Lin & Murray (2015). Participants sat 50 cm from the monitor and used a chin rest to stabilize the head. Stimuli were presented with the MATLAB Psychophysics Toolbox (Brainard, 1997). Figure 2.5 shows the behavioral paradigm. Each trial began with a black central fixation point on a gray background, presented for 1 second. Participants were instructed to fixate during the trial. This was followed by a blank, gray screen for 200 ms. This was followed by a cue period lasting 2 refresh cycles (~ 33 ms). During this period, on some trials (as described later in this section) a cue stimulus was presented. The cue consisted of either a red or black ring, 3° in diameter, presented either 6° to the left or 6° to the right of fixation. Following the cue period, a target period was presented for 5 refresh cycles (~ 84 ms). During this period, three black rings were presented, one in the center, one 6° to the left of fixation, and one 6° to the right of fixation, each 3° in diameter. On 80% of trials, a target stimulus also appeared, a black spot with a diameter of 1.9° . On 40% of trials, the target appeared inside the ring on the left, and on 40% of trials, it appeared inside the ring on the right. In 20% of trials, no target appeared. The location and presence of the target was randomized across trials. Participants were instructed to press the spacebar as fast as possible when they detected a target, but not when no target was present. If participants incorrectly pressed the spacebar on a trial

with no target, they received feedback in the form of two tones and a 5 second pause. Finally, following the target period, the three black rings remained on the screen for 12 refresh cycles (~200 ms). Each trial concluded with a blank, gray screen for 1 second, after which the next trial began.

The presence and location of the cue stimulus depended on the block, in the following manner. The experiment began with a ‘fixed-cue’ block, consisting of 12 runs of 40 trials each. During this block, the cue appeared in a fixed location on every trial, on the right for half of participants, and on the left for the other half of participants, randomly assigned. This block was followed by a ‘no-cue’ block, consisting of 3 runs of 40 trials each. During this block, no cue was presented.

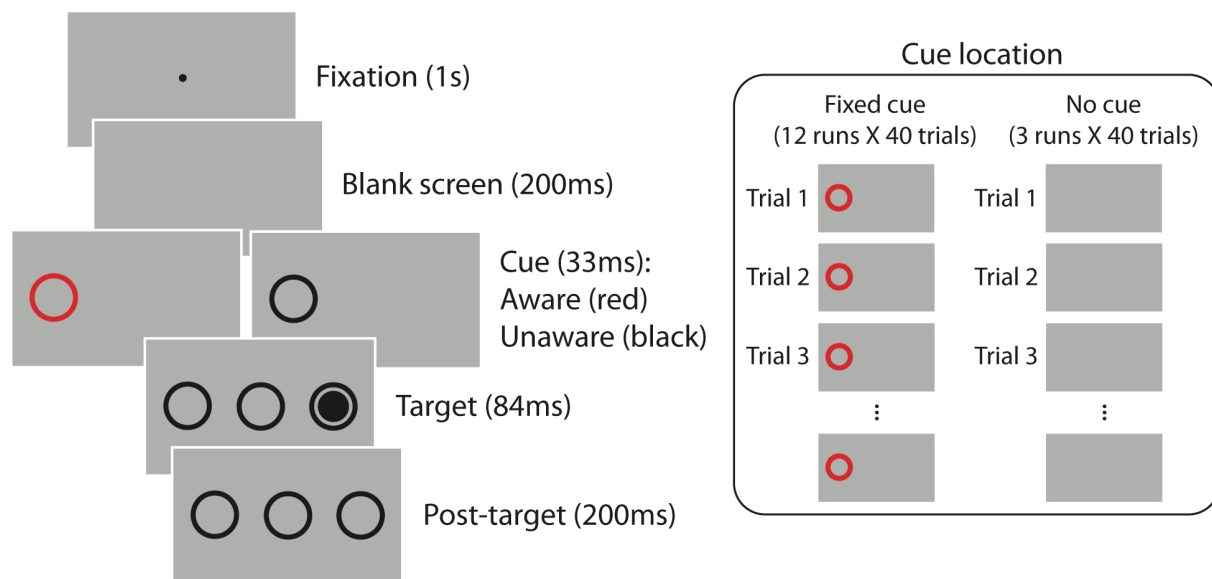


Figure 2.5: Behavioral paradigm for measuring effects of awareness on attention to a task-irrelevant stimulus, replicated from Lin & Murray (2015).

The color of the cue stimulus depended on assignment to one of two conditions. 28 participants were assigned to the ‘aware’ condition, in which the cue stimulus was a bright, red ring. 27 participants were assigned to the ‘unaware’ condition, in which the

cue stimulus was a black ring. Because this was followed by the array of three black rings, this condition was equivalent to one of these three rings appearing ~33 ms earlier than the other two, a subtle temporal difference that participants were intended not to be aware of.

Importantly, the cue did not predict the location of the target, and participants were not asked about the cue on a trial-by-trial basis. This was critical to ensuring that the cue was truly task-irrelevant. However, participants were asked a series of questions at the end of the experiment to determine whether they had noticed the cue, as described in the next section.

2.4.3 Analysis

Following the experiment, participants were first asked, ‘Did you notice anything about the experiment that was not described in the instructions?’. In the ‘aware’ condition, 26 out of 28 participants reported seeing the red ring. The two participants who did not report seeing the red ring were excluded from analysis. None of the participants in the ‘unaware’ condition reported anything about the cue stimulus. These participants were asked a series of increasingly specific follow-up questions, culminating in an animation depicting the black cue stimulus at a slower speed than it had been presented during the experiment. Even after this thorough debriefing, participants in the unaware condition still did not report having seen anything like what was described, suggesting that they were truly unaware of the cue.

We analyzed the effect of the cue on attention at the block level, as in Lin & Murray (2015). For the fixed-cue block, the effect of the cue on attention was computed as a difference of response latencies: $\Delta t = [\text{average target RT on cue/target mismatch trials}] - [\text{average target RT on cue/target match trials}]$. Because the location of the cue

was fixed across this entire block, the difference between these two trial types consisted only of a different target location. For the no-cue block, a similar measure was computed. The primary question, however, was whether the cue presented *during the previous block* would have an effect on attention during the no-cue block. Thus, cue/target mismatch trials consisted of trials in which the target appeared in a location opposite the cue location in the fixed-cue block (despite the fact that no cue actually appeared in the no-cue block), and the cue/target match trials consisted of trials in which the target appeared in the same location as the cue in the fixed-cue block.

We also supplemented this block-level analysis with a trial-by-trial analysis of the cue's effect on attention, in the following manner. First, the data from all participants was pooled within each of 6 conditions (aware vs. unaware condition; cue/target match vs. cue/target mismatch trials; fixed-cue vs. no-cue blocks). Then, a time course of target RT was constructed using a Gaussian kernel with a standard deviation of 15 trials. The time course of the cue's effect on attention was computed as a difference of response latencies: [RT time course for cue/target mismatch trials] - [RT time course for cue/target match trials]. 95% bootstrap confidence intervals were constructed for the resulting time courses, and the time courses for the fixed-cue and no-blocks were concatenating for plotting (though they were computed separately).

2.4.4 Results

The results are shown in Figure 2.6. At the block level (Figure 2.6A), the results replicate those reported by Lin & Murray (2015). First, in the aware condition, when averaging over the entire fixed-cue block, the cue did not have a significant effect on attention (one-sample t-test, $t_{25} = 1.00$, $p = 0.328$). The trial-by trial analysis (Figure 2.6B) further revealed that attention was actually directed toward the cue for the beginning of this

block, but fell back to zero over the course of approximately 40 trials. During the no-cue block, there was a significant aftereffect such that attention was directed opposite to the location where the cue had appeared in the previous block (one-sample t-test, $t_{25} = -6.84$, $p = 2 \times 10^{-7}$). In the unaware condition, the cue had a significant effect on attention during the fixed-cue block (one-sample t-test, $t_{26} = 3.91$, $p = 0.0006$), but no significant after-effect was found during the no-cue block (one-sample t-test, $t_{26} = -0.95$, $p = 0.350$). The trial-by-trial analysis revealed a small after-effect in the unaware condition, at the beginning of the no-cue block, but this after-effect was smaller in both duration and magnitude than the after-effect in the aware condition.

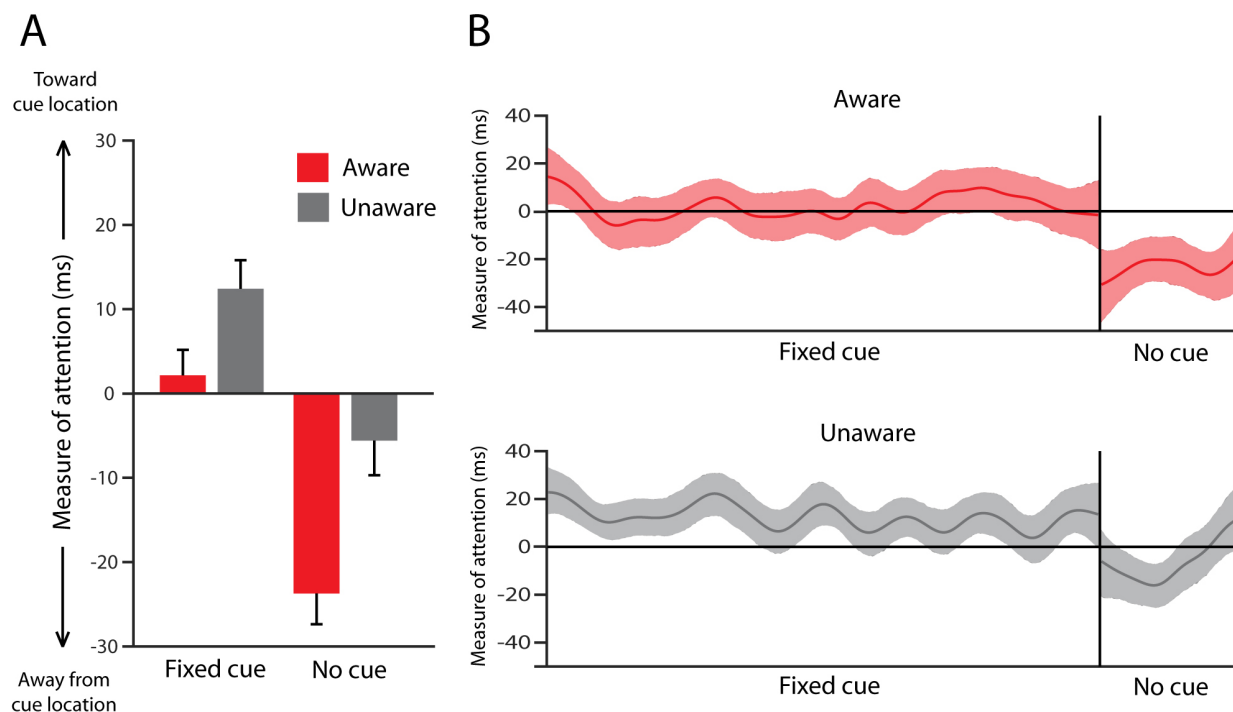


Figure 2.6: A. Results at the block level for replication of Lin & Murray (2015). B. Trial-by-trial time courses for the aware (top) and unaware (bottom) conditions.

2.4.5 Discussion

One possibility is that, in the aware condition, the cue biased attention initially, as confirmed by the trial-by-trial analysis, but was actively suppressed after some number

of trials because of its task-irrelevance. This active suppression would explain the negative effect on attention that resulted when the cue was removed during the no-cue block. This resembles the classic ‘inhibition of return’ effect (Posner, 1980), but on a timescale of many trials, instead of milliseconds within a single trial. As in Behavioral Experiment 2, the results suggest that this process may depend on awareness.

2.5 Behavioral Experiment 4

In Behavioral Experiments 1-3, we tested how awareness interacts with the task-relevance of a stimulus. The purpose of Behavioral Experiment 4 was to test how awareness interacts with the bottom-up signal strength of the stimulus.

2.5.1 Subjects

We tested 26 participants (19 females, 18–50 years old, normal or corrected-to-normal vision). All participants provided informed consent, and all procedures were approved by the Princeton Institutional Review Board.

2.5.2 Behavioral Paradigm

Figure 2.7 shows the behavioral paradigm. At the start of each trial, a central fixation cross and two black boxes with their centers displaced 5° to the left and right of fixation appeared on a gray screen. Participants were instructed to fixate centrally during the trial. After 1500 ms, a 50-ms cue period began. The cue was a single pixel presented at one of two luminance levels. These two luminance levels were tailored to each participant based on performance in practice runs. For both luminance levels, the cue was near perceptual threshold such that on some trials participants reported seeing it and on some trials reported not seeing it. This trial-by-trial variation provided the “aware” and “unaware” trials without having to rely on a mask. With respect to the gray

background (RGB of 190/190/190), the higher contrast cue was chosen from the luminance range of RGB between 50/50/50 and 110/110/110, and the lower contrast cue was chosen from the luminance range of RGB between 145/145/145 and 170/170/170. The cue was presented at low signal strength in the center of the left box on 1/5 of trials, at low signal strength in the center of the right box on 1/5 of trials, at high signal strength in the center of the left box on 1/5 of trials, and at high signal strength in the center of the right box on 1/5 of trials. On 1/5 of trials, no cue was presented.

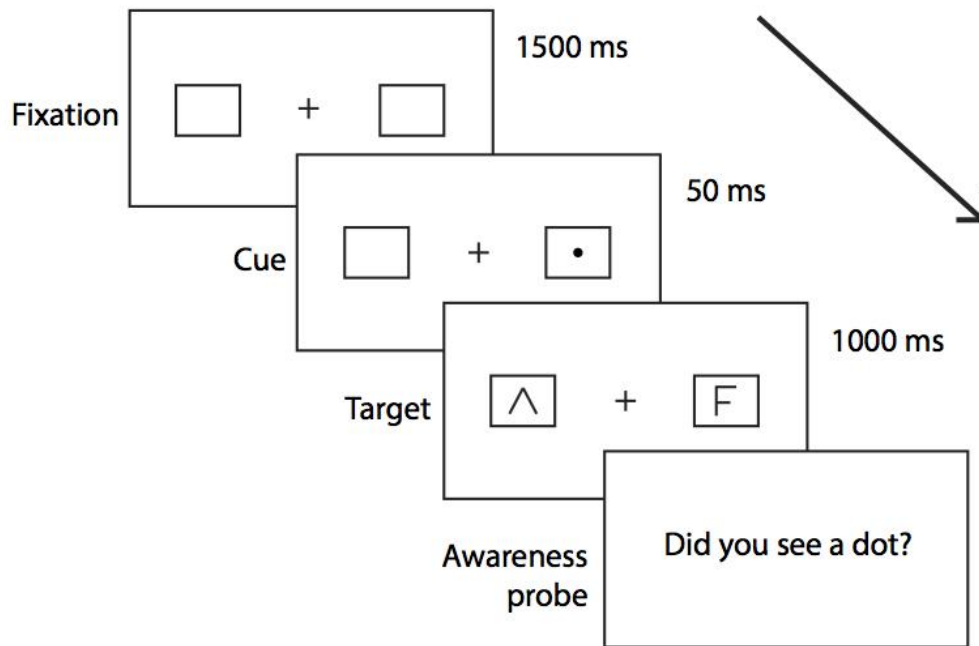


Figure 2.7: Behavioral paradigm for investigating interaction between awareness and bottom-up signal strength. 'Cue' is presented at two levels of visual contrast.

The cue period was followed by a 1-second target period during which two items were presented, one in each box. The items consisted of one target character, an "A" or "F," and one distractor character, identical to an "A" or "F" but with the omission of a

horizontal line. Participants were instructed to ignore the distractor and identify the target using the “A” and “F” keys on a keyboard as quickly as possible.

Following the target period, the screen displayed the question: “Did you see a dot?” Participants rated on a scale of 1–7 how certain they were that they saw a dot before the target. A 1 indicated confidence a dot was not seen. A 7 indicated confidence a dot was seen. Participants responded with either 1 or 7 on most trials (81%, 40% 1 and 41% 7). To ensure that we analyzed only trials in which participants were definitely aware or not aware of the dot, we limited analysis to trials in which participants indicated either 1 or 7.

Each participant performed 24 runs of 20 trials each (480 trials total) completed in a single session. Trial types were counterbalanced and randomized within each run. The side on which the cue appeared, the signal strength of the cue, and the side on which the target appeared were randomized and the location of the cue did not predict the location of the target.

Participants performed the cue awareness response with a low false alarm rate (rate at which they reported 7 when no cue was present, mean = 2%, SD = 2%) and performed the target discrimination task with high accuracy (mean = 95%, SD = 3.4%). The same Δt metric as in Behavioral Experiments 1 and 2 was used as a measure of attention. For each participant, four Δt scores were computed corresponding to a 2×2 design: Aware vs. unaware $\times$ Low contrast vs. high contrast. The data were then analyzed with a 2×2 repeated-measures ANOVA.

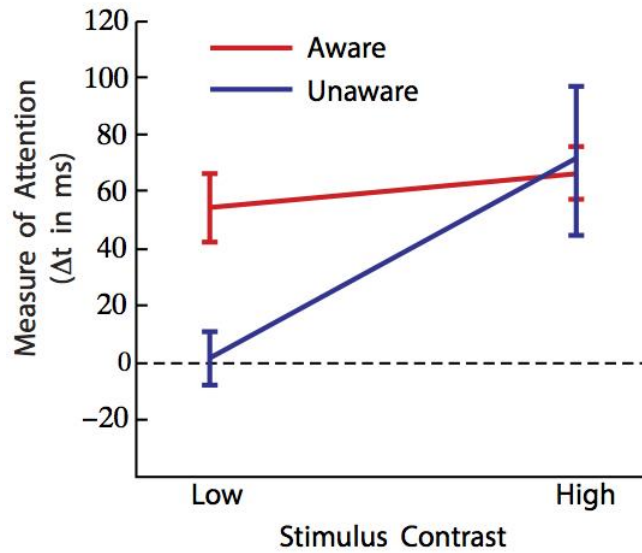


Figure 2.8: Results for Behavioral Experiment 3, showing interaction between awareness and stimulus contrast.

2.5.3 Results

Behavioral Experiment 4 tested stimulus-driven attention 50 ms after the onset of the stimulus. The effect of stimulus contrast on attention was measured. Figure 2.8 shows the results. When participants reported being definitely aware of the cue (red line), attention to the cue was weakly modulated by contrast. When participants reported being definitely unaware of the cue (blue line), attention to the cue was strongly modulated by contrast. A 2×2 repeated-measures ANOVA showed no significant main effect of awareness ($F_{1,25} = 1.2, p = .28$) indicating that awareness was not associated with an overall increase or decrease in attention. The main effect of stimulus contrast was significant ($F_{1,25} = 9.1, p = .005$), confirming that attention was overall greater to the higher contrast stimulus. Crucially, the interaction was significant ($F_{1,25} = 5.2, p = .03$), confirming that stimulus contrast had a greater effect on attention in the unaware condition than in the aware condition. Because of this interaction, the results cannot

easily be explained in terms of simple main effects such as an overall increase of attention or signal strength in aware trials.

2.5.4 Discussion

Behavioral Experiments 1-3 focused on interactions between awareness and task-relevance, suggesting that awareness plays a role in helping to control attention in line with task demands. In Behavioral Experiment 4, by contrast, we explored the interaction between awareness and bottom-up signal strength. The results showed that attention was driven by external signal strength more when participants were unaware of the stimulus.

2.6 General Discussion

In this chapter, the results of four behavioral experiments were presented. The goal of these experiments was to precisely map the psychophysical relationship between attention and awareness. This goal served two purposes. First, this relationship is interesting in its own right. A number of theories of consciousness make specific predictions about this relationship. Some theories claim that attention and awareness are literally the same process (Posner, 1994; De Brigard & Prinz, 2010), though these theories would appear untenable now on the basis of evidence for the dissociation between these two processes. (McCormick, 1997; Lamme, 2004; Tsushima et al., 2006; Jiang et al., 2006; Koch & Tsuchiya, 2007; Kentridge et al., 2008; Hsieh et al., 2011; Norman et al., 2013; Lin & Murray, 2015; Webb et al., 2016a).

Some theories make the counterintuitive claim that awareness plays no functional role whatsoever (Lau & Rosenthal, 2011), which would rule out the possibility of awareness playing a specific role in attention. By contrast, some theories claim that

awareness, though a distinct process from attention, interacts with attention in important ways. One such theory is the ‘attention schema theory’ (Graziano & Kastner, 2011; Graziano, 2013; Graziano & Webb, 2014; Webb & Graziano, 2015), according to which awareness is a schematic model of attention, similar to the relationship between the body schema and the body. Just as the body schema plays a role in facilitating the control of the body, a key prediction of the attention schema theory is that awareness should play a role in facilitating the control of attention. The experiments presented in this chapter are consistent with this prediction. Experiments 1-3 show that, in the presence of awareness, attention to a task-relevant stimulus is more effectively sustained, and attention to a task-irrelevant stimulus is more effectively inhibited. By contrast, Experiment 4 shows that attention in the absence of awareness is likely a function of the bottom-up signal strength of the stimulus. Together, these results suggest that awareness plays a specific role in the top-down control of attention, but that bottom-up attention is still possible in the absence of awareness.

It is worth noting, however, that the attention schema theory is not the only theory which predicts a relationship between awareness and top-down attention. The global workspace theory (Baars, 1997; Dehaene & Naccache, 2001) also predicts such a relationship. Future experiments should focus on mitigating between the more specific predictions of these theoretical perspectives. One promising avenue for doing so might be to explore different signatures for ‘model-based’ vs. ‘model-free’ control, a notion which has been fruitfully applied to the domain of motor control (Haith & Krakauer, 2013).

In addition to the relevance of the present experiments for theories of consciousness, they are also of methodological significance for research on the brain

basis of attention and awareness. Though it is possible to dissociate attention and awareness in the laboratory, they are nevertheless highly correlated under normal circumstances. Unless care is taken to explicitly measure and control for interactions between these two processes, studies of the brain basis of awareness risk being confounded by effects of attention. By systematically mapping the relationship between attention and awareness, the experiments reported in this paper lay the foundation for investigating the extent to which separate brain networks underlie these processes, an approach pursued in Chapter 5.

Chapter 3

Parcellation of TPJ involvement in multiple cognitive domains

A large body of research has investigated the brain networks that underlie subjective awareness (Lumer et al., 1998; Dehaene et al., 2001; Rees et al., 2002; Tong, 2003; Crick & Koch, 2005; Carmel et al., 2006; Lau & Passingham 2006; Webb et al. 2016b). Though this literature has considered many brain regions, ranging from primary sensory cortex (Tong, 2003), to prefrontal cortex (Lau & Passingham 2006), and even subcortical regions (Crick & Koch, 2005), many studies now point to a network of frontal and parietal regions that are consistently engaged by subjective awareness (Lumer et al., 1998; Dehaene et al., 2001; Rees et al., 2002; Carmel et al., 2006; Lau & Passingham 2006; Webb et al., 2016b), especially when controlling for lower level perceptual processes. One region of special interest is the temporoparietal junction (TPJ), both because this region is consistently engaged by subjective awareness in neuroimaging studies (Lumer et al., 1998; Dehaene et al., 2001; Rees et al., 2002; Webb et al. 2016b), and because lesions to this region often produce a devastating disorder of awareness known as hemispatial neglect (Vallar & Perani, 1986; Corbetta & Shulman, 2011).

In addition to its established involvement in subjective awareness, the TPJ is also thought to play a role in a number of other cognitive domains. These include bottom-up attention (Corbetta & Shulman, 2000; Corbetta & Shulman, 2002), social cognition

(Saxe & Kanwisher, 2003; Saxe & Powell, 2006), episodic memory (Konishi et al., 2000), and target detection (Stevens et al., 2000). Correspondingly, a number of cortical networks pass through the TPJ, including the so-called ‘default mode’ network (DMN; Buckner et al., 2008), the ‘ventral attention’ network (VAN; Corbetta & Shulman, 2002), the ‘frontoparietal control’ network (FPN; Vincent et al., 2008), and the ‘cingulo-opercular’ network (CON; Dosenbach et al., 2007). Ideally, an account of TPJ involvement in subjective awareness will specify which of these specific networks play a role, and how this relates to the involvement of these networks in other task domains. An important first step in this direction is to systematically map out the involvement of these networks in the many task domains that are known to engage the TPJ. In this chapter, a series of experiments are reported which map out the involvement of 4 distinct cortical networks, each of which passes through the TPJ, in social cognition, episodic memory, bottom-up attention and target detection.

Given the spatial resolution of functional MRI (fMRI), signals from adjacent cortical networks are often spatially overlapping, presenting a technical challenge for any attempt to separately map out their involvement in different tasks. Independent Component Analysis (ICA) provides a powerful solution to this problem. ICA allows the ‘unmixing’ of overlapping signals, and is thus an appealing tool for resolving distinct, but spatially overlapping networks in fMRI and other brain imaging data (Brown et al., 2001).

ICA has previously been used to demonstrate the presence of 5 independent components within the TPJ, each connected to distinct cortical networks (Igelström et al., 2015). These include a dorsally located component (TPJd), functionally connected to the frontoparietal control network (FPCN), an anteriorly located component (TPJa),

functionally connected to the cingulo-opercular network (CON), a ventrally located component (TPJv), functionally connected to an auditory network, a posteriorly located component (TPJp) functionally connected to the default mode network (DMN), and a centrally located component (TPJc), functionally connected to the ventral attention network (VAN).

The experiments in this chapter investigate which of these TPJ subcomponents are engaged by different cognitive domains (Igelström et al., 2016). Two tasks are used to investigate social cognition. One is a standard ‘false belief’ task, commonly used to study ‘theory of mind’, the attribution of beliefs to another agent (Saxe & Kanwisher, 2003; Dodell-Feder et al., 2011). Another is a task recently developed to investigate the attribution of attention, rather than beliefs, to another agent (Kelly et al., 2014). Two tasks investigate the closely related processes of reorienting attention (Corbetta & Shulman, 2002) and target detection (Stevens et al., 2000). A fifth task investigates episodic memory retrieval (Konishi et al., 2000). The results show that TPJa, TPJc and TPJp are activated in a domain-specific manner, while TPJd is activated across all of the tested task domains.

My collaborator, Dr. Igelström was the primary author on four out of five of the experiments presented in this chapter, but they are presented here because they provide crucial contextual information that aids in the interpretation of the results presented in subsequent chapters, on which I was the principal experimenter.

3.1 Behavioral Paradigms

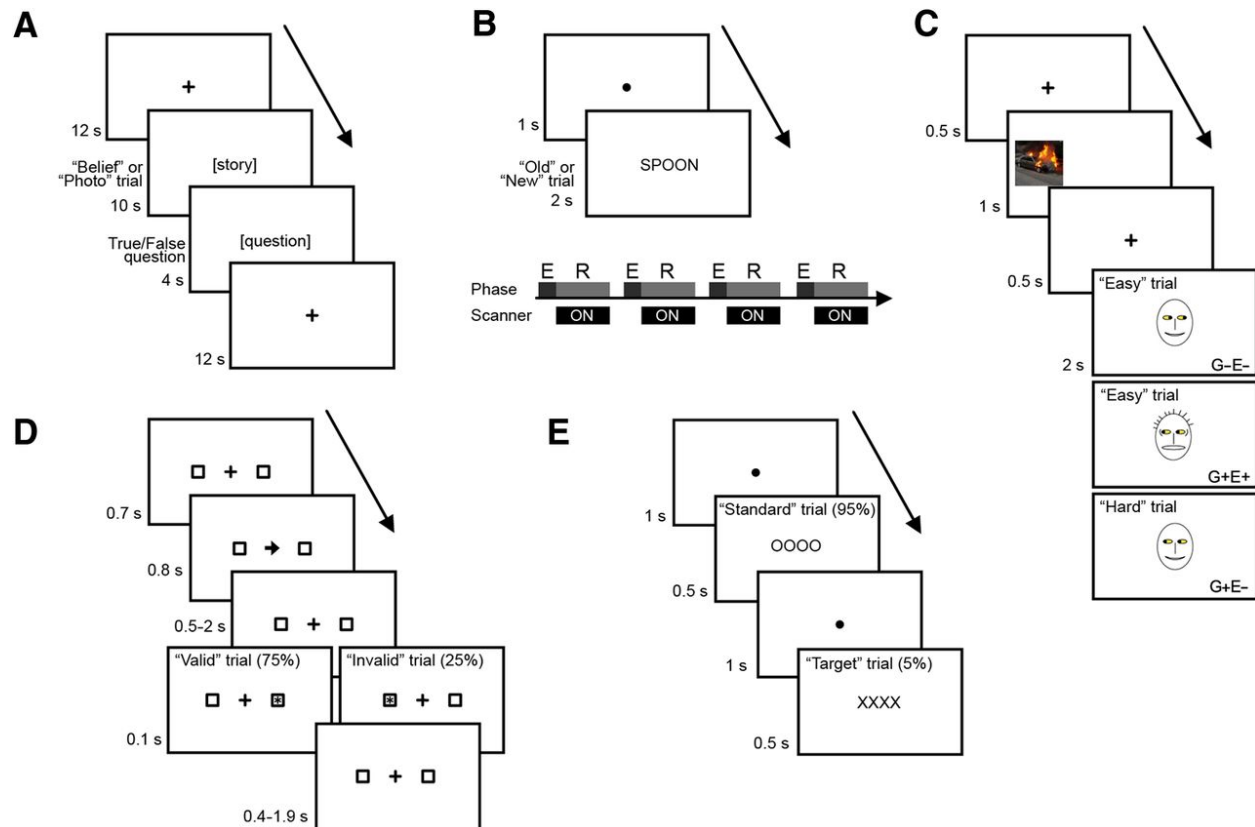


Figure 3.1: Behavioral paradigms used to study TPJ involvement in multiple cognitive domains. A. Theory-of-mind task. B. Episodic memory retrieval task. C. Attribution-of-attention task. D. Attentional reorienting task. E. Target detection ("oddball") task.

The Behavioral paradigms are shown in Figure 3.1. These paradigms were chosen because they represent a relatively broad cross-section of the cognitive domains that are known to engage the TPJ. In each case, we used a paradigm which has already been shown to reliably produce TPJ activity. Detailed descriptions of those paradigms can be found in other publications, but are described briefly in this section.

3.1.1 Theory-of-mind Task

Twenty participants performed the theory-of-mind task (12 females; mean age, 22.6 ± 0.8 years old). All participants provided informed consent, and all procedures were approved by the Princeton Institutional Review Board.

We used a theory-of-mind localizer task from a study by Dodell-Feder et al. (2011). In this task (Figure 3.1A), participants performed trials of two types: “belief” trials, in which participants made judgments about the beliefs of another agent, and “photo” trials, in which participants made judgments about the content of photographs. Belief trials produce greater TPJ activation than photo trials, motivating a hypothesized role for the TPJ in reasoning about the beliefs of others, a process referred to as ‘theory of mind’ (Saxe & Powell, 2006; Scholz et al., 2009; Dodell-Feder et al., 2011).

Each trial began with the presentation of a story for 10 seconds, followed by the presentation of a true/false question for 4 seconds, and an intertrial interval (ITI) of 12 seconds. The task consisted of two runs of 10 trials each, and the order of stories was counterbalanced and equally distributed across the two runs. The BOLD response was modeled with a 14 s boxcar convolved with a standard hemodynamic response function, and the standard contrast of [belief trials - photo trials] was applied.

3.1.2 Episodic Memory Task

Twenty participants performed the episodic memory task, one of whom was excluded due to movement in the MRI scanner (10 females; mean age, 22.3 ± 1 years old). All participants provided informed consent, and all procedures were approved by the Princeton Institutional Review Board.

We used an episodic memory task based on the one used by Konishi et al. (2000). In this task (Figure 3.1B), participants performed four runs. Each run began with a memory-encoding phase, during which participants were presented with 33 words to memorize (presented on screen for 2 seconds each, with a 1-second ITI). This was followed by a memory-retrieval phase, during which participants were presented with 25 words from the encoding phase, and 25 new words, in a random order. Participants were asked to judge whether each word was ‘new’ or ‘old’ (from the list presented during the encoding phase). Trials with successfully retrieved old words produce more activation in the TPJ than trials with new words (which are correctly identified as new words), motivating a hypothesized role for the TPJ in episodic memory retrieval (Konishi et al., 2000; Hutchinson et al., 2009, Cabeza et al., 2012).

The word list consisted of five-letter words, and was taken from the SUBTLEXus 1.00 word frequency database (Brysbaert & New, 2009). The BOLD response was modeled with a boxcar time course convolved with a standard hemodynamic response function, and the standard contrast of [old trials - new trials] was applied.

3.1.3 Attribution-of-attention Task

Kelly et al. (2014) studied brain activations during a task requiring the attribution of attention, rather than beliefs (as in the theory-of-mind task), to another agent, finding that this process also evokes TPJ activity. We used data collected as a part of that study. Though 50 participants participated in that study, we used only the data from the first 20 participants in the dataset (8 females; mean age, 19.4 ± 0.4 years old), both as an unbiased way of balancing the sample size with the other 4 tasks, and because of computational limits on performing ICA for large group datasets. All

participants provided informed consent, and all procedures were approved by the Princeton Institutional Review Board.

In this task (Figure 3.1C), in each trial, participants were presented with an image of an object for 1 second, followed by fixation for 0.5 seconds, followed by an image of a cartoon face (referred to as “Kevin”) for 2 seconds. Participants were asked to judge whether Kevin is aware of the object in question, on a scale of 1 (“not aware”) to 3 (“very aware”). Participants have two cues to Kevin’s state of awareness: the direction of Kevin’s gaze, which can either be aligned (looking toward; “G+” in Figure 3.1C) or misaligned with (looking away from; “G-” in Figure 3.1C) the object, and Kevin’s facial expression, which can also be either aligned (matching in valence with the object, such as smiling for an image of a cake; “E+” in Figure 3.1C) or misaligned with (such as smiling for an image of a car accident; “E-” in Figure 3.1C) the object. Participants performed 8 runs of 48 trials each, with gaze direction and facial expression randomized and counterbalanced across trials.

Two trial types exist. On ‘easy’ trials, both sources of information were aligned (either “G+E+” or “G-E-”), meaning that a judgment about Kevin’s state of awareness was relatively easy. On ‘hard’ trials, Kevin’s gaze and facial expression were misaligned with each other (either “G+E-” or “G-E+”), making a decision about Kevin’s state of awareness more difficult. Kelly et al. (2014) found that TPJ activity was greater on “hard” trials than “easy” trials, motivating a hypothesized role for the TPJ in integrating disparate sources of information to determine another agent’s state of attention. We convolved hard and easy trials with a hemodynamic response function and applied the same contrast ([hard trials - easy trials]).

3.1.4 Attentional Reorienting Task

Twenty participants performed the attentional reorienting task (11 females; mean age, 21.6 ± 0.6 years old). All participants provided informed consent, and all procedures were approved by the Princeton Institutional Review Board.

Participants performed a Posner spatial attention task (Posner, 1980). In this task (Figure 3.1D), each trial began with central fixation for 700 ms, followed by the presentation of a ‘cue’ stimulus for 800 ms, a centrally presented arrow pointing to the left or the right, intended to inform participants about the most probable location of a subsequent ‘target’ stimulus (an asterisk). This was followed by a variable delay (500 ms-2 seconds), and the presentation of the target for 100 ms on either the left or the right. Participants responded to the appearance of the target as quickly as possible. Each trial concluded with a variable ITI of 400 ms-1.9 seconds is presented. Participants performed 5 runs of 40 trials each.

The target appeared in the cued location on 75% of trials (“valid” trials) and the uncued location on 25% of trials (“invalid” trials). The common interpretation is that participants direct attention toward the cued location, and therefore must ‘reorient’ attention on invalid trials, when the target appears on the opposite side. Invalid trials produce more TPJ activity than valid trials, motivating a hypothesized role for the TPJ in reorienting attention (Corbetta & Shulman, 2002; Corbetta et al., 2008; Mitchell, 2007), though others have argued that this result instead reflects the adjustment and updating of top-down expectations (Geng & Vossel, 2013).

We convolved stimulus onset times with a gamma function and applied the contrast [invalid trials - valid trials]. The reaction time for invalid trials was significantly

longer than the reaction time for valid trials (paired t test, $p = 0.00007$), reflecting the effect of the cue on attention.

3.1.5 Target Detection Task

Twenty participants performed the target detection task, one of whom was excluded due to poor behavioral performance (10 females; mean age, 22.9 ± 1.0 years old). All participants provided informed consent, and all procedures were approved by the Princeton Institutional Review Board.

In this task (Figure 3.1E), each trial began with 1 second of fixation, followed by the presence of a stimulus for 500 ms. On 93-96% of trials, this stimulus was four O's (a "standard" trial), but on 4-7% of trials, this stimulus was four X's (a "target" trial). At the end of each trial, participants reported how many target trials they saw. Because of the infrequency of the 'target' trials, this paradigm is commonly referred to as an 'oddball' task. Target trials produce more TPJ activation than standard trials, motivating a hypothesized role for the TPJ in target detection (Stevens et al., 2000).

Participants performed 4 runs of 120 trials each. We convolved stimulus onset times with a gamma function and applied the contrast [target trials - standard trials].

3.1.6 Distribution of Subjects Across Tasks

For reasons of feasibility, every subject did not perform every task. Fourteen subjects performed both the theory-of-mind task and the attentional reorienting task. Six subjects performed only the attentional reorienting task, and six subjects performed only the theory-of-mind task. Eighteen subjects performed both the episodic memory retrieval task and the target detection task. One subject performed only the episodic memory retrieval task, and one subject performed only the target detection task. For the

social attribution-of-attention task, data from 20 subjects were used. In total across the five tasks, 66 subjects were tested. Because of the mixed within-subjects and between-subjects design, statistical analysis was performed separately for each task, and only qualitative comparisons are made between tasks.

3.2 MRI Methods

3.2.1 MRI Data Collection

MRI images covering the whole cerebral cortex were acquired with a 20-channel receiver head coil on a Siemens Skyra scanner. Functional imaging used a gradient echo, echoplanar pulse sequence with a 64×64 matrix [27 axial slices; 4 mm thick; in-plane resolution, 3×3 mm; TR, 1.5 s; TE, 28 ms; flip angle (FA), 64° ; generalized GRAPPA iPAT = 2. Anatomical imaging used an MP2RAGE sequence (256×240 matrix; TR, 5 seconds; TE, 2.98 ms; FA, 4° ; 1 mm^2 resolution; GRAPPA iPAT = 3). The reanalyzed data from Kelly et al. (2014) were acquired on the same scanner. The functional data were acquired with a 64×64 matrix (35 axial slices; 3 mm thick; in-plane resolution, 3×3 mm; TR, 2 seconds; TE, 30 ms; FA, 77°), and the anatomical data were acquired with an MPRAGE sequence (256×224 matrix; TR, 2.3 seconds; FA, 9° ; and with 1 mm^2 [TE, 2.98 ms], 0.9 mm^2 [TE, 3.08 ms], or 1.1 mm^2 [TE, 2.93 ms] resolution).

3.2.2 MRI Preprocessing

Preprocessing was performed with AFNI (Cox, 1996) and FSL (Jenkinson et al., 2012). The functional data were slice time corrected and motion corrected with FSL (Jenkinson et al., 2002), and then detrended (linear and quadratic) with AFNI. The data were spatially normalized to the FSL MNI-152 template with AFNI, and spatially smoothed with a Gaussian kernel (5 mm FWHM). We used an ICA-based strategy for automatic

removal of motion artifacts (ICA-AROMA; Pruim et al., 2015, Beckmann & Smith, 2004).

3.2.3 Independent Component Analysis

For each of the five tasks, we performed group ICA by temporally concatenating the data from all runs and all subjects into one matrix (MELODIC toolbox in FSL; Beckmann & Smith, 2004). We performed the ICA decomposition separately for each task (instead of grouping the tasks into one ICA) because we did not want to assume that the ICA decomposition would be the same across task conditions.

We applied the ICA to the voxels within a region of interest (ROI) mask that included the TPJ and surrounding cortex to ensure that all relevant independent components (ICs) were detected in their entirety. The mask was constructed from the standard surface `cvs_avg35_inMNI152` in Freesurfer, using `mri_label2vol` to combine multiple labels from the `aparc.a2009s` atlas into one mask (`G_pariet_inf-Supramar`, `G_pariet_inf-Angular`, `G_temp_sup-Plan_tempo`, `G_temp_sup-Lateral`, `G_temp_sup-G_T_transv`, `S_interm_prim-Jensen`, `S_temporal_sup`, `S_temporal_transverse`) and trimming temporal cortex voxels anterior to the postcentral sulcus. The fMRI data were decomposed into 20 ICs, a dimensionality that is known to isolate the major functional processes in the region (Igelström et al., 2015).

Task-related ICs were identified using a mixed-effects multiple regression (subjects as random effects) in R version 3.0.3 (nlme package version 3.1-113; R Core Team, 2014), with the IC time courses as dependent variables and the predicted BOLD responses for each condition as independent variables (two trial types per task). ICs were only included as task-related if both 1) they had a statistically significant relationship with the positive regressor of interest (e.g. ‘belief’ trials in the theory-of-

mind task) *and* 2) there was a statistically significant contrast between that regressor and the corresponding control condition (e.g. a statistically significant contrast between ‘belief’ and ‘photo’ trials).

3.2.4 Functional Connectivity Analysis

We performed seed-to-voxel functional connectivity analysis (Biswal et al., 1995), using the CONN toolbox 15.c in SPM 12 (Whitfield-Gabrieli & Nieto-Castanon, 2012), with the subject-specific IC time courses as seed time courses. Conventional bivariate correlation analysis was used, with a voxelwise threshold of $p < 0.001$ uncorrected and a cluster extent threshold of $p < 0.05$ (false discovery rate corrected).

3.3 Results

3.3.1 Theory-of-mind Task

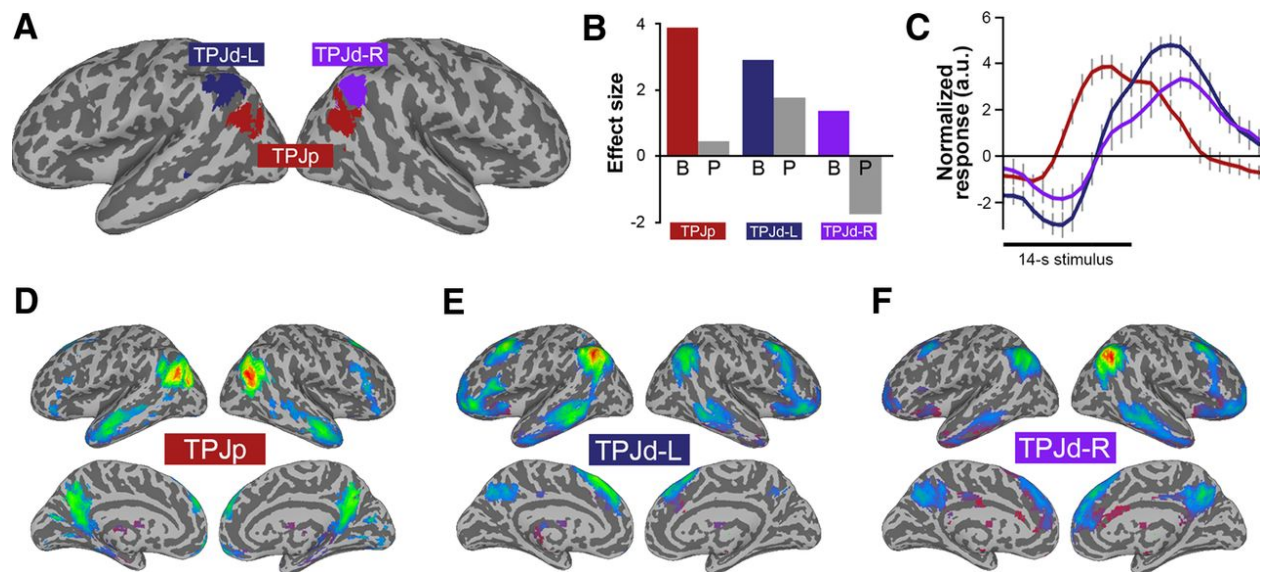


Figure 3.2: Results for the theory-of-mind task. A. Winner-take-all maps for significant task-related ICs. B. Regression coefficients for the belief (B) and photo (P) conditions. C. Time courses for the photo condition. D. Functional connectivity map for TPJp. E. Functional connectivity map for TPJd-L. F. Functional connectivity map for TPJd-R.

The results for the theory-of-mind task are shown in Figure 3.2. The contrast [belief - photo] was significant for a bilateral, posterior component consistent with the location of TPJp, and functionally connected to the default mode network (Figure 3.2D), as expected for TPJp. This contrast was also significant for two lateralized, dorsal components, consistent with the location of left and right TPJd (TPJd-L and TPJd-R), and functionally connected to the frontoparietal control network (Figures 3.2E and 3.2F), consistent with the standard pattern of connectivity for TPJd.

3.3.2 Episodic Memory Task

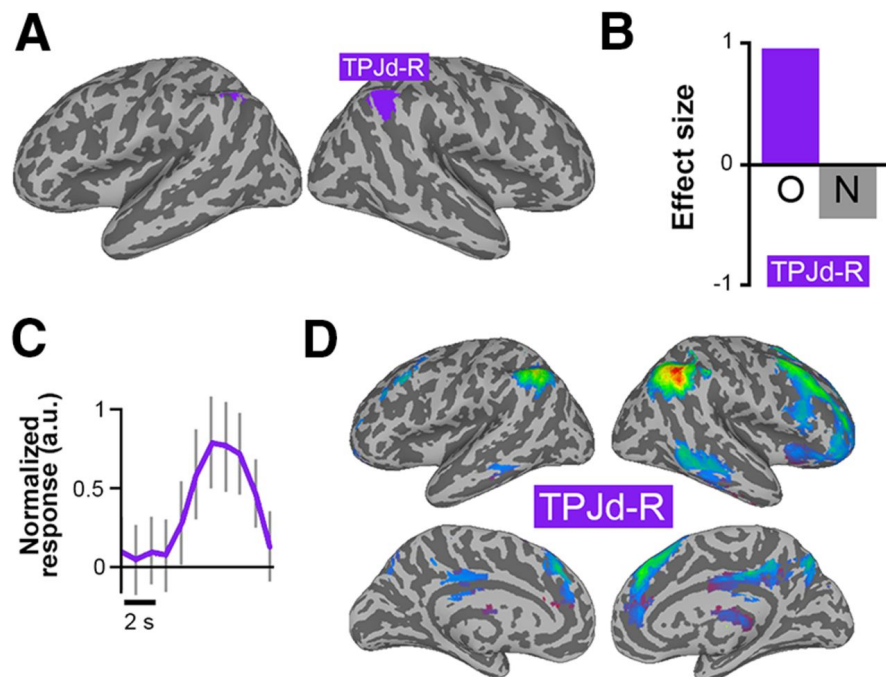


Figure 3.3: Results for the episodic memory task. A. Thresholded map for TPJd-R, the only significant task-related IC. B. Regression coefficients for the old (O) and new (N) conditions. C. Time course for the old condition. D. Functional connectivity map for TPJd-R.

The results for the episodic memory task are shown in Figure 3.3. The contrast [old - new] was significant for a right-lateralized, dorsal component consistent with the location of TPJd-R, and functionally connected to the frontoparietal control network

(Figures 3.3D). A marginally significant left-lateralized, dorsal component was also present ($p = 0.026$), consistent with the location of TPJd-L.

3.3.3 Attribution-of-attention Task

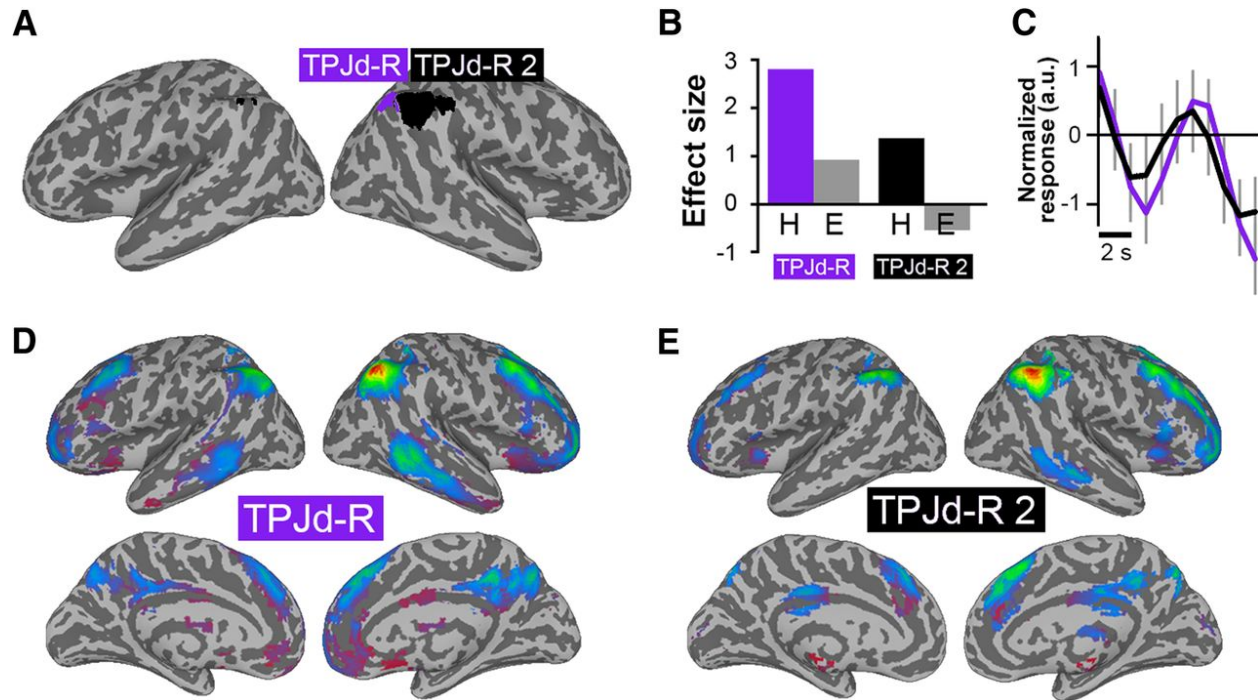


Figure 3.4: Results for the attribution-of-attention task. A. Winner-take-all maps for significant task-related ICs. B. Regression coefficients for the hard (H) and easy (E) conditions. C. Time courses for the hard condition. D. Functional connectivity map for TPJd-R. E. Functional connectivity map for TPJd-R2.

The results for the attribution-of-attention task are shown in Figure 3.4. The contrast [hard - easy] was significant for two right-lateralized, dorsal components, both consistent with the general location of TPJd-R. Both of these components were functionally connected to the frontoparietal control network (Figures 3.4D and 3.4E). Because of this common pattern of connectivity, they are labelled TPJd-R and TPJd-R2.

3.3.4 Attentional Reorienting Task

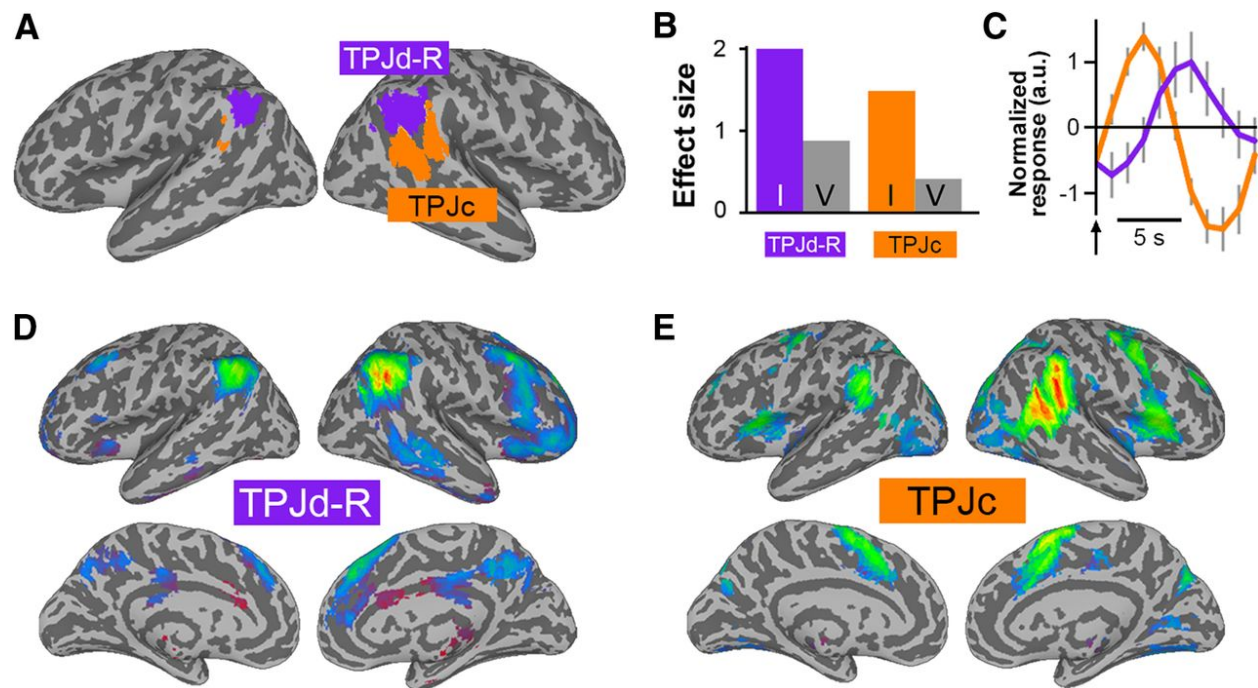


Figure 3.5: Results for the attentional reorienting task. A. Winner-take-all maps for significant task-related ICs. B. Regression coefficients for the invalid (I) and valid (V) conditions. C. Time courses for the invalid condition. D. Functional connectivity map for TPJd-R. E. Functional connectivity map for TPJc.

The results for the attentional reorienting task are shown in Figure 3.5. The contrast [invalid - valid] was significant for a right-lateralized, dorsal component consistent with the location of TPJd-R, and functionally connected to the frontoparietal control network (Figure 3.5D). This contrast was also significant for a right-lateralized, central component, consistent with the location of right TPJc (TPJc-R), and functionally connected to the ventral attention network (Figure 3.5E), as expected for TPJc-R.

3.3.5 Target Detection Task

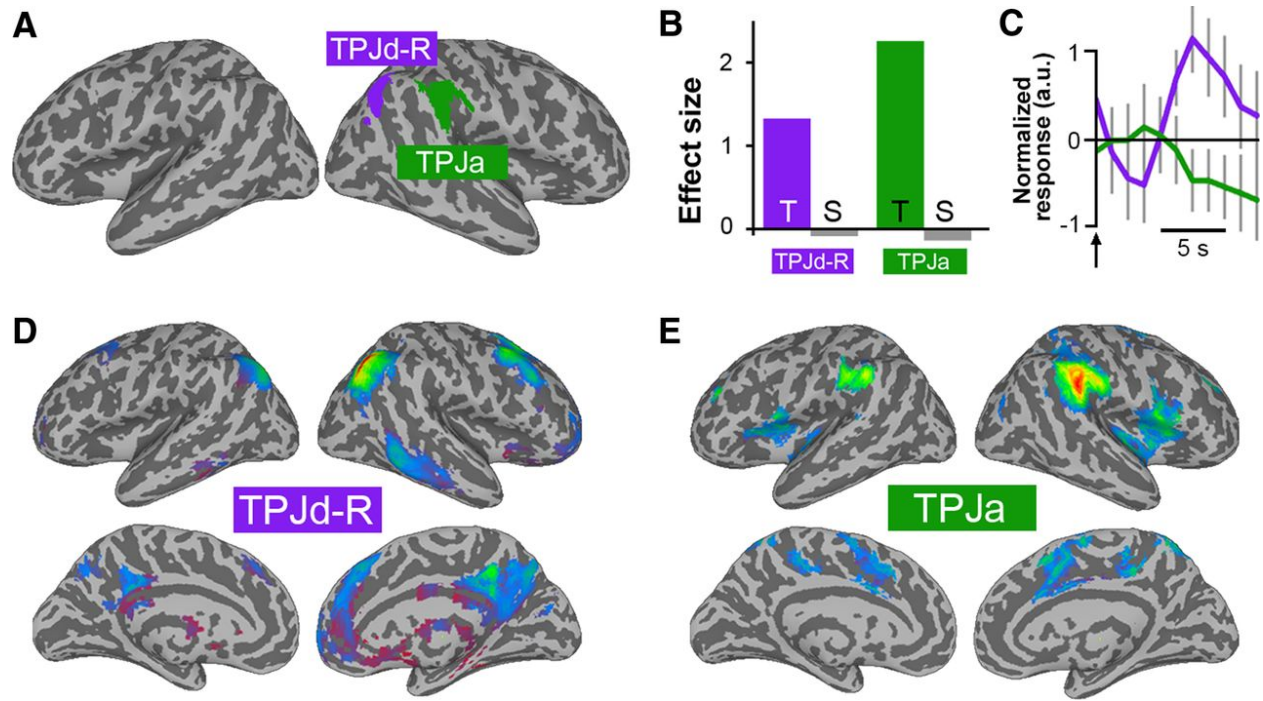


Figure 3.6: Results for the target detection task. A. Winner-take-all maps for significant task-related ICs. B. Regression coefficients for the target (T) and standard (S) conditions. C. Time courses for the target condition. D. Functional connectivity map for TPJd-R. E. Functional connectivity map for TPJa.

The results for the target detection task are shown in Figure 3.6. The contrast [target - standard] was significant for a right-lateralized, dorsal component consistent with the location of TPJd-R, and functionally connected to the frontoparietal control network (Figure 3.6D). This contrast was also significant for a bilateral, anterior component consistent with the location of TPJa, and functionally connected to the cingulo-opercular network (Figures 3.6E), consistent with previous characterizations of this component.

3.4 Comparison Between Tasks and Discussion

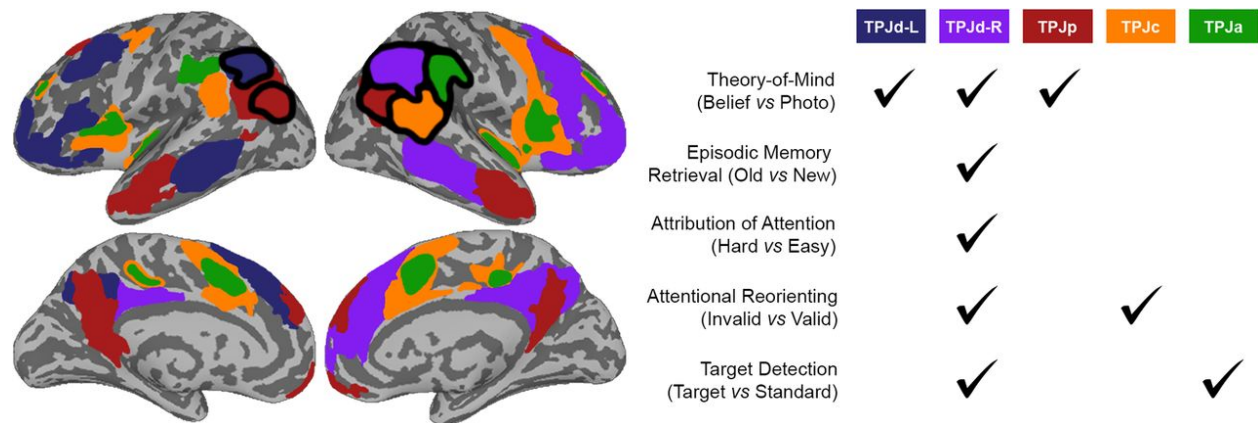


Figure 3.7: Comparison of results from all tasks. Left. Winner-take-all maps for significant task-related ICs and their corresponding patterns of functional connectivity. Right. Summary of ICs activated by each task.

Figure 3.7 shows a summary of the results from all five tasks. The results show a combination of domain-specificity and domain-generality within the TPJ. Some components played a domain-specific role in the cognitive domains tested. TPJa participated only in the target detection task, TPJc participated only in the attentional reorienting task, and TPJp participated only in the theory-of-mind task. TPJd, however, played a domain-general role, participating in all five of the tasks, ranging from social cognition, to episodic memory, to attentional reorienting and target detection.

Previous studies have asked whether the many task domains that engage the TPJ reflect distinct neural processes, or some common underlying computation. Specifically, previous studies have proposed that distinct parts of the TPJ may underlie memory and attention functions (Hutchinson et al., 2009), or attention and social cognition (Decety & Lamm, 2007; Scholz et al., 2009), while others have argued that these processes are largely overlapping (Mitchell, 2007), and can be accounted for by common mechanisms (Corbetta et al., 2008; Cabeza et al., 2012). The results presented in this chapter suggest

that both of these views may be partially correct, with specific functional components for some of the tested cognitive domains, but a significant degree of overlap in TPJd and the frontoparietal control network.

The present results are consistent with a range of proposals that the TPJ plays an important integrating role, either as a nexus or convergence point for disparate streams of information (Shimamura, 2011; Carter & Huettel, 2013), as an episodic buffer (Baddeley, 2000; Vilberg & Rugg, 2008) which integrates over a longer timescale than most other cortical regions (Baldassano et al., 2016), or by integrating multiple sources of information into a deeper model of one's own body (Blanke et al., 2015) and state of awareness (Graziano & Kastner, 2011, Graziano, 2013). The present results, however, suggest that this integrating role may be played more specifically by the dorsal TPJ and the frontoparietal control network to which it is connected, including the dorsolateral prefrontal cortex, the medial parietal cortex, and the middle temporal gyrus. This specific result is also consistent with the finding that the frontoparietal control network shows a more dynamic profile of functional connectivity with other cortical networks, switching based on task demands in a manner that suggests it might play a role in coordinating information flow between these networks (Cole et al., 2013).

These results also provide important context for research on the role of the TPJ in subjective awareness. By mapping out the involvement of the TPJ in other cognitive domains, they lay the groundwork for a more precise characterization of its involvement in subjective awareness, and, ultimately, an account of how these processes interact and emerge from neural mechanisms in the TPJ. With this foundation established, the subsequent chapters directly address the involvement of the TPJ in subjective awareness.

Chapter 4

Transcranial magnetic stimulation to the TPJ produces neglect-like effects

In the previous chapter, the involvement of the TPJ in multiple cognitive domains was assessed, showing that the dorsal TPJ (TPJd) was a common node for all of the tested domains. In this chapter, the results of an experiment are reported in which transcranial magnetic stimulation (TMS) is applied to this same region of the TPJ, resulting in an effect resembling the neuropsychological condition ‘hemispatial neglect’.

Hemispatial neglect usually takes the form of impaired detection of stimuli on the left side of visual space, typically following a right-lateralized brain injury (Brain, 1941). Neglect can emerge following damage to many different brain regions, including in the parietal lobe (Brain, 1941; Critchley, 1953; Mort et al., 2003), the temporal lobe (Karnath et al., 2001), the frontal lobe (Heilman & Valenstein, 1972), and even subcortical regions such as the pulvinar nucleus of the thalamus or parts of the basal ganglia (Karnath et al., 2002). A common lesion site for neglect however, perhaps the most common lesion site, is the TPJ (Vallar & Perani, 1986), likely involving some interaction with more dorsal regions of the parietal lobe in the intraparietal sulcus (IPS) and superior parietal lobule (SPL) (Corbetta & Shulman, 2011).

In the previous chapter, it was shown that TPJd was a point of convergence for multiple, distinct cognitive processes, including attention, episodic memory, and social cognition. A natural question is whether this specific node also underlies some of the

deficits observed in hemispatial neglect. One previous study has investigated neglect-like effects by applying TMS to the TPJ (Meister et al., 2006), but the site of stimulation in that study was more ventral and anterior than TPJd. In this chapter, the results of an experiment (Kelly et al., 2014) are reported in which TMS to TPJd produced a neglect-like impairment in visual detection. Furthermore, in this experiment, TPJd was localized using the social ‘attribution-of-attention’ task described in the previous chapter, allowing a characterization of the overlap between social cognitive functions and the visual detection impairments observed in hemispatial neglect. We discuss a theoretical perspective that provides a potential explanation for this puzzling pattern of overlap (Graziano & Kastner, 2011; Graziano, 2013).

4.1 Methods

4.1.1 Subjects

50 participants (28 females, 18-49 years old, normal or corrected-to-normal vision) performed the fMRI task that was used to localize sites for TMS. One to eight weeks later, participants performed the visual detection task while TMS was applied. Not all participants returned to perform the TMS experiment, so a large number of participants were tested in the MRI phase to result in a useable number of participants in the TMS phase. 20 participants returned to perform the TMS phase of the experiment (9 females, 18-45 years old, normal or corrected-to-normal vision). All participants provided informed consent, and all procedures were approved by the Princeton Institutional Review Board.

4.1.2 Visual Detection Task

Participants performed a simple visual detection task while undergoing TMS. In this task, participants had to detect, on each trial, whether a small dot appeared. On any given trial, the dot could appear on the right or the left, or no dot could appear at all. The task is shown in Figure 4.1.

Each block of trials (experimental block or control block) had a 2×3 design: [TMS to left or right hemisphere] $\times$ [target dot presented in left visual field, presented in right visual field, or absent]. Trials when the dot was absent were used to measure false alarm rates. Within a block, each of the six trial types was repeated 20 times for a total of 120 trials, and trial types were randomly interleaved.

Visual stimuli were presented on a monitor 100 cm away, using MATLAB Psychophysics Toolbox (Brainard, 1997). On each trial, a central fixation cross was presented for 1.5 seconds. Then, on a third of trials, a dark dot (one pixel) was presented 5° to the left of fixation for 200 ms; on a third of trials, the dot was presented 5° to the right of fixation; and on a third of trials, no dot was presented during the same 200 ms interval. Then a black and white random-dot mask covered the screen for a variable intertrial interval of 4-6 seconds. The subject was instructed to press a response key as quickly as possible after seeing the dot.

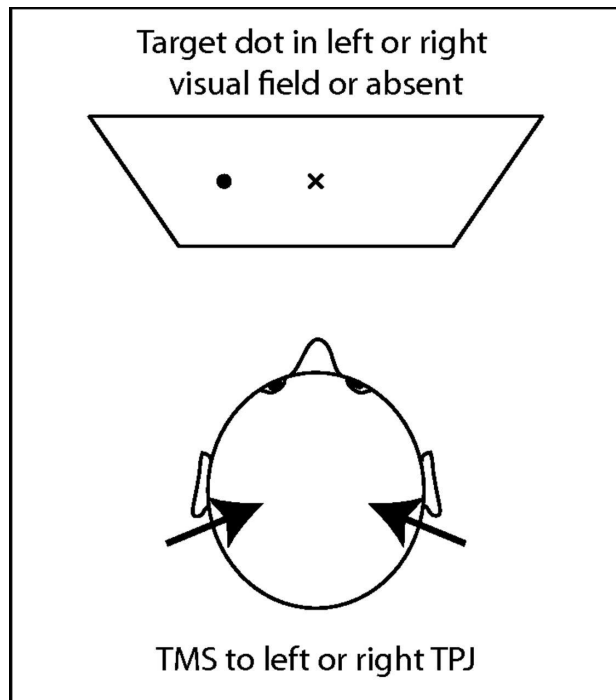


Figure 4.1: TMS visual detection paradigm. Participants performed a simple dot detection task, in which a dot could appear on the right or the left. Participants also received TMS to the left or the right TPJ.

Initial thresholding runs were performed in which the contrast of the target stimulus against the gray background was varied until the subject could detect the target with a hit rate between 40% and 60%. False alarm rates were computed from the no-dot trials, and most subjects performed below a 10% false alarm rate. When false alarm rates rise, it can be an indication that subjects are not motivated and that the performance is a result of guessing. Two subjects were excluded from further analysis for this reason (false alarm rates of 64% and 94%), leaving a total of 18 subjects for the analysis reported here. Once the behavioral threshold was obtained, TMS testing began. The TMS pulse was presented to the right or left TMS site on interleaved trials, 200 ms after the onset of the target dot. This asynchrony was successful in the past for disrupting parietal visual areas (Szczepanski & Kastner, 2013).

4.1.3 Localization of TPJd and TMS Control Site

To assess the overlap between mechanisms for social cognition and those involved in hemispatial neglect, TPJd was localized using a social cognition task. Specifically, participants performed the ‘attribution-of-attention’ task described in Chapter 3 (section 3.1.3). Briefly, participants viewed images of objects followed by images of a cartoon face (named ‘Kevin’) and were asked to determine whether Kevin was aware of the object in question. The behavioral paradigm is exactly as described in section 3.1.3.

Two trial types existed. In the ‘hard’ condition, participants had conflicting cues to Kevin’s state of awareness (his gaze and facial expression), making the task of determining his state of awareness difficult. In the ‘easy’ condition, cues to Kevin’s state of awareness were aligned, making the task easier. As demonstrated in Chapter 3 (section 3.3.3), the contrast [hard - easy] trials in this task shows significant activity in TPJd. We used this contrast to localize the TPJd site used in our TMS task.

The MRI data collection and preprocessing methods are exactly as described in Chapter 3 (sections 3.2.1 and 3.2.2). To localize TPJd, in each subject, the contrast [hard - easy] was performed at the whole brain level using a general linear model (Friston et al., 1995) in AFNI (Cox, 1996). The results were thresholded using a Monte Carlo simulation to achieve a corrected significance of $p < 0.05$ adjusted for a cluster size of 15 adjacent voxels. For each subject, the ‘experimental’ TMS site was defined as the peak of statistically significant activity in the left and right TPJ. The ‘control’ TMS site was defined as 2 cm anterior to the left and right ‘experimental’ TPJ site, at a site that did not show any statistically significant activity in the attribution-of-attention task.

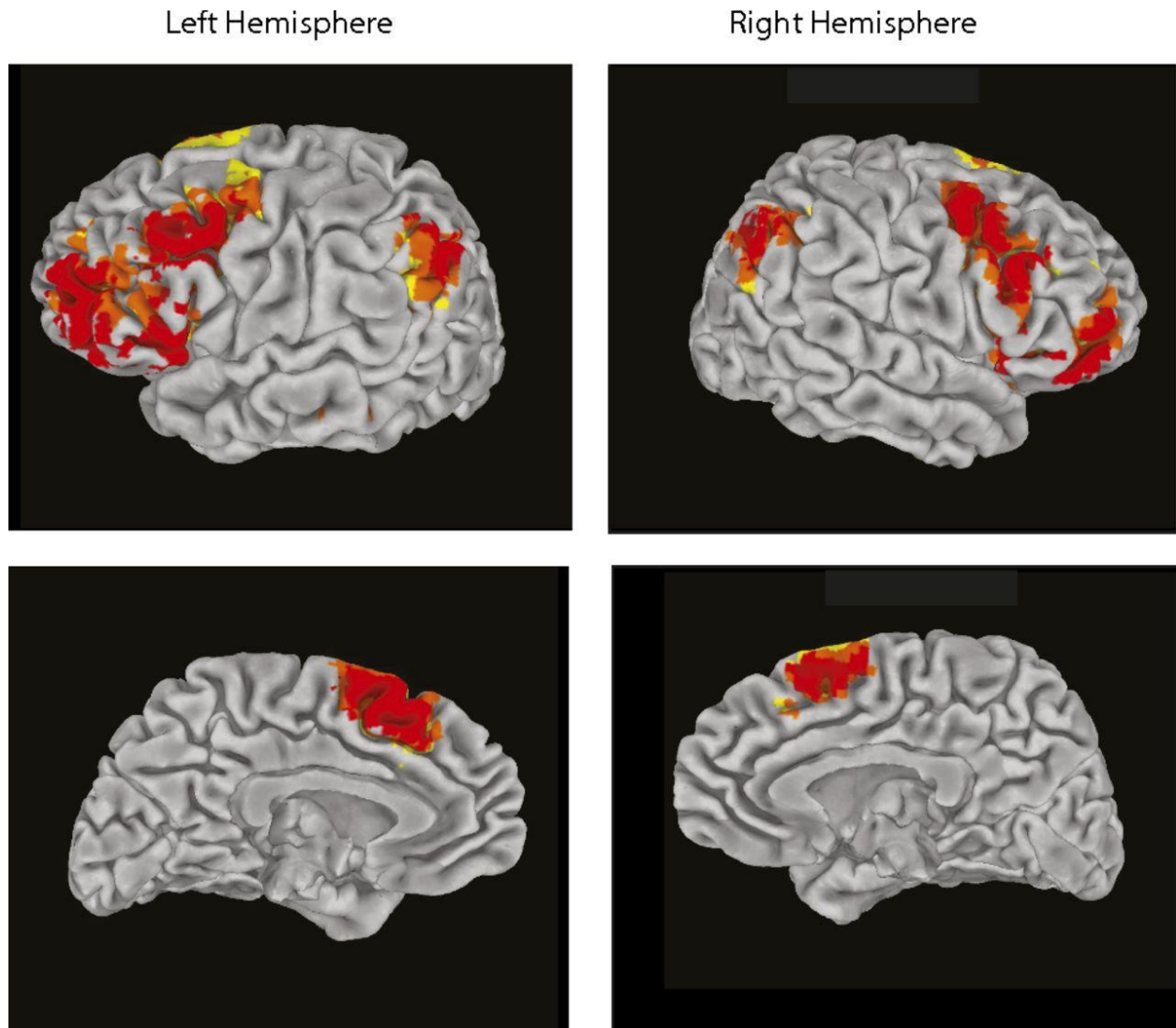


Figure 4.2: Group fMRI data from 50 subjects, showing the contrast [hard - easy] trials in a task requiring the social attribution of attention. TMS was targeted to the peak of activity in the left and right TPJ on an individual basis.

The fMRI results were also visualized at the group level. The nine-parameter landmark method (Talairach & Tournoux, 1988) available through AFNI was used to align the functional activation maps of all subjects to a single reference. A mixed-effects meta-analysis (Chen et al., 2012) was used to find regions of significant activation thresholded using a Monte Carlo simulation to achieve a corrected significance of $p < 0.05$ adjusted for a cluster size of 15 adjacent voxels.

The result of the [hard - easy] contrast, for all 50 participants is shown in Figure 4.2. The result shows significant clusters of activity in the right and left TPJ, consistent with the location of TPJd. The result also shows significant bilateral clusters of activity in the dorsolateral prefrontal cortex (DLPFC), consistent with the connectivity pattern of the frontoparietal control network.

4.1.4 TMS Procedures

Subjects sat with head position stabilized by a chin rest. Single-pulse TMS was applied using two independently triggered figure-of-eight Magstim 200 Mono Pulse TMS coils. The coils were targeted to specific cortical loci by means of theBrainsight 2.2.7 targeting software, using a Polaris camera to track the 3D position and orientation of the coils and the subject's head. After confirmation that a muscle twitch could be evoked from the hand region of the motor cortex, the TMS coils were clamped into position on either side of the subject's head to target the left and right TPJ with a signal strength set at 70% of the maximum for the coil drivers.

Subjects were tested with an experimental block of trials and a control block of trials, the order randomized between subjects. This blocked design was necessary because of the mechanical difficulty of repositioning the coils to target the new cortical site. For five subjects, we did not obtain a statistically significant result in the TPJ at the individual level, preventing us from being able to define an 'experimental' site for those subjects. For four subjects, TMS-induced nerve pain or twitches prevented us from testing the anterior 'control' site. Thus, of 18 subjects (after excluding 2 for high false alarm rates), the experimental site was tested in 13 and the control site was tested in 14.

4.2 Results

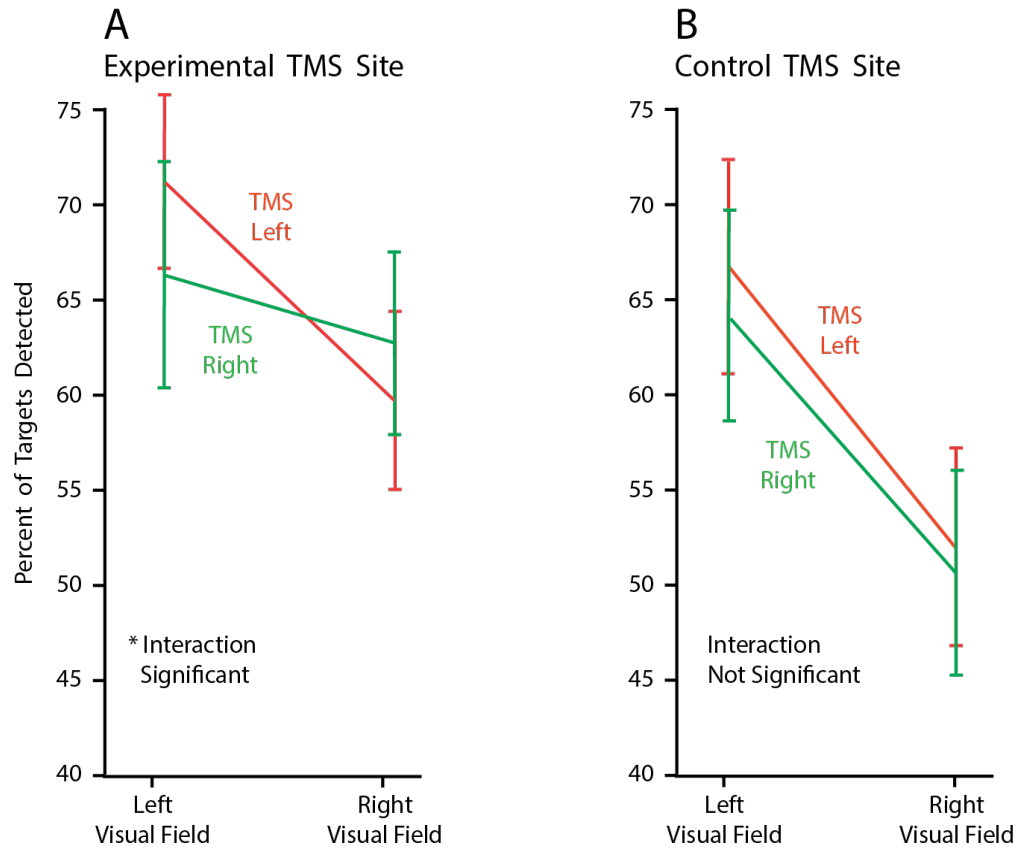


Figure 4.3: Results of the TMS visual detection task. Error bars are standard error of the mean. A. Detection rates for experimental trials, significant interaction between TMS left/right and left/right visual field ($p = 0.03$). B. Detection rates for control trials, interaction not significant ($p = 0.75$).

Figure 4.3A shows the results of TMS to the experimental site (in the dorsal TPJ) for 13 participants. The graph shows visual detection rates (percentage of targets detected) in the left and right visual field, during TMS to the left and right side of the brain. Because single-pulse TMS is a minimal perturbation, a sensitive threshold detection was used to measure performance. Because of the sensitivity of the threshold detection task, a large number of factors can shift baseline performance. It has been reported that visual detection is slightly better in the left visual field (Umiltà et al., 1979); hence, it is not surprising that the left side of the graph is shifted up. The noise of the stimulator can

distract and reduce detection performance; thus, comparison with non-stimulation controls is not useful. For these reasons, main effects are not informative about the hypothesis. Only the interaction term can test the hypothesis. If TMS induces symptoms of visual neglect, then this dataset should show a significant interaction indicating that TMS to the right or left TPJ differentially affected detection in the left or right visual field. A 2×2 repeated-measures ANOVA was applied to this dataset. Crucially, the interaction between the two variables was significant in a direction consistent with TMS reducing visual detection in the contralateral visual field (interaction: $F_{1,12} = 5.76, p = 0.03$; main effect of visual field: $F_{1,12} = 1.06, p = 0.32$; main effect of TMS hemisphere: $F_{1,12} = 0.08, p = 0.78$).

Figure 4.3B shows the results of TMS to the control site (located 2 cm anterior to the experimental site) for 14 participants. For this site, there was no evidence of an interaction between TMS to the left/right and left/right visual field (2×2 repeated-measures ANOVA; interaction: $F_{1,13} = 0.11, p = 0.75$; main effect of visual field: $F_{1,13} = 2.59, p = 0.13$; main effect of hemisphere: $F_{1,13} = 0.52, p = 0.48$), suggesting that the interaction observed in Figure 4.3A was not a non-specific effect of TMS, such as one that might be produced by distraction from the noise of the stimulator.

4.3 Discussion

The results presented in this chapter show that effects resembling hemispatial neglect emerge following TMS to the TPJ. TMS was targeted to the TPJ based on the peak of activation in the ‘attribution-of-attention’ task, a task designed to probe the social attribution of attention. In Chapter 3, the TPJ activity associated with this task was shown to peak in the dorsal TPJ, in a region functionally connected with the

frontoparietal control network. That conclusion is also reflected in the whole-brain analysis presented in Figure 4.1. This suggests that, in addition to being an important node for the social attribution of attention, TPJd may also be an important node for the processes that underlie hemispatial neglect.

Why might there be an overlap between the mechanisms underlying one's own awareness of space and the mechanisms underlying social cognition? According to one view, social cognition involves not only tracking the beliefs, desires and personalities of others, but also involves maintaining a model of their state of attention (Graziano & Kastner, 2011; Graziano, 2013). Such a model is useful, for instance, in predicting their behavior, since attention is a powerful determinant of behavior. Similarly, subjective awareness may depend on constructing a model of one's own state of attention, such that one can explicitly report on the contents of that model. Thus, the attribution of 'awareness' to others and oneself may in some ways depend on common mechanisms.

If this view is correct, in addition to overlap in neural mechanisms, one might also expect to find correlations at the behavioral level between one's own spatial awareness and the attribution of awareness to others. This was shown in a recent behavioral experiment (Bio et al., 2018). Normal participants, without brain damage, often show a subtle spatial bias to the left or right, a property termed 'pseudoneglect'. Bio et al. showed that this subtle spatial bias correlates with subtle reaction time biases in a spatial version of the attribution-of-attention task. Taken together with the TMS results presented in this chapter, these results suggest that attributing awareness to oneself and to others may indeed depend at least partially on common mechanisms, though the exact nature of these common mechanisms remains to be elucidated by future experiments.

One additional question concerns whether hemispatial neglect is a disorder of spatial attention or spatial awareness. As reviewed in previous chapters, attention and awareness are highly correlated processes (Posner, 1994; De Brigard & Prinz, 2010), but are dissociable under certain conditions (McCormick, 1997; Lamme, 2004; Tsushima et al., 2006; Jiang et al., 2006; Koch & Tsuchiya, 2007; Kentridge et al., 2008; Hsieh et al., 2011; Norman et al., 2013; Lin & Murray, 2015; Webb et al., 2016a). It may be, of course, that neglect is a disorder of attention in some cases, a disorder of awareness in other cases, and a mixture of both in still other cases, and there are reasons to suspect that disruptions to awareness should cause disruptions to attention, as detailed in Chapter 2. Nevertheless, it is important to understand how different brain networks contribute to these related, yet distinct processes. The paradigm used in this chapter does not allow for the dissociation between attention and awareness. To answer this question, in the next chapter, the results of an fMRI experiment are presented that dissociate cortical networks involved in subjective awareness and attention.

Chapter 5

Cortical networks involved in visual awareness independent of visual attention

In Chapter 2, interactions between attention and awareness at the behavioral level were investigated. One result of these experiments was the identification of conditions under which attention to a visual stimulus was balanced, despite a difference in awareness for that stimulus. In Chapter 3, a method was introduced for parcellating task-based activity in distinct subregions of the TPJ, and the cortical networks to which they're connected. In this chapter, these approaches are combined. The results from an fMRI experiment investigating the cortical networks involved in visual awareness, independent of attention, are reported (Webb et al., 2016b).

Many studies have found awareness related activity in the frontal and parietal lobes (Lumer et al., 1998; Dehaene et al., 2001; Rees et al., 2002; Carmel et al., 2006; Lau & Passingham 2006). One concern, however, is that much of this broad frontoparietal network is also implicated in the closely related process of attention (Kastner & Ungerleider, 2000; Corbetta & Shulman, 2002), raising concerns about the specificity of these activations (Naghavi & Nyberg, 2005; Lau, 2008; Bor & Seth, 2012; Tallon-Baudry, 2012; Tsuchiya et al., 2015). As reviewed in previous chapters, attention and awareness are dissociable processes (McCormick, 1997; Lamme, 2004; Tsushima et al., 2006; Jiang et al., 2006; Koch & Tsuchiya, 2007; Kentridge et al., 2008; Hsieh et al., 2011; Norman et al., 2013; Lin & Murray, 2015; Webb et al., 2016a), but they covary

under normal conditions (Posner, 1994; De Brigard & Prinz, 2010). Unless care is taken to explicitly control for this confound, significantly more attention may be drawn to a stimulus when subjects are aware of it.

In this chapter, a behavioral paradigm similar to the one presented in Chapter 2 is used, combined with the fMRI parcellation methods introduced in Chapter 3. In this paradigm, based on the Posner cueing paradigm (Posner, 1980), attention is drawn to a spatially localized cue. Awareness for the cue is controlled using metacontrast masking (Breitmeyer & Ögmen, 2006), producing one condition in which participants are relatively aware of the cue, and another condition in which participants are relatively unaware of the cue. Crucially, we measure the amount of attention drawn to the cue and find that, consistent with the results from this paradigm presented in Chapter 2, the overall amount of attention drawn to the cue is not significantly different between these two conditions. The results reveal that two regions of the TPJ, a dorsal region (TPJd) connected to the frontoparietal control network, and an anterior region (TPJa) connected to the ventral attention network, are responsive to awareness, independent of attention.

5.1 Methods

5.1.1 Subjects

Twenty-five participants were tested (14 females, 18–48 years old, normal or corrected-to-normal vision). All participants provided informed consent, and all procedures were approved by the Princeton Institutional Review Board.

5.1.2 Behavioral Paradigm

Stimuli were projected with the Hyperian MRI Digital Projection System (Psychology Software Tools) at the end of the scanner bore. Each subject lay face-up on the scanner bed, with foam surrounding the head to reduce head movements and earplugs to reduce noise. All stimuli were developed and presented with the MATLAB psychophysics toolbox (Brainard, 1997).

Figure 5.1 shows the behavioral paradigm. Each trial began with a white central fixation point on a black background. Participants were instructed to fixate during the trial. After 1 second, the cue period began and lasted three refresh cycles (~ 50 ms). Throughout the cue period, the cue (a white spot 1.1° in diameter) was presented 6° to the left of fixation (1/3 of trials), to the right of fixation (1/3 of trials), or not presented (1/3 of trials).

The cue was followed by a mask. In the long cue/mask interval condition ('aware' condition; 1/2 of trials), a cue/mask interval of three refresh cycles (~ 50 ms) was inserted between the cue and the mask. Thus, the time from cue onset to mask onset was ~ 100 ms, which was intended to allow the cue to be subjectively visible. The mask then remained on the screen for five refresh cycles (~ 80 ms). In the short cue/mask interval condition ('unaware' condition; 1/2 of trials), no time was inserted between the cue and the mask. The time from cue onset to mask onset was thus ~ 50 ms, which was intended to render the cue subjectively invisible. The mask then remained on the screen for eight refresh cycles (~ 130 ms). In both conditions, the time from cue onset to mask offset was 11 refresh cycles (~ 180 ms). The mask consisted of two white metacontrast rings (Breitmeyer & Öğmen, 2006) with an outer diameter of 1.5° and an inner diameter

of 1.1° , presented 6° to the left and right of fixation. The mask therefore did not preferentially attract bottom-up attention to one side.

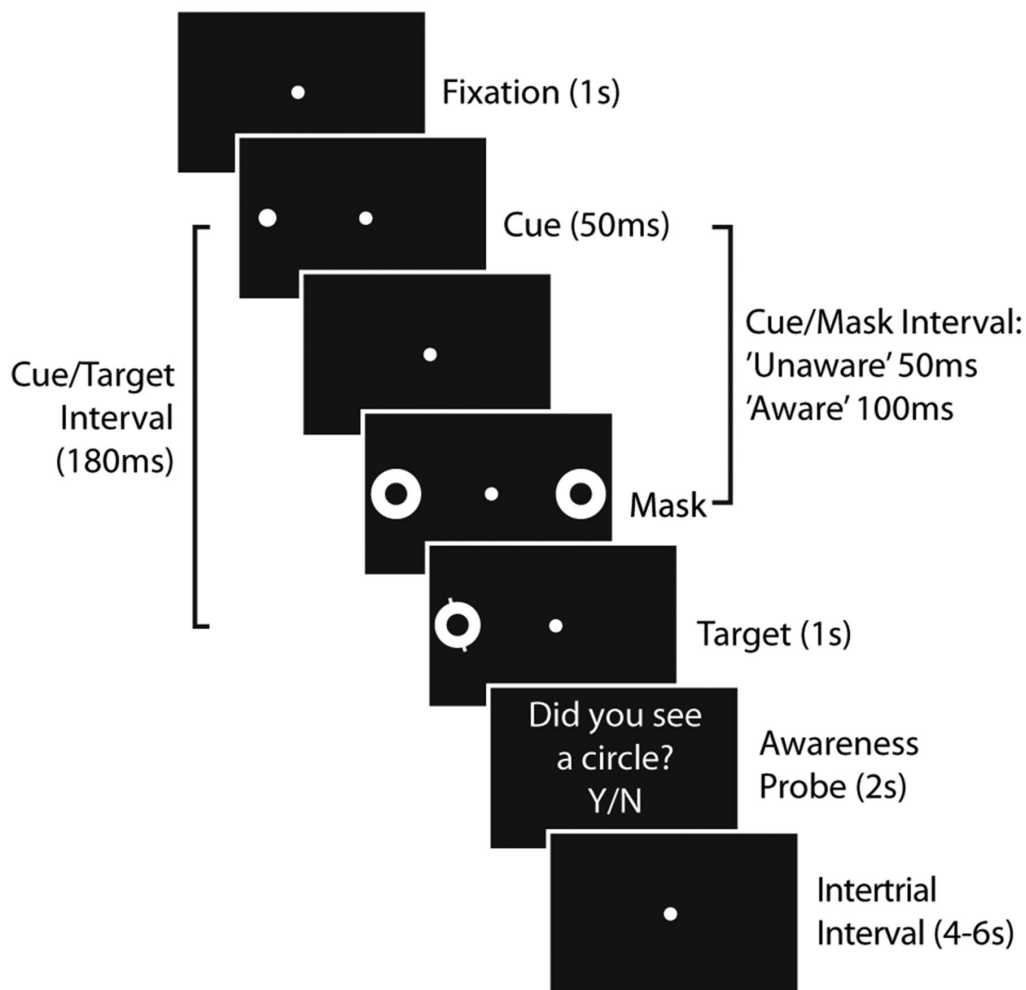


Figure 5.1: Behavioral paradigm for measuring attention to, and manipulating awareness of, spatial cue.

After the mask period, a target was added to one of the already-present circles that composed the mask. The target was either on the left (1/2 of trials) or right (1/2 of trials) of fixation. The location of the cue did not predict the location of the target. The interval between cue onset and target onset was 11 refresh cycles (~ 180 ms). The target consisted of one white line segment extending from the top of the mask ring and one white line segment extending from the bottom of the mask ring. The segments were

collinear, forming an implied line through the ring. The line was tilted toward the left (1/2 of trials) or toward the right (1/2 of trials) by 3°. Participants were required to discriminate the orientation of the line, using a button box with the right hand. The target lasted 1 second, and participants were instructed to respond as quickly as possible during that period. The 20% of trials with incorrect or no response during the 1-second response window were excluded from the behavioral analysis. Trials with a latency <300 ms were also excluded from the behavioral analysis because the reaction time was too short to be a plausible response to the discrimination task.

On each trial, after the 1-second target period, a question appeared on the screen: “Did you see a circle? Y/N.” Participants were instructed that this question referred to the cue stimulus presented at the beginning of some trials. The question remained on the screen for 2 seconds, and participants were instructed to respond during that time, using a button box with the right hand.

After the question period, a 4–6-second intertrial interval began, during which the fixation point was present. Participants were instructed to maintain fixation during the intertrial interval.

All trial types were randomly interleaved. Each participant performed practice trials followed by nine runs of 24 trials each (216 trials total). Target orientation (tilted left or right), target location (left or right), cue condition (left, right, or no cue), and cue/mask interval (long or short) were randomized and counter-balanced within each run.

5.1.3 MRI Data Collection and Preprocessing

MRI images covering the whole cortex were acquired with a 20-channel receiver head coil on a Siemens Skyra scanner. Functional imaging used a gradient echo, echoplanar

pulse sequence with a 64×64 matrix [35 axial slices, 3 mm thick, field of view (FOV), 192×192 mm; repetition time (TR), 2 seconds; echo time (TE), 30 ms; flip angle (FA), 77° ; in-plane resolution, 3×3 mm]. Functional images were aligned with a high-resolution anatomical scan (MPRAGE) taken at the end of the session (FOV, 256×224 mm; TR, 2.3 seconds; TE, 2.98 ms; FA, 9° ; 256×224 matrix; 1 mm^3 resolution).

Preprocessing was done with AFNI (Cox, 1996) and FSL (Jenkinson et al., 2012) software packages. The functional data were slice time corrected, motion corrected (to the image acquired closest in time to the anatomical scan), and detrended (linear and quadratic) with AFNI. Single-session ICA was applied to each subject's unsmoothed functional data, using the MELODIC toolbox in FSL (Beckmann & Smith, 2004), and components that represented noise were regressed out using the FSL tool `fsl_regfilt` (Beckmann & Smith, 2004; Kelly et al., 2010). The following spatial or temporal features were considered to represent noise: 1, spatial association with white matter, ventricles, or background voxels; 2, a lack of cluster formation; 3, large spikes in the time course; 4, high-frequency noise; or 5, temporal sawtooth patterns likely to reflect aliasing of cardiac or respiratory signals exceeding the Nyquist frequency (Kelly et al., 2010). To ensure that all neural activity would remain untouched for local ICA at the group level, ICs that appeared to contain a mixture of noise and signal were not filtered out in the denoising step. The denoised data were spatially smoothed with a Gaussian kernel (full-width half-maximum 4 mm), using AFNI.

5.1.4 GLM Analysis of fMRI Data

Statistical analyses were performed using multiple regression within the framework of the GLM (Friston et al., 1994) with AFNI. An event-related design was used in which cue onset was the relevant event and differing responses to aware vs. unaware cues were

analyzed. Separating the response to the cue onset from the response to other elements of the trial such as fixation onset or target onset was not necessary in this design, as these other trial elements were balanced across the aware and unaware conditions. The blood oxygenation level dependent response was modeled by convolving the stimulus timing (cue onset) with a gamma function. Regressors of non-interest were included to account for head motion and linear drift in scanner signal. These regressions produced statistical maps at the individual subject level for two conditions of interest: long cue/mask interval trials on which a cue was presented (aware condition), and short cue/mask trials on which a cue was presented (unaware condition). Subject-level statistical maps were spatially normalized to FSL's Montreal Neurological Institute (MNI)-152 template, using AFNI (Talairach & Tournoux, 1988). A mixed-effects analysis was then computed at the group level (awareness condition as fixed effect, subject as random effect), using AFNI (Chen et al., 2012), producing the contrast [aware condition – unaware condition]. Coefficients for this contrast were thresholded using a Monte Carlo simulation to achieve a corrected significance of $p < 0.01$ adjusted for a cluster size of 5 adjacent voxels.

5.1.5 Local ICA of TPJ Activity

The fMRI data from all subjects were temporally concatenated and then subjected to probabilistic ICA, using the MELODIC toolbox in FSL (Beckmann & Smith, 2004). ICA decomposition was applied to the voxels within a region of interest mask containing the TPJ and the intraparietal sulcus. The mask was constructed from the standard surface `cvs_avg35_inMNI152` in Freesurfer, using `mri_label2vol` to combine multiple labels from the `aparc.a2009s` atlas into one mask (`G_pariet_inf-Supramar`, `G_pariet_inf-Angular`, `G_temp_sup-Plan_tempo`, `G_temp_sup-Lateral`, `G_temp_sup-G_T_transv`,

S_interm_prim-Jensen, S_temporal_sup, S_temporal_transverse, S_intrapariet_and_P_trans) and trimming temporal cortex voxels anterior to the postcentral sulcus. The fMRI data were decomposed into 20 ICs. ICs were thresholded at $Z = 4$ for the creation of winner-take-all maps for the figures.

The time courses from all 20 ICs were separated into time courses for each individual subject and entered into a multiple regression analysis, using AFNI. This analysis was similar to the voxel-wise GLM analysis described in the previous section, except that instead of using the time courses of voxels as dependent variables, the time courses of ICs were used. This analysis produced coefficients for the aware and unaware conditions, one for each of 25 subjects, for each of 20 ICs. These coefficients were then entered into a mixed-effects analysis at the group level (awareness condition as fixed effect, subject as random effect), using MATLAB (MathWorks).

5.1.5 Functional Connectivity Analysis

Functional connectivity analysis was performed to identify the brain-wide networks to which each IC was connected. IC time courses were used as the independent variables in multiple regression analyses at the individual subject level, using AFNI. These regressions produced brain-wide statistical maps for each of 25 subjects. Subject-level statistical maps were spatially normalized to FSL's MNI-152 template, using AFNI (Talairach & Tournoux, 1988), and entered in a mixed-effects analysis at the group level (subject as random effect) using AFNI (Chen et al., 2012), producing group-level statistical maps. These maps were thresholded using a Monte Carlo simulation to achieve a corrected significance of $p < 0.01$ adjusted for a cluster size of 5 adjacent voxels.

5.2 Results

5.2.1 Behavioral Results

Participants performed the target discrimination task (orientation of tilted line) with a mean accuracy of 84% (SD = 12%) and a mean latency of 712 ms (SD = 124 ms). The responses on catch trials (one third of trials on which no cue was presented) were analyzed to determine the false-positive rate, the rate at which participants indicated that a cue was present when none was. This rate was low (mean = 18%, SD = 14%), suggesting participants were generally not guessing about the presence of the cue. Responses on the remaining two-thirds of trials in which a cue was presented were analyzed to determine the true-positive rate, the rate at which participants correctly indicated a cue was present. The true-positive rate was compared with the false-positive rate to calculate d' , a measure of sensitivity to the presence or absence of the cue. The results showed that participants were not aware of the cue on the majority of trials with a 50-ms cue/mask interval (27% true-positive rate; SD = 15%; $d' = 0.29$), and were aware of the cue on the majority of trials with a 100-ms cue/mask interval (80% true-positive rate; SD = 13%; $d' = 2.01$). A paired t-test confirmed a highly significant difference between the two cue/mask interval conditions (paired t-test, $t_{24} = 13.7$; $p = 7 \times 10^{-13}$). The metacontrast masking manipulation thus successfully separated most trials into “aware” and “unaware.”

A standard measure of attention drawn by the cue is $\Delta t = [\text{latency when target and cue are misaligned} - \text{latency when target and cue are aligned}]$. To the extent that this measure is significantly above zero, the cue can be said to have pulled spatial attention to one side at the expense of the other side. This was the case for both aware

and unaware trials (aware condition: $\Delta t = 25$ ms; $SD = 28$; one-sample t-test, $t_{24} = 4.5$; $p = 0.0001$; unaware condition: $\Delta t = 17$ ms, $SD = 31$; one-sample t-test, $t_{24} = 2.8$; $p = 0.01$). These two measures were not significantly different from each other (planned comparison, paired t-test, $t_{24} = 1.3$; $p = 0.21$), indicating that the cue did not draw significantly more attention in the aware condition than in the unaware condition. There was also no significant difference in discrimination accuracy between the aware and unaware conditions (planned comparison, paired t-test, $t_{24} = 0.65$; $p = 0.52$), indicating that the two conditions were balanced for task difficulty.

5.2.2 Whole-Brain Analysis

Figure 5.2A shows the result of the contrast [aware condition – unaware condition], at the whole-brain level. The largest area of activity was in the bilateral TPJ. Smaller clusters of activity were scattered through the left prefrontal cortex, cingulate cortex, occipitotemporal cortex, left insula, and left precuneus. Some of these clusters extend into sulci and are thus partially occluded in the surface view. To see the extent of these clusters, Figure 5.3 shows the results of the contrast [aware condition – unaware condition] with coronal slice views of each cluster. Table 5.1 gives coordinates for the peaks of activity.

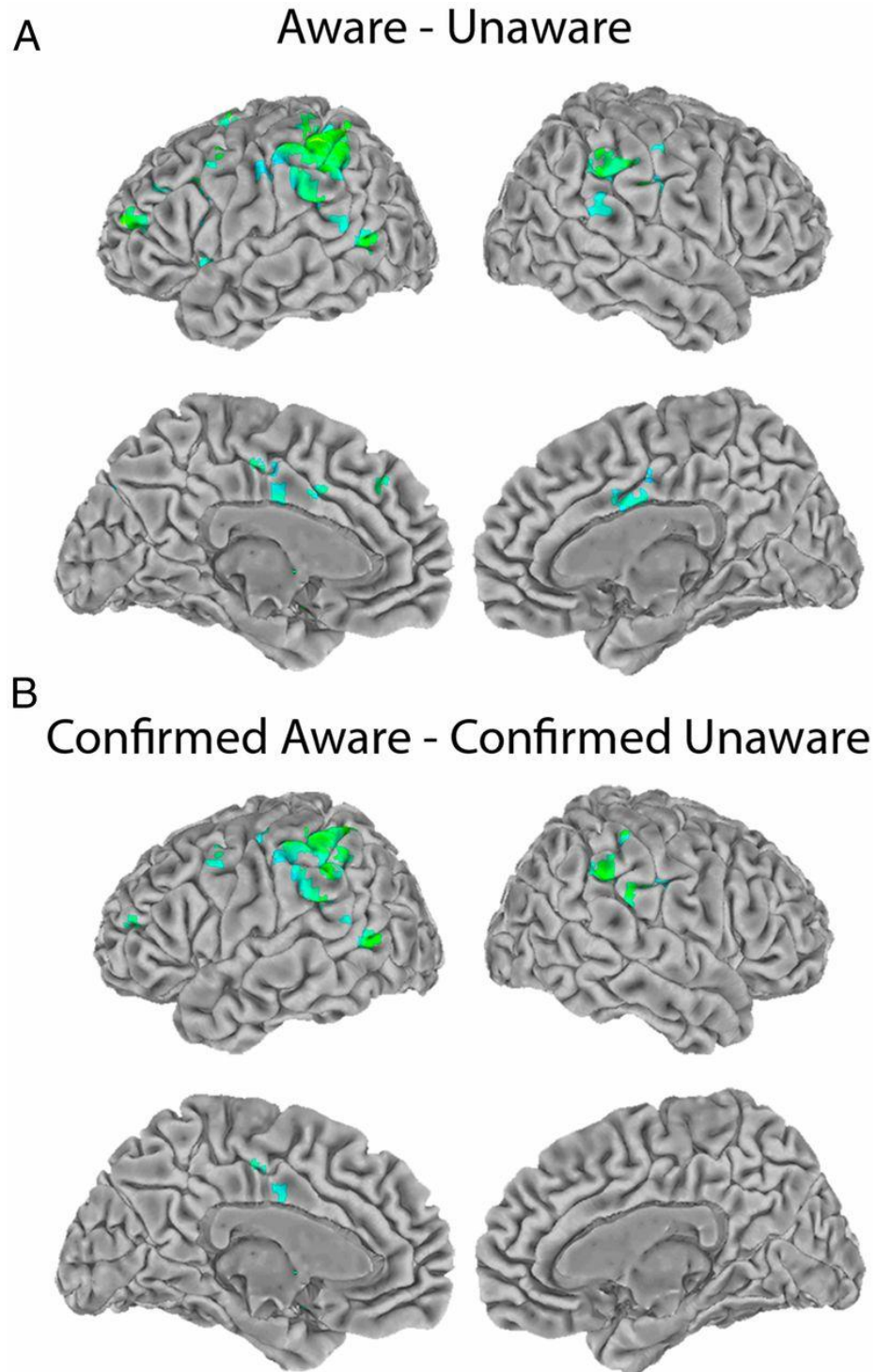


Figure 5.2: Whole-brain group fMRI results. A. Results of [aware condition - unaware condition] contrast. B. Results of [confirmed aware condition - confirmed aware condition] contrast.

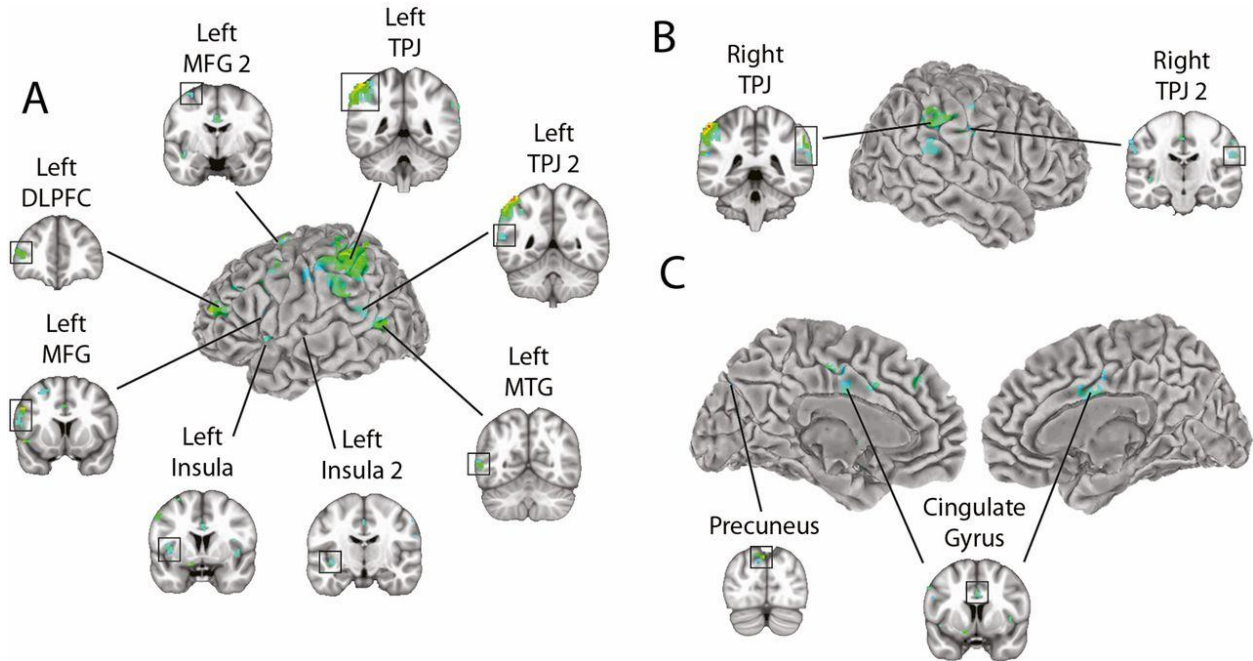


Figure 5.3: Coronal slice views for the [aware condition - unaware condition] contrast. A. Left lateral surface. B. Right lateral surface. C. Medial surfaces.

One potential concern with these results is that the metacontrast masking manipulation did not perfectly separate trials into “aware” and “unaware” conditions. To address this concern, we performed a separate analysis in which the dataset was restricted to the 80% of long cue/mask interval trials on which participants reported being aware of the cue, and the 73% of short cue/mask interval trials on which participants reported being unaware of the cue. We refer to these as “confirmed aware” and “confirmed unaware” trials. Figure 5.2B shows the result of this more restrictive analysis [confirmed aware trials – confirmed unaware trials]. The results closely resemble those shown in Figure 5.2A, but with fewer scattered areas of activity outside the TPJ.

Area	Volume	Peak		
		LR	PA	IS
Left TPJ	390	-55	-46	57
Cingulate gyrus	46	-1	8	40
Left DLPFC	35	-46	49	20
Left MFG	30	-49	11	37
Right TPJ	29	66	-32	41
Left MTG	21	-58	-65	0
Right TPJ	21	63	-29	25
Left insula	15	-40	3	7
Left precuneus	13	-4	-85	42
Left insula	12	-40	-15	-7
Left TPJ	11	-64	-53	14
Left MFG	10	-28	-3	72

Table 5.1: MNI coordinates for areas of activation in Figures 5.2A and 5.3

5.2.3 Local ICA Results

The activity shown in Figure 5.2 covers a broad region of the TPJ. To fractionate the TPJ into more specific components, we used a local independent component analysis (ICA) analysis, similar to the approach taken in Chapter 3. In this approach, the TPJ can be reliably parcellated into distinct spatiotemporal components, each connected to distinct, brain-wide networks (Igelström et al., 2015; Igelström et al., 2016), and task-related activity in these components can be separately analyzed.

Using the fMRI data from the present experiment, we performed a local ICA within a cortical mask that included the TPJ, a margin of surrounding cortex, and cortical areas dorsal to the TPJ that include the intraparietal sulcus. The local ICA decomposed the fMRI data into 20 independent components (ICs) within the mask, each IC defined by its pattern of activity over time. The ICs were further analyzed to determine which ICs were significantly more active in the aware condition than in the unaware condition. Essentially, the ICs were treated like voxels in a standard GLM analysis. Because 20 ICs were tested in this manner, the analysis was Bonferroni corrected for 20 comparisons.

The analysis revealed that three TPJ components responded significantly more to the aware condition than to the unaware condition. These components included a left-lateralized, dorsal IC ($F_{1,24} = 39.2$; $p < 0.05$ Bonferroni corrected for 20 comparisons) consistent with the location of left TPJd (dorsal TPJ); a right-lateralized, dorsal IC ($F_{1,24} = 11.4$; $p < 0.05$ Bonferroni corrected for 20 comparisons) consistent with the location of right TPJd, and a bilateral, anterior IC ($F_{1,24} = 11.5$; $p < 0.05$ Bonferroni corrected for 20 comparisons) consistent with the location of TPJa (anterior TPJ). No other ICs responded significantly to this comparison.

Winner-take-all spatial maps for the significant ICs are shown in Figure 5.4A. The right TPJd component has some bilateral representation. It includes a large area on the right and a smaller area on the left. The left TPJd is confined to the left hemisphere. The TPJa is approximately equally represented in both hemispheres. Regression coefficients for the aware and unaware conditions for each component are shown in Figure 5.4B. Figure 5.5 shows the time courses for these components.

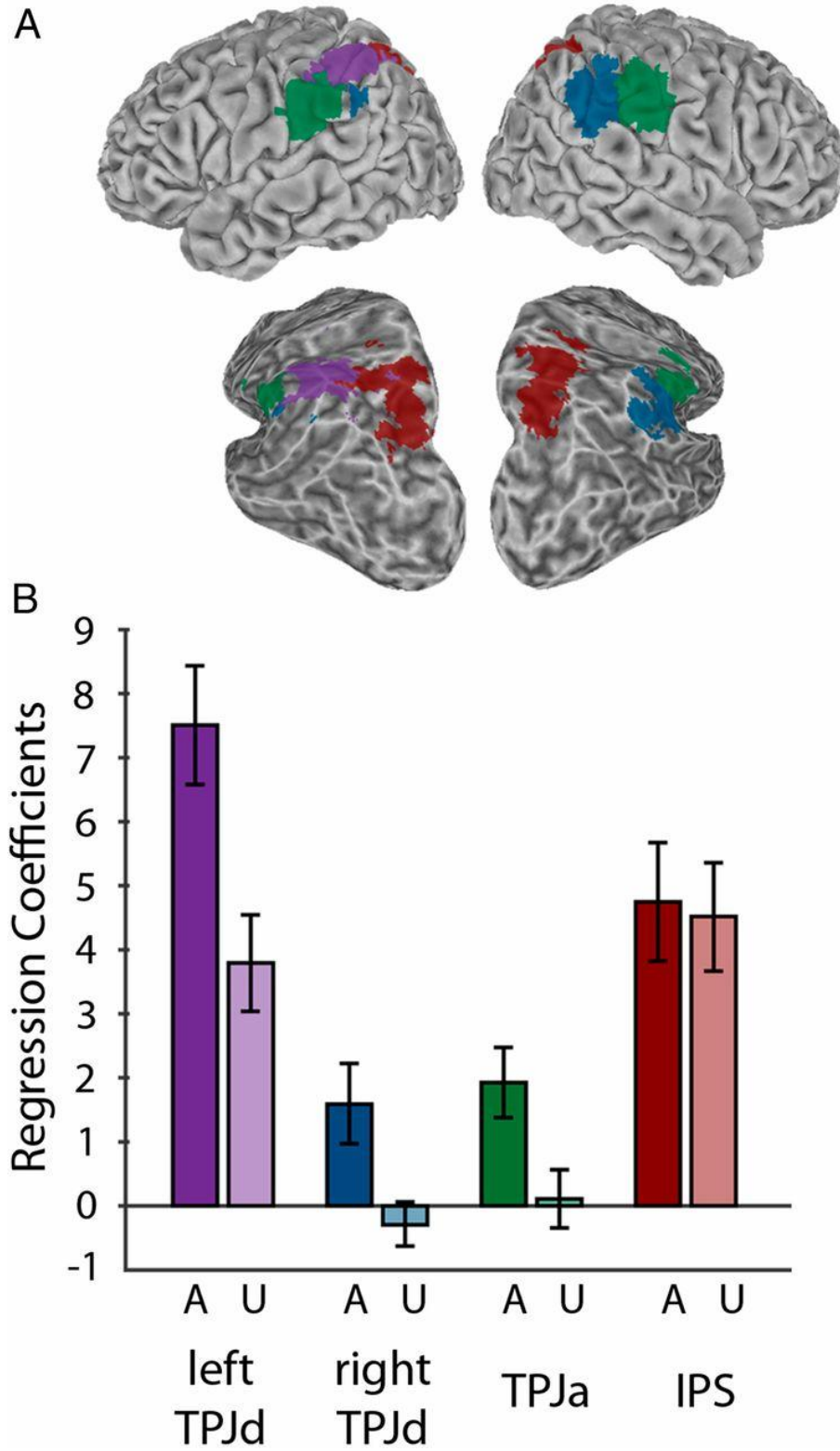


Figure 5.4: Local ICA results for the [aware condition (A) - unaware condition (U)] contrast. A. Winner-take-all maps for the significant ICs. B. Regressions coefficients.

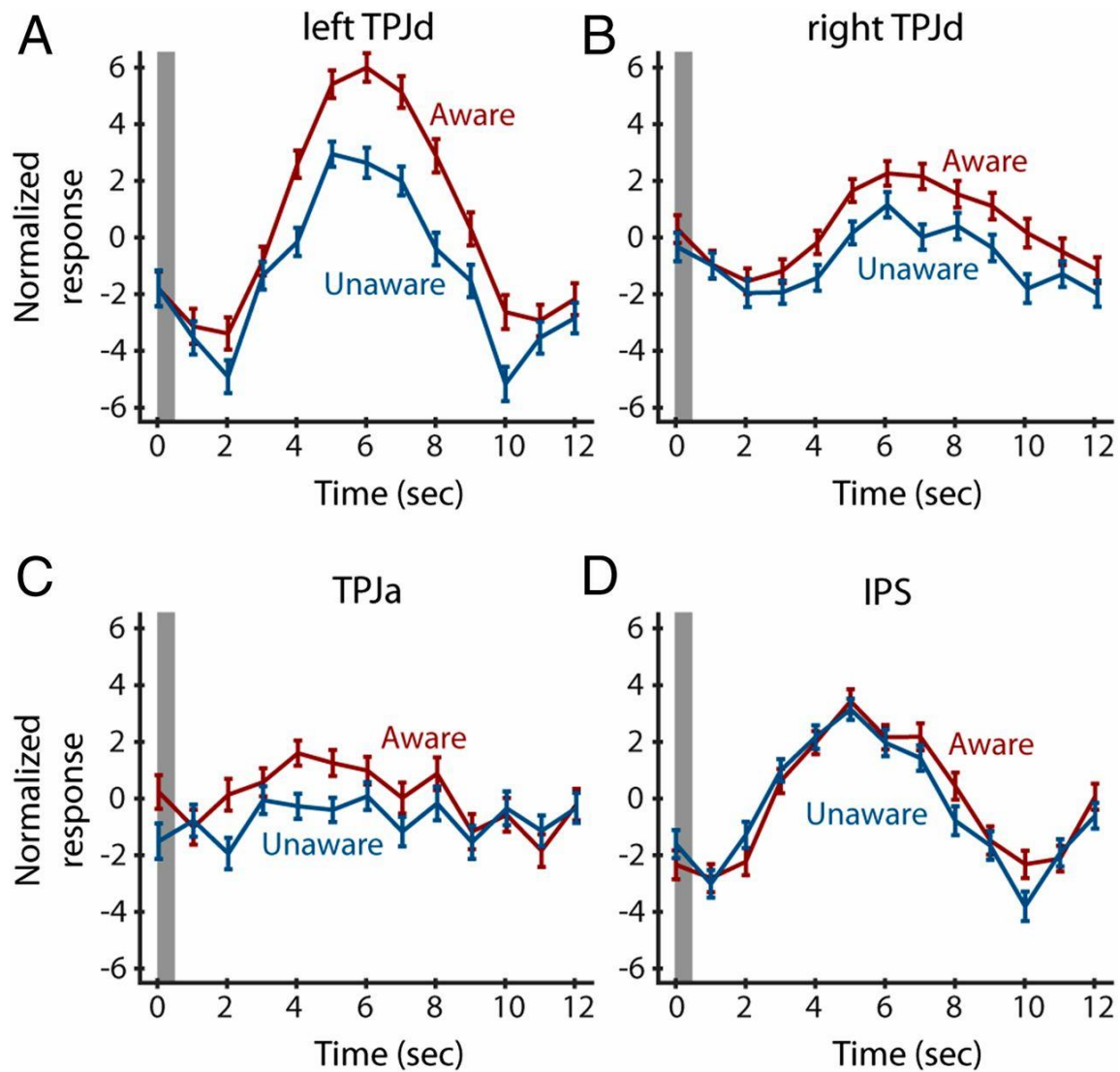


Figure 5.5: Time courses for the aware and unaware conditions in specific ICs. A. Left TPJd. B. Right TPJd. C. TPJa. D. IPS (not significant).

5.2.4 Functional Connectivity Results

We performed a functional connectivity analysis (Figure 5.6) on the three TPJ components identified in the ICA-based regression. The bilateral TPJa component was functionally connected to the bilateral anterior insula, bilateral precuneus, bilateral cingulate gyrus, bilateral inferior frontal gyrus, and bilateral middle temporal gyrus, a pattern consistent with the connectivity of TPJa in our previous studies (Igelström et al.,

2015; Igelström et al., 2016) and partially overlapping the known connectivity of the ventral attention network (Corbetta & Shulman, 2002), although the ventral attention network is typically right biased. It also shared similarities with a previously described “salience” network and a “cingulo-opercular” network (Dosenbach et al., 2007; Seeley et al., 2007). The left TPJd component was functionally connected to the left dorsolateral prefrontal cortex, left medial prefrontal cortex, left precuneus, left middle temporal gyrus, and left insula, with smaller corresponding clusters on the right side, matching the pattern of connectivity obtained in our previous studies of the TPJd (Igelström et al., 2015; Igelström et al., 2016) and partially overlapping the known connectivity of the frontoparietal control network (Vincent et al., 2008). The right TPJd component was functionally connected to the right dorsolateral prefrontal cortex, right medial prefrontal cortex, right inferior frontal gyrus, right precuneus, and right insula, with smaller corresponding clusters on the contralateral side. This pattern showed some similarity to the frontoparietal control network and to the ventral attention network.

We also performed a quantitative comparison between these results and known cortical networks. We computed the spatial overlap of our functional connectivity results with network templates obtained from a previously reported seven-network parcellation of resting-state fMRI data (Yeo et al., 2011). Specifically, we compared our results with three templates identified as the frontoparietal control network, the dorsal attention network, and the ventral attention network (Yeo et al., 2011).

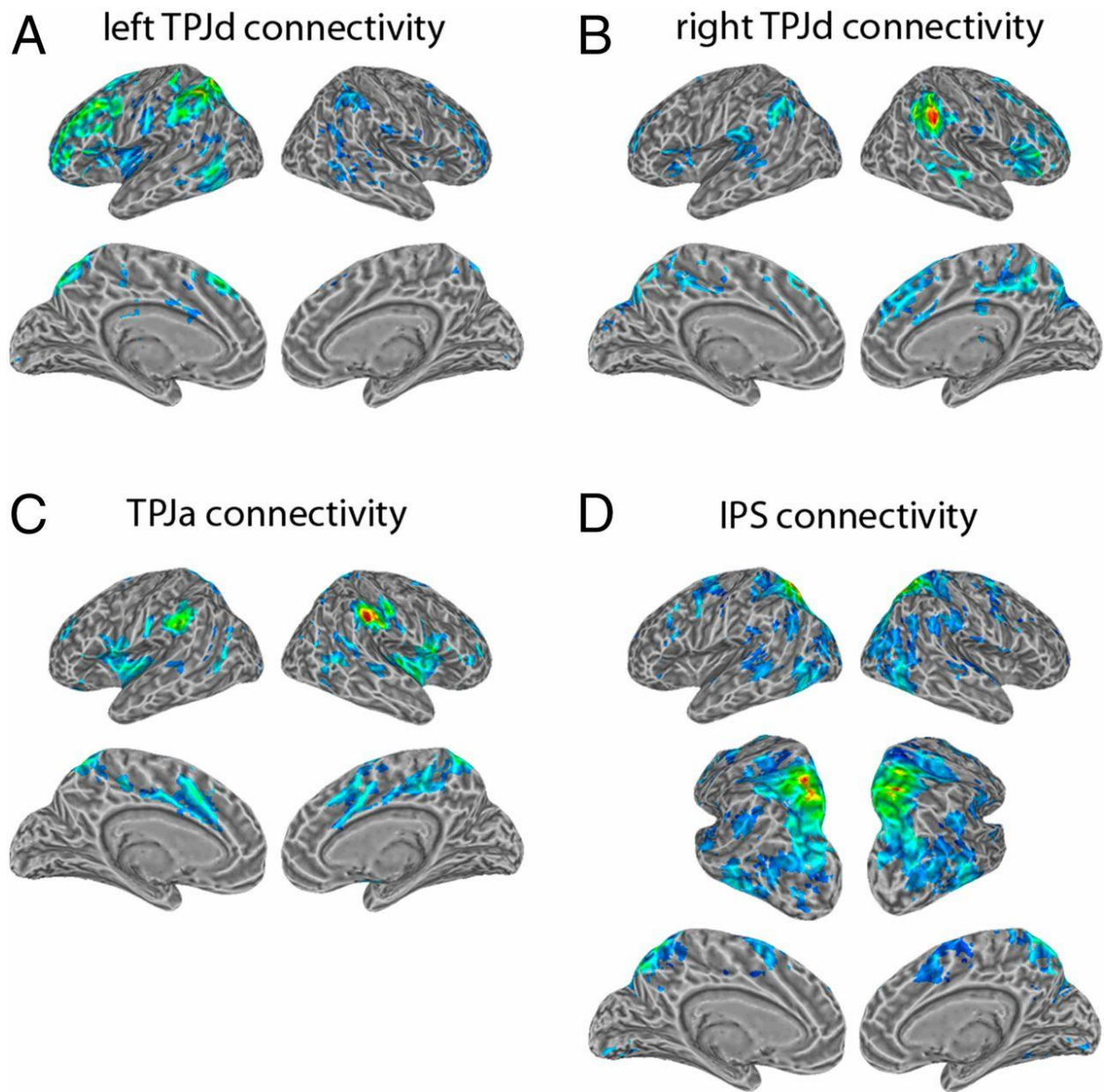


Figure 5.6: Functional connectivity results for specific ICs. A. Left TPJd. B. Right TPJd. C. TPJa. D. IPS (not significant).

Figure 5.7 shows the results of this comparison. The network seeded by the TPJa component showed significant overlap with the ventral attention network template. The network seeded by the left TPJd component overlapped mostly with the frontoparietal control network template. The network seeded by the right TPJd component overlapped partially with both the frontoparietal control network template and the ventral attention

network template, although the overlap was stronger with the frontoparietal control network template.

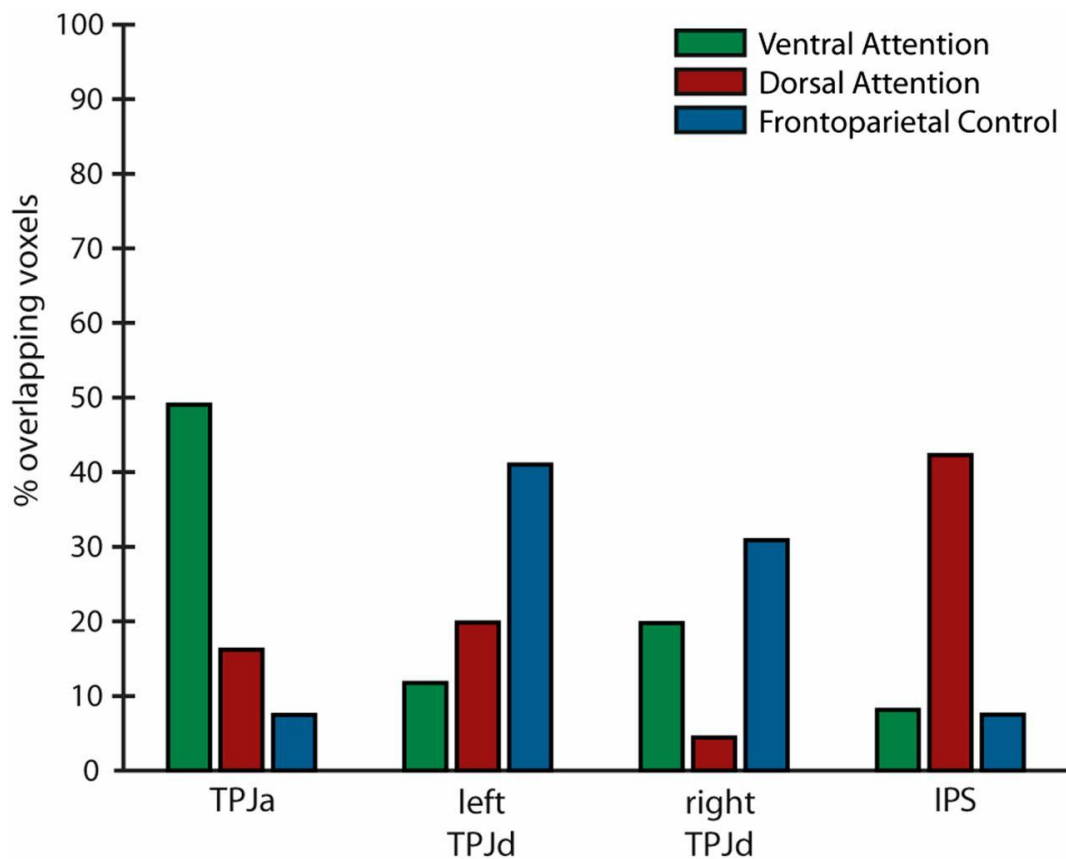


Figure 5.7: Quantitative comparison between functional connectivity results and known cortical networks.

It should be noted that the network templates used for this comparison were derived from data-driven parcellation of resting-state data, and therefore may not correspond exactly to the same networks when derived using task-based activations. The ventral attention network, for instance, has been reported as a right-lateralized network when defined using attention tasks (Corbetta & Shulman, 2002), whereas the ventral attention network template used here is bilateral, and may reflect a combination of the ventral attention network and the “salience” or “cingulo-opercular” network (Dosenbach et al., 2007; Seeley, et al., 2007). Future work is needed to directly compare the network

activity in our task with task-based definitions of known cortical networks, but the present analysis suggests that areas in the temporoparietal cortex, those connected with the ventral attention and frontoparietal control networks, play an important role in subjective awareness, whereas more dorsal regions, connected with the dorsal attention network, do not.

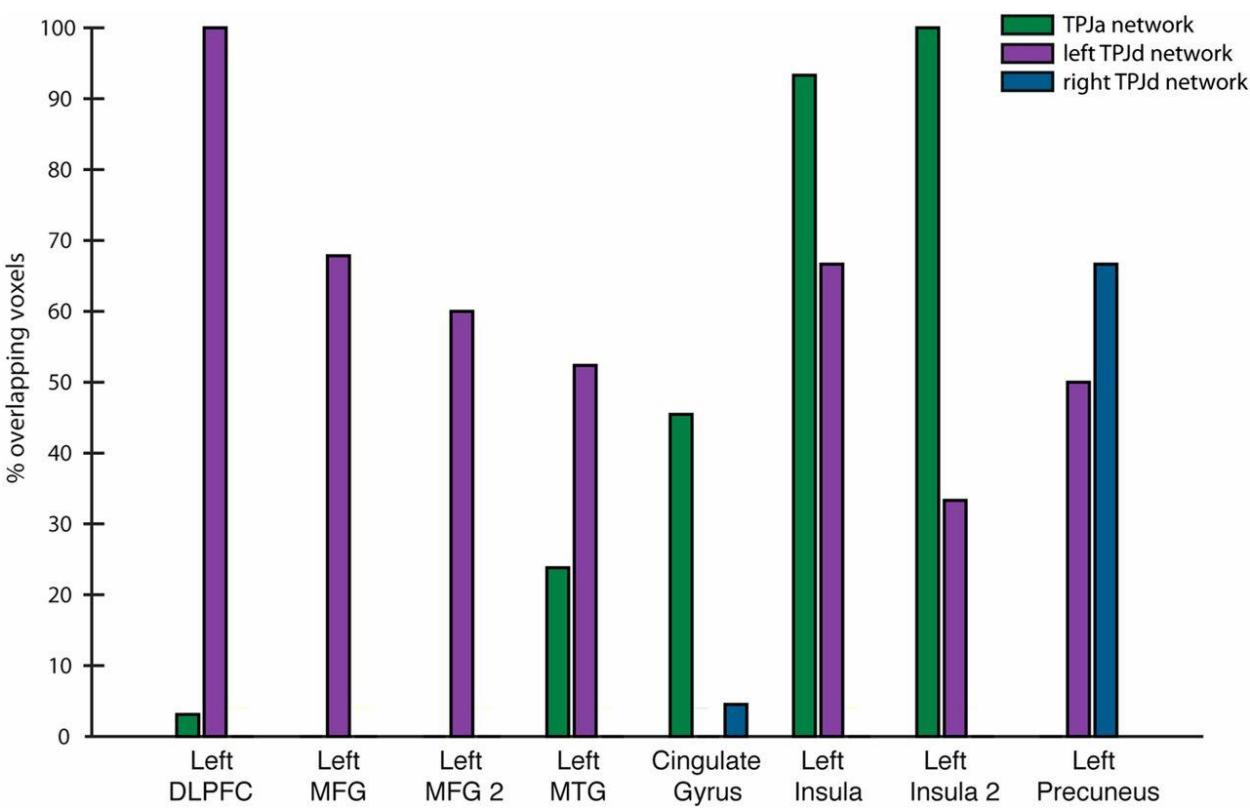


Figure 5.8: Overlap of functional connectivity results with task-related activity outside of TPJ.

We also computed the spatial overlap between the functional connectivity results shown in Figure 5.6 and the task-related activity observed outside the TPJ shown in Figures 5.2 and 5.3 and reported in Table 5.1.

The results of this comparison are shown in Figure 5.8. The clusters in the frontal lobe, including the left DLPFC and two clusters in the MFG and the cluster in the left MTG, mostly overlapped the network seeded by the left TPJd. The clusters in the

cingulate gyrus and left insula mostly overlapped the network seeded by the TPJa. The cluster in the precuneus overlapped with the networks seeded by the TPJa and the right TPJd.

5.2.5 Involvement of the Dorsal Attention Network

We examined the functional connectivity pattern for each of the 20 components identified by the ICA and found one component that showed connectivity matching the dorsal attention network. As expected, this component was located bilaterally, extending over both banks of the intraparietal sulcus (IPS). It was functionally connected to the bilateral superior parietal lobule, bilateral frontal eye fields, bilateral supplementary eye fields, and bilateral middle and inferior temporal gyri (Figure 5.6D), closely matching the dorsal attention network, as identified by previous task-based studies of attention networks (Kastner & Ungerleider, 2000; Corbetta & Shulman, 2002) and previous parcellations of cortical networks (Fox et al., 2006; Yeo et al., 2011). The IPS component also had a high degree of overlap with the dorsal attention network as quantified in Figure 5.7. The winner-take-all spatial map for this component is shown in Figure 5.4A.

This intraparietal component did not respond significantly more in the aware condition than in the unaware condition ($F_{1,24} = 0.2$; $p > 0.05$ Bonferroni corrected for 20 comparisons). The bar graphs in Figure 5.4B show regression coefficients for the aware and unaware conditions for the left TPJd, right TPJd, TPJa, and the intraparietal component. The awareness manipulation had an effect on the activity in the three TPJ components, but not in the intraparietal component.

Figure 5.5 shows the time courses for all four components. Again, the awareness manipulation affected activity in the TPJ components, but not the intraparietal component.

5.2.6 Further Consideration of Potential Attention Confound

A central goal of this study was to disambiguate the roles of attention and awareness in cortical networks. This consideration is important, given the substantial overlap between attention and awareness under normal conditions, both in behavior and in the involvement of cortical networks. Our task was designed specifically to address this potential confound by using conditions under which aware stimuli and unaware stimuli have balanced effects on attention. As expected, we did not observe a significant difference in our measure of attention (Δt = [response latency when target and cue are misaligned – response latency when target and cue are aligned]) between the aware condition and the unaware condition (planned comparison, paired t-test, $t_{24} = 1.3$; $p = 0.21$).

However, one concern might be that the data lack the necessary statistical power to identify a difference between the unaware and aware conditions. Perhaps a subtle difference in attention was present, but was too small for us to measure. The experimental results showed more activity in the TPJ during aware trials than during unaware trials (aware condition – unaware condition). If on average the subjects were paying subtly more attention to the cue in the aware trials than in the unaware trials, the difference in MRI signal might be explained by a difference in attention, rather than by a difference in awareness. To address this possibility, we performed an additional analysis that took advantage of the differences among subjects in performance on the behavioral task.

For some subjects, the measured amount of attention (Δt) was greater for the aware trials than for the unaware trials. For other subjects, the measured amount of attention was greater for the unaware trials than for the aware trials. We performed a

correlation analysis across subjects, comparing their behavioral performance with their MRI signal. For a behavioral measure, we used the difference in attention between the aware and unaware conditions (Δt for the aware condition – Δt for the unaware condition). For the MRI measure, we used the difference in activity (aware condition – unaware condition) in each of the three TPJ components that were significantly activated by the task. We found no significant correlations between the behavioral performance and the MRI signal (TPJa, $r = 0.15$, $p = 0.49$; left TPJd, $r = 0.29$, $p = 0.16$; right TPJd, $r = -0.14$, $p = 0.52$). Participants who paid more attention to the cue in the aware condition did not have a greater MRI signal in the TPJ in the aware condition. The observed activity in the TPJ is thus unlikely to have resulted from a subtle difference in attention between the aware and unaware conditions.

5.3 Discussion

Previous studies suggested that a broad frontoparietal network may be involved in subjective awareness (Lumer et al., 1998; Dehaene et al., 2001; Rees et al., 2002; Carmel et al., 2006; Lau & Passingham 2006). In this study, we asked which parts of this network participate specifically in subjective awareness, and which parts are involved in the closely related process of attention. Because the unaware and aware stimuli in our task had balanced effects on attention, the present experiment provides an opportunity to disambiguate these processes. The results suggest that more ventral regions of this network, especially in the TPJ, may participate in subjective awareness. More dorsal regions in the banks of the intraparietal sulcus did not respond in association with awareness, despite their established involvement in attention.

When local ICA (Igelström et al., 2015; Igelström et al., 2016) was used to specify subcomponents of the TPJ, three specific subcomponents showed greater activity in the aware than in the unaware conditions. These subcomponents included left TPJd, right TPJd, and a bilateral component in TPJa. Their functional connectivity partly resembled the previously reported connectivity of the frontoparietal control network and the ventral attention network (Corbetta & Shulman, 2002; Vincent et al., 2008). The functional connectivity of these three TPJ components overlapped with the smaller foci of awareness-related activity obtained in our GLM analysis throughout the prefrontal, cingulate, and temporal cortex (Figures 5.2, 5.3, and 5.8). Thus, although the largest area of activity obtained in the present study was in the TPJ, it was likely acting in the context of larger cortical networks.

In contrast, a component in the banks of the intraparietal sulcus did not show significantly greater activity in the aware than in the unaware conditions. This component was functionally connected with the dorsal attention network identified by previous studies (Kastner & Ungerleider, 2000; Corbetta & Shulman, 2002; Fox et al., 2006; Yeo et al., 2011).

These results are consistent with the proposal that subjective awareness is associated with the TPJ (Graziano & Kastner, 2011). In that proposal, awareness serves as an internal model to help the brain monitor its own state of attention. Awareness and attention therefore normally work together, and interactions between the two are to be expected, as reported in Chapter 2. However, when they are separated, as in the present task, then awareness should be more closely associated with the TPJ than with the dorsal attention network.

These results also have implications for hemispatial neglect. The TPJ is a major lesion site associated with neglect (Vallar & Perani, 1986, Corbetta & Shulman, 2011), and the results presented in Chapter 4 show that neglect-like effects can result from TMS to the TPJ. Neglect has not typically been probed with standard paradigms for distinguishing attention from awareness, but the present results suggest it may be worth testing for a dissociation between awareness-related and attention-related neglect caused by damage to different, adjacent networks.

Chapter 6

Conclusion

How does the brain become aware of its own processes, and what functional purpose might this serve? The experiments presented in this dissertation contribute to our emerging understanding of these issues in two ways. The experiments presented in Chapter 2 suggest that awareness plays a functional role in helping to control attention, perhaps analogous to the role the body schema plays in motor control. The experiments presented in Chapters 3, 4, and 5 suggest that a dorsal region of the TPJ (TPJd) plays a domain general role in a diverse range of processes, including attention, social cognition, and episodic memory, and that this region is also an important node in the cortical networks that underlie awareness.

The experiments in Chapter 2 tested a specific idea regarding the functional relationship between attention and awareness (Graziano & Webb, 2014; Webb & Graziano, 2015). The ‘attention schema theory’ proposes that the brain computes a simple, schematized model of its own ongoing processes, and uses this model to help control attention. This idea is explicitly inspired by the notion of a ‘body schema’, a schematic model of the body that plays an important role in motor control. The experiments in Chapter 2 showed that attention to a visual stimulus is more susceptible to bottom-up influences, and less well-controlled in line with task demands, when participants are unaware of that stimulus (Webb, 2016a). These results imply that

awareness plays a role in the control of attention, and that, therefore, ‘epiphenomenalism’ - the persistent idea that awareness has no function - is incorrect.

The exact nature of the functional relationship between awareness and attention, however, remains to be elucidated. Specifically, other theories propose the existence of such a relationship, such as the global workspace theory (Baars, 1997; Dehaene & Naccache, 2001). Future experiments should focus on dissociating these distinct accounts. For example, the attention schema theory predicts that awareness should play a role in ‘model-based’ control of attention, but not ‘model-free’ control of attention (Webb & Graziano, 2015), whereas the global workspace theory proposes a more general role in the control of attention (Dehaene & Naccache, 2001). One way forward may be to focus on signatures of model-based vs. model-free control, as has been done in the literature on motor control (Haith & Krakauer, 2013).

The experiments in Chapter 3 showed that distinct subcomponents of the TPJ play domain-specific roles in attention, social cognition, episodic memory, and target detection, while TPJd plays a more domain-general role (Igelström et al., 2016). One appealing idea is that TPJd, and the frontoparietal network to which it is connected, serves as a ‘global hub’, integrating diverse information processing streams, and helping to coordinate the activity of distinct cortical networks (Cole et al., 2013). The experiments in Chapter 4 and 5 showed that TPJd also plays an important role in awareness. Together, these findings are consistent with a broad range of proposals that awareness has a special link to integration (Baars, 1997; Dehaene & Naccache, 2001; Tononi, 2008; Graziano, 2014).

The experiment presented in Chapter 4 showed that TMS to TPJd caused a transient impairment of visual detection reminiscent of hemispatial neglect -

participants were less likely to detect visual stimuli contralateral to the stimulated hemisphere (Kelly et al., 2014). The experiment presented in Chapter 5 showed that the involvement of the TPJ in visual awareness could not be accounted for as an effect of visual attention (Webb et al., 2016b). This raises the question of whether hemispatial neglect is a disorder of attention, as it has traditionally been conceived, or is a disorder of awareness. Future studies might address this question by testing hemispatial neglect patients using paradigms that allow the dissociation of attention and awareness. One possibility, of course, is that neglect is in some cases a disorder of attention, and in other cases a disorder of awareness, perhaps depending on the exact lesion site (Graziano & Kastner, 2011).

The experiment in Chapter 5 showed that, when the effects of attention are controlled for, the TPJ still plays a role in awareness; however, given the interactions observed between attention and awareness in Chapter 2, we should expect that these mechanisms will interact. This may also play an important role in the deficits observed in hemispatial neglect. Indeed, it has been found that neglect patients with lesions in the TPJ show impaired function in more dorsal regions of the parietal and frontal cortex traditionally associated with attention (Corbetta et al., 2005). This suggests that, in many cases, it may not be appropriate to ask *whether* neglect is a disorder of attention or awareness, and may be more fruitful instead to ask how these processes interact in both impaired and normally functioning brains.

References

Baars, B. J. (1997). In the theatre of consciousness. Global workspace theory, a rigorous scientific theory of consciousness. *Journal of Consciousness Studies*, 4(4), 292-309.

Baddeley, A. (2000). The episodic buffer: a new component of working memory?. *Trends in cognitive sciences*, 4(11), 417-423.

Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering event structure in continuous narrative perception and memory. *Neuron*, 95(3), 709-721.

Beckmann, C. F., & Smith, S. M. (2004). Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE transactions on medical imaging*, 23(2), 137-152.

Bio, B. J., Webb, T. W., & Graziano, M. S. (2018). Projecting one's own spatial bias onto others during a theory-of-mind task. *Proceedings of the National Academy of Sciences*, 201718493.

Biswal, B., Zerrin Yetkin, F., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magnetic resonance in medicine*, 34(4), 537-541.

Blanke, O., Slater, M., & Serino, A. (2015). Behavioral, neural, and computational principles of bodily self-consciousness. *Neuron*, 88(1), 145-166.

Bor, D., & Seth, A. K. (2012). Consciousness and the prefrontal parietal network: insights from attention, working memory, and chunking. *Frontiers in psychology*, 3, 63.

Brain, W. R. (1941). A form of visual disorientation resulting from lesions of the right cerebral hemisphere. *Proceedings of the Royal Society of Medicine*, 34, 771–776.

Brainard, D. H., & Vision, S. (1997). The psychophysics toolbox. *Spatial vision*, 10, 433-436.

Breitmeyer, B., & Öğmen, H. (2006). *Visual masking: Time slices through conscious and unconscious vision* (No. 41). Oxford University Press.

Brown, G. D., Yamada, S., & Sejnowski, T. J. (2001). Independent component analysis at the neural cocktail party. *Trends in neurosciences*, 24(1), 54-63.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior research methods*, 41(4), 977-990.

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network. *Annals of the New York Academy of Sciences*, 1124(1), 1-38.

Cabeza, R., Ciaramelli, E., & Moscovitch, M. (2012). Cognitive contributions of the ventral parietal cortex: an integrative theoretical account. *Trends in cognitive sciences*, 16(6), 338-352.

Carmel, D., Lavie, N., & Rees, G. (2006). Conscious awareness of flicker in humans involves frontal and parietal cortex. *Current biology*, 16(9), 907-911.

Carter, R. M., & Huettel, S. A. (2013). A nexus model of the temporal–parietal junction. *Trends in cognitive sciences*, 17(7), 328-336.

Chen, G., Saad, Z. S., Nath, A. R., Beauchamp, M. S., & Cox, R. W. (2012). FMRI group analysis combining effect estimates and their variances. *Neuroimage*, 60(1), 747-765.

Cole, M. W., Reynolds, J. R., Power, J. D., Repovs, G., Anticevic, A., & Braver, T. S. (2013). Multi-task connectivity reveals flexible hubs for adaptive task control. *Nature neuroscience*, 16(9), 1348.

Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P., & Shulman, G. L. (2000). Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nature neuroscience*, 3(3), 292.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature reviews neuroscience*, 3(3), 201.

Corbetta, M., Kincade, M. J., Lewis, C., Snyder, A. Z., & Sapir, A. (2005). Neural basis and recovery of spatial attention deficits in spatial neglect. *Nature neuroscience*, 8(11), 1603.

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron*, 58(3), 306-324.

Corbetta, M., & Shulman, G. L. (2011). Spatial neglect and attention networks. *Annual review of neuroscience*, 34, 569-599.

Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical research*, 29(3), 162-173.

Crick, F., & Koch, C. (1990). Towards a neurobiological theory of consciousness. In *Seminars in the Neurosciences* (Vol. 2, pp. 263-275). Saunders Scientific Publications.

Crick, F. C., & Koch, C. (2005). What is the function of the claustrum?. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1458), 1271-1279.

Critchley, M. (1953). The parietal lobes. Hafner Press, London.

De Brigard, F., & Prinz, J. (2010). Attention and consciousness. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(1), 51-59.

Decety, J., & Lamm, C. (2007). The role of the right temporoparietal junction in social interaction: how low-level computational processes contribute to meta-cognition. *The Neuroscientist*, 13(6), 580-593.

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition*, 79(1-2), 1-37.

Dehaene, S., Naccache, L., Cohen, L., Le Bihan, D., Mangin, J. F., Poline, J. B., & Rivière, D. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature neuroscience*, 4(7), 752.

Dodell-Feder, D., Koster-Hale, J., Bedny, M., & Saxe, R. (2011). fMRI item analysis in a theory of mind task. *Neuroimage*, 55(2), 705-712.

Dosenbach, N. U., Fair, D. A., Miezin, F. M., Cohen, A. L., Wenger, K. K., Dosenbach, R. A., ... & Schlaggar, B. L. (2007). Distinct brain networks for adaptive and stable task control in humans. *Proceedings of the National Academy of Sciences*, 104(26), 11073-11078.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D., & Frackowiak, R. S. (1994). Statistical parametric maps in functional imaging: a general linear approach. *Human brain mapping*, 2(4), 189-210.

Fox, M. D., Corbetta, M., Snyder, A. Z., Vincent, J. L., & Raichle, M. E. (2006). Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proceedings of the National Academy of Sciences*, 103(26), 10046-10051.

Geng, J. J., & Vossel, S. (2013). Re-evaluating the role of TPJ in attentional control: contextual updating?. *Neuroscience & Biobehavioral Reviews*, 37(10), 2608-2620.

Graziano, M. S., & Botvinick, M. M. (2002). How the brain represents the body: insights from neurophysiology and psychology. *Common mechanisms in perception and action: Attention and performance XIX*, 136-157.

Graziano, M. S., & Kastner, S. (2011). Human consciousness and its relationship to social neuroscience: a novel hypothesis. *Cognitive neuroscience*, 2(2), 98-113.

Graziano, M. S. (2013). *Consciousness and the social brain*. Oxford University Press.

Graziano, M. S. (2014). Speculations on the evolution of awareness. *Journal of cognitive neuroscience*, 26(6), 1300-1304.

- Graziano, M. S., & Webb, T. W. (2014). A mechanistic theory of consciousness. *International Journal of Machine Consciousness*, 6(02), 163-176.
- Haider, M., Spong, P., & Lindsley, D. B. (1964). Attention, vigilance, and cortical evoked-potentials in humans. *Science*, 145(3628), 180-182.
- Haith, A. M., & Krakauer, J. W. (2013). Model-based and model-free mechanisms of human motor learning. In *Progress in motor control* (pp. 1-21). Springer, New York, NY.
- Heilman, K. M., & Valenstein, E. (1972). Frontal lobe neglect in man. *Neurology*, 22(6), 660-660.
- Hsieh, P. J., Colas, J. T., & Kanwisher, N. (2011). Pop-out without awareness: Unseen feature singletons capture attention only when top-down attention is available. *Psychological science*, 22(9), 1220-1226.
- Hutchinson, J. B., Uncapher, M. R., & Wagner, A. D. (2009). Posterior parietal cortex and episodic retrieval: convergent and divergent effects of attention and memory. *Learning & Memory*, 16(6), 343-356.
- Huxley, T. (1874). On the hypothesis that animals are automata, and its history. Available in: Huxley, T. (1898). *Collected Essays, Vol. 1*. Macmillan

Igelström, K. M., Webb, T. W., & Graziano, M. S. (2015). Neural processes in the human temporoparietal cortex separated by localized independent component analysis. *Journal of Neuroscience*, 35(25), 9432-9445.

Igelström, K. M., Webb, T. W., Kelly, Y. T., & Graziano, M. S. (2016). Topographical organization of attentional, social, and memory processes in the human temporoparietal cortex. *eNeuro*, 3(2), ENEURO-0060.

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*, 17(2), 825-841.

Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., & Smith, S. M. (2012). Fsl. *Neuroimage*, 62(2), 782-790.

Jiang, Y., Costello, P., Fang, F., Huang, M., & He, S. (2006). A gender-and sexual orientation-dependent spatial attentional effect of invisible images. *Proceedings of the National Academy of Sciences*, 103(45), 17048-17052.

Karnath, H. O., Ferber, S., & Himmelbach, M. (2001). Spatial awareness is a function of the temporal not the posterior parietal lobe. *Nature*, 411(6840), 950.

Karnath, H. O., Himmelbach, M., & Rorden, C. (2002). The subcortical anatomy of human spatial neglect: putamen, caudate nucleus and pulvinar. *Brain*, 125(2), 350-360.

Kelly Jr, R. E., Alexopoulos, G. S., Wang, Z., Gunning, F. M., Murphy, C. F., Morimoto, S. S., ... & Hoptman, M. J. (2010). Visual inspection of independent components: defining a procedure for artifact removal from fMRI data. *Journal of neuroscience methods*, 189(2), 233-245.

Kelly, Y. T., Webb, T. W., Meier, J. D., Arcaro, M. J., & Graziano, M. S. (2014). Attributing awareness to oneself and to others. *Proceedings of the National Academy of Sciences*, 111(13), 5012-5017.

Kentridge, R. W., Nijboer, T. C., & Heywood, C. A. (2008). Attended but unseen: Visual attention is not sufficient for visual awareness. *Neuropsychologia*, 46(3), 864-869.

Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: two distinct brain processes. *Trends in cognitive sciences*, 11(1), 16-22.

Konishi, S., Wheeler, M. E., Donaldson, D. I., & Buckner, R. L. (2000). Neural correlates of episodic retrieval success. *Neuroimage*, 12(3), 276-286.

Lamme, V. A. (2001). Blindsight: the role of feedforward and feedback corticocortical connections. *Acta psychologica*, 107(1-3), 209-228.

Lamme, V. A. (2004). Separate neural definitions of visual consciousness and visual attention; a case for phenomenal awareness. *Neural networks*, 17(5-6), 861-872.

Lau, H. C., & Passingham, R. E. (2006). Relative blindsight in normal observers and the neural correlate of visual consciousness. *Proceedings of the National Academy of Sciences*, 103(49), 18763-18768.

Lau, H. C., & Passingham, R. E. (2007). Unconscious activation of the cognitive control system in the human prefrontal cortex. *Journal of Neuroscience*, 27(21), 5805-5811.

Lau, H. (2008). Are we studying consciousness yet. *Frontiers of Consciousness: Chichele Lectures, 2008*, 245.

Lau, H., & Rosenthal, D. (2011). Empirical support for higher-order theories of conscious awareness. *Trends in cognitive sciences*, 15(8), 365-373.

Lin, Z., & Murray, S. O. (2015). More power to the unconscious: Conscious, but not unconscious, exogenous attention requires location variation. *Psychological Science*, 26(2), 221-230.

Lumer, E. D., Friston, K. J., & Rees, G. (1998). Neural correlates of perceptual rivalry in the human brain. *Science*, 280(5371), 1930-1934.

Marshall, J. C., & Halligan, P. W. (1988). Blindsight and insight in visuo-spatial neglect. *Nature*, 336(6201), 766.

McCormick, P. A. (1997). Orienting attention without awareness. *Journal of Experimental Psychology: Human Perception and Performance*, 23(1), 168

Meister, I. G., Wienemann, M., Buelte, D., Grünewald, C., Sparing, R., Dambeck, N., & Boroojerdi, B. (2006). Hemiextinction induced by transcranial magnetic stimulation over the right temporo-parietal junction. *Neuroscience*, 142(1), 119-123.

Mitchell, J. P. (2007). Activity in right temporo-parietal junction is not selective for theory-of-mind. *Cerebral cortex*, 18(2), 262-271.

Morsella, E. (2005). The function of phenomenal states: supramodular interaction theory. *Psychological review*, 112(4), 1000.

Mort, D. J., Malhotra, P., Mannan, S. K., Rorden, C., Pambakian, A., Kennard, C., & Husain, M. (2003). The anatomy of visual neglect. *Brain*, 126(9), 1986-1997.

Naghavi, H. R., & Nyberg, L. (2005). Common fronto-parietal activity in attention, memory, and consciousness: shared demands on integration?. *Consciousness and cognition*, 14(2), 390-425.

Norman, L. J., Heywood, C. A., & Kentridge, R. W. (2013). Object-based attention without awareness. *Psychological science*, 24(6), 836-843.

Persaud, N., Davidson, M., Maniscalco, B., Mobbs, D., Passingham, R. E., Cowey, A., & Lau, H. (2011). Awareness-related activity in prefrontal and parietal cortices in blindsight reflects more than superior visual performance. *Neuroimage*, 58(2), 605-611.

Posner, M. I. (1980). Orienting of attention. *Quarterly journal of experimental psychology*, 32(1), 3-25.

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

Pruim, R. H., Mennes, M., van Rooij, D., Llera, A., Buitelaar, J. K., & Beckmann, C. F. (2015). ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. *Neuroimage*, 112, 267-277.

Rees, G., Kreiman, G., & Koch, C. (2002). Neural correlates of consciousness in humans. *Nature Reviews Neuroscience*, 3(4), 261.

Rosenthal, D. M. (2004). Varieties of higher-order theory. *Advances in Consciousness Research*, 56, 17-44.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: the role of the temporo-parietal junction in “theory of mind”. *Neuroimage*, 19(4), 1835-1842.

Saxe, R., & Powell, L. J. (2006). It's the thought that counts: specific brain regions for one component of theory of mind. *Psychological science*, 17(8), 692-699.

Scholz, J., Triantafyllou, C., Whitfield-Gabrieli, S., Brown, E. N., & Saxe, R. (2009). Distinct regions of right temporo-parietal junction are selective for theory of mind and exogenous attention. *PloS one*, 4(3), e4869.

Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., ... & Greicius, M. D. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *Journal of Neuroscience*, 27(9), 2349-2356.

Seitz, A. R., Kim, D., & Watanabe, T. (2009). Rewards evoke learning of unconsciously processed visual stimuli in adult humans. *Neuron*, 61(5), 700-707.

Shimamura, A. P. (2011). Episodic retrieval and the cortical binding of relational activity. *Cognitive, Affective, & Behavioral Neuroscience*, 11(3), 277-291.

Stevens, A. A., Skudlarski, P., Gatenby, J. C., & Gore, J. C. (2000). Event-related fMRI of auditory and visual oddball tasks. *Magnetic resonance imaging*, 18(5), 495-502.

Szczepanski, S. M., & Kastner, S. (2013). Shifting attentional priorities: control of spatial attention through hemispheric competition. *Journal of Neuroscience*, 33(12), 5411-5421.

Talairach, J., & Tournoux, P. (1988). Co-planar stereotaxic atlas of the human brain. 3-Dimensional proportional system: an approach to cerebral imaging.

Tallon-Baudry, C. (2012). On the neural mechanisms subserving consciousness and attention. *Frontiers in psychology*, 2, 397.

Tong, F. (2003). Cognitive neuroscience: primary visual cortex and visual awareness. *Nature Reviews Neuroscience*, 4(3), 219.

Tononi, G. (2008). Consciousness as integrated information: a provisional manifesto. *The Biological Bulletin*, 215(3), 216-242.

Tsuchiya, N., Wilke, M., Frässle, S., & Lamme, V. A. (2015). No-report paradigms: extracting the true neural correlates of consciousness. *Trends in cognitive sciences*, 19(12), 757-770.

Tsushima, Y., Sasaki, Y., & Watanabe, T. (2006). Greater disruption due to failure of inhibitory control on an ambiguous distractor. *Science*, 314(5806), 1786-1788.

Umiltà, C., Salmaso, D., Bagnara, S., & Simion, F. (1979). Evidence for a right hemisphere superiority and for a serial search strategy in a dot detection task. *Cortex*, 15(4), 597-608.

Ungerleider, S. K. A. L. G. (2000). Mechanisms of visual attention in the human cortex. *Annual review of neuroscience*, 23(1), 315-341.

Vallar, G., & Perani, D. (1986). The anatomy of unilateral neglect after right-hemisphere stroke lesions. A clinical/CT-scan correlation study in man. *Neuropsychologia*, 24(5), 609-622.

Vilberg, Kaia L., and Michael D. Rugg. "Memory retrieval and the parietal cortex: a review of evidence from a dual-process perspective." *Neuropsychologia* 46.7 (2008): 1787-1799.

Vincent, J. L., Kahn, I., Snyder, A. Z., Raichle, M. E., & Buckner, R. L. (2008). Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. *Journal of neurophysiology*, 100(6), 3328-3342.

Webb, T. W., & Graziano, M. S. (2015). The attention schema theory: a mechanistic account of subjective awareness. *Frontiers in psychology*, 6, 500.

Webb, T. W., Kean, H. H., & Graziano, M. S. (2016a). Effects of awareness on the control of attention. *Journal of cognitive neuroscience*, 28(6), 842-851.

Webb, T. W., Igelström, K. M., Schurger, A., & Graziano, M. S. (2016b). Cortical networks involved in visual awareness independent of visual attention. *Proceedings of the National Academy of Sciences*, 113(48), 13923-13928.

Weiskrantz, L. (1986). *Blindsight: A case study and implications*. Oxford University Press.

Whitfield-Gabrieli, S., & Nieto-Castanon, A. (2012). Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain connectivity*, 2(3), 125-141.

Yeo, BT Thomas, et al. "The organization of the human cerebral cortex estimated by intrinsic functional connectivity." *Journal of neurophysiology* 106.3 (2011): 1125-1165.