Face perception is category-specific: Evidence from normal body perception in acquired prosopagnosia

Tirta Susilo \(^{a, *}\), Galit Yovel \(^b\), Jason J.S. Barton \(^c\), Bradley Duchaine \(^a\)

\(^a\) Department of Psychological and Brain Sciences, Dartmouth College, United States
\(^b\) School of Psychological Sciences & Sagol School of Neuroscience, Tel Aviv University, Israel
\(^c\) Departments of Medicine (Neurology) & Ophthalmology and Visual Sciences, University of British Columbia, Canada

**A B S T R A C T**

Does the human visual system contain perceptual mechanisms specialized for particular object categories such as faces? This question lies at the heart of a long-running debate in face perception. The face-specific hypothesis posits that face perception relies on mechanisms dedicated to faces, while the expertise hypothesis proposes that faces are processed by more generic mechanisms that operate on objects we have extended experience with. Previous studies that have addressed this question using acquired prosopagnosia are inconclusive because the non-face categories tested (e.g., cars) were not well-matched to faces in terms of visual exposure and perceptual experience. Here we compare perception of faces and bodies in four acquired prosopagnosics. Critically, we used face and body tasks that generate comparable inversion effects in controls, which indicates that our tasks engage orientation-specific perceptual mechanisms for faces and bodies to a similar extent. Three prosopagnosics were able to discriminate bodies normally despite their impairment in face perception. Moreover, they exhibited normal inversion effects for bodies, suggesting their body perception was carried out by the same mechanisms used by controls. Our findings indicate that the human visual system contains processes specialized for faces.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

A fundamental issue in cognitive science concerns the extent to which the human mind consists of processes specialized for particular object categories. This issue motivates the long-running debate about the nature of face perception. According to the face-specific hypothesis, face perception is carried out by mechanisms specialized for faces (Pitcher, Charles, Devlin, Walsh, & Duchaine, 2009; Tanaka & Farah, 1993; Yin, 1969). According to the expertise hypothesis, faces are analyzed by more generic mechanisms for objects with which we have extended experience (Diamond & Carey, 1986; Gauthier & Tarr, 1997; McGugin, Gatensby, Gore, & Gauthier, 2012). Here we contrast the two hypotheses by examining body perception when face perception is impaired in acquired prosopagnosia (Bodamer, 1947).

Previous studies that have investigated the nature of face processing using acquired prosopagnosia have typically compared perception of faces and a variety of non-face objects (e.g., Busigny, Graf, Mayer, & Rossion, 2010; Farah, Klein, & Levinson, 1995; Moscovitch, Winocur, & Behrmann, 1997; Riddoch, Johnston, Bracewell, Boutsen, & Humphreys, 2008). While dissociations between perception of faces and non-faces suggest that faces are processed differently than most objects, they do not distinguish between the face-specific hypothesis and the expertise hypothesis because (i) both hypotheses agree that faces are processed by mechanisms different from those used for objects (i.e., most of us are experts with faces but not objects), and (ii) the non-face categories tested (e.g., cars, chairs) are not matched to faces in terms of perceptual...
experience. To discriminate between the two hypotheses, faces need to be compared with an object category for which participants have similar amounts of perceptual experience. Only then will the expertise hypothesis predict an association between faces and non-faces in all prosopagnosics, while the face-specific hypothesis suggest dissociations can occur in some prosopagnosics.

Here we used bodies as a comparison category, because faces and bodies share two theoretically important characteristics. First, both faces and bodies produce inversion effects (i.e. worse discrimination of visual stimuli presented upside-down) larger than those for other objects (the face inversion effect, Yin, 1969; the body inversion effect, Reed, Stone, Bozova, & Tanaka, 2003). This is important because inversion effects indicate orientation-specific processing and are considered a marker of perceptual expertise. Most critical for our study, inversion effects for faces and bodies can be similar in size (Robbins & Coltheart, 2012a; Yovel, Pelc, & Lubetzky, 2010), which indicates that faces and bodies can engage orientation-specific mechanisms to a similar extent. Second, faces and bodies exhibit consistent first-order configurations (i.e., fixed spatial relations between eyes, nose, and mouth for faces; arms, torso, and legs for bodies), which have been suggested to be necessary for the development of visual expertise for particular object categories (Diamond & Carey, 1986).

Our study consisted of three steps. We first confirmed that our face and body tasks generate comparable inversion effects in controls. This step ensured that our tasks engage orientation-specific processing of faces and bodies to a similar extent, a critical factor in contrasting the face-specific and the expertise hypotheses. Next we compared how the prosopagnosics discriminate among exemplars of upright faces and of upright bodies. Finally we examined whether the prosopagnosics who were able to discriminate upright bodies as accurately as controls also showed normal-sized inversion effects for bodies, which would suggest that their body perception was generated by the same mechanisms used by controls. The status of the body inversion effect in acquired prosopagnosia is of additional interest because there is some evidence that the body inversion effect might involve face-selective rather than body-selective neural mechanisms (Brandman & Yovel, 2010).

2. Method

2.1. Participants

We tested four acquired prosopagnosics, namely Florence, Sandy, Grace, and Galen, as part of our broader investigation of prosopagnosia. Table 1 shows their performance on tests of face recognition.

Florence is a right-handed nurse born in 1982. She was 29 years old when tested. In 2006, Florence noticed problems with face recognition following a right amygdalo-hippocampectomy. Functional MRI scans showed bilateral activations in her fusiform face area, occipital face area, and superior temporal sulcus. In 2008 she underwent a second surgery that removed the anterior third of her right temporal lobe, sparing the core face areas previously identified. Florence did not complain of visual impairments other than prosopagnosia, and she performed normally on within-class recognition of objects including hairstyles, cars, and abstract paintings. In Fox, Hanif, Iaria, Duchaine, and Barton (2011), Florence was referred to as R-AT1.

Sandy is a right-handed woman born in 1975. She was 36 years old when tested. Sandy became prosopagnostic after a right hippocampal resection in 2003 during which she had a stroke in the occipital lobe and she lost her left visual field completely. She complained of severe difficulties recognizing faces, including herself in the mirror and her children in the school, and reported that she relies heavily on gait and walking sound to identify people. Sandy also complained of object recognition problems, such as finding her car when other cars in the parking lot have similar colors. Sandy was impaired on tests of visual closure, eye gaze perception, facial expression recognition, and hairstyle recognition. Sandy reported no general memory problems.

Grace is a right-handed pharmacist born in 1968. She was 43 years old when tested. Grace acquired prosopagnosia after a brain biopsy of the right temporal lobe in 1982 to treat herpes simplex viral encephalitis. Grace complained about difficulties in face recognition and relies on non-face cues like voice, hairstyle, glasses, and gait to identify people. In addition to her prosopagnosia, Grace was also impaired on tests of color perception, visual closure, and basic object recognition from line drawings. Structural MRI scans showed a lesion in the right anterior temporal lobe extending to the middle fusiform and inferior temporal gyri, as well as a small lesion in the middle aspect of the left fusiform gyrus. She was referred to as B-OT/AT1 in Dalrymple et al. (2011).

Galen is a right-handed physician born in 1982. He was 29 years old when tested. Galen became prosopagnostic in 2004 following a craniotomy for an arteriovenous malformation in his right temporal lobe. He complained of difficulties recognizing faces, including celebrities and people who are related or have similar appearances and reported using contextual cues to identify people. Galen previously experienced a left-superior quadrantanopia, but a recent examination showed his low-level abilities in the
left-superior visual field are in the normal range. Galen did not complain of visual agnosia in general, and he scored normally on recognition tests involving hairstyles, cars, and abstract paintings. Functional scans showed an absence of right fusiform face area and right occipital face area.

Twenty people from the Dartmouth College community (13 female, age range 18–27 years) participated as controls.

2.2. Stimuli and procedure

The main experiment used a task developed by Yovel et al. (2010), in which participants made same/different judgments on 144 sequentially presented pairs of headless bodies, faceless bodies, and faces, shown in different blocks (Fig. 1). Body pairs differed in terms of the position of the arms, legs, and heads (in the faceless bodies). Face pairs differed in terms of eyes, nose, and mouth. For each of the three categories, upright and inverted trials (72 each) were interleaved in a pre-determined random order. Headless bodies were tested first, faceless bodies second, and faces last to ensure that poor face discrimination was not due to unfamiliarity with the paradigm and that normal body discrimination was not confounded by practice effects. Dependent measures were d’-prime and response time. Inversion effects were computed as [upright d’-prime – inverted d’-prime] and as [upright RT – inverted RT]. (Note that we also computed inversion effects in a relative manner: [(upright d’ – inverted d’)/(upright d’ + inverted d’)] and [(upright RT – inverted RT)/(upright RT + inverted RT)]; as presented in Supplementary Figure, we found similar results for all prosopagnosics and thus came to the same conclusion.)

2.3. Statistical analysis

We used the t-test for single-case analysis (Crawford & Howell, 1998) to compare a case score against the control mean in a particular condition (e.g., Florence’s discrimination of upright faces). To compare each case’s difference scores (e.g., the difference between Florence’s discrimination of faces and her discrimination of faceless bodies) against the difference scores in controls, we used the Bayesian Standardized Difference Test (Crawford & Garthwaite, 2007). For all statistical analyses we report the estimated percentage of the control population that would perform worse than a case score or would exhibit a larger difference score in the predicted direction. Note that these estimated percentages directly correspond to p-values. Scores below a 5% cut-off were classified as abnormal.

3. Results

3.1. Did faces and bodies show comparable inversion effects in controls?

Fig. 2A shows that all conditions produced inversion effects in controls. Computed using the absolute index (i.e. [(upright d’ – inverted d’)]), the inversion effect for faces (M = 1.28, SE = 0.21) was comparable to that for faceless bodies (M = 1.04, SE = 0.18), F(1,19) = 3.19, p = 0.09, but larger than that for headless bodies (M = 0.55, SE = 0.18), F(1,19) = 24.91, p < 0.001.1 Computed using the relative index (i.e., [(upright d’ – inverted d’)/(upright d’ + inverted d’)], the inversion effect for faces (M = 0.31, SE = 0.04) was again comparable to that for faceless bodies (M = 0.26, SE = 0.03), F(1,19) = 1.52, p = 0.23, and larger than that for headless bodies (M = 0.16, SE = 0.07), F(1,19) = 4.9, p = 0.04. Fig. 2B shows that there was no speed–accuracy trade-off: participants were slower to discriminate inverted than upright stimuli. This result replicates a previous finding (Yovel et al., 2010), and indicates that faces and faceless bodies, but not headless bodies, engage orientation-specific processes to a similar degree.

3.2. How did the prosopagnosics discriminate faces and bodies in the upright orientation?

Table 2 (condition scores) shows that all prosopagnosics, except Sandy, were statistically impaired with faces but normal with faceless bodies and headless bodies on both d’-prime and RT. However, given that a statistically

---

1 The inversion effect for faceless bodies was trending smaller than that for faces (p = 0.09), but two previous studies using the same task found no such trend (p = 0.54 in Brandman & Yovel, 2012; p > 0.3 in Yovel et al., 2010). Based on all available data we would argue that our task generates statistically comparable inversion effects for faces and faceless bodies.

---

Fig. 1. Experimental task showing example stimuli for faces, faceless bodies, and headless bodies.
impaired score and a non-impaired score may not be significantly different (Crawford & Garthwaite, 2007), we next assessed whether the differences between discrimination of faces and bodies in the prosopagnosics were statistically abnormal compared to the same differences in controls.

### 3.3. Was the difference between discrimination of faces and faceless bodies abnormal?

The difference scores in Table 2 show that the d-primes for Florence, Grace, and Galen were significantly worse for faces than for faceless bodies. Florence and Galen were also significantly different on RT. Critically, all three prosopagnosics exhibited normal-sized inversion effects for faceless bodies on both d-prime and RT (Table 2 inversion effects, Fig. 3). Normal discrimination of and normal-sized inversion effects for faceless bodies indicates that despite their prosopagnosia, Florence, Grace, and Galen processed faceless bodies like controls did.

### 3.4. Was the difference between discrimination of faces and headless bodies abnormal?

The difference scores in Table 2 also show that all prosopagnosics performed worse with faces than with headless bodies on d-prime. As above, Florence and Galen also
showed a dissociation for RT. Florence, Sandy, and Grace showed normal-sized inversion effects for headless bodies on both d-prime and RT (Table 2 inversion effects, Fig. 3C). These results indicate that perception of headless bodies can be spared in prosopagnosia, although it does not distinguish between the face-specific and the expertise hypotheses because the inversion effect for headless bodies in controls was smaller than that for faces to start with.

3.5. Did the prosopagnosics who show normal body inversion effects also show normal face inversion effects?

The idea that body inversion effects might rely on mechanisms for face rather than body perception (Brandman & Yovel, 2010) predicts that prosopagnosics who showed normal body inversion effects should also show normal face inversion effects. Is this the case? Our data suggest not. While Florence, Grace, and Galen showed face inversion effects in the normal range on d-prime, Florence and Galen exhibited a clear speed/accuracy trade-off: they were much slower with upright faces than with inverted faces (Table 2 inversion effects, Fig. 3A). As a result, their data are difficult to interpret. In contrast, Florence and Galen showed normal-sized inversion effects for faceless bodies on both d-prime and RT, thus indicating a dissociation between their inversion effects for faces and for faceless bodies.

3.6. Was discrimination of bodies easier than discrimination of faces?

- None of our results can be accounted by easier discrimination of bodies than of faces. In fact controls were better at discriminating faces than both faceless bodies, $t(19) = 2.90$, $p < .01$, and headless bodies, $t(19) = 8.08$, $p < .0001$. This means that three of four prosopagnosics performed normally on body tasks that are more challenging than the face task they had impairments with.
4. Discussion

In this study we addressed whether faces are processed by specialized mechanisms (face-specific hypothesis) or by more generic mechanisms for objects we have extended experience with (expertise hypothesis). We did so by comparing perception of faces and bodies in acquired prosopagnosia, because faces and bodies engage orientation-specific perceptual mechanisms and exhibit consistent first-order configurations among their parts. Controls exhibited comparable inversion effects for faces and faceless bodies. Three of four prosopagnosics were able to discriminate bodies as well as controls and showed normalized body inversion effects. Their results indicate body perception can be normal when face perception is impaired, consistent with the face-specific hypothesis.

Our findings add to the literature on acquired prosopagnosics who performed normally with non-faces. Termed "pure" prosopagnosics (for a review of 13 existing cases see Busigny et al., 2010), these participants are often considered evidence that faces are processed by dedicated mechanisms, especially when the face and non-face tasks are matched for task demands, equated in difficulty, and free of speed/accuracy trade-offs. However, such dissociations do not discriminate between the face-specific and the expertise hypotheses because it is unclear whether the prosopagnosics had extensive experience with the non-face categories tested. In contrast, our use of bodies as a comparison category allows us to tease apart the two hypotheses because faces and bodies share theoretically-important characteristics mentioned above.

Our study is the first to report a dissociation between face and body perception in acquired prosopagnosia. Case FM was impaired with perception of both faces and bodies (Moro et al., 2012). The well-known case PS showed typical fMRI activation in body-selective areas to emotional face-related body stimuli, but her behavioral performance with bodies was impaired. Three of four prosopagnosics were able to discriminate the parts of the body as well as controls and showed normalized body inversion effects. Their results indicate body perception can be normal when face perception is impaired, consistent with the face-specific hypothesis.

Our findings add to the literature on acquired prosopagnosics who performed normally with non-faces. Termed “pure” prosopagnosics (for a review of 13 existing cases see Busigny et al., 2010), these participants are often considered evidence that faces are processed by dedicated mechanisms, especially when the face and non-face tasks are matched for task demands, equated in difficulty, and free of speed/accuracy trade-offs. However, such dissociations do not discriminate between the face-specific and the expertise hypotheses because it is unclear whether the prosopagnosics had extensive experience with the non-face categories tested. In contrast, our use of bodies as a comparison category allows us to tease apart the two hypotheses because faces and bodies share theoretically-important characteristics mentioned above.

Our study is the first to report a dissociation between face and body perception in acquired prosopagnosia. Case FM was impaired with perception of both faces and bodies (Moro et al., 2012). The well-known case PS showed typical fMRI activation in body-selective areas to emotional body stimuli, but her behavioral performance with bodies was not assessed (Peelen, Lucas, Mayer, & Vuilleumier, 2009). A few studies examined face and body perception in developmental prosopagnosia and found mixed results (e.g., Duchaine, Yovel, Butterworth, & Nakayama, 2006; Righart & De Gelder, 2007). Crucially, however, none of these studies demonstrated that the body tasks used generated face-size inversion effects in controls. It thus remains possible that normal body perception observed in developmental prosopagnosics did not depend on orientation-specific mechanisms to a similar extent as faces did.

Our result agrees with evidence from single-cell, functional imaging (fMRI), and transcranial magnetic stimulation (TMS) studies. The existence of face-selective neurons has long been reported (Gross, Rocha-Miranda, & Bender, 1972; Perrett, Rolls, & Caan, 1982), and recent investigations have established that these neurons are functionally organized in a network of face-selective patches (Moeller, Freiwald, & Tsao, 2008). Functional imaging studies have found separate cortical areas selective for faces and bodies in humans (de Gelder et al., 2010; Peelen & Downing, 2007), and TMS studies have indicated the causal involvement of some of these areas in discrimination tasks only for their preferred category (Pitcher et al., 2009; Urgesi, Berlucchi, & Aglioti, 2004). Our finding complements these data by showing a cognitive dissociation between face and body perception. While direct mapping between behavioral performance and lesion location is beyond the scope of the present study, future studies of acquired prosopagnosia are likely to benefit from obtaining functional scans to body and body part stimuli.

What is the nature of the orientation-specific mechanisms for faceless bodies that were spared in our prosopagnosics? Given their face-size inversion effects, these mechanisms might perform holistic computations similar to those in face perception. Consistent with this possibility, perception of body parts benefits from the presence of the whole body (Seitz, 2002), just like perception of face parts benefits from the context of the whole face (i.e., the part-whole effect, Tanaka & Farah, 1993). Perception of one-half of bodies can be influenced by the unattended half (Robbins & Coltheart, 2012b, but see Soria Bauser, Suchan, & Daum, 2011), similar to the composite face effect (Young, Hellaowell, & Hay, 1987). Future studies can use these paradigms to further clarify whether normal body perception in acquired prosopagnosia is holistic in nature.

A recent study however suggests that the face-size inversion effect for faceless bodies might not be driven by body perception mechanisms, but instead by face detection mechanisms (Brandman & Yovel, 2012). This study compared the size of inversion effects for several body conditions including faceless bodies, heads with shoulders, heads only, and bodies from the back. Only faceless bodies and heads with shoulders generated face-size inversion effects; other conditions produced smaller inversion effects. In a second experiment, different body conditions were flashed for 17 ms each, and participants were asked whether they saw a face. Interestingly, participants were more likely to report seeing a face in the same two conditions that produced face-size inversion effects, namely faceless bodies and heads with shoulders. The authors interpreted these results as evidence that the face-size inversion effects in bodies are generated by face detection mechanisms.

Although our data are not inconsistent with the potential involvement of face detection mechanisms (we did not systematically assess face detection ability of the prosopagnics), it is worth noting that face-size inversion effects were obtained for faceless bodies and bodies with shoulders, not for heads only (Brandman & Yovel, 2012). This means two things: (i) there has to be some body parts in the stimuli (shoulders at the minimum) for the face-size inversion effects to emerge, and (ii) these body parts have to be processed normally. A participant with impaired shoulder perception, for example, would be expected to process faceless bodies abnormally, and thus would fail to exhibit normal inversion effects. The fact that our prosopagnics exhibited normal inversion effects for faceless bodies implies that their ability to process all aspects of faceless bodies was normal.

Regardless of the underlying mechanisms, the body inversion effect indicates that bodies, unlike most non-face objects, are processed by perceptual mechanisms that are very sensitive to orientation and are therefore a suitable category for distinguishing between the expertise and the
face-specific hypotheses. No prosopagnosia studies to date have used a task where the non-face category is comparable to faces in terms of sensitivity to orientation, and thus our study offers a critical piece of evidence that is inconsistent with the notion that prosopagnosia is an impairment affecting the processing of objects with which we have extended experience. Rather, our findings suggest prosopagnosia can be a category-specific deficit that is restricted to faces, which indicates the human mind contains processes specialized for particular object categories.

Acknowledgements

We thank Florence, Sandy, Grace, and Galen for their participation, and Constantin Rezlescu for testing Sandy and Grace. This work was supported in part by CIHR Grant MOP-102567 and the Hitchcock Foundation. J.J.S.B. was supported by a Canada Research Chair.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cognition.2013.06.004.

References

Right Parietal Cortex Plays a Critical Role in Change Blindness

There is increasing evidence from functional magnetic resonance imaging (fMRI) that visual awareness is not only associated with activity in ventral visual cortex but also with activity in the parietal cortex. However, due to the correlational nature of neuroimaging, it remains unclear whether this parietal activity plays a causal role in awareness. In the experiment presented here we disrupted activity in right or left parietal cortex by applying repetitive transcranial magnetic stimulation (rTMS) over these areas while subjects attempted to detect changes between two images separated by a brief interval (i.e. 1-shot change detection task). We found that rTMS applied over right parietal cortex but not left parietal cortex resulted in longer latencies to detect changes and a greater rate of change blindness compared with no TMS. These results suggest that the right parietal cortex plays a critical role in conscious change detection.

Keywords: change detection, humans, parietal cortex, rTMS, visual awareness

Introduction

It is well accepted that activity in the occipitotemporal cortex plays a role in visual awareness. Lesions to areas of extrastriate cortex specialized for analysis of a visual attribute give rise to corresponding deficits in awareness for that attribute (for review, see Farah, 1990). These same visual areas have also been associated with reports of awareness in neurologically intact subjects. For example, it has been reported that activity in the lateral occipital complex (LOC) is correlated with subjects' ability to recognize masked objects (Grill-Spector et al., 2000), and left fusiform gyrus and extrastriate areas have also been found to be more active in association with seen versus masked unseen words (Dehaene et al., 2001). Indeed, Milner and Goodale (1995) have argued that ventral stream activation is a necessary condition of awareness. Based on the abilities of a visual agnosia patient, they further suggest that information in the dorsal system can be processed without reaching consciousness.

More recently, however, evidence from fMRI has suggested that the dorsal stream may play a role in visual awareness. In fact, parietal activity has been implicated in visual awareness with a variety of paradigms. For example, Dehaene et al. (2001) found that parietal areas were active when subjects were aware of words but not when the same words were masked and thus unseen. Using a binocular rivalry paradigm in which subjects alternately report seeing only one of two incompatible images presented to the two eyes, Lumer et al. (1998) not only found that activity in ventral extrastriate areas was correlated with the seen image, in accordance with other research (e.g. Tong et al., 1998), but also that activity in the inferior and superior parietal lobules was enhanced during the perceptual transitions from one image to another. In other words, they found parietal activity when the contents of awareness changed. Similar parietal activation was found during the perceptual transitions that occur when viewing bistable images such as the Necker cube and Rubin’s face/vase (Kleinschmidt et al., 1998).

Additionally, using the phenomenon of change blindness, Beck et al. (2001) found enhanced activity in the parietal lobe when subjects consciously detected a change but not when they were blind to it. Change blindness can be produced by introducing a brief flicker between successive views of a visual scene (Rensink et al., 1997). Even though the change is large enough to be detected immediately under normal viewing conditions (i.e. without the screen flicker), the flicker paradigm results in long RTs or a complete failure to detect the change. Beck et al. (2001) used the flicker paradigm and the resulting change blindness to probe the neural correlates of visual awareness using event-related fMRI. A comparison of trials in which subjects consciously detected a change to trials in which they were blind to it not only revealed enhanced activity in category-selective regions of the ventral visual cortex, as predicted by many neural theories of visual awareness (Logothetis, 1998; Zeki, 2003), but also enhanced activity in bilateral parietal cortex.

However, because fMRI can only reveal an association between activity in a brain region and behavior, it is unclear whether the parietal regions implicated in these experiments play any causal role in awareness. In all these studies, it remains possible that the parietal activity found was a consequence of subjects’ awareness and did not play either a necessary or causal role in producing that awareness. One can imagine, for instance, that parietal activity found in the change blindness experiment reflects attention being drawn to the change after the change has been detected and the subject has become aware of it. In order to assess whether parietal activity is in fact critically involved in change detection, we applied rTMS over either right or left parietal cortex (Fig. 1) to transiently disrupt neural activity in these regions (Walsh and Pascual-Leone, 2003), while subjects attempted to detect changes across a brief blank interval (Fig. 2). The peak coordinates of the parietal regions associated with change detection in the fMRI study of Beck et al. (2001) were located on each individual anatomical MRI and co-registered with the coil using BrainSight software (Rogue Research, Montreal Canada). If parietal activity plays a causal role in the conscious detection of change, subjects’ performance on a change detection task should be selectively worse during rTMS stimulation than during no rTMS stimulation.

Diane M. Beck¹, Neil Muggleton²,³, Vincent Walsh²,³ and Nilli Lavie²,³

¹Department of Psychology, Princeton University, Green Hall, Princeton, NJ 08544, USA, ²Department of Psychology, University College London, London, UK and ³Institute of Cognitive Neuroscience, University College London, London, UK
Materials and Methods

Subjects
Nine right-handed subjects (three females; age 23-40 years) participated in the study. All subjects were in good health with no past history of psychiatric or neurological diseases and gave their written and informed consent. Subjects had normal or corrected-to-normal vision.

Stimuli and Procedure
Subjects viewed two successive displays of faces in what has been termed the 1-shot task. A trial consisted of a 500 ms presentation of a fixation cross, followed by two 200 ms displays separated by a 100 ms blank interval. Subjects had 2.8 s to respond whether or not one of the faces changed by pressing one of two buttons on a response pad. In this example, the lower right face changed. All four locations were equally likely, but randomly chosen, to contain the change. rTMS stimulation was time-locked to the onset of the first display and continued for the duration of the visual presentation (500 ms).

On two-thirds of the trials, one of the four faces was replaced by another face in the second display (Fig. 1). All four locations were equally likely, but randomly chosen, to contain the change. The remaining one third of trials did not contain a change.

RTMS Parameters and Design
A Magstim SuperRapid TMS machine, set at 65% of output, and a 70 mm figure of eight coil were used to deliver 500 ms trains of 10 Hz pulses that were time-locked to the onset of the four faces. rTMS was administered in blocks of 60 trials, and subjects received two blocks of each of the following conditions: right parietal TMS, left parietal TMS, and no TMS. Block order was randomized with the constraint that each condition occurred once in the first half of the experiment and once in the second half, and subjects began the session with 10 practice trials in which no TMS was delivered. The stimulation site was co-registered with each subject’s MRI scan. Brainsight software (Rogue Research, Montreal Canada) was used to identify, in each subject, the Talairach coordinates of the peak parietal activity (i.e. in posterior parietal cortex) previously reported to be associated with conscious change detection (Fig. 2). It should be noted that these are the peak co-ordinates from the random effects analysis, which included data from the first and second experiment, and the extent of the activations...
was large enough that the area stimulated by TMS (at the intensities used and with the large figure of eight coil) bears a close correspondence to the activation areas in both experiments.

Results
Mean RTs (to correct trials only) and error rates were computed for each subject as a function of TMS condition (right TMS, left TMS and no TMS) and side of change (right versus left) and entered into a two-way repeated-measures analysis of variance (ANOVA). Figure 3 presents the average of these mean RTs and error rates across participants. The ANOVA on the RTs revealed a significant main effect of TMS condition \( t(2,16) = 43.95, P < 0.05 \), no effect of side of change or interaction of TMS condition and side of change, \( F < 1 \). To better understand the main effect of TMS, \( t \) tests comparing right versus no TMS, left versus no TMS, and right versus left TMS were conducted. These analyses revealed that right parietal TMS produced slower change detection responses (819 ms) than no TMS (706 ms) \( t(8) = 2.76, P < 0.05 \). Left parietal TMS (714 ms), on the other hand, yielded no impairment relative to no TMS \( t < 1 \), and right parietal TMS resulted in significantly slower responses than left parietal TMS \( t(8) = 3.23, P < 0.05 \). The direct comparison of the effects of right versus left parietal TMS indicates that the deficit in performance during right parietal stimulation cannot be due to unspecific effects of TMS, because the same stimulation that produced a deficit on the right produced no effect when administered on the left.

Disrupting activity in the right parietal lobe had an adverse effect on the speed with which subjects detected changes, but, more important for the issue of awareness, similar effects were found on accuracy of change detection. The initial ANOVA on the accuracy data, involving all three levels of the TMS factor, revealed a significant main effect of TMS condition \( F(2,16) = 6.16, P < 0.05 \), no effect of side of change \( F(1.8) = 1.8, P > 0.20 \) or interaction of TMS condition and side of change, \( F < 1 \). The subsequent \( t \) tests revealed that right parietal TMS increased the rate of change blindness compared with no TMS \( t(8) = 2.86, P < 0.05 \); subjects missed more changes during right parietal stimulation (37%) than during no TMS (28%). As with the RTs, left TMS failed to produce more change blindness (30% misses) than the no TMS condition \( t < 1 \), and right TMS resulted in greater change blindness than left TMS \( t(8) = 2.34, P < 0.05 \). These analyses included all misses, including those in which subjects failed to respond at all. The same pattern of results, however, was obtained when failures to respond were removed from the analysis. Subjects missed 36% of changes during right TMS, and 30 and 28% of changes during left TMS and no TMS conditions, respectively. In order to assess the effects of TMS on sensitivity of change detection, taking into account false alarm rate as well as hit rate, we transformed the accuracy data into \( d' \). An analysis of the \( d' \) data revealed the same pattern of results as the accuracy and RT data. Right parietal stimulation significantly decreased sensitivity to change \( d' = 1.55 \) relative to no TMS \( d' = 1.88 \), \( t(8) = 2.38, P < 0.05 \) and relative to left TMS \( t(8) = 2.50, P < 0.05 \), whereas left TMS did not result in a deficit \( d' = 1.88 \) compared to no TMS \( t < 1 \). Thus, once again the significant effect of right TMS cannot be due to unspecific effects of TMS, because the same stimulation over the left parietal lobe produced no effect.

Eye movement artifacts cannot explain our data either. TMS over posterior parietal cortex (PPC) does not induce eye movements in humans as shown in several TMS experiments on PPC function (e.g. Ashbridge et al., 1997; for review, see Walsh and Pascual-Leone, 2003). Exploratory eye movements on the part of our subjects were also unlikely since the presentation of each array was too brief (200 ms) to complete an eye movement, and anticipatory eye movements between the first and second display to one of the four image positions would have been detrimental to performance: given that the location of the change (if there was one) was random, subjects would have a 75% chance of saccading to the wrong image. Moreover, there is no reason why eye movements of any kind should be present during right posterior parietal stimulation and not left, and yet our performance effects are restricted to the right TMS condition.

Finally, our effects cannot be the result of interference with response selection mechanisms. All subjects made both present and absent responses with their right hand, and given that our effects were also present in the \( d' \) measure it would be difficult to explain them in terms of a bias in response. Moreover, although activity in the parietal cortex has been associated with response selection, this activity is lateralized to the left parietal cortex for both left and right-handed responses (Rushworth 2000).
conscious change detection was associated with bilateral activation of parietal cortex, but in the current experiment disrupting left PPC had no effect on detection performance. The lack of effect for left parietal TMS, however, is in accordance with neuropsychological studies of the phenomenon of visual neglect in which dramatic deficits in visual awareness predominantly follow right, but not left, parietal lesions (Vallar and Perani, 1986). A number of researchers have suggested that while left parietal cortex directs attention to contralateral space (i.e. the right visual field), the right parietal cortex has a bilateral representation of visual space (Heilman and Van Den Abell, 1980; Mesulam, 1981). This bilateral representation of space would allow the right parietal cortex to take over and compensate for the disruption during rTMS of the left parietal lobe. However, when both hemispheres are intact (as they were in the fMRI experiments) there would be no need for the right parietal lobe to take over the role of the left, and thus bilateral activation would be expected.

**Right Parietal TMS Effects on Both Sides of Space**

The right hemisphere dominance theory can account for the lateralized effect of TMS, but our data are also in accordance with spatially non-specific functions associated with the right PPC, such as the ability to sustain attention (for review, see Robertson, 2001) and VSTM functions (Duncan et al., 1999; Wojciulik et al., 2001; Pisella et al., 2004; Malhotra et al., 2005). According to the right hemisphere dominance theory, during right parietal disruption, the left hemisphere should still be able to direct attention to the right side of space. However, although performance was slightly better on the right than on the left during right parietal TMS, our data clearly show a pronounced effect of TMS on both the right and left sides of space, implicating these non-spatially lateralized functions of the right PPC. Using a change detection paradigm, Pisella et al. (2004) showed that neglect patients with right PPC lesions not only displayed poorer detection performance for changes that occurred on the left than the right (in keeping with a spatial orienting deficit) but they also showed a general deficit in detecting a change in location on either side of space.

A deficit in the ability to sustain attention may also play a role in increasing change blindness, particularly if this interacts with the ability to spatially orient attention (Robertson et al., 1998; Robertson, 2001). If right parietal TMS interferes with subjects’ ability to sustain attention over time, they may fail to direct their attention to the location of the change, irrespective of the side of the change. Alternatively, a deficit in sustained attention may interfere with the VSTM component of the change detection task, again irrespective of the side of change.

Because change detection is a complicated process it is unclear which of the functions associated with right PPC are responsible for the effects shown here, indeed it seems likely that they may all play a role. Directing attention to the change may be necessary for encoding the individual images to the level needed to support change detection, or it may be necessary for transfer to VSTM, or for comparing the images in memory. The abilities to maintain the items in VSTM and sustain attention throughout the experiment also seem to be important for change detection. Although future studies, using briefer intervals of TMS at different points in the change detection trial, will help to determine which specific processes are being disrupted by right parietal TMS, it is clear from the current experiment that right PPC plays a critical role in consciously detecting change.

**Right Parietal TMS Effects on Both Sides of Space**

The right hemisphere dominance theory can account for the lateralized effect of TMS, but our data are also in accordance with spatially non-specific functions associated with the right PPC, such as the ability to sustain attention (for review, see Robertson, 2001) and VSTM functions (Duncan et al., 1999; Wojciulik et al., 2001; Pisella et al., 2004; Malhotra et al., 2005). According to the right hemisphere dominance theory, during right parietal disruption, the left hemisphere should still be able to direct attention to the right side of space. However, although performance was slightly better on the right than on the left during right parietal TMS, our data clearly show a pronounced effect of TMS on both the right and left sides of space, implicating these non-spatially lateralized functions of the right PPC. Using a change detection paradigm, Pisella et al. (2004) showed that neglect patients with right PPC lesions not only displayed poorer detection performance for changes that occurred on the left than the right (in keeping with a spatial orienting deficit) but they also showed a general deficit in detecting a change in location on either side of space.

A deficit in the ability to sustain attention may also play a role in increasing change blindness, particularly if this interacts with the ability to spatially orient attention (Robertson et al., 1998; Robertson, 2001). If right parietal TMS interferes with subjects’ ability to sustain attention over time, they may fail to direct their attention to the location of the change, irrespective of the side of the change. Alternatively, a deficit in sustained attention may interfere with the VSTM component of the change detection task, again irrespective of the side of change.

Because change detection is a complicated process it is unclear which of the functions associated with right PPC are responsible for the effects shown here, indeed it seems likely that they may all play a role. Directing attention to the change may be necessary for encoding the individual images to the level needed to support change detection, or it may be necessary for transfer to VSTM, or for comparing the images in memory. The abilities to maintain the items in VSTM and sustain attention throughout the experiment also seem to be important for change detection. Although future studies, using briefer intervals of TMS at different points in the change detection trial, will help to determine which specific processes are being disrupted by right parietal TMS, it is clear from the current experiment that right PPC plays a critical role in consciously detecting change.

**Discussion**

Hitherto it has been unclear whether the parietal activity found in functional imaging experiments on change blindness was critically involved in the detection of change or whether it was merely a consequence of the detection. The data presented here indicate that the right posterior parietal lobe plays a causal role in the conscious detection of change in a ‘change blindness’ paradigm. When rTMS was applied over right PPC, subjects not only took longer to detect changes in either visual field they also exhibited a greater rate of change blindness. By *causal* we do not mean that the right PPC gives rise to conscious detection of changes, but rather that it is a critical component. We believe conscious change detection results from an interaction of occipitotemporal cortex and right dorsolateral prefrontal cortex (DLPF) and the PPC (Beck et al., 2001), all of which are necessary components of conscious change detection. It is well established that regions of the occipitotemporal cortex specialized for analysis of a particular visual attribute are necessary for the perception of that attribute, and recently a similar rTMS paradigm to the one used here has shown that right DLPF also plays a causal role in change detection (Turatto et al., 2004).

A causal role for parietal cortex in change detection is in accordance with what is known about change blindness. Change blindness is thought to occur because the transients that would normally be induced by a change in an image are swamped by the larger transient induced by the offset and onset of the displays (Rensink et al., 1997). Thus, in the absence of a clear bottom-up signal, top-down control must be exerted to find the change. There are at least two processes that one might readily associate with detecting a change in a top-down fashion, both of which are compatible with parietal activity. First, a number of researchers have suggested that in order to detect a change in these paradigms the observer must move attention around the image, selecting candidate objects in turn. Indeed, cuing the location of the image that might change dramatically reduces change blindness (Rensink et al., 1997; Scholl, 2000; Driver et al., 2001). In keeping with this, the region in parietal cortex that we stimulated has been consistently implicated in the allocation of attention to stimuli in the visual field (Kastner and Ungerleider, 2000; Corbetta and Shulman, 2002). Second, it also seems necessary that some form of visual short-term memory (VSTM) is involved in change detection. Although the timing used in change blindness paradigms, including the one presented here, is typically shorter than that used for VSTM tasks, the onset of the second display appears to disrupt iconic memory and so a more durable longer lasting form memory is needed. Parietal cortex has been implicated in VSTM tasks, and in particular two groups (Todd and Marois, 2004; Vogel and Machizawa, 2004) found that activity in the PPC is tightly correlated with an individual’s storage capacity limit. These data suggest that PPC may be involved in determining what does and does not enter VSTM from a visual scene.

**Unilateral Right Parietal Effects**

The previous neuroimaging data on this paradigm showed that conscious change detection was associated with bilateral...
A Role for PPC in Awareness in General

As mentioned in the Introduction, the parietal cortex has been implicated in a variety of tasks that examine visual awareness including binocular rivalry, reversible figures, masking, and change blindness (Kleinschmidt et al., 1998; Lumer et al., 1998; Beck et al., 2001; Dehaene et al., 2001), suggesting that the parietal cortex may play a role in visual awareness in general.

Indeed the processes associated with the PPC are often attributed to awareness more generally, and not just to change detection. If the PPC is responsible for visual short-term storage capacity it is responsible for those items that are ultimately reportable, one of the hallmarks of conscious awareness (Dennett, 1991). Selective attention has also been implicated in visual awareness. It has been known since the ‘early selection’ experiments of Cherry (1953) and Neisser and Becklen (1975) that selectively attending to a portion of the sensory input renders the observer, for the most part, unaware of the unattended information. More recently, similar conclusions have been drawn from inattentive blindness paradigms (Mack and Rock, 1998). But perhaps the strongest evidence for the involvement of the parietal cortex in visual awareness comes from work with neglect and Balint's syndrome patients (Robertson et al., 1997; Kim and Robertson, 2001). Both of the syndromes typically involve lesions to the parietal lobes and result in dramatic deficits in awareness.

It is important to note that we are not arguing that the parietal cortex is the neural locus of consciousness, but rather that the functions associated with parietal cortex, such as attention and VSTM, may be necessary prerequisites to visual awareness. Although the data presented here provide evidence that the parietal lobe is critically involved in consciously detecting a change, given the evidence from neglect patients and the fact that the parietal lobe has been repeatedly implicated (although not causally) in other awareness paradigms (Kleinschmidt et al., 1998; Lumer et al., 1998; Dehaene et al., 2001) it seems likely that parietal activity may be necessary for awareness in general. As such, the data presented here are the first to imply that parietal lobe plays a critical role in determining awareness in neurologically intact subjects. Future research into whether this role extends to other paradigms, designed to investigate awareness, will reveal whether parietal activity is in fact a necessary condition for visual awareness in general.

Notes

This research was supported by a Medical Research Council (UK) grant to N.L. and a Royal Society Research Fellowship (UK) to V.W., and an equipment grant from the Wellcome Trust (UK) to the Institute of Cognitive Neuroscience.

Address correspondence to Dr Diane Beck, Department of Psychology, University of Illinois, 603 East Daniel Street, Champaign, IL 61820, USA. E-mail: dbeck@heschl.cns.uiuc.edu.

References


Rensink RA, O’Regan JK, Clark JJ (1997) To see or not to see: the need for attention to perceive changes in scenes. Psychol Sci 8:368–373.


Prefrontal cortex and basal ganglia contributions to visual working memory

Bradley Voytek1,2* and Robert T. Knight3,4*

1Helen Wills Neuroscience Institute and 2Department of Psychology, University of California, Berkeley, CA 94720

Edited by Robert Desimone, Massachusetts Institute of Technology, Cambridge, MA, and approved September 7, 2010 (received for review May 25, 2010)

Visual working memory (VWM) is a remarkable skill dependent on the brain’s ability to construct and hold an internal representation of the world for later comparison with an external stimulus. The prefrontal cortex (PFC) and basal ganglia (BG) interact within a cortical and subcortical network supporting VWM. We used scalp electroencephalography in groups of patients with unilateral PFC or BG lesions to provide evidence that these regions play complementary but dissociable roles in VWM. PFC patients show behavioral and electrophysiological deficits manifested by attenuation of extrastriate attention and VWM-related neural activity only for stimuli presented to the contralesional visual field. In contrast, patients with BG lesions show behavioral and electrophysiological VWM deficits independent of the hemifield of stimulus presentation but have intact extrastriate attention activity. The results support a model wherein the PFC is critical for top-down intrahemispheric modulation of attention and VWM with the BG involved in global support of VWM processes.

Even a seemingly simple action such as determining which of two bananas is riper requires us to compare real world visual information, such as the color of the banana you are currently looking at in the store, with your memory of the yellowness of the other banana you just put down. This relies in part on visual working memory (VWM), a remarkable ability wherein we construct and hold an internal model of a real-world visual stimulus that we then later compare against another stimulus. In essence, we construct and hold a model of the visual world and compare that model against subsequent inputs from the external world. VWM relies upon an intact and functioning prefrontal cortex (PFC), and damage to this region, such as from stroke, causes VWM impairments (1–3). However, cognitive processes do not localize to specific brain regions per se and a behavior as complex as VWM recruits a distributed network of cortical and subcortical structures (4–8), including the basal ganglia (BG) (9, 10) and visual extrastriate regions (11–13).

Most computational models of VWM rely upon intercommunication between the PFC and the striatum such that memories are maintained via recurrent activation in fronto-striatal loops (14–16). In vivo, working memory maintenance is associated with sustained delay-period activity in the PFC (5, 17) and BG (18), although the BG are thought to play a role in gating information into the PFC to allow it to update representations where necessary (19). Although neurons in both visual extrastriate and the PFC maintain VWM representations during delay periods, PFC neurons encode more information about the stimuli and are more resistant to distractors than visual extrastriate neurons (20). Animal research shows that the BG rapidly learn task-relevant rules and may send relevant, preprocessed information to the PFC for subsequent selection and further processing (21). Anatomically, the BG are situated in an ideal position to mediate cognitive behavior modulated via reinforcement learning (22, 23). Each striatum receives bilateral inputs from many cortical regions including the PFC and visual extrastriate cortex (24), and these inputs converge with dopaminergic afferents from the substantia nigra (25). The striatum is organized in parallel interconnected loops (24, 26, 27) with frontal cortical regions (including the PFC) via the globus pallidus, thalamus, and subthalamic nucleus. From a neuromorphological perspective, each striatum receives PFC input bilaterally from both hemispheres (28) and thus both BG have connections to both PFC hemispheres. The BG are anatomically situated such that they receive inputs from many cortical regions, which may allow them to integrate broadly distributed cortical information such as from the PFC and visual extrastriate cortices (29).

Patients with BG pathology, such as from stroke or Parkinson disease, have deficits in a variety of cognitive learning and switching tasks (30–35) similar to the profile observed in patients with lateral PFC lesions (2). The BG deficits are proposed to be due to a general deficit in the manipulation of internally represented stimuli (36). Human neuroimaging shows that activity in the BG and PFC is associated with individual differences in VWM capacity and that BG activity is specifically associated with filtering out irrelevant distracting information (9, 37), consistent with gating models of BG function and stimulus manipulation. Scalp electroencephalography (EEG) studies show that extrastriate activity increases with the number of items held in VWM up to an individual’s VWM capacity limit and that this activity correlates with individual VWM capacity differences (11). Although sustained PFC activity is associated with working memory maintenance, the role of attention in working memory—both to external stimuli and internal representations of the same—cannot be ignored (38–40). This attention/working memory interrelationship has lead to theories of PFC function that highlight the role of the PFC in information integration (41), with interactions between the PFC and BG necessary to build models of complex rules and behavior from discrete components (42).

Lesion studies in human and nonhuman primates have provided the strongest evidence for a causal relationship between anatomy and function (1, 43). For example, because PFC lesions lead to working memory deficits, the PFC can be said to play an important, necessary role in working memory networks. Research has shown that unilateral PFC lesions cause lateralized deficits in top-down modulation of visual attention (44, 45). These deficits manifest as errors in target detection specifically to targets that appear in the contralesional hemisphere. These target-detection errors are associated with attenuation of visual extrastriate event-related potentials (ERPs), including the early visual N1. This early latency ERP (100–200 ms after stimulus onset) is modulated by attentional state and is enhanced in the stimulated visual hemisphere during lateralized attentional allocation (46) and attenuated in the damaged hemisphere in the presence of a unilateral PFC lesion (44). Because EEG studies provide a direct neural measure of working memory load (11) and attentional allocation (46, 47), we used

[Author contributions: B.V. and R.T.K. designed research; B.V. performed research; B.V. analyzed data; and B.V. and R.T.K. wrote the paper.]
EEG to assess top-down cognitive deficits associated with unilateral lesions on a within-hemisphere basis. We hypothesized that the BG plays a visual-field independent role in VWM updating and learning. Conversely, we predicted that the PFC has an executive role in VWM maintenance, attentional control, and top-down facilitation of visual extrastrate cortices on a within-hemisphere basis. Thus, we examined two groups of patients with either unilateral PFC or BG lesions (Fig. 1) performing a lateralized VWM task (Fig. 2A) while recording scalp EEG. By making use of a lateralized visual design, we took advantage of the inherent contralateral organization of the mammalian visual system wherein visual input from the right visual field enters the left visual cortex and vice versa. In Fig. 2B we illustrate how a patient with a left PFC lesion viewing a stimulus in the left visual hemifield would receive the visual input into the intact cerebral hemisphere; that same patient viewing a right hemifield stimulus would receive the information in the damaged hemisphere, leading to behavioral deficits mainly in the contralesional visual field. By combining a lateralized VWM design with scalp electrophysiology in patients with unilateral brain lesions, we reveal distinct contributions of the PFC and BG to VWM maintenance and examine the role of each region in top-down modulation of extrastrate activity.

Results

Behavioral Effect of Lesions. In a three-way ANOVA including all three groups, we found a main effect of load on accuracy such that all groups were less accurate with increasing memory load ($F_{2,42} = 344.45, P < 0.0005$). There was also a three-way interaction between group, memory load, and hemifield of presentation ($F_{4,42} = 12.47, P < 0.0005$). We performed ANOVAs comparing performance between and within the patient groups to examine the nature of this three-way interaction. Accuracy results are summarized by the group x hemifield effect (collapsed across load) in Fig. 2C ($F_{2,21} = 10.17, P = 0.001$; Table S1 contains all accuracy results).

In a comparison between controls and PFC patients, we found a three-way interaction ($F_{2,32} = 14.41, P < 0.0005$). Consistent with our hypothesis, there was a significant group x hemifield interaction ($F_{1,16} = 16.17, P = 0.001$). The PFC patients showed a significant hemifield x load interaction ($F_{1,5} = 37.46, P = 0.002$) and a main effect of hemifield ($F_{1,5} = 29.21, P = 0.003$) wherein they were less accurate overall for contralesional stimuli. There was no effect of hemifield in the control group ($P > 0.5$). These results suggest that the hemifield x group interactions were driven by deficits in the PFC group in response to contralesional stimuli. This was confirmed in an analysis comparing accuracy by hemifield between groups wherein PFC patients were impaired for contralesional stimuli compared with controls ($P = 0.026$). In comparing controls and BG patients, we also found a three-way interaction ($F_{2,32} = 5.40, P = 0.010$). Unlike the PFC group, BG patients showed no main effect of hemifield on performance ($F_{1,5} < 1.0$) and were impaired compared with control subjects in both hemifields (ipsi: $P = 0.046$; contra: $P = 0.025$). Analyses of other behavioral measures, including response bias, reaction times, and hit rates (SI Results), indicate that the patient behavioral deficits arise from errors in working memory rather than from motoric deficits or systematic response biases.

Research suggests that the BG are critical in learning behavioral requirements (8, 21, 32, 47, 48). Therefore, we examined the temporal evolution of behavioral performance across the first 100 trials (Materials and Methods). In comparing controls to PFC patients, there was a main effect of trial on performance ($F_{3,48} = 3.14, P = 0.034$) and a main effect of group ($F_{1,16} = 15.88, P = 0.001$) but no group x trial number interaction, which suggests that both groups improved across the first 100 trials and that the PFC group performed worse than controls. In contrast, when we compared the BG group to controls, we found a significant group x trial number interaction ($F_{3,48} = 3.64, P = 0.019$). Although both the BG and control groups showed a main effect wherein behavior improved across trials (BG: $F_{3,15} = 5.13, P = 0.012$; controls: $F_{3,33} = 2.95, P = 0.047$), only the BG group showed a significant deficit during the initial trials (Fig. 2D, trials 1-25 compared with 26-51, $P = 0.001$; $P > 0.05$ for all other pair-wise comparisons between successive trial bins for both BG and control groups). It is important to note that although the behavioral deficits in the BG group were exaggerated during the first 25 trials, they continued to perform worse than controls in all time bins examined ($P < 0.05$ for all other binned analyses). This accuracy deficit was not due to prolonged reaction times extending through the end of the trial, as there was no effect of trial number on number of misses ($F_{3,35} < 1.0$).

Electrophysiological Effects of Lesions. We examined the effects of PFC and BG lesions on delay period EEG activity. We replicated previous findings that in normal subjects (11) the amplitude of contralateral delay activity (CDA) (Materials and Methods, Fig. 3, and Fig. S1) increases with memory load in a three-way ANOVA including all three groups ($F_{2,42} = 18.84, P < 0.0005$); visual inspection of the CDA time courses (Fig. 3) showed that patient CDA amplitudes for contralesional stimuli are abnormal for both groups and that this is reflected in a different scalp topogra-

---

Fig. 1. Patient lesion reconstruction. Structural MRI slices illustrating the lesion overlap across the two patient groups (color represents number of subjects with a lesion at that voxel). For the PFC group ($n = 6$), mean lesion volume was 58.6 cm$^3$ and maximal lesion overlap ($\sim$50%) was in Brodmann areas 6, 8, 9, and 46 centered in the middle frontal gyrus and including portions of the inferior and middle frontal gyrus in some patients. For the BG group ($n = 6$), mean lesion volume was 9.7 cm$^3$ and maximal lesion overlap was in the putamen and encompassed the head and body of the caudate as well as the globus pallidus in some patients. All lesions are normalized to the left hemisphere for comparison; however, two patients in each group had right hemisphere lesions. Software reconstructions were performed using MRicrop (53).
phy and a general loss of top-down facilitation as indexed by increased alpha power in posterior electrodes in the lesioned hemisphere (detailed analyses are in SI Results; see Fig. S2). For this reason, we will refer to the abnormal patient visual cortical ERPs as “sustained negativity” and not CDA. In the three-way ANOVA, there was a significant quadratic three-way interaction between group, memory load, and hemifield of presentation ($F_{2,21} = 3.74, P = 0.041$), driven by the effects of the lesion leading to the abnormal patient contralesional sustained negativity. This was reflected in a significant group × hemifield effect ($F_{2,21} = 6.65, P = 0.006$; Table S1 contains all CDA results).

In comparing PFC patients to controls, there was a significant group × hemifield interaction ($F_{1,16} = 7.45, P = 0.015$), although neither group showed a significant effect of hemifield in separate ANOVAs of each group (controls: $F_{1,11} = 2.95, P = 0.11$; PFC: $F_{1,5} = 3.21, P = 0.13$). This interaction was driven by a crossover effect wherein CDA amplitude is reduced in the PFC group for ipsilesional stimuli ($P < 0.0005$) but is higher for contralesional stimuli ($P < 0.05$). In separate planned contrasts, we examined the effects of hemifield of presentation on CDA amplitude within the patient groups for ipsilesional and contralesional stimuli. When this analysis was done in the control group, effect of load was significant for both hemifields (left: $F_{2,22} = 7.37, P = 0.004$; right: $F_{2,22} = 6.44, P = 0.006$). In the PFC group there was a significant effect of load for ipsilesional stimuli ($F_{2,10} = 4.17, P = 0.048$), driven by an effect wherein CDA amplitude increased from one to two items ($P = 0.003$) but not from two to three items ($P = 0.69$), similar to the pattern seen in control subjects (one to two: $P < 0.0005$; two to three: $P = 0.13$). As predicted due to the loss of top-down facilitation, for contralesional stimuli there was no effect of load ($F_{2,10} < 1.0$) in the PFC group.

In an analysis comparing CDA between the BG and control groups, there was also a significant group × hemifield interaction ($F_{1,16} = 13.20, P = 0.002$), although neither group showed a significant effect of hemifield in separate ANOVAs of each group (controls: $F_{1,11} = 2.95, P = 0.11$; BG: $F_{1,5} = 3.39, P = 0.13$). Just as with the comparison between controls and PFC patients, this interaction appears to be driven by a crossover effect wherein CDA amplitude is reduced in the BG group for ipsilesional stimuli ($P < 0.0005$) but is higher for contralesional stimuli ($P < 0.0005$). In contrast to PFC patients, in an analysis of hemifield of presentation on CDA amplitude within the BG group there was no effect of load for either ipsilesional or contralesional stimuli (ipsilesional: $F_{1,5} = 1.52, P = 0.27$; contralesional: $F_{1,5} < 1.0$).

In a final analysis, we examined the effects of lesions on the attention-related N1. Because of the relatively rapid nature of our task and the brief stimulus presentation time (180 ms), we hypothesized that the observed behavioral deficits in the patient groups could be partly due to the effects of the lesion on attentional control. In a three-way ANOVA including all three groups, we found a main effect of load on N1 amplitude such that increasing perceptual load lead to more negative N1 amplitude ($F_{2,42} = 23.54, P < 0.0005$). There was also a three-way interaction between

---

**Fig. 2.** Behavioral paradigm and performance. (A) Diagram of task design. (B) For a patient with a left unilateral PFC lesion, as illustrated here, stimuli that appear in the left visual hemifield are ipsilesional, and the visual information selectively enters the intact cerebral hemisphere, whereas stimuli that appear in the right visual hemifield are contralesional and selectively enter the damaged hemisphere. (C) Plots of average behavior by group and hemifield. Patients with unilateral PFC lesions performed as well as controls when stimuli were presented ipsilesionally but were impaired for contralesional stimuli. In contrast, patients with unilateral BG lesions performed more poorly overall, regardless of the hemifield of stimulus presentation. (*$P < 0.05$ compared with controls, **$P < 0.0005$, error bars represent SEM). (D) Control subjects and PFC patients performed equally well across trials. BG patients were significantly impaired in early trials.

---

**Fig. 3.** Electrophysiological analyses (group grand averages). (A) Average CDA for control subjects collapsed across hemifield. For controls, CDA amplitude increases with memory load (*main effect of load, $P < 0.0005$). (B) Summary of CDA findings for ipsilesional stimuli in the two patient groups (shown in detail in C-F) and for left hemifield stimuli for controls. For ipsilesional stimuli (C and E), both controls and the PFC group show a significant effect of memory load on CDA (*$P < 0.05$, error bars represent SEM) that is not seen in the BG group ($ns$, not significant). For contralesional stimuli (D and F), the relationship between CDA and load is abolished in both patient groups. Both patient groups generated a sustained negative shift for contralesional stimuli that was not sensitive to VWM load (SI Results).
group, load, and hemifield of presentation ($F_{4,42} = 5.63, P = 0.001$; Table S1 contains all N1 results). The N1 results are summarized by the group x hemifield effect in Fig. 4. In separate analyses comparing controls with PFC patients and controls with BG patients, we also observed significant three-way interactions in both comparisons (PFC: $F_{2,32} = 8.89, P = 0.001$; BG: $F_{2,32} = 5.78, P = 0.007$). The control versus BG interaction arose from a group x load interaction ($F_{2,32} = 8.01, P = 0.002$) that was mediated by group differences for one-item arrays wherein BG patients had lower N1 amplitudes ($P = 0.024$). These differences disappeared for higher loads (two items: $P = 0.41$; three items: $P = 0.23$). In a post hoc analysis of the control versus PFC interaction, we examined the a priori hypothesis that PFC patients would have attention deficits in response to contralesional stimuli. Looking across all memory loads, there was no significant difference in N1 amplitude between groups for ipsilesional stimuli ($P = 0.43$). However, N1 amplitude was attenuated in the PFC group for contralesional stimuli ($P = 0.003$). As a comparison, there were no differences between controls and BG patients for either hemifield (ipsilesional: $P = 0.42$; contralesional: $P = 0.24$).

**Discussion**

These results highlight the distinct roles of the PFC and BG in VWM maintenance. We tested two separate groups of patients with either unilateral PFC or unilateral BG lesions, and age-matched controls while they performed a lateralized VWM task. By making use of a lateralized VWM design with a scalp EEG, we were able to take advantage of the anatomical separation of visual inputs into the neocortex by visual hemifield of presentation and examine the effects of lesions on top-down VWM maintenance. This lesion by hemifield design allowed us to assess behavioral and electrophysiological responses on a within- and between-subjects basis. That is, because patients’ lesions were unilateral, we could assess differences in response to contralesional stimuli versus ipsilesional stimuli. Previous studies have shown this to be an effective means in highlighting top-down attention deficits associated with PFC lesions (44).

We found that patients with unilateral PFC lesions performed just as well as controls for ipsilesional stimuli and that accuracy dropped only when stimuli were lateralized to the contralesional hemifield. When we examined the evolution of performance over time, we found that PFC patients performed as well in the first few trials as they did in later trials, similar to the results of normal control subjects. In contrast to PFC patients, the BG group performed worse than controls regardless of the hemifield of stimulus presentation. Furthermore, BG patients performed worse during the initial 25 trials than they did in later trials. This was despite the fact that subjects were able to explicitly restate the rules and requirements of the task when questioned before the experiment began. The fact that the number of misses did not change across early trials argues against the possibility that this learning effect is an artifact due to BG patients making more responses outside of the response window. Interestingly, although patients in the BG group understood the task, they had difficulties initially engaging the neural mechanisms necessary to correctly perform it. The stabilization of behavioral performance at ~30 trials suggests that the BG group adopted a new strategy for performing the task.

Previous EEG research using a paradigm similar to ours in normal subjects has shown that delay-period CDA activity increases in magnitude with increasing memory load up to a subject’s VWM capacity (11). We replicated this scaling effect for VWM load in our control group and extended this work to show that individuals’ CDA amplitudes at each load correlate with their later behavioral performance (SI Results and Fig. S3). These results suggest that CDA accurately indexes behavioral performance. Within our PFC group, we found similar CDA effects for ipsilesional stimuli only. That is, the PFC group, as with controls, showed an increase in CDA from one- to two-item loads. CDA amplitude in response to ipsilesional stimuli also correlated with later behavioral performance. Similar to their behavioral performance, patients with unilateral PFC lesions showed no scaling of CDA amplitude in response to contralesional stimuli nor did CDA amplitude correlate with later behavioral outcomes. In contrast to BG patients and controls, we found that PFC patients also had attenuated attention-dependent N1 amplitudes within the lesioned hemisphere only for contralesional stimuli. Previous studies have shown that posterior visual association cortex N1 amplitude is modulated by voluntary attention under top-down PFC control (46). Combined with the impaired CDA to contralesional stimuli, these electrophysiological results suggest that PFC lesions lead to an overall executive functioning deficit affecting multiple cognitive domains within the damaged hemisphere. That is, PFC damage results in a loss of top-down facilitation of visual extrastriate cortex during the working memory delay period, resulting in attention and VWM maintenance deficits contributing to poorer behavioral performance. Although we observed a strong brain/behavior correlation (SI Results and Fig. S3), previous research has found that the best predictor of behavioral performance is the load difference in CDA amplitudes rather than the actual amplitudes themselves (49).

Notably, both patient groups showed a pronounced sustained negativity for all contralesional stimuli that was independent of VWM load. Contrary to our findings in the PFC group, patients with unilateral BG lesions showed no load-dependant scaling of CDA amplitudes for either ipsilesional or contralesional stimuli. This was despite the fact that N1 amplitudes within the BG group were intact, even in the lesioned hemisphere. Although patients with unilateral BG neuropathology show deficits in attentional set shifting and general cognitive flexibility (19, 30, 50), the BG do not appear to play a critical role in the rapid allocation of visual attention. Rather, our BG patients show intact electrophysiology related to attentional allocation, whereas our PFC group has attentional impairments for contralesional stimuli. This suggests that the BG play a critical visual-field independent role in VWM maintenance but are not critical for top-down facilitation of early visual extrastriate cortex attentional processes. This adds further support to the specificity of the PFC in intrahemispheric control of top-down visual attention in the visual extrastriate cortex. The behavioral and VWM maintenance impairments in the BG group cannot be explained by a general effect of larger lesion volumes, as overall lesion volumes were significantly smaller in the BG group compared with PFC patients ($P = 0.024$). The fact that BG patients are especially impaired during the first 25 trials provides support for the hypothesis that the BG are critical for rule-based learning and implementation (31).
We hypothesize that unilateral BG lesions lead to a deficit in updating VWM representations, which in turn leads to a degradation in the fidelity of the VWM representation in fronto-extrastriate networks. The deficits may also be due in part to a failure to filter out irrelevant information (9, 37). Even though our protocol had no explicit distractors, the BG have been reported to play an important role in filtering out irrelevant information, and, thus, the stimulus information that is to be retained may be degraded over time due to increased ambient noise from the visual world. These results suggest that the PFC plays a broader role in executive functioning including both top-down attentional control and VWM maintenance, whereas the BG are more directly related to global VWM maintenance processes, extending the role of the BG outside the motor domain. Several studies have reported VWM deficits after lateral PFC damage (1–3). In contrast, BG lesions lead to a VWM behavioral impairment associated with maintenance deficits despite intact attention mechanisms. It is important to note that, although patients performed worse than controls in our study, the NI and CDA deficits we report were from our examination of correct trials only. Thus, despite their pathological electrophysiological responses, patients performed the task well, albeit with impairments. This suggests that there are other mechanisms related to correct behavioral outcomes, possibly including functional reorganization, whereby the unilaterality of the lesions allows other intact cortical structures to compensate for the damaged regions (52).

Materials and Methods

Participants. All subjects gave informed consent approved by the University of California, Berkeley, CA, Committee for Protection of Human Subjects and the Department of Veterans Affairs Northern California Health Care System Human Research Protection Program. Control subjects were matched to patients by age and education. Because there were neither age nor education differences between PFC and BG groups (P > 0.50 both comparisons), we compared the results of each group separately to the combined group of 12 controls. For both patient groups, testing took place at least 6 mo after the date of the stroke; lesion etiology was either cerebrovascular accident or lesion of 25 trials moving in one-trial steps looking at overall behavioral performance regardless of memory load or hemifield of stimulus presentation. For analyses on learning, we ran a repeated measures ANOVA with trial number as the within-subjects factor and memory load and hemifield as the between-subjects factors. Comparisons between control and patient results were such that responses to left hemifield stimuli in controls were compared against ipsilesional responses in patients and right hemifield stimuli were compared with contralesional responses. To test the effects of learning on behavioral performance, we calculated a sliding window d’ measure across blocks of 25 trials moving in one-trial steps looking at overall behavioral performance regardless of memory load or hemifield of stimulus presentation. For analyses on learning, we ran a repeated measures ANOVA with trial number as the within-subjects factor and memory load and hemifield as the between-subjects factors. Comparisons between control and patient results were such that responses to left hemifield stimuli in controls were compared against ipsilesional responses in patients and right hemifield stimuli were compared with contralesional responses. To test the effects of learning on behavioral performance, we calculated a sliding window d’ measure across blocks of 25 trials moving in one-trial steps looking at overall behavioral performance regardless of memory load or hemifield of stimulus presentation. For analyses on learning, we ran a repeated measures ANOVA with trial number as the within-subjects factor and memory load and hemifield as the between-subjects factors. Comparisons between control and patient results were such that responses to left hemifield stimuli in controls were compared against ipsilesional responses in patients and right hemifield stimuli were compared with contralesional responses.

Electrophysiological Recording. Subjects were tested in a sound-attenuated EEG recording room at the University of California, Berkeley, CA. EEG data were collected using a 64 × 8 channel BioSemi ActiveTwo (51) amplifier sampled at 1,024 Hz. Horizontal eye movements (HEOG) were recorded at both external canthi, and vertical eye movements (VEOG) were monitored with a left inferior eye electrode and a fronto-polar electrode. Subjects were instructed to maintain central fixation and to respond using the thumb of their unaffected, ipsilesional hand. All data were referenced offline to the average potential of two earlobe electrodes and analyzed in MATLAB (R2009a) using custom scripts and the EEGLAB toolbox (52) and SPSS (Rel. 18; SPSS Inc.). Only correct trials were included in EEG analyses.

Behavioral Task. Subjects were presented with a memory array consisting of a set of one, two, or three colored squares (180 ms presentation; equiprobable presentation of each set size to either the left or right visual hemifield). After a 900-ms delay, a test array of the same number of colored squares appeared in the same spatial location. Subjects were instructed to manually respond to indicate whether the test array was the same color as the initial (memory) array. Behavioral accuracy was assessed using a d’ measure of sensitivity, which takes into account false alarm rate to correct for response bias. To avoid mathematical constraints in the calculation of d’, we applied a standard correction procedure wherein, for any subjects with a 100% hit rate or 0% false alarm rate, performance was adjusted such that 1/(2N) false alarms were added or 1/(2N) hits subtracted where necessary.

Data Analysis. All statistical analyses on behavior and ERP were first assessed using repeated-measures ANOVAs with group membership (control, PFC, or BG) as the between-subjects factor and memory load and hemifield as the within-subjects factors. Comparisons between control and patient results were such that responses to left hemifield stimuli in controls were compared against ipsilesional responses in patients and right hemifield stimuli were compared with contralesional responses. To test the effects of learning on behavioral performance, we calculated a sliding window d’ measure across blocks of 25 trials moving in one-trial steps looking at overall behavioral performance regardless of memory load or hemifield of stimulus presentation. For analyses on learning, we ran a repeated measures ANOVA with trial number as the within-subjects factor and memory load and hemifield as the between-subjects factors. Comparisons between control and patient results were such that responses to left hemifield stimuli in controls were compared against ipsilesional responses in patients and right hemifield stimuli were compared with contralesional responses.

ERPs were analyzed on bandpass filtered (0.1–20 Hz) data resampled to 256 Hz using a 100-ms prestimulus baseline. Blinks and saccades were identified on raw VEOG and HEOG channels, respectively, and verified with scalp topographies. Events with incorrect or no response, blinks, or saccades were removed from all analyses. CDA values were calculated as the mean amplitude difference from 300 to 900 ms between a group of extrastriate electrodes contralateral to the stimulus and a group ipsilateral to the stimulus. Thus, for controls, CDA for a right hemifield stimulus was calculated as the average of left minus right extrastriate activity from 300 to 900 ms. For patients, CDA was calculated in the same manner but was analyzed relative to the lesion such that, for patients with left hemisphere lesions, CDA for right hemifield stimuli was classified as contralesional and CDA for left hemifield stimuli was classified as ipsilesional (and vice versa). We classified patient behavioral data in the same manner. NI amplitude was calculated as the maximum negative amplitude over the extrastriate cortical contralateral to the hemifield of stimulus presentation from 100- to 200-ms poststimulus onset.

ACKNOWLEDGMENTS. We thank Matar Davis and Lisa Tseng for assistance with data collection and study design, Richard B. Ivry and Michael Silver for scientific comments on previous versions of this manuscript, Clay Lawson for lesion reconstruction, and Donatella Scabini for patient recruitment and delineation. This work was supported by American Psychological Association Diversity Program in Neuroscience Grant 5-T32-MH18882 (to B.V.) and National Institute of Neurological Disorders and Stroke Grants NS21135, NS21135-2251, and PO40813 (to B.V. and R.T.K.).

Unreinforced Spatial (Latent) Learning is Mediated by a Circuit That Includes Dorsal Entorhinal Cortex and Fimbria Fornix

Stephane Gaskin* and Norman M. White

ABSTRACT: The relationship of the entorhinal cortex (EC) and fimbria-fornix (FF) in unreinforced spatial (latent) learning was studied using the conditioned-cue-preference task on an eight-arm radial maze. The maze was turned before every trial to eliminate the use of local cues. During three pre-exposure sessions, food-deprived rats explored the center platform and two adjacent arms of the maze. Since most of the same cues were visible from both arm locations, discriminating them required spatial learning. The rats were then alternately confined to the end of each arm over several days: one arm always contained food, the other was empty. Finally, the rats were allowed free access to both arms with no food present. Normal rats spent more time in their food-paired than in their unpaired arms showing that they learned to discriminate between the arm locations. Bilateral micro-injections of muscimol into the dorsal, but not into the ventral EC, given before the pre-exposure sessions only, impaired the discrimination. The discrimination was also impaired in rats with unilateral lesions of FF and contra-lateral injections of muscimol into the dorsal EC given before the pre-exposure sessions. Ipsilateral FF lesions and entorhinal inactivation had no effect. These results indicate that the acquisition of information during unreinforced exploration of a novel environment requires an intact circuit involving the dorsal EC and fimbria fornix. Together with previous reports, that this form of learning does not require a functional hippocampus, (Gaskin et al. (2005) Hippocampus 15:1085–1093) the findings also suggest that the acquisition of certain kinds of unreinforced information by this circuit is independent of the hippocampus.

KEY WORDS: radial maze; discrimination; inactivation; muscimol; conditioned-cue-preference; neural systems; parahippocampal

INTRODUCTION

Although the hippocampus is known to be important for spatial learning and memory (O’Keefe and Dostrovsky, 1971; O’Keefe and Nadel, 1978; Olton and Papas, 1979; Morris et al., 1982; Barnes, 1988), findings that latent spatial learning is not impaired by lesions (Kimble and BreMiller, 1981) or inactivation (Gaskin et al., 2005) of the dorsal hippocampus suggests that further examination of this generalization is required. In latent learning paradigms (Blodget, 1929; Tolman and Honzik, 1930), rats explore a maze of some kind with no reinforcers present. When a reinforcer is introduced, rats that have explored the maze learn to find the reinforcer faster than rats that have not previously explored the maze. This effect shows that some form of learning occurs during unreinforced exploration. Since no reinforcers are present, no conditioned or instrumental responses can be acquired. Rather, information is acquired about the relationships among the stimuli in the maze environment, possibly taking the form of stimulus-stimulus (S-S) associations. The aggregation of these relationships may constitute a “spatial map” of the maze environment (O’Keefe and Dostrovsky, 1971; O’Keefe and Nadel, 1978).

Although lesions (Kimble and BreMiller, 1981) or inactivation (Gaskin et al., 2005) of the dorsal hippocampus have no effect on unreinforced learning of the spatial map, they do impair learning the location of a reinforcer when it is introduced into the maze (Kimble and BreMiller, 1981; Chai and White, 2004; White and Gaskin, 2006). Such differences between latent acquisition of the spatial map and learning the location of the reinforcer can only be observed when the two kinds of learning are explicitly separated by the experimental procedure, as in the latent learning paradigm. In most spatial learning situations the reinforcer is always available so the two kinds of information are acquired simultaneously. Since both kinds of information are required to observe spatial learning, performance on such tasks is impaired by hippocampal lesions.

Interestingly, the effects of fimbria-fornix (FF) lesions are opposite to those of the hippocampus. FF lesions impair latent spatial learning but have no effect on learning to find the reinforcer if made after latent spatial learning (Chai and White, 2004). Therefore, FF lesions also impair performance on spatial learning tasks in which the two kinds of learning occur simultaneously (Hirsh and Segal, 1972). The latent learning paradigm reveals that the impairments produced by FF and hippocampus lesions are due to interference with learning different kinds of information.

Since it is unlikely that the FF mediates spatial learning by itself, the present study addressed the question of which other nonhippocampal brain area(s) might work with FF to acquire this kind of information. One way to approach this question is to ask which brain area(s) other than the hippocampus has direct connections with FF and is known to be involved in spatial learning.

One such possibility is the entorhinal cortex (EC). A subset of FF fibers originating in the medial septal area (Swanson, 1977) course through the hippocampus.
and terminate in EC (Gaykema et al., 1990). However, lesions of the EC have resulted in spatial memory deficits in some studies, but not in others. Two factors may account for these differences. First, the studies reporting impaired spatial learning in rats with lesions of the EC used electrolytic, radio-frequency, or aspiration lesions, which destroy both local neurons and fibers of passage (Olton et al., 1978; Schenk and Morris, 1985; Hagan et al., 1992; Cho and Kesner, 1996; Good and Honey, 1997; Parron and Save, 2004). The studies that found no effect used neurotoxic lesions (Bannerman et al., 2001; Galani et al., 2002; Oswald et al., 2003). Second, the lesions in studies reporting a deficit included both the medial and lateral parts of the EC and extended into the subiculum. In contrast, most studies in which no deficits in spatial learning were observed after EC lesions reported no damage, or very little damage to the subiculum. These findings suggest that spatial learning is intact in animals with EC lesions provided that information flow through the EC is unaffected.

Recently, Fyhn et al. (2004) reported that neurons involved in processing spatial information are located in the dorsolateral band (Dolorfo and Amaral, 1998) of the EC. This area is also the primary recipient of visuospatial information from the neocortex (Burwell, 2000). This finding led Steffenach et al. (2005) to lesion the dorsolateral band of EC. These workers found that fiber-sparing (NMDA) lesions of the dorsolateral but not ventromedial band of the EC impaired spatial learning.

In the present study, we examined the contributions to latent spatial learning of the EC and FF using a conditioned cue preference (CCP) paradigm in a radial maze. The paradigm includes three phases. In the pre-exposure phase rats explore the maze with no food in either arm. This is when they are thought to acquire the latent spatial information. Since there is no food on the maze, they obviously cannot learn its location.

In the training phase, the rats are confined on the ends of the two adjacent-arms on alternate days. One of the arms always contains food; the other is always empty. In this phase the rats learn the location of the food. However, spatial learning is limited because they are confined in the maze arms and prevented from moving around on the maze (Sutherland and Dyck, 1984; Sutherland, 1985; Alyan, 1994; White and Ouellet, 1997; Terrazas et al., 2005) making it unlikely that the acquisition of a spatial map of the environment occurs during this phase.

In the test phase, the rats are placed on the center platform of the maze and allowed free access to both arms, neither of which contains food. Normal, hungry rats forage for food, choosing to spend more time in the food-paired than in the unpaired arm, showing that they have learned to discriminate between the two spatial locations.

In Experiment 1, we tested the effects of temporary inactivation with muscimol of either the ventral or dorsal EC during the pre-exposure sessions. In Experiment 2, we used a disconnection paradigm to test the hypothesis that latent spatial learning during pre-exposure is mediated by a neural circuit that includes the dorsal EC and FF.
enamel insulation removed from 0.8 mm at the tip by careful scraping. To destroy the entire FF unilaterally, two lesions were made in each rat, both on the same side of the brain. The coordinates were 1.2 mm posterior, 0.7 and 2.4 mm lateral to bregma, and 5.0 mm below the skull surface. The ground was a rectal probe. A 1.5-mA current was passed for 20 s. For eight of these rats the entorhinal cannulas and FF lesions were on opposite sides of the brain (the contralateral disconnection group). In the eight remaining rats the two lesions were on the same side of the brain (the ipsilateral control group). The side of the brain on which the cannulas were implanted and the lesions made were counterbalanced in both groups.

Procedure

All rats were allowed to recover from surgery for 7 days. At the start of the procedure food was removed from all home cages. All rats were handled daily for 4 days. During each daily session, groups of 5–8 rats were put into a large plastic box with sawdust on the floor. Each rat was picked up in turn five times and handled for 1 min each time. The rats were then returned to their home cages where they were given 2–3 food pellets and 10 pieces of Kellogg’s Froot Loops cereal. After the handling period, the rats received only food pellets in their home cages for the rest of the experiment. They were weighed daily and given sufficient food to maintain 80–85% of their initial free-feeding weights.

The radial maze was rotated by two arm positions to the left before the start of training each day to prevent learning based on local cues. The maze was cleaned with germicidal deodorant solution before each trial for each rat. The rats’ home cages were placed in a mobile cage holder and wheeled to a waiting area outside the testing room. For each trial, the experimenter placed a rat on the maze, left the room and observed the rat’s behavior on the video monitor.

Each rat was assigned a unique pair of adjacent arms on the radial maze. One of these arms was randomly designated as the food-paired arm and the other as the unpaired arm. The procedure had three phases: pre-exposure, training, and testing. During the pre-exposure phase, each rat’s assigned arms were open and the other arms were blocked. There was no food on the maze. On each of the three pre-exposure days, each rat received bilateral injections of muscimol or saline (see below) and was placed on the center platform of the maze 30 min later. Each rat explored the center platform and its two assigned arms for 10 min on each of 3 consecutive days. On the next day the training trials began. No injections were given. On each trial, the rats were confined using a wooden block with a panel attached on the end of their assigned food-paired arm for 30 min with 50 Froot Loops on one day and on the end of their unpaired arm for 30 min with no food on the next day. The order of exposure to the food-paired and unpaired arms was counterbalanced within each group and remained the same for each rat over all training trials. All rats received four 2-day trials.

On the day following the final training trial the test trial was given. No injections were given. Each rat was placed on the center platform with its two assigned arms open and the other arms blocked for 15 min. There was no food on the maze. The times at which each rat entered and exited each arm were recorded. A rat was considered to be in an arm or to have left the arm if its front feet crossed the threshold between the arm and the center platform. The total time spent in each arm was calculated.

Intracranial Injections

All rats were given intracranial injections of muscimol or saline through their implanted guide cannulas before unreinforced pre-exposure trials only. Each rat was taken individually in its cage from the waiting area to an injection room. The obdurator wires were removed and injection cannulas 13 mm long (protruding 2 mm beyond the ends of the 11 mm guide cannulas) were inserted into the guide cannulas. Some rats received injections of 0.5 μl of muscimol hydrobromide (1 μg/ml dissolved in 0.9% sterile saline) at the rate of 0.3 μl/min. Saline injections consisted of the 0.9% sterile saline vehicle injected using the same parameters. Following the injections the rats were returned to their home cages in the waiting area for 30 min. They were then placed on the maze for the pre-exposure trial.

Histology

After all behavioral testing was complete, the rats were euthanized with a lethal dose of 30% chloral hydrate. Methylene blue (0.5 μl) was injected through the same guide and inner cannulas used for the infusions of muscimol or saline. The brains were then perfused with 0.9% saline followed by a 10% formalin-saline solution. The brains were excised and stored in 10% formalin for at least 3 days. The brains were then frozen and cut into sections 30 μm thick. Sections through the area of the cannulas were mounted on glass slides, stained with cresyl violet and examined microscopically.

EXPERIMENT 1

In this experiment, the effect on arm preference during the test trial produced by temporary inactivation of dorsal and ventral EC during unreinforced exploration of the maze (pre-exposure) was examined. Of the 31 rats in this experiment, 7 were assigned to the dorsal entorhinal group (DENT) and 8 to the ventral entorhinal group (VENT). The remaining 15 rats were assigned to a combined saline control group (SAL) consisting of 7 rats with dorsal and 8 rats with ventral EC cannulas.

Results and Discussion

Figure 1 shows the cannula placements for the rats that were used in the statistical analysis in this experiment. There is considerable controversy about the radius of neural inactivation...
produced by muscimol injections. Estimates range from 1.7 mm around the tip of the injection cannula in the cerebral cortex using 2-deoxyglucose uptake as a measure of neural activity (Martin, 1991), to more than 3.0 mm in nucleus basalis magnocellularis using electrophysiological and autoradiographic methods (Edeline et al., 2002). Although these values do not necessarily apply to the present injections, it is clear that our procedure does not permit a distinction between the dorsolateral and ventromedial bands of the EC as originally described by Dolorfo and Amaral (1998) and functionally differentiated with respect to spatial learning by Steffenach et al. (2005). However, since the same volume was used for both the ventral and dorsal injections, it is likely that they affected the same volume of tissue. In the case of the dorsal placement it seems likely that the inactivated area included the dorsal posterior pole of the EC, a part of the dorsolateral band thought to be critical for spatial learning (Steffenach et al., 2005). The ventral placements probably did not affect this area. Therefore, although they are anatomically less precise, the present findings are consistent with the difference between the effects of dorsolateral vs. ventromedial band lesions on spatial learning found by Steffenach et al. (2005). It can also be noted that although it lacks anatomical precision, the temporary inactivation technique permits discrimination of function during the acquisition of different kinds of information at different times in the experiment.

The behavioral results are shown in Figure 2. Planned comparisons based on a two-way analysis of variance (ANOVA) (Keppel, 1991, p 112; Tabachnick and Fidell, 2001, p 50; Howell, 2002, p 375–376) showed that rats in the SAL and VENT groups preferred their food-paired arms \( F(1, 28) = 18.871, P < 0.001 \) and \( F(1, 28) = 4.21, P < 0.05 \), respectively. However, the rats in the DENT group did not exhibit a preference \( F(1, 28) = 0.39 \).

The hypothesis that rats acquire spatial information while exploring the maze during the pre-exposure sessions implies that the absence of a CCP in the rats that received the dorsal injections could have been due to reduced exploration during pre-exposure. To examine this possibility two measures of exploration, the total number of entries into the arms (Fig. 3a) and the total time spent on the arms (Fig. 3b) during the three pre-exposure sessions by the rats in all three groups were compared. A one-way ANOVA on the numbers of arm entries was not significant \( F(2,26) = 2.205, P = 0.130 \). A similar analysis on the total time in the arms revealed a significant difference among groups \( F(2,26) = 5.360, P < 0.02 \). Fisher’s LSD test showed that this effect was attributable to the increased arm time for the ventral muscimol group, which was significantly different from both the saline \( (P < 0.01) \) and the dorsal muscimol \( (P < 0.02) \) groups. The mean arm times for the dorsal muscimol and saline groups were not significantly different \( (P = 0.93) \). Since the dorsal muscimol group was the only one of the three that failed to exhibit a CCP and neither of the exploration measures for this group differed from those for the saline group, there is no basis for attributing the effect of dorsal inactivation on the CCP to differences in behavior during pre-exposure.

The pre-exposure sessions were the rats’ first contact with the maze and no food was present during these sessions. Normal rats explore this novel, open environment very cautiously at first. Accordingly, reported decreases in responding to fear-producing stimuli and faster habituation in rats with entorhinal lesions (Maren and Fanselow, 1997; Galani et al., 1998; Burwell et al., 2004) could have been the cause of the increased time spent in the arms by the rats in the ventral muscimol group.
The elimination of behavioral differences during pre-exposure as a cause of the failure of rats with inactivation of dorsal EC to express a CCP makes it likely that the deficit was due to impairment of the normal processing of visual-spatial information (Burwell, 2000; Fyhn et al., 2004) in that part of the structure. The results suggest that acquisition of this information during unreinforced exploration is essential to the ability to express the CCP in the present paradigm.

EXPERIMENT 2

The results of Experiment 1, taken together with previous evidence (Chai and White, 2004) that an intact FF is required for the acquisition of latent spatial information, suggest that this form of learning may be mediated by a circuit that includes both the FF and the dorsal EC. To test this hypothesis a “disconnection” paradigm, as previously used in monkeys (Keating and Horel, 1971; Gaffan and Harrison, 1987) and in rats (Everitt et al., 1991) was used. The dorsal EC and FF were each disabled unilaterally. The hypothesis that the two structures on opposite sides of the brain during unreinforced pre-exposure to the maze would impair learning the adjacent arms discrimination. The control for this manipulation was a group that received the same two treatments on the same side of the brain. Since this manipulation left the hypothesized circuit intact on one side of the brain, it was predicted to have no effect on learning the discrimination.

Subjects were 16 rats similar to those used in Experiment 1. Eight of the rats had unilateral FF lesions and dorsal entorhinal cannulas on opposite sides of the brain (CONTRA) and another eight rats had the lesions and cannulas on the same side of the brain (IPSI). The apparatus and procedure were identical to those of Experiment 1. All rats received unilateral muscimol or saline injections before all pre-exposure trials.

Results and Discussion

Drawings of the lesions and cannula placements for the rats used in the statistical analysis are shown in Figure 4.

The behavioral results are shown in Figure 5. The rats with ipsilateral impairments preferred their food-paired arms \[F(1, 14) = 9.93, P < 0.01\] but the rats with contralateral impairments did not exhibit this preference \[F(1, 14) = 1.11, P = 0.33\].

The data for exploration during pre-exposure are shown in Figure 6. Two-tailed independent \(t\)-test showed no significant differences between the means for arm entries \([t(14) = 1.29, P = 0.21]\) or for time in the arms \([t(14) = 0.73, P = 0.47]\). These observations make it unlikely that the lack of a CCP in the contralateral group were due to effects of the treatments on the rats’ behavior during the pre-exposure sessions.

FIGURE 4. Drawings of unilateral FF lesions and dorsal entorhinal cannula placements for rats used in the statistical analysis in Experiment 2. Half of the rats had a unilateral FF lesion (A) on the right, the other half had a unilateral lesion on the left. In each case the lesion was paired with a unilateral cannula in the dorsal EC (B), also evenly distributed between left and right. In half of the rats the lesions and cannula were on opposite sides (the contralateral disconnection group), in the other half they were on the same side (the ipsilateral control group). The largest lesion (A) on each side is in gray, the smallest is in black.
The results of Experiment 2 suggest that the information acquired during pre-exposure trials depends on a circuit involving the EC and the FF. This conclusion is consistent with a previous demonstration (Olton et al., 1982) that contralateral lesions of EC and FF impair spatial learning using the win-shift task on a radial maze. However, the present results suggest that the deficit produced by this disconnection is specific to the acquisition of unreinforced (or “pure”) spatial information, and does not include learning related to the location of the reinforcers, which depends on the hippocampus (Chai and White, 2004; White and Gaskin, 2006).

A number of studies have shown that bilateral FF lesions produce spatial learning deficits (Olton and Werz, 1978; Walker and Olton, 1979; Aggleton et al., 1995; Galani et al., 2002), which are generally attributed to a the elimination of communication between the septum and the hippocampus (Galey et al., 1989; Marighetto et al., 1989; Ammassari-Teule and Maho, 1992; Izquierdo and Medina, 1995). The present findings, taken together with those of Chai and White (2004) and Gaskin et al. (2005) suggest that spatial learning may involve two independent processes: latent spatial learning, mediated by a FF-entorhinal circuit, and learning about reinforcement, mediated by the hippocampus. This suggests that deficits in spatial learning produced by FF lesions may be due to impaired acquisition of unreinforced (pure) spatial information which does not involve the hippocampus when no reinforcers are present.

The FF contains cholinergic inputs to both the hippocampus (Wainer et al., 1985; Senut et al., 1989; Kiss et al., 1990; Naumann et al., 1992) and the EC (Mitchell et al., 1982; Alonso and Kohler, 1984; Gaykema et al., 1990). A small number of cholinergic fibers originating in the medial septum project to EC on both sides of the brain (Alonso and Kohler, 1984); however, the fact that the present contralateral impairments eliminated the CCP suggest that these projections do not contribute significantly to unreinforced spatial learning.

The cholinergic projections that reach EC via FF have been implicated in spatial learning (Mitchell et al., 1982), and spatial memory impairments induced by septal inactivation with muscimol can be reversed by inhibition of AChsterase in the EC (Degroot and Parent, 2000). Lesions of the FF result in the reduction of c-fos expression in the both the septum and EC (Vann et al., 2000). These findings are all consistent with the suggestion that unreinforced spatial learning is mediated by a circuit that includes FF and EC, but not the hippocampus.

The FF contains cholinergic inputs to both the hippocampus (Wainer et al., 1985; Senut et al., 1989; Kiss et al., 1990; Naumann et al., 1992) and the EC (Mitchell et al., 1982; Alonso and Kohler, 1984; Gaykema et al., 1990). A small number of cholinergic fibers originating in the medial septum project to EC on both sides of the brain (Alonso and Kohler, 1984); however, the fact that the present contralateral impairments eliminated the CCP suggest that these projections do not contribute significantly to unreinforced spatial learning.

The cholinergic projections that reach EC via FF have been implicated in spatial learning (Mitchell et al., 1982), and spatial memory impairments induced by septal inactivation with muscimol can be reversed by inhibition of AChsterase in the EC (Degroot and Parent, 2000). Lesions of the FF result in the reduction of c-fos expression in the both the septum and EC (Vann et al., 2000). These findings are all consistent with the suggestion that unreinforced spatial learning is mediated by a circuit that includes FF and EC, but not the hippocampus.

A number of studies have shown that bilateral FF lesions produce spatial learning deficits (Olton and Werz, 1978; Walker and Olton, 1979; Aggleton et al., 1995; Galani et al., 2002), which are generally attributed to a the elimination of communication between the septum and the hippocampus (Galey et al., 1989; Marighetto et al., 1989; Ammassari-Teule and Maho, 1992; Izquierdo and Medina, 1995). The present findings, taken together with those of Chai and White (2004) and Gaskin et al. (2005) suggest that spatial learning may involve two independent processes: latent spatial learning, mediated by a FF-entorhinal circuit, and learning about reinforcement, mediated by the hippocampus. This suggests that deficits in spatial learning produced by FF lesions may be due to impaired acquisition of unreinforced (pure) spatial information which does not involve the hippocampus when no reinforcers are present.

The FF contains cholinergic inputs to both the hippocampus (Wainer et al., 1985; Senut et al., 1989; Kiss et al., 1990; Naumann et al., 1992) and the EC (Mitchell et al., 1982; Alonso and Kohler, 1984; Gaykema et al., 1990). A small number of cholinergic fibers originating in the medial septum project to EC on both sides of the brain (Alonso and Kohler, 1984); however, the fact that the present contralateral impairments eliminated the CCP suggest that these projections do not contribute significantly to unreinforced spatial learning.

The cholinergic projections that reach EC via FF have been implicated in spatial learning (Mitchell et al., 1982), and spatial memory impairments induced by septal inactivation with muscimol can be reversed by inhibition of AChsterase in the EC (Degroot and Parent, 2000). Lesions of the FF result in the reduction of c-fos expression in the both the septum and EC (Vann et al., 2000). These findings are all consistent with the suggestion that unreinforced spatial learning is mediated by a circuit that includes FF and EC, but not the hippocampus.
GENERAL DISCUSSION

Experiment 1 showed that inactivation of the dorsal but not ventral EC results in an impairment of the acquisition and/or storage of the unreinforced spatial information necessary for the development of a CCP based on the discrimination of ambiguous spatial cues. Experiment 2 showed that inactivating the dorsal EC on one side of the brain and lesioning the FF on the other results in impairment in the development of an ambiguous-cue CCP.

Combined with the previous findings that inactivation of the dorsal hippocampus during training (White and Gaskin, 2006) but not during unreinforced pre-exposure (Gaskin et al., 2005) impairs the development of an adjacent-arm CCP, the results of Experiment 1 suggest that learning about the location of food on a maze may involve at least two distinct processes: One is a hippocampus independent, entorhinal dependent process involved in the representation of spatial relationships among environmental cues. This kind of representation is often called a spatial map (O’Keefe and Nadel, 1978). The second process is hippocampus dependent, representing the location of the reinforcer (food) within the previously acquired spatial map.

As pointed out in the Introduction, these processes normally occur simultaneously. The latent learning paradigm allows them to be studied separately. When studied in this way they are revealed as different kinds of learning mediated by different (although closely related) parts of the brain.

This hypothesis raises the question of how the two kinds of information are combined in normal rats to produce the observed preference for the food-paired arm. It has recently been suggested (Hafting et al., 2005) that spatial information may be represented in the EC in the form of a grid anchored to spatial regularities of the environment. It is possible that this representation is activated by exposure to the environment during the training trials, and that the information about the location of the food is mapped onto this grid by a process that involves the hippocampus.

This suggests that a spatial representation in the EC can be reactivated after it is acquired. This could occur during the training trials when the rats are re-exposed to the represented environment. Since a functional hippocampus is required during training, one possibility is that the hippocampus retrieves the spatial information from the EC, combines it with information about the food and restores the combined information in either the hippocampus, EC, or elsewhere. This hypothesized process could be an instance of “systems reconsolidation” (Misanin et al., 1968; Lewis, 1979; Nader et al., 2000), an hypothesis that suggests re-activated memories are labile and therefore subject to change before being consolidated again. Although reconsolidation has been considered as a local, cellular phenomenon (Nader et al., 2000) the present findings suggest the possibility that it could function across related structures to combine information mediated in each of them (Alvarez and Squire, 1994; Squire and Alvarez, 1995).

Reactivation of the spatial representation may require movement, just as its initial acquisition does. If this is the case, reactivation would not occur during the training trials because the rats are confined at the ends of the arms during these trials, but would occur on the test trial when the rats are allowed to move around freely. In this case, reactivation of the spatial map and retrieval of information about stimuli associated with the food would occur simultaneously. Some combination of the two kinds of information would produce the observed preference.

One implication of this hypothesis is that the failure to observe long-term consolidation of spatial memory tasks (Clark et al., 2005a,b; Winocur et al., 2005) may be due to the fact that the tasks used in these experiments did not permit a dissociation between learning pure spatial information and learning about the location of the reinforcer. For example, in the in-switch task, even though rats learn concurrently about the spatial environment surrounding the maze and its relationship with the presence of the food, the spatial map and information about the location of food within that map may still be acquired by different neural systems. If performance of the task requires retrieval of the two kinds of information, and if either of them must be retrieved by the hippocampus from another structure, this retrieval function would not be subject to a consolidation gradient.

The findings of the present and our previous experiments (Gaskin et al., 2005; White and Gaskin, 2006) may be of relevance to the debate concerning semantic memory (for facts) and episodic memory (for personal experiences) in humans. Some reports suggest that semantic memories can be acquired in the absence of a functional hippocampus (Tulving, 1983; Vargha-Khadem et al., 1997; Tulving and Markowitsch, 1998). Our data are consistent with this suggestion because they show that entorhinal-dependent spatial information can be acquired in the absence of a functional hippocampus. Furthermore, the data suggest that the hippocampus becomes involved when a reinforcer is available, making the situation motivationally or emotionally relevant to the individual. This observation may provide a clue to the situational parameters that engage hippocampus-based memory in humans as well as in rats.

Acknowledgments

We thank Maliha Naeem for assistance with behavioral testing and Melissa Latourelle for assistance with surgery and histology.

REFERENCES


Hippocampus DOI 10.1002/hipo


Degroot A, Parent MB. 2000. Increasing acetylcholine levels in the hippocampus or entorhinal cortex reverses the impairing effects of septal GABA receptor activation on spontaneous alternation. Learn Mem 7:293–302.


Galani R, Weiss I, Cassel JC, Kelche C. 1998. Spatial memory, habituation, and reactions to spatial and nonspatial changes in rats with selective lesions of the hippocampus, the entorhinal cortex or the subiculum. Behav Brain Res 96:1–12.


Specialization and semantic organization: Evidence for multiple semantics linked to sensory modalities

J. FREDERICO MARQUES
University of Lisbon, Lisbon, Portugal

The present article reviews the case for multiple systems in semantic memory and empirically evaluates a multiple semantics proposal based on sensory modalities. In the experiments, a conceptual switching cost paradigm was used (Pecher, Zeelenberg, & Barsalou, 2003), in which participants verified properties for concepts and verification time was compared for target trials (e.g., a dog can bark) that were preceded by context trials of either the same or a different modality (e.g., a bee can buzz or a horse can have spots). Experiment 1 involved a modality switch while controlling for the concept's category and demonstrated that when modalities were switched, a processing cost was incurred that could not be attributed to the latter dimension. Experiment 2 further supported this conclusion by showing that, in a reverse situation, the cost incurred when category was switched was at least smaller. The results are discussed by considering possible alternative amodal explanations and other data that have shown the influence of sensory information in conceptual processing.

The organization of concepts in semantic memory has been a central topic in the field of cognitive psychology for several decades (see, e.g., Chang, 1986, for a review). Since the early 1980s, this topic has received even wider attention stemming from research in cognitive neuropsychology. In fact, a huge number of diverse cases of brain-lesioned patients who have been held to have specific domain or category impairments involving the semantic system have been reported (for reviews, see, e.g., Capitani, Laiacona, Mahon, & Caramazza, 2003; Forde & Humphreys, 1999). In order to explain these cases and, ultimately, normal functioning, very different theoretical models of semantic memory organization have been proposed (for reviews, see, e.g., Caramazza & Mahon, 2003; Forde & Humphreys, 1999). One of the major discussions between models is whether the semantic system is better conceptualized as a unitary amodal store or as composed of multiple stores that are more or less independent and have been linked to different sensory modalities (including sensorimotor information) and/or domains of knowledge. Regarding this discussion, three positions can be distinguished that correspond to the three main current theories of semantic organization: the modality-specific theory, the conceptual structure theory, and the domain-specific theory.

The modality-specific theory offers an explanation of cases of impairment that maintains that they reflect an underlying organization of specific modality-processing channels. The idea of multiple semantic stores linked with different sensory modalities can be traced more systematically to the work of Warrington and associates (e.g., Crutch & Warrington, 2003; McCarthy & Warrington, 1988). Building initially from the verbal–visual distinction, Warrington and associates proposed a multiple processing channels view of semantic memory. The salience of the different sensory and motor channels for the acquisition of knowledge would explain the organization of meaning (McCarthy & Warrington, 1988; see also Shallice, 1988), and fine-grained categorical dissociations would arise as a consequence of different weighting values between the information from each of the channels. A second view, developed by Shallice and associates (e.g., Lauro-Grotto, Piccini, & Shallice, 1997; Shallice, 1988, 1993), proposes a highly interconnected multimodal network. In this distributed network, the primary form of specialization was defined in terms of the different types of semantic operation that would occur predominantly with the respective complementary types of input (Lauro-Grotto et al., 1997; Shallice, 1988, 1993).

The conceptual structure theory offers an explanation of cases of impairment by assuming random damage to a unitary amodal system in which the correlation between features (especially the correlation between perceptual and function features) and their degree of distinctiveness govern the organization of information (e.g., Moss, 2006).
associated with different modalities in successive trials

sensory modalities. They showed that verifying features
normal subjects in favor of a specialization in terms of
berg, and Barsalou (2003) have provided evidence with
was used to evaluate the question mainly in
Caramazza & Shelton, 1998; Shallice, 1993).
full representation (and also with respect to the contribu-
imation of presemantic processing; for discussions, see, e.g.,
unitary semantics are very general, it is extremely difficult
were used and all of the features were verbal, although
related to different modalities (i.e., vision, motor action,
audition, taste, touch, and smell). This switching cost was
thus conceptual and, as such, provided a more decisive
argument for multiple semantics. Pecher et al. concluded
that conceptual processing is grounded in different sen-
sory systems and that the cost results from processing
switches from one modality system to another.

Finally, the domain-specific theory offers a straightforward explanation of cases of impairment that states that they reflect an underlying organization in terms of domains of knowledge (Caramazza & Mahon, 2003; Caramazza & Shelton, 1998). This theory proposes a direct organization in terms of evolutionarily salient object domains (especially animals, plant life, and tools) that would have been established on the basis of survival and reproductive advantages (Caramazza & Mahon, 2003; Caramazza & Shelton, 1998). In its first formulation (Caramazza & Shelton, 1998), the preference was to link the theory with a unitary amodal account of semantic memory, the organized unitary content hypothesis (Caramazza, Hil-lis, Rapp, & Romani, 1990), which maintained that items with similar properties would be stored in adjacent neural areas. More recently, Caramazza and Mahon (2003) have pointed out that the domain-specific theory is also commensurable with modality specificity (this possibility was only alluded to in Caramazza & Shelton, 1998). Available evidence would be consistent with an organization of conceptual knowledge by modality-specific constraints within domain-specific processing channels or with the reverse (Caramazza & Mahon, 2003). The former possibility is, however, explicitly preferred considering that the first-order constraint on the organization of conceptual knowledge is object domain (Caramazza & Mahon, 2003).

Since the proposals for multiple semantics and, also, for unitary semantics are very general, it is extremely difficult to differentiate between them with regard to cases of impairment. In fact, some cases could be readily interpreted in terms of both of the alternative proposals as reflecting either a subsystem impairment or problems in accessing a full representation (and also with respect to the contribution of presemantic processing; for discussions, see, e.g., Caramazza & Shelton, 1998; Shallice, 1993).

Evidence from cognitive psychology and behavioral data has been used to evaluate the question mainly in terms of the verbal–visual distinction and is not relevant to the present debate. Recently, however, Pecher, Zeelenberg, and Barsalou (2003) have provided evidence with normal subjects in favor of a specialization in terms of sensory modalities. They showed that verifying features associated with different modalities in successive trials (e.g., a target trial of blender–loud preceded by context trials of cranberries–tart) produced a switching cost in relation to successive same-modality trials (e.g., a target trial of blender–loud preceded by a context trial of leaves–rustling). In fact, verification time was longer for the former case, which is similar to the switching costs produced by detection signals from different modalities at the perceptual level (Spence, Nicholls, & Driver, 2001). Moreover, this result could not be attributed to other factors, such as associative priming between features of the same modality (Pecher et al., 2003, Experiment 2), expectancy effects (successive same- and different-modality trials presented in mixed blocks and with a high ratio of filler trials), or any presemantic factors, since only words were used and all of the features were verbal, although related to different modalities (i.e., vision, motor action, audition, taste, touch, and smell). This switching cost was thus conceptual and, as such, provided a more decisive argument for multiple semantics. Pecher et al. concluded that conceptual processing is grounded in different sensory systems and that the cost results from processing switches from one modality system to another.

However, this study did not manipulate or control for the concept's domain and category, and thus, it may be disputed whether modality makes an independent contribution to the results. In fact, by considering the parallel with perceptual switching costs, it is also possible to find at this level situations that reflect processing costs of changing from one type of stimulus to another within the same presentation modality (e.g., Kavcic, Krar, & Doty, 1999; Los, 1999). It is thus possible that the fact that trials also differed in terms of domain/category may have contributed to the switching costs. Moreover, if the domain-specific theory is taken into consideration, it is important to ensure that there is no possible confusion between modality and domain effects.

The main purpose of the present study was to disentangle these two possible effects in Pecher et al.'s (2003) experiments, evaluating their respective contributions within a modality-specific versus a domain-specific theory. Replicating the experiment while controlling for domain (and category) also makes it possible, using also the results in Pecher et al., to complete a possible parallel between these normal switching costs at the semantic level and those associated with cases of impairment described at the modality level (i.e., the different aphasias) and more specifically confined at the domain level within a certain modality. Finally, a replication of the experiment in which domain (animals vs. nonliving) is changed while sensory modality information is controlled for can also complement this evaluation and make possible a further test of the influence of the conceptual domain on property verification, following the same switching cost principle.

These questions were addressed in two experiments: Experiment 1, in which a modality switch was evaluated while concept domain was controlled for, and Experiment 2, in which a concept domain switch was evaluated while feature modality was controlled for.
EXPERIMENT 1

This first experiment replicated Pecher et al.‘s (2003) study while controlling for concept domain, thus allowing an evaluation of the possibility of modality-switching costs in a situation in which successive trials were composed of items from the same conceptual domain—either both animals or both nonliving things.4

To ensure that feature selection and manipulation reflected modalities in terms of the salience of the latter to feature acquisition, we previously had had a larger set of features rated in terms of how audition, vision, and touch contributed to their acquisition and knowledge. This is an important evaluation, since some features can be acquired by or, at least, simultaneously activate different modalities (e.g., a horse’s mane can activate/be acquired by vision and touch). This multimodal nature of stimulus evaluation has been extensively demonstrated at the perceptual level in terms of crossmodal links in both spatial attention (e.g., Spence, Pavani, & Driver, 2000) and tactile perception (e.g., Sathian & Zangaladze, 2002). At the level of semantic organization and in relation to category-specific dissociations, this aspect has been considered less. Authors either have contemplated the overall perceptual nature of features and emphasized visual versus verbal semantics (e.g., Farah & McClelland, 1991; Warrington & Shallice, 1984) or have considered the contribution of different perceptual modalities independently (Cree & McRae, 2003; McCarthy & Warrington, 1988).

On the basis of the general switching costs demonstrated by Pecher et al. (2003) and taking into consideration the evidence in favor of a multiple semantics in terms of sensory modalities, it was expected that in this situation, in which the conceptual domain was constant, modality-switching costs would also be observed in successive different-modality trials, in comparison with the same-modality trials.

Method

Participants. Sixty-two undergraduate students from the University of Lisbon participated for partial fulfillment of an introductory psychology course requirement. Twenty-five participated in the preliminary feature-rating task, and 37 participated in the main experiment. From the latter group, 1 participant was eliminated due to technical problems with the response box during the experiment. All analyses of the main experiment were based on a total sample of 36 participants.

Materials. Ninety-five concept plus feature pairs (e.g., a rocket can be large) were considered for the rating task (all the materials were in Portuguese). In an adaptation of Tranel, Logan, Frank, and Damasio’s (1997) procedure, the participants were asked to rate, on a 5-point scale, the degree to which audition, vision, and touch contributed to the knowledge of that concept’s feature (i.e., large for rocket). Each participant rated the 95 pairs separately for the same modality and the filler pairs were randomly intermixed for each participant. Also, the concepts and features were used only once. The practice trials were similar in nature to the experimental trials, consisting of 12 true items and 12 false items. Materials were arranged in two separate blocks by conceptual domain (animals and nonliving things) and were mounted on computer-presented slides, using the SuperLab for Windows software.

Procedure. Each trial began with a 500-msec fixation stimulus (** * * * *), which was followed by a concept-feature pair that was framed in a sentence “a CONCEPT can be/have FEATURE.” The sentence remained on the screen until the participant responded by pressing a two-button response box, with the preferred hand being assigned to the yes button and the other to the no button (the right hand was preferred by 95% of the participants). The initial instructions emphasized that a yes/no decision should be based on whether the feature was “usually true” of the concept and that the participant should try to respond as quickly as he or she could. The participants received feedback for 600 msec after pressing the wrong key (“ERROR”) or after taking 2,000 msec or longer to respond (“TOO

<table>
<thead>
<tr>
<th>Target Trial</th>
<th>Same Modality</th>
<th>Different Modality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dog–bark</td>
<td>Bee–buzz</td>
<td>Horse–pinto</td>
</tr>
<tr>
<td>Worm–flaccid</td>
<td>Lobster–rough</td>
<td>Deer–brown</td>
</tr>
<tr>
<td>Telephone–ring</td>
<td>Clock–tick-tock</td>
<td>Mirror–reflect</td>
</tr>
<tr>
<td>Lamp–light</td>
<td>TV–colors</td>
<td>Mattress–damp</td>
</tr>
</tbody>
</table>

Thus, each target trial appeared with both same-modality and different-modality context trials, counterbalanced across lists (all the features were true of their respective concepts). These 64 critical trials were also controlled for in terms of concept familiarity between animals and nonliving things (Portuguese familiarity norms from Marques, 1997) and, more important, were controlled for in terms of association, so that neither the concepts nor the features were associated between consecutive context and target trials (Portuguese free association norms from Marques, 2002). In addition to the 64 critical trials, the experiment included 124 filler trials, designed to mask the nature of the experiment. Within the filler trials, 92 were false (i.e., the feature was not true of the concept), and 32 were true (i.e., the feature was true of the concept). The filler trials were arranged in pairs (like the context + target trials) as follows: 30 false + false trial pairs, 16 false + true pairs, and 16 true + false pairs, consisting, in each case, half of animal + animal items and half of nonliving things + nonliving things items. Thus, true and false responses were almost equally likely overall (49% false and 51% true items). The features in the fillers referred to a dominant modality or to multiple modalities. As in Pecher et al. (2003), to ensure that the participants actually verified the features, the concept and feature in many of the false items were semantically related (e.g., squid–fish, sofa–rest). The critical (i.e., context + target trials) and the filler pairs were randomly intermixed for each participant.
sions to a situation in which processing switches can be attributed only to modality. These results are also consistent with a domain-specific theory in which modality-specific systems are organized by first-order domain-specific constraints (Caramazza & Mahon, 2003). This interpretation was further evaluated in Experiment 2.

**EXPERIMENT 2**

In this experiment, the possibility of a conceptual domain switch when modality is kept constant was examined. As was previously mentioned, the results from Experiment 1 are consistent with an organization by first-order domain-specific processing channels and with an organization by modality within those channels. From this hypothesis, it follows that we should also expect conceptual domain switches when modality is kept constant, since these would correspond to a switch between major domain-specific processing channels. If this domain organization corresponds to a specialization that is orthogonal to modality and constitutes a first-order constraint, we should expect even larger switch costs in this situation, which is the reverse of that in Experiment 1.

**Method**

**Participants.** Thirty-eight undergraduate students from the University of Lisbon participated for partial fulfillment of an introductory psychology course requirement. Two participants were eliminated for presenting more than 25% errors in the target trials. All the analyses were based on a total sample of 36 participants.

**Materials.** The same database as that established for Experiment 1 was used to build the critical context and target trials. Twenty concept–feature pairs were selected for the target trials (i.e., the trials following a trial of the same or a different feature modality), of which half consisted of animals and another half of nonliving things (arms, clothes, furniture, and tools). Twenty concept–feature pairs were then selected for the context trials (i.e., the trials that preceded the target trial of the same or a different feature modality), which consisted of items with a feature of the same modality but which were of either the same category or a different category from that of the target (i.e., animals vs. nonliving things). As in Experiment 1, two lists were created in such a way that each target trial had a same-category context trial in one list but a different-category context trial in the other, as is shown in Table 3.

Thus, each target item appeared with both same-category and different-category contexts, counterbalanced across lists (all the features were true of their respective items). These 40 critical trials were also controlled for in terms of familiarity between animals and artifacts (Portuguese familiarity norms from Marques, 1997) and, more important, were controlled for in terms of association, so

| Table 2 Mean Reaction Times (RTs, in Milliseconds) and Error Rates (in Percentages) of Attribute Verification on Target Trials by Context Trial and Conceptual Domain (With Standard Errors) |
|-----------------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                                        | Animals         | Nonliving Things |
| Context Trial                          | RT (M) SE       | RT (M) SE       | Errors (M) SE   | Errors (M) SE   |
| Same modality                          | 1,083.97 30.70  | 1,107.06 26.10  | 17.31 2.26      | 17.63 2.22      |
| Different modality                     | 1,125.18 28.73  | 1,137.72 24.41  | 13.44 2.29      | 14.97 2.65      |

Thus, each target item appeared with both same-category and different-category contexts, counterbalanced across lists (all the features were true of their respective items). These 40 critical trials were also controlled for in terms of familiarity between animals and artifacts (Portuguese familiarity norms from Marques, 1997) and, more important, were controlled for in terms of association, so
that between consecutive context and target trials, neither items nor features were associated (Portuguese free association norms from Marques, 2002). In addition to the 40 critical trials, the experiment included 80 filler trials, designed to mask the nature of the experiment. Within the filler trials, there were 60 filler trials in which the feature was not true of the item and 20 trials in which the feature was true of the item. The filler trials were arranged in pairs (like the context + target trials) as follows: 20 false + false trial pairs, 10 false + true pairs, and 10 true + false pairs, consisting, in each case, half of animal + animal items and half of artifact + artifact items. Thus, true and false responses were equally likely overall. The features in the fillers referred to a dominant modality or to multiple modalities and also included many false related items. The critical (i.e., context + target trials) and filler pairs were randomly intermixed for each participant. Also, the items and features were used only once. The practice trials were the same as those in Experiment 1. All the materials were mounted in a single block of trials on computer-presented slides, using the SuperLab for Windows software.

**Procedure.** The procedure was the same as that in Experiment 1 with respect to general instructions, slide sequence on each decision, answer mode, and recording.

**Results and Discussion**

RTs were considered only for target trials with a correct response on both the target trial and the preceding context trial. RTs exceeding 2,000 msec (the too slow limit) were eliminated from the analyses (outliers corresponded to 3% of the responses) but were not counted as errors. Median RTs and percentages of errors for the same-category versus the different-category trials on target trials were then calculated separately by item category (living and nonliving) for each participant and then averaged for the total group (N = 36). The results are presented in Table 4 by category switch condition (switch vs. no switch) and target item category (animals vs. nonliving things).

The results were analyzed by participant with a 2 x 2 repeated measures ANOVA with switch condition and item category as factors, but no significant effects were found for RTs or errors. For the category switch crucial dimension, switch cost differences are in the “right” direction but are not significant \[ F(1,35) = 0.10, MS_e = 18,748, p < .74 \] for RTs, and \[ F(1,35) = 1.33, MS_e = 209.21, p < .25 \] for errors. Taking into consideration from Experiment 1 that a comparable effect size of .40 could be expected according to Cohen’s (1988) terminology, a power analysis showed that power was .92 (for n = 36 and \( \rho = .05 \)), which is clearly above the conventional .80 specified by Cohen (1988) to detect an effect. However, if a medium size effect was to be posited, power would be reduced to .39, which clearly is below Cohen’s recommendation.

In summary, the results from Experiment 2 do not support an organization of semantic memory by domain as a first-order constraint. Although conceptual cost switches by domain were observed, they were neither significant nor larger than the modality switches in Experiment 1. With regard to the contribution of the conceptual domain to property verification, all that can be said from the present results is that any possible effect of this dimension seems to be smaller than the one related to modality-associated feature acquisition. As such, the results are not inconsistent with an organization by domain-specific subsystems within modality-specific processing channels, a possibility admitted by Caramazza and Mahon (2003). This possibility will be further analyzed in the General Discussion section.

**GENERAL DISCUSSION**

In the present article, the evidence for multiple systems of semantic memory has been analyzed in terms of the role of sensory modalities in the organization of semantic memory, and this proposal has been evaluated using behavioral data obtained with normal participants. This evaluation was done with a feature verification task and a switching cost framework considered at the conceptual level. The first experiment replicated Pecher et al.’s (2003) results for modality switch costs in a situation in which the concept’s domain (and also the category) was kept constant. The second experiment gave additional support to this interpretation in terms of modality. In fact, by failing to show a comparable category switch cost effect in a reverse situation (i.e., a category switch, with modality kept constant), the results clearly separate the contribution of sensory modality from a possible smaller contribution of the conceptual domain to conceptual processing in this situation. Overall, the results are more in accordance with a view of semantic memory organization that proposes

### Table 3

<table>
<thead>
<tr>
<th>Target Trial</th>
<th>Same Category</th>
<th>Different Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dog–bark</td>
<td>Lion–roar</td>
<td>TV–low</td>
</tr>
<tr>
<td>Telephone–ring</td>
<td>Clock–tick-tock</td>
<td>Donkey–hee-haw</td>
</tr>
<tr>
<td>Worm–flaccid</td>
<td>Octopus–slippery</td>
<td>Sweater–cottony</td>
</tr>
<tr>
<td>Skirt–rough</td>
<td>Scarf–smooth</td>
<td>Bear–furry</td>
</tr>
</tbody>
</table>

### Table 4

<table>
<thead>
<tr>
<th></th>
<th>Animals</th>
<th>Nonliving Things</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RT</td>
<td>Errors</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>SE</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Context Trial</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Same category</td>
<td>1,083.47</td>
<td>29.28</td>
</tr>
<tr>
<td>Different category</td>
<td>1,088.69</td>
<td>29.08</td>
</tr>
<tr>
<td>Switching cost</td>
<td>−5.22</td>
<td>−1.67</td>
</tr>
</tbody>
</table>
multiple processing channels, corresponding to a specialization of the semantic system in terms of types of information associated with sensory modalities (e.g., Crutch & Warrington, 2003; Lauro-Grotto et al., 1997; McCarthy & Warrington, 1988; Shallice, 1988, 1993).

As was previously mentioned, the results are also consistent with an organization of semantic memory by domain-specific subsystems within modality-specific processing channels. This would explain the larger modality-specific switch costs in Experiment 1 and the absence of significant category switch costs in Experiment 2 that, corresponding to switches between subchannels, could have easily blurred together. This possibility would, however, be inconsistent with what is known about how the brain processes information. For example, fMRI evidence for subregions of the visual-form–processing pathway that are more active for one domain than for another does not imply that those areas are dedicated to processing that domain. In fact, the same areas are active, to some degree, for all domains (e.g., Martin, Wiggs, Ungerleider, & Haxby, 1996; Moore & Price, 1999). Moreover, this possibility would also mean that relatively pure domain deficits (e.g., Patient E.W. for animals; reported by Caramazza & Shelton, 1998) would have to be explained by assuming a selective impairment of that domain subsystem within each major modality-specific channel, which is certainly not a very plausible situation in terms of brain damage.6

Another alternative account of the results would be to suppose that the kind of sentences used (“a CONCEPT can be/have FEATURE”) favored modality over category switches. This could have happened, since the sentence frame allowed a stronger spreading of activation from feature to modality than from concept to category and/or because features had the greatest impact on brain activation, since they came second in the trial sequence. Thus, one could argue that the two manipulations (maintaining category and manipulating modality or the reverse) are not directly comparable. Although a direct test of this possibility with other kinds of sentences was not made, different evidence would argue against it. In relation to the stronger spreading activation explanation, data from feature production norms seem to disconfirm it. When the stimuli are at the concept level (e.g., Cree & McRae, 2003, feature norms), many participants give categorical information as features (e.g., is a fruit, is a reptile, etc.). However, when the stimuli are at the feature level (e.g., Underwood & Richardson, 1956), participants usually select concepts that have those features, and not modalities with which those features can be associated. In relation to the kind of recency as impact explanation, it must be first said that in the language tested, as in many others, this feature after concept structure is that of the usual common sentence, not the reverse. This fact can also lead us to argue that the reverse format could enhance the item (and thereafter the category) just because it is an unusual language format. Moreover, in recent neuroimaging research using the same kind of sentences as those employed here and evaluating brain activation, Simmons, Pecher, Hamann, Zeelenberg, and Barsalou (2003) observed that concept profiles in terms of modality experience (how much of your experience with Item X involves Modality Y) were better predictors of brain activation (also observed to be multimodal) than were feature profiles (how much of your experience with Feature X involves Modality Y). These results seem to show that brain activation in this situation is not influenced by sentence position in a way that favors modality over category switches. Instead, what these results suggest is that the absence of a large category switch may have to do with multimodal activation stemming from the concepts that overlaps, at least partially, between categories. This explanation, however, needs to be further evaluated. In more general terms, evidence from neuroimaging studies has been presented either as not conclusive (e.g., Devlin et al., 2002; Price & Friston, 2002) or as supporting a multiple semantics account in terms of sensory modalities (Barsalou, Simmons, Barbe, & Wilson, 2003; Thompson-Schill, 2003), but not in terms of the verbal–visual distinction (Bright, Moss, & Tyler, 2004; Thompson-Schill, 2003). Nevertheless, as different authors have emphasized (e.g., Caramazza & Mahon, 2003; Devlin et al., 2002), there are still significant differences in task, materials, and techniques between studies that do not allow for a more definitive conclusion about a modality-specific organization in the semantic system.

Finally, although the present results seem to show that a view of semantic memory organization that does not contemplate a specialization in terms of sensory modalities is probably wrong, they do not exclude the possibility that some part of the information may correspond to an amodal store. Plaut (2002) recently presented a model in which these two possibilities are articulated and provided support for the model in terms of a computational simulation that could account for the pattern of performance observed in optic aphasias. In this model, specialization is influenced by input modality, by the nature of the information, and by a topographic bias favoring short connections. The result would be that regions equidistant from multiple modalities would learn to function in a relatively amodal way, whereas regions near a particular modality would serve functions that are more modality specific. One possibility for testing this hypothesis with the present paradigm would be to use features whose acquisition and representation do not seem dependent (or seem equally dependent) on the multiple modalities evaluated. The main difficulty may be to identify these features. In the initial rating task in Experiment 1, the items that had a nondominant modality of acquisition mainly indicated a similar dominance of touch and vision. However, this may just be a conceptual extension of the crossmodal links of stimulus evaluation found at the perceptual level (e.g., Sathian & Zangaladze, 2002).

Whereas the possibility of an account of multiple semantics in terms of a combination of modality-specific and amodal systems cannot be excluded on the basis of the present results, a defense of an account in which separate systems of amodal symbols represent the information from each modality does not seem tenable. In the first place, evidence that sensory variables influence concep-
tual processing (e.g., Stanfield & Zwaan, 2001; Zwaan, Stanfield, & Yaxley, 2002) is contrary to the idea that concepts are represented by amodal symbols. In the second place, even if this were to be true, it would undermine the associated idea of a single amodal system of semantic knowledge.

In conclusion, there is growing evidence from cognitive psychology, neuropsychology, and neuroimaging for a specialization of semantic organization in terms of sensory modalities. This evidence should now be integrated into more well defined theoretical proposals that can be evaluated at these different levels.

REFERENCES


NOTES

1. In accordance with the more recent consideration of category-specific deficits (e.g., Caramazza & Mahon, 2003; Cree & McRae, 2003), domain is used to refer to the tripartite distinction of animals, fruits/vegetables, and nonliving things, whereas category is used to further distinctions within domains (e.g., vehicles vs. tools; birds vs. insects).

2. The verbal–visual distinction (e.g., Paivio, 1971) does not really correspond to different sensory modalities but, rather, reflects a combination of content, context, and format differences (Bright, Moss, & Tyler, 2004; Plaut, 2002).

3. As will be apparent from the description of method, the number of trials needed for the manipulation of one variable was so large (especially in terms of fillers) that we preferred to use two experiments with a one-independent-variable design, instead of only one experiment with a two-variable (i.e., modality and category) design.

4. Plant life (mainly fruits and vegetables), the third conceptual domain generally considered (Caramazza & Mahon, 2003; Cree & McRae, 2003), was not tested, because there were not enough familiar concepts in this domain to allow the manipulations needed.

5. Inasmuch as the same critical target trials were used for the two switch conditions through a counterbalanced design, item analyses were not performed, since in this situation the traditional $F_1$ is the correct test statistic (Raaijmakers, Schrijnemakers, & Gremmen, 1999).

6. I thank one anonymous reviewer for raising this point.

APPENDIX

Target and Context Concept–Feature Pairs (English Equivalents)
Used in Experiments 1 and 2

<table>
<thead>
<tr>
<th>Target Trials</th>
<th>Context Trials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Living</td>
<td>Nonliving</td>
</tr>
<tr>
<td>Living</td>
<td>Nonliving</td>
</tr>
<tr>
<td>Experiment 1</td>
<td></td>
</tr>
<tr>
<td>Dog–bark</td>
<td>Revolver–bang</td>
</tr>
<tr>
<td>Duck–quack</td>
<td>Telephone–ring</td>
</tr>
<tr>
<td>Tiger–growl</td>
<td>Wardrobe–creak</td>
</tr>
<tr>
<td>Cat–silky</td>
<td>Skirt–rough</td>
</tr>
<tr>
<td>Worm–flaccid</td>
<td>Blouse–satin</td>
</tr>
<tr>
<td>Canary–yellow</td>
<td>Lamp–light</td>
</tr>
<tr>
<td>Kangaroo–beige</td>
<td>Saw–flash</td>
</tr>
<tr>
<td>Zebra–stripes</td>
<td>Screw–brass</td>
</tr>
<tr>
<td>Owl–beige</td>
<td>Machine gun–noisy</td>
</tr>
<tr>
<td>Monkey–screech</td>
<td>Bell–resonate</td>
</tr>
<tr>
<td>Parrot–talk</td>
<td>Bed–squeak</td>
</tr>
<tr>
<td>Rabbit–soft</td>
<td>Sweater–cottony</td>
</tr>
<tr>
<td>Snail–viscous</td>
<td>Scarf–smooth</td>
</tr>
<tr>
<td>Panther–black</td>
<td>Fridge–white</td>
</tr>
<tr>
<td>Peacock–colorful</td>
<td>Knife–silvery</td>
</tr>
<tr>
<td>Bear–dark</td>
<td>Scissors–shiny</td>
</tr>
</tbody>
</table>

Experiment 2

<table>
<thead>
<tr>
<th>Target Trials</th>
<th>Context Trials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Living</td>
<td>Nonliving</td>
</tr>
<tr>
<td>Living</td>
<td>Nonliving</td>
</tr>
<tr>
<td>Chicken–cluck</td>
<td>Revolver–bang</td>
</tr>
<tr>
<td>Duck–quack</td>
<td>Telephone–ring</td>
</tr>
<tr>
<td>Mouse–gray</td>
<td>Lamp–light</td>
</tr>
<tr>
<td>Giraffe–spots</td>
<td>Scissors–shiny</td>
</tr>
<tr>
<td>Worm–flaccid</td>
<td>Skirt–rough</td>
</tr>
<tr>
<td>Pig–grunt</td>
<td>Machine gun–noisy</td>
</tr>
<tr>
<td>Dog–bark</td>
<td>Door–creak</td>
</tr>
<tr>
<td>Peacock–colorful</td>
<td>Chair–blue</td>
</tr>
<tr>
<td>Camel–cream</td>
<td>Screw–brass</td>
</tr>
<tr>
<td>Turtle–hard</td>
<td>Blouse–satin</td>
</tr>
</tbody>
</table>

(Manuscript received July 27, 2004; revision accepted for publication December 12, 2004.)
Syntactic language processing: ERP lesion data on the role of the basal ganglia

SONJA A. KOTZ,1,2 STEFAN FRISCH,1,3 D. YVES VON CRAMON,1,2 AND ANGELA D. FRIEDERICI1

1Max-Planck-Institute of Cognitive Neuroscience, Leipzig, Germany
2Day Care Clinic of Cognitive Neurology, University of Leipzig, Germany
3Institute of Linguistics, University of Potsdam, Germany

(Received January 15, 2003; Revised April 28, 2003; Accepted May 2, 2003)

Abstract
The role of the basal ganglia in syntactic language processing was investigated with event-related brain potentials in fourteen neurologically impaired patients. Seven of these patients had basal ganglia lesions while 7 other patients primarily had lesions of the left temporo–parietal region excluding the basal ganglia. All patients listened to sentences that were either correct or included a verb argument structure violation. In previous experiments this type of violation elicited a biphasic pattern of an N400–P600 complex in young healthy participants. While the N400 may result from incorrect semantic-thematic role assignment, the P600 reflects the fact that verb information does not license the syntactic structure at present. Results of the patient experiment revealed a double dissociation: patients with left temporo–parietal lesions only show a P600, whereas patients with lesions of the basal ganglia showed no P600, but a negativity with extended duration that resembled an N400. The latter pattern not only confirms previous reports that the basal ganglia modulate the P600 but extends these results by showing that the N400 as a late semantic–thematic integration process appears partially modulated by the basal ganglia. (JINS, 2003, 9, 1053–1060.)

Keywords: Basal ganglia, Event-related brain potentials (ERPs), E(L)AN, P600, N400

INTRODUCTION
Basal ganglia, Event-related brain potentials (ERPs), E(L)AN, P600, N400

While the role of cortical structures in language processing has been confirmed by numerous investigations (see Friederici, 1999; Goodglass, 1993, for reviews), the participation of subcortical structures, such as the basal ganglia, in language is highly controversial. In particular, functional implications of subcortical structures are at stake. Nadeau and Crosson (1997) and Crosson (1999) postulate that the thalamus rather than the basal ganglia is engaged during lexical–semantic processing. This argument is supported by data that do not reveal lexical–semantic deficits in patients with basal ganglia lesions (Gotham et al., 1988; Mortimer et al., 1982; Tyler & Marslen-Wilson, 1986). However, other empirical data suggest that lexical–semantic and prosodic processes do engage the basal ganglia (e.g., Lieberman, 2001). Neuroimaging studies with healthy subjects and Parkinson patients (PD), who suffer from a neurodegenerative disorder of the basal ganglia, have confirmed a correlation of the basal ganglia function with the perception of emotional prosody (e.g., Breitenstein et al., 1998, 2001; Kotz et al., in press; Pell, 2002), as well as with lexical–semantic processes (e.g., Cappa & Abutalebi, 1999; Lieberman, 2001; Wallesch & Papagno, 1988).

In addition, the basal ganglia have been linked to language production in general (Alexander et al., 1987; Robin & Schienberg, 1990) or to processing of syntactic information in both language production and comprehension (e.g., Grossman et al., 1993; Lieberman et al., 1992). Three general cognitive processes have been implicated in syntactic comprehension of sentences: (1) regulation of attention, (2) working memory and (3) speed of information processing. Several authors investigated syntactic complexity (e.g., subject–object relative clauses) during sentence comprehension in basal ganglia patients (e.g., Grossman et al., 1991, 1992, 1993; Lieberman et al., 1990, 1992; Natsopoulos et al., 1992).
initially argued that syntactic comprehension deficits result from attentional rather than syntactic deficits, Lieberman et al. (1990; 1992) proposed that repeated errors on syntactically complex sentences cannot be attributed to an attention deficit, but to a working memory deficit. Recently, Grossman et al. (2002) attributed syntactic comprehension deficits in PD patients to slowed lexical access.

Finally, Ullman (2001) and Ullman et al. (1997) proposed that a fronto–striatal network engages in the computation of procedural knowledge, which reflects the implicit rules and operations of syntax. Utilizing a verb participle production paradigm that allowed one to separate regular verb forms (rule-based; e.g., walk,_ed) and irregular verb forms (lexically based; e.g., teach, taught), the authors reported that patients with anterior lesions and PD patients at a late stage of their disease progression show a selective deficit for regular verb forms, while patients with posterior lesions or Alzheimer’s disease cannot produce irregular verb forms. This dissociation of verb-specific deficits was taken as evidence that the whole fronto–striatal loop plays a role in implicit rule based syntactic processing.

In summary, while it is clear from the literature that the basal ganglia are engaged during language processing, a clear functional specification of the role of the basal ganglia in language remains open. In particular, the claim that the basal ganglia play a specific role during syntactic processing is controversial as the mental operations proposed to underlie or correlate with syntactic processing are diverse. Furthermore, there is no clarity as to whether lexical–semantic processes engage the basal ganglia or not. Most authors consider that the deficit arising from basal ganglia damage is not one that is solely “automatic” in nature (but see Ullman et al., 1997). The latter statement can be best clarified by analogy to the motor control hypothesis introduced by Marsden and Obeso (1994). These authors suggested that the primary role of the basal ganglia is a controlled response to changes in cortically regulated automatic behavior. If one applies this proposition to syntactic processes in language comprehension one could speculate that the basal ganglia engage in the controlled reordering or altering of cortically driven automatic syntactic processes.

This dissociation of automatic and controlled syntactic processing is also made explicit in a recent model on auditory sentence processing by Friederici (2002). The model describes that in a first step, a simple syntactic structure is built on the basis of word-category information (e.g., noun, verb). As will be described below, this first processing phase is highly automatic. In a second phase, which is controlled, lexical–semantic information is processed to realize thematic role assignment. If initial syntactic information and lexical–semantic information do not map onto each other, as in the case of some syntactic violations, the sentence structure needs to be reanalyzed in a third phase which is also a controlled process.

A number of event-related brain potential (ERP) studies investigating syntactic processes in healthy participants have shown that automatic and controlled syntactic processes can be separated (Friederici, 1995; Friederici, 2002; Hahne & Friederici, 1999). For example, phrase structure violations (e.g., violating the expectancy of a word class as in, *The fish was in the_ caught rather than The fish was in the pond caught; literal German translation) elicit an early anterior negativity (E(L)AN), followed by a late positivity (P600). Adhering to the sentence processing model described above, the early anterior negativity has been correlated with automatic syntactic processes, as the component does not vary as a function of manipulations that implicate control. This was shown by manipulating the proportion of violations that do not modulate the E(L)AN component (Hahne & Friederici, 1999). Furthermore, the E(L)AN is not influenced by additional violations of lexical-semantic information (Frisch et al., 2000; Hahne & Friederici, 1999).

On the other hand, the P600 has been linked to controlled syntactic processes (e.g., Frisch et al., 2002; Kaan et al., 2000). Finally, a third component, the N400 with a maximal centro-parietal distribution has been linked to the processing of lexical–semantic information. At the sentence level, the N400 is discussed as a component reflecting controlled, integrative processing of lexical–semantic information.

In three investigations with patients we explored the role of automatic versus controlled syntactic processes in lesion patients and PD patients. Friederici et al. (1999) reported that patients with anterior lesions show no early anterior negativity, but a P600 elicited by phrase structure violations, while patients with basal ganglia lesions show an early anterior negativity, but a strongly reduced P600. The authors take this evidence in support of the hypothesis that anterior cortical areas, but not subcortical regions such as the basal ganglia, are engaged in automatic syntactic processes, while the basal ganglia seem to modulate controlled syntactic processes. In a study with early PD patients a similar pattern emerged: PD patients showed an early anterior negativity, but barely any P600 effect (Friederici et al., 2003). Thus, unilateral focal vascular lesions as well as PD patients with unilateral functional deficits result in a comparable syntactic deficit as evidenced in the reduction of the P600 effect. Furthermore, these data support a functional as well as a structural separation of the two syntactic processes. While automatic syntactic processes seem to be regulated in anterior cortical regions, controlled late syntactic processes appear to be modulated by the basal ganglia. One question that these studies left open was to clarify whether the P600 reflects late syntactic processes or rather varies as a function of attentional demands. This question is also reflected in an ongoing debate whether the P600 is language-specific or just a P300-like effect, indicating the attention driven detection of an unexpected, task-relevant target (Coulson et al., 1998; Gunter et al., 1997). To test this question, Frisch et al. (2003) tested patients with focal vascular basal ganglia lesions and patients without basal ganglia lesions. They were presented with correct and incorrect sentences that included a morphosyntactic violation.
It is central for language processing research that verbs can be classified with respect to the number and type of constituents that they take as (syntactic and semantic) arguments. For example, an intransitive verb such as to grin can only take a subject argument expressing who is doing the grinning (e.g., The little boy grins). In contrast to a transitive verb which allows or demands an object besides the subject, adding a direct object to a verb such as to grin would render a sentence ungrammatical as well as semantically anomalous (see, *The little boy grins the old man). In the experiment, participants listened to grammatical and ungrammatical sentences. Ungrammatical sentences contained violations of the verb-argument structure (see above). In a number of studies that explored this syntactic violation type a biphasic ERP pattern of a negativity, resembling an N400, followed by a positivity (P600) was reported in healthy young subjects (Friederici & Frisch, 2000; Frisch et al., 2000; Osterhout et al., 1994).

**METHODS**

**Research Participants**

Fourteen brain damaged patients (4 female, all right-handed) in a chronic state participated in the current study after giving informed consent. Lesions primarily resulted from left hemisphere ischemic (n = 12) strokes, but 2 patients of the basal ganglia group had suffered a left-hemisphere hemorrhage. The average time since lesion in the basal ganglia group was: 2.85 years (range: 2–4 years) and in the patient group without basal ganglia lesions: 5.71 years (range: 3–9 years). Lesion sites were determined by (T1- and T2-weighted) anatomical MRI datasets from a 3.0 T system (Bruker 30/100 Medspec) and evaluated by an experienced neuroanatomist. The individual patient information is listed in Table 1.

**Materials**

All sentences were German passive constructions. In contrast to English, German allows passivization of intransitive verbs (such as arbeiten/to work). In this case, however, the sentence initial position can only be filled with an expletive (such as es/there), a prepositional phrase (such as im Zimmer/in the room) or an adverb (such as gestern/yesterday). Filling the initial position with a subject argument (such as das Zimmer/the room) creates an argument-structure violation since the subject can neither be syntactically nor semantically integrated. Thus, we realized an argument-structure violation by using sentences with an intransitive verb and a subject NP (Das Zimmer wurde gearbeitet/The room was worked). In the correct conditions, the sentence initial element was a prepositional phrase (Im Zimmer wurde gearbeitet/In the room it was worked). This allowed us to keep the critical word (verb participle) identical across correct and incorrect conditions.

In order to exclude possible confounds with a sentence final wrap up effect (e.g., Friederici & Frisch, 2000; Osterhout, 1997), the critical verb participle was always fol-
ollowed by a conjunction ‘and’ and a second verb participle which was transitive and therefore always correct. Forty sentences per condition, resulting in 80 critical sentences were created on the basis of 80 noun-(intransitive) verb sets. In addition, 80 filler sentences (half of them ungrammatical) with a similar sentence structure were created.

A female native speaker of German spoke the sentences at a normal speech rate. The sentences were recorded onto digital audio tape and digitized at a sampling rate of 44.1 KHz. In order to ensure a precise time locking of the ERP in each individual sentence, the onset of the critical word was marked by way of a careful visual and auditory inspection of the auditory speech signal.

**Procedure**

Patients listened to all 160 sentences that were presented via loudspeakers in a pseudorandomized order. A visual cue on the center of a computer screen indicated the onset of each sentence. 800 ms after the offset of the sentence, subjects judged whether the sentence was acceptable or not by pressing one of two response buttons. The next trial started 1000 ms after the subject’s button press. In the non-linguistic task patients heard standard tones (600 Hz) with a probability of .8 and deviants (660 Hz), with a probability of .2. The two-tone block contained a total of 50 auditory stimuli. All stimuli had a duration of 200 ms (including 10-ms rise and 40-ms fall time; sound pressure level (SPL) 75 dB) and were presented with a constant offset-to-onset interval of 600 ms.

ERPs were recorded from 19 scalp sites by means of Ag/AgCl electrodes with a NEUOSCAN 4.1 amplifier. C2 served as ground electrode. Recordings were referenced to the left mastoid and were re-referenced to linked mastoids off-line. Electrode impedances were kept below 5 kΩ. In order to control for eye movement artifacts, a horizontal and a vertical electro-oculogram (EOG) were recorded. Eye artifact control measures were applied to the raw data of each patient to increase the number of critical trials in each condition (Pfeifer et al., 1995). Then individual EEG recordings were scanned for additional artifacts on the basis of visual inspection. The average percentage of trials rejected due to behavioral performance and additional artifacts was 24.9%.

**Data analysis**

Accuracy in the behavioral task was calculated as the percentage of incorrectly performed trials in one condition relative to all trials in that condition. An ANOVA with lesion as a between-subjects factor and grammaticality as a within-subjects factor was conducted. ERPs were computed for each of the critical conditions for each electrode and each subject. All ERP averages were aligned to a 200 ms baseline relative to the onset of the auxiliary verb preceding the critical verb. Only trials with correct responses and without movement and amplifier saturation artefacts entered the averages. Separate repeated-measures ANOVAs were conducted separately for midline electrodes (FZ, CZ, PZ), for anterior lateral electrodes (FC3/4, F3/4, F7/8) and for pos-

<table>
<thead>
<tr>
<th>Patient group</th>
<th>Lesion site/ left hemisphere</th>
<th>Classification</th>
<th>Age (years)</th>
<th>Sex</th>
<th>Token Test</th>
<th>AAT test scores (AUD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients with basal ganglia lesions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Fronto–lateral, insula, caud, put</td>
<td>Broca</td>
<td>55</td>
<td>F</td>
<td>5</td>
<td>47/60</td>
</tr>
<tr>
<td>2</td>
<td>Fronto–lateral, insula, caud, put</td>
<td>Amnesic</td>
<td>38</td>
<td>F</td>
<td>21</td>
<td>43/60</td>
</tr>
<tr>
<td>3</td>
<td>Fronto–lateral, insula, caud, put</td>
<td>Residual</td>
<td>62</td>
<td>M</td>
<td>6</td>
<td>59/60</td>
</tr>
<tr>
<td>4</td>
<td>Caud, put</td>
<td>Amnesic</td>
<td>50</td>
<td>M</td>
<td>27</td>
<td>46/60</td>
</tr>
<tr>
<td>5</td>
<td>Caud, put</td>
<td>Amnesic</td>
<td>45</td>
<td>M</td>
<td>1</td>
<td>51/60</td>
</tr>
<tr>
<td>6</td>
<td>Put</td>
<td>Residual</td>
<td>60</td>
<td>M</td>
<td>0</td>
<td>52/60</td>
</tr>
<tr>
<td>7</td>
<td>Pall</td>
<td>Non-aphasic</td>
<td>57</td>
<td>M</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Patients without basal ganglia lesions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Multiple (bilateral), white matter</td>
<td>Non-aphasic</td>
<td>51</td>
<td>F</td>
<td>0</td>
<td>—</td>
</tr>
<tr>
<td>9</td>
<td>Parieto–lateral</td>
<td>Non-aphasic</td>
<td>50</td>
<td>F</td>
<td>0</td>
<td>—</td>
</tr>
<tr>
<td>10</td>
<td>Temporo–parietal–lateral</td>
<td>Amnesic</td>
<td>61</td>
<td>M</td>
<td>15</td>
<td>50/60</td>
</tr>
<tr>
<td>11</td>
<td>Temporo–parietal–lateral</td>
<td>Residual</td>
<td>39</td>
<td>M</td>
<td>0</td>
<td>59/60</td>
</tr>
<tr>
<td>12</td>
<td>Temporo–parietal–lateral</td>
<td>Non-aphasic</td>
<td>61</td>
<td>M</td>
<td>0</td>
<td>—</td>
</tr>
<tr>
<td>13</td>
<td>Fronto–lateral, insula</td>
<td>Residual</td>
<td>43</td>
<td>M</td>
<td>3</td>
<td>52/60</td>
</tr>
<tr>
<td>14</td>
<td>Fronto–lateral, insula thalamus (bilateral)</td>
<td>Non-aphasic</td>
<td>41</td>
<td>M</td>
<td>0</td>
<td>—</td>
</tr>
</tbody>
</table>
terior lateral electrodes (P3/4, P7/8, O1/O2) in order to capture potential distributional differences. The ANOVA for the midline analysis included lesion as the between-subjects factor (lesions including vs. excluding the basal ganglia) and two within-subjects factors grammaticality (grammatical vs. ungrammatical) and electrode (FZ vs. CZ vs. PZ). The ANOVA for the two lateral regions of interest were calculated with lesion as the between-subjects factor, and with two within-subjects factors grammaticality, and hemisphere (left vs. right), respectively. The statistical analyses were computed in two time windows relative to the critical word (verb), selected on the basis of visual inspection: 300 to 700 ms for the N400 and 800 to 1200 for the P600. Main effects of grammaticality in the respective time windows will reflect a N400 effect and a P600 effect, respectively. Results will be reported as statistically significant for p values .05 or less. Furthermore, to ensure that any modulation of the patient data was of linguistic nature, the P300 oddball paradigm was applied as a non-linguistic test. The statistical analyses followed the same ANOVA design as presented for the linguistic experiment with a within-subjects factor probability (rare vs. often) and a between-subjects factor lesion on the averages in a time window between 300 and 600 ms.

RESULTS

Accuracy

We found a main effect of lesion \([F(1,12) = 13.55, p < .01]\) due to more errors made by the basal ganglia group (33.2%) as compared to the group without a basal ganglia lesion (16.6%) and of grammaticality \([F(1,12) = 8.76, p < .05]\) showing that on average patients made more errors in the violation condition (35.0%; \(SD = 17.88\)) than in the correct condition (14.8%; \(SD = 9.64\)). There was no grammaticality \times\ lesion interaction (\(p > .1\)).

ERPs

Figure 1 displays the ERP patterns from the onset of the critical verb up to 1500 ms for each of the two lesion groups at selected electrode-sites. It is apparent that the patients without basal ganglia lesions show a clear P600 effect for verb-argument violations, but no apparent N400 effect (B). On the other hand, patients with basal ganglia lesions do not show a P600 effect, but an extended negativity resembling an N400 (A).

![Fig. 1.](image)

**A** Patients with basal ganglia lesions

**B** Patients without basal ganglia lesions

--- correct  ······· incorrect

Negativity

ErPs from the onset of the critical verb (onset at zero ms/vertical bar) up to 1500 ms show that patients with basal ganglia lesions display an extended N400-like negativity effect, but no P600 effect, while patients without basal ganglia lesions show no N400-like negativity effect, but a clear P600 effect.
The global ANOVA for the N400-like negativity in the selected time window (300–700 ms) at midline sites revealed no main effects of lesion nor grammaticality (all Fs < 1), but a marginal interaction of grammaticality × lesion \([F(1,12) = 3.98, p = .07]\). Follow-up analyses by lesion group did not confirm a statistically significant N400-like negativity in either patient group at midline sites (Fs < 1).

Analyses of lateral sites showed no significant effects for lesion, grammaticality, hemisphere nor an interaction of any of the three factors (all Fs < 1) in the anterior region. In the posterior region we found an interaction of grammaticality × lesion \([F(1,12) = 7.77, p < .01]\), but no main effects of lesion nor grammaticality. Follow-up analyses by lesion group revealed that patients with basal ganglia lesions show a N400-like negativity at posterior electrode sites \([F(1,6) = 7.37, p < .05]\), but the patients without basal ganglia lesions did not \((p > .1)\).

Analyses of the N400-like negativity preceding the P600 time window resulting from verb–argument structure violations indicate that patients with unilateral basal ganglia lesions show a bilaterally distributed extended N400-like negativity effect at posterior sites, while patients without basal ganglia lesions do not show such an N400-like negativity effect.

**P600**

Analyses of the midline sites revealed a main effect of grammaticality \([F(1,12) = 4.42, p < .05]\), but not of lesion \((F < 1)\). However, an interaction of grammaticality × lesion \([F(1,12) = 9.99, p < .01]\) as well as an interaction of grammaticality × lesion × electrode \([F(2,24) = 3.91, p < .03]\) can be reported. These interactions resulted from the fact that patients without basal ganglia lesions display a P600 effect \([F(1,6) = 15.24, p < .001]\) and an interaction of grammaticality × electrode \([F(2,12) = 5.39, p < .05]\) at midline sites, but patients with basal ganglia lesions do not show comparable effects (all effects: \(F < 1\)).

Grammaticality was significant at all three electrode sites for patients without basal ganglia lesions: Fz \([F(1,6) = 6.67, p < .05]\); Cz \([F(1,6) = 15.24, p < .001]\); Pz \([F(1,6) = 24.63, p < .001]\).

A similar picture emerged for the analyses of lateral sites. Analyses of anterior sites showed no main effects nor any critical interactions for any factor (all Fs < 1). Analyses of posterior sites displayed a main effect of grammaticality \([F(1,12) = 8.00, p < .01]\), but not of lesion \((F < 1)\). However, a significant interaction of grammaticality × lesion \([F(1,12) = 12.82, p < .001]\) was found. Follow-up analyses by patient group revealed that patients without lesions of the basal ganglia showed a P600 effect at posterior sites \([F(1,6) = 14.66, p < .001]\), but not at anterior sites \((F < 1)\). This effect was not qualified by hemisphere \((F < 1)\). Patients with basal ganglia lesions showed no main effect of grammaticality nor any interaction at either anterior or posterior electrode sites (all Fs < 1).

The data show that patients with basal ganglia lesions show no P600 effect following the preceding extended N400-like negativity effect, while patients without basal ganglia lesions show only a P600 effect.

Statistical analyses of repeated-measures ANOVA on the P300 effect did not reveal any significant differences as a function of lesion, but a main effect of probability indicating that patients of both groups showed a normal P300 effect. Detailed data analyses of the P300 effect were reported elsewhere (Frisch et al., 2003).

In summary, while patients with unilateral basal ganglia lesions show an extended negativity comparable to an N400 effect at posterior electrode sites, but no P600 effect as a result of verb–argument violations, patients without lesions of the basal ganglia show no N400-like negativity effect, but a P600 effect to this type of violation. These results are in contrast to the biphasic pattern of an N400 followed by a P600 in younger healthy participants (see Frisch et al., 2000).

**DISCUSSION**

Taken together, the data from the current experiment further support the role of the basal ganglia in controlled syntactic processing by replicating the lack of a P600 effect in patients with focal unilateral basal ganglia lesions (Friederici et al., 1999; Frisch et al., 2003) and extending it by reporting evidence on the role of the basal ganglia in a second controlled process that relates to semantic–thematic processing.

As it was hypothesized, a P600 effect due to verb–argument structure violations was only found in patients without basal ganglia lesions. However, these patients showed an extended N400-like negativity effect preceding the P600 time window which resembles the N400 effect in young healthy participants (see Frisch et al., 2000), while patients without basal ganglia lesions did not show such an N400-like negativity effect. Furthermore, a P300 effect in response to rarely occurring auditory stimuli was shown in both groups of patients and was comparable to healthy controls (see Frisch et al., 2003).

With respect to our hypothesis that the basal ganglia only regulate controlled syntactic processes, the current results support the fact that the basal ganglia play a necessary role in the mediation of the P600 effect. Thus, the present results are in agreement with recent findings that the P600 effect is strongly reduced in Parkinson patients (Friederici et al., 2003) and in patients with unilateral lesions of the basal ganglia (Friederici et al., 1999; Frisch et al., 2003; Kotz & Friederici, 2003).

The question of whether there is a functional correlation of the basal ganglia and lexical–semantic processes can be partially answered. The fact the basal ganglia group show

---

1 In a pilot study with a sample of 14 age-, gender- and education-matched controls for the patients tested in the current experiment we also found a biphasic pattern of an N400 followed by a P600.
an N400-like negativity effect, but with an extended duration points to the modulatory role of the basal ganglia in lexical–semantic processing such as thematic role assignment. This extended N400-like negativity adds to controversial previous evidence as the data imply that speed of information processing affecting lexical–semantic information might be modulated by the basal ganglia (e.g., Crosson, 1999; Nadeau & Crosson, 1997; but see Cappa & Abutalebi, 1999; Wallesch & Papagno, 1988).

In particular, it needs to be noted that in comparison to the N400 seen in younger and older healthy participants (e.g., Frisch et al., 2000) the duration of the N400-like negativity effect in the basal ganglia patients differs. While the negativity in the patients shows a similar onset to the one seen in healthy controls, the duration of the N400-like negativity effect in this study extends up to 700 ms post-stimulus onset in the patients. We would like to argue that this effect is due to global cognitive slowing that impairs lexical–semantic processing during language comprehension. As this duration difference only occurred for the N400-like negativity effect that reflects semantic–thematic role assignment, but was not visible in the non-linguistic P300 effect (see Frisch et al., 2003), it is plausible that the rate of lexical–semantic information processing in the broader sense is changed as a result of unilateral basal ganglia lesions. In a recent paper, Grossman et al. (2002) discussed evidence that the striatum may play a critical role in information processing speed (see also Rao et al., 2001; Schubotz et al., 2000). In accordance, it appears that while the time course of semantic–thematic processing is hampered by basal ganglia lesions, the process is still realized in an extended time window. This is clearly not the case for controlled syntactic processes reflected in the P600 effect which is absent in the basal ganglia patients.

In conclusion, our study demonstrates that patients with focal lesions of the basal ganglia show a selective deficit of controlled syntactic processes as reflected in the P600 effect, while controlled semantic–thematic processes as reflected in a preceding N400-like negativity effect are present but its duration was extended. These results show that the basal ganglia play a mediating role in controlled syntactic processes during comprehension and may also play a role in the rate of controlled semantic processes.

ACKNOWLEDGMENTS

We want to thank Anja Hahne, Korinna Eckstein, and Angelika Wolf for their assistance in preparing the sentence materials, Claudia Misch for speaking the sentences, Ina Koch for recording the ERP data and Juliane Hofmann for statistical support. Thanks also to Bruce Crosson, Arturo Hernandez and Natalie Phillips for helpful comments on an earlier version of the paper.

REFERENCES


The time-averaged pulsed flux density of the pulsars is about 1.8 mJy (1 Jy = 10^{-26} W m^{-2} Hz^{-1}) at 1390 MHz (Table 1), compared with a total flux density at this frequency of 7 mJy (3). The ~5-mJy unpulsed emission probably arises in the impact region described above. We find it remarkable, with much of the magnetosphere of B blown away by the wind of A, that B still works as a pulsar. This suggests that the radio emission is probably generated close to the neutron star, providing a direct constraint on the emission height.

**Conclusion.** We have detected the binary companion of the millisecond pulsar J0737–3039 as a pulsar, making this the first known double-pulsar system. This discovery confirms the neutron-star nature of the companions to recycled pulsars in eccentric binary systems and validates the suggested evolutionary sequences in which a companion star, having spun up the pulsar, forms a young pulsar in a supernova explosion (31). The highly relativistic nature of this compact system opens up opportunities for much more stringent tests of relativistic gravitation than have been possible previously. Not only have we already measured four quantities attributable to, and consistent with, general relativity, but the mass ratio $R$ is a new high-precision constraint that is independent of gravitational theories. Within a year or so, we expect to measure the orbital decay due to emission of gravitational radiation. If the intrinsic value due to gravitational-wave damping can be extracted, it will allow tests of radiative aspects of gravitational theories mixed with strong-field effects. On somewhat longer time scales of a few years, we expect to detect several other relativistic effects, such as geodetic precession of the pulsars’ spin axes, spin-orbit coupling, and other deviations, making this a superb test bed for relativity.

Human survival depends on the ability to understand what others feel, be it an experience of pain, touch, or tickling (3). Empathy involves being able to understand others’ intentions and beliefs. This capacity is fundamental to the social fabric of human society, enabling us to understand what others think and feel. Empathy is a complex and multifaceted phenomenon that involves both emotional and cognitive processes.

**References and Notes**

5. Receiver bandwidths were used 64 MHz at 680 MHz, 256 MHz at 1390 MHz, and 576 MHz at 3030 MHz.
14. This is correct up to so-called first post-Newtonian $\ell$ order $\ell$, and any intersection of the PK-parameter lines must be located on the $\ell$ line, which will only deviate from that shown in Fig. 3 by corrections of order $\ell v/c^2$.
15. This is correct because of the usage of the DD timing model in our analysis and its definition of mass (11, 13).
33. We would like to thank J. Sarkissian and other members of the Parkes multibeam team for their kind help with making the observations described in this paper and N. Wex and G. Schäfer for useful discussions. Extensive use was made of the PSRCHIVE pulsar analysis system developed by A. Hotan and colleagues (see http://astronomy.swin.edu.au/pulsar/). The Parkes radio telescope is part of the Australia Telescope, which is funded by the Commonwealth of Australia for operation as a National Facility managed by CSIRO, M.B.A., and N.D. acknowledge financial support from the Italian Ministry of University and Research (MIUR) under the national program Cofin 2001. F.C. is supported by NSF, NASA, and National Radio Astronomy Observatory. D.R.L. is a University Research Fellow funded by the Royal Society.

**Empathy for Pain Involves the Affective but not Sensory Components of Pain**

**Tania Singer, Ben Seymour, John O’Doherty, Holger Kaube, Raymond J. Dolan, Chris D. Frith**

Our ability to have an experience of another’s pain is characteristic of empathy. Using functional imaging, we assessed brain activity while volunteers experienced a painful stimulus and compared it to that elicited when they observed a signal indicating that their loved one—present in the same room—was receiving a similar pain stimulus. Bilateral anterior insula (AI), rostral anterior cingulate cortex (ACC), brainstem, and cerebellum were activated when subjects received pain and also by a signal that a loved one experienced pain. AI and ACC activation correlated with individual empathy scores. Activity in the posterior insula/secondary somatosensory cortex, the sensorimotor cortex (SI/MI), and the caudal ACC was specific to receiving pain. Thus, a neural response in AI and rostral ACC, activated in common for “self” and “other” conditions, suggests that the neural substrate for empathic experience does not involve the entire “pain matrix.” We conclude that only part of the pain network associated with its affective qualities, but not its sensory qualities, mediates empathy.

Human survival depends on the ability to function effectively within a social context. Central to successful social interaction is the ability to understand others’ intentions and beliefs. This capacity to represent mental states is referred to as “theory of mind” (1) or the ability to “mentalize” (2). Empathy, by contrast, broadly refers to being able to understand what others feel, be it an emotion or a sensory state. Accordingly, empathic experience enables us to understand what it feels like when someone else experiences sadness or happiness, and also pain, touch, or tickling (3).

Even though empathy has been extensively discussed and investigated by philosophers and social scientists, only recently has it become a focus for neuroscience (3–8). Influenced by perception-action models of motor behavior and imitation (9),
Preston and de Waal (8) proposed a model of empathy that incorporates most theoretical accounts of, as well as empirical findings on, empathy. The key suggestion is that observation or imagination of another person in a particular emotional state automatically activates a representation of that state in the observer, with its associated autonomic and somatic responses ("automatic" refers to a process that does not require conscious and effortless processing but can nevertheless be inhibited or controlled). The philosopher Susanne Langer has described it as an involuntary breach of individual separateness (10).

Recent neuroimaging studies have explored the neural correlates of empathic experience by measuring brain activity while subjects watched video clips showing actors telling personal stories with neutral or sad content accompanied by sad and happy facial expressions (11), by comparing brain activity associated with the imitation of and the mere observation of pictures showing different emotional facial expressions (12), and by comparing neural responses elicited by watching videos of faces with emotional expressions of disgust and pleasure with responses induced by smelling aversive and pleasant odors (13). Whereas the first study did not permit the identification of shared networks underlying emotions in self and others, the latter studies report activation in areas previously identified in the perception and production of facial emotional expressions (action representation and emotional systems) as well as during the smelling of aversive odors (insula). These results suggest that regions associated with feeling an emotion can be activated by seeing the facial expression of the same emotion, a phenomenon described as emotional contagion.

A paradigm to study empathy in vivo. The present study extends previous findings in a number of ways. First, we tested for shared and unique networks for experienced and empathic pain. More specifically, we provide evidence for pain-related empathic responses and demonstrate that empathic experience does not involve activation of an entire pain matrix, but only of that component associated with the affective dimension of pain experience. Moreover, we show a relation between empathy-related brain activity and individual difference in empathy as assessed by commonly used empathy scales. We also demonstrate—in contrast to accounts of emotional contagion—that empathic responses can be elicited automatically in the absence of an emotional cue (such as facial emotional expressions) through mere presentation of an arbitrary cue that signals the feeling state of another person.

We used functional magnetic resonance imaging (fMRI) to explore both common and unique brain circuitry involved when we experience pain in ourselves, as well as when we observe someone else feeling pain. Previous studies on pain have revealed that aversive stimuli consistently activate the secondary somatosensory cortex (SII), insular regions, the anterior cingulate cortex (ACC), the movement-related areas such as the cerebellum and supplementary motor areas and, less robustly, the thalamus and the primary somatosensory cortex (SI) (14, 15). This pain-related network, which is commonly referred to as the pain matrix, served to define regions of interest in the present study.

To investigate pain-related empathy, we developed a paradigm that allows the investigation of empathic experience in vivo, with the object of empathy being a real person present in the same room. More specifically, we investigated pain-related empathy in 16 couples, under an assumption that couples are likely to feel empathy for each other. We assessed brain activity in the female partner while painful stimulation was applied to her or to her partner’s right hand through an electrode attached to the back of the hand. The partner was seated next to the MRI scanner and the right hand of each subject was placed on a tilted board, allowing the female partner with help of a mirror system to see her and her partner’s right hand. On a large screen situated behind the board, cues were presented in random order indicating whether she (self) or her partner (other) would get low (no pain condition) or high (pain condition) stimulation. We were especially interested in comparing pain-related brain activity (assessed by the difference between trials involving painful and nonpainful stimulation) in the context of “self” and “other.” Questionnaires administered after scanning served to validate measurements of individual pain threshold made before scanning, to obtain subjective evidence for empathic experience during scanning, and to assess stable individual differences in empathy in order to determine whether these scores predict the amplitude of empathy-related brain activity.

Analysis of pain intensity ratings after...
scanning confirmed the individual thresholds for nonpainful and painful stimulation determined before scanning (fig. S1) (16). In addition, the unpleasantness ratings indicated empathic involvement of the subjects. Nonpainful trials were rated as being significantly less unpleasant than painful trials, irrespective of whether the pain was applied to themselves or to the partner [main effect of pain: F (1, 15) = 19.93, P < 0.001; main effect of self/other: F (1, 15) = 0.12, P = 0.73]).

Does empathizing with pain in others activate the entire pain matrix? Comparison of brain activity associated with painful and nonpainful trials in the “self” condition revealed increased activity in contralateral SI/MI, in bilateral SII with a peak activation in contralateral posterior insula extending into SII, in bilateral mid and anterior insula, in ACC [caudal and posterior rostral zones, using Picard and Strick’s terminology, (17)], in right ventrolateral and mediodorsal thalamus, brainstem, and mid and right lateral cerebellum (Fig. 1, A to D, green). These regions have all been identified as responding to painful stimuli in previous imaging studies (14, 15). Many structures in this pain matrix were also activated when pain was applied to the partner, that is, in the absence of somatosensory stimulation (Fig. 1, A to D, red).

When comparing painful with nonpainful trials in the context of “other,” increases in brain activity were observed in the ACC (anterior and posterior rostral zones), the anterior insula (AI) bilaterally with an extension into inferior prefrontal cortex, the cerebellum, and the brainstem. In addition, significant activations were observed in the ventral and dorsal visual stream, including bilateral fusiform cortex, lateral occipital and right posterior superior temporal sulcus, the left inferior parietal cortex, and the left superior frontal cortex.

To test for common networks activated by pain in the self and other conditions, we performed a conjunction analysis (Fig. 2, A and B), as well as a more conservative masking procedure in which we masked the simple contrast pain–no pain in others by the simple contrast pain–no pain in self. Both analyses revealed a network common to pain in self and other conditions that comprised ACC (caudal and posterior rostral zones), bilateral middle insula and AI (with a peak activation in the right AI), brainstem, and lateral cerebellum. We also tested for an interaction in order to identify regions where there were stronger pain-related activations in the self as compared to the other condition. These analyses confirmed that the contralateral activations in SI/MI and SII/posterior insula, as well as a region in caudal ACC, were specific to the pain experience in the self condition (Fig. 3). In contrast, extrastriate visual related activations were specific to the empathy condition.

Inspection of the time courses of pain-related activity for self and others suggest two peaks, probably as the result of an instant response to the anticipation cue followed by another response to the delivery of the pain stimulus 3.5 s later. Accordingly, pain-related activations for self in ACC and AI peak first early, around 2 to 4 s, and again around 8 to 12 s (Fig. 2C). Activation specific to pain in self in SI, SII, and caudal ACC seems only to peak late, around 8 to 12 s, probably registering the actual receipt of the painful stimulus (Fig. 3, A to C). Although the present study was not designed to test differences between anticipation and receipt of pain in self and others, the present data suggest that both anticipation and experience of pain involve ACC and AI and that activation of these networks is involved in understanding the pain of others.

Do people scoring higher in empathy show higher pain-related brain activity? To explore further the role of anterior insular cortex and ACC—the two major regions of the pain matrix identified as being also involved when empathizing with the pain of others—we determined whether individual differences in empathy, assessed by two questionnaires, covary with brain activity elicited in the empathy conditions (pain—no pain in the “other” condition). As Fig. 4 illustrates, individual differences in empathy as measured by two empathy scales, the Balanced Emotional Empathy Scale (18, 19) and the Empathic Concern Scale (a subscale of the Interpersonal Reactivity Index) (20), showed significant covariation with activity in ACC (anterior rostral zone) and left AI, but not right AI. Thus, subjects who scored higher on general empathy scales showed stronger activations in areas significantly activated when the subjects perceived their partner as being in pain. In addition, an anterior part of ACC (anterior rostral zone adjacent to paracingulate sulcus) and lateral right cerebellum showed significant covariation with self-rated individual differences in empathy.
These findings underscore the crucial role of rostral ACC and anterior insula cortices for empathic experience related to pain. They are also in line with a report based on single-neuron recordings in a precuneus-lobeotomy patient that provided evidence that neurons in the ACC can respond not only when a person receives a painful stimulus but also when a person observes or anticipates a potentially painful stimulus delivered to an experimenter (21). Note that these regions are fundamentally different from sites subserving mirror neurons that respond when performing or watching a conspecific performing particular movements (22).

Numerous imaging studies have investigated the neural correlates of pain experience (14, 15), but none have explored the empathic experience of pain. Our results confirm previous findings of pain-related activation in SI, SII, bilateral insula cortex, ACC, thalamus, brainstem, and cerebellum. In addition, our data show that only part of the network mediating pain experience is shared when empathizing with pain in others. Empathizing with someone else’s pain elicited activity principally in left and right AI, ACC, lateral cerebellum, and brainstem. This activity was elicited without an explicit focus on empathy insofar as subjects were not aware of the study aims, nor were they required to make an explicit judgement of what their loved one was feeling at the time of the scan. The finding of empathy-related activation suggests an automatic engagement of empathic processes when perceiving pain in others. Moreover, our analysis demonstrates that pain-related activation in contralateral SI, SII/posterior insula, and caudal ACC are specific to self-experienced pain, as opposed to perceived pain in others.

Recent neuroimaging studies on pain have demonstrated contralaterally biased representations of painful stimulus in SI, distinct parts of SII, and posterior insula, as well as in lateral thalamus, which suggests that these structures provide sensory-discriminative information concerning nociceptive stimuli such as location, quality, and intensity (23, 24). In contrast, AI and ACC do not show such a contralateral bias, supporting a role in coding the autonomic and affective dimension of pain such as the subjectively perceived unpleasantness (15, 25–27). Using hypnosis as a tool to disassociate sensory-discriminative from affective pain components, activation in ACC (posterior rostral zone) was shown to be modulated by perceived unpleasantness, whereas activation in SI and SII was unaffected (28). Similarly, attentional manipulations modulate ACC (posterior rostral zone) and right AI, but not posterior insular/SII cortices (29). The role of ACC (posterior rostral zone) and right AI in coding the subjective affective dimension of pain is also highlighted by findings showing that activation of right AI was correlated with subjective intensity ratings of thermal sensation in a manner that is distinct from a linear representation of stimulus temperature in posterior insula (30). Indeed, anticipation of pain activates more anterior insular regions, whereas the actual experience of pain activates more posterior insula, which suggests that the latter is associated with the actual sensory experience of pain and the former with affective dimensions such as the anticipatory arousal and anxiety of pain (31). Furthermore, subjective reduction of pain associated with placebo and opioid analgesia is associated with increased activity in rostral ACC and right AI (32).

**Fig. 3.** Brain activity specific to the experience of pain in oneself. (A to C) illustrate results of the interaction between the two factors (pain–no pain and self/other). Statistical maps (threshold set at $P < 0.001$) are overlaid on axial (A), coronal (B), and sagittal (C) sections from the mean structural scan. Coordinates refer to peak activations and are in mm. (A) Increased pain-related activity [(−21, −33, 75); (−30, −24, 72); (−27, −36, 60)] in SI/MI. (B) Activity in left posterior insula/SII [−39, −27, 24]. (C) Activity in a part of caudal ACC [6, 6, 42]. On the bottom, time courses are displayed for peak voxels of contralateral SI, SII, and caudal ACC. Green lines reflect parameter estimates for pain-related activity (pain–no pain) in self; red lines reflect pain-related activity in others.

**Conclusion.** The strong anatomical connections between regions constituting the pain matrix suggest that these regions do not function independently in encoding different aspects of pain but are highly interactive. Nevertheless, the results of the present study and previous investigations suggest a segregation of sensory-discriminative and autonomic-affective attributes of the pain experience. Rostral ACC and AI appear to reflect the emotional experience that evokes our reactions to pain and constitutes the neural basis for our understanding of the feelings of others and ourselves.

The above proposal is consistent with a recent model by Craig based on detailed anatomical observations that conceives of pain as one of the homeostatic emotions that reflects the internal (interoceptive)
condition of the body, similar to temperature, sensual touch, itch, hunger, or thirst (25, 33). More specifically, it is assumed that an image of the body’s internal state is mapped to the brain by afferents that provide input by way of the ventromedial thalamic nucleus to area 3a (sensorimotor cortex), as well as to the mid/posterior dorsal insula. In humans, this high-resolution, modality-specific sensory representation of the physiological condition of the body in the posterior insula is initially re-represented in the AI on the same side of the brain, and then, by way of a callosal pathway, remapped to the other side of the brain in the right AI. Such a second-order re-representation in the right AI is assumed to subserve subjective feelings and the awareness of a physical self as a feeling entity. At the same time, afferents also project by way of the medial dorsal thalamic nucleus to produce behavioral drive in ACC. Thus, direct activation of both the insula (limbic sensory cortex) and the ACC (the limbic motor cortex) may correspond to a simultaneous generation of both a feeling and an affective motivation with its attendant autonomic effects (25, 33). Indeed, imaging studies focusing on the relation between peripheral measures of arousal and brain activity give robust evidence for the crucial role of rostral ACC and AI cortices in the representation of internal bodily states of arousal, as well as emotional awareness (26, 27). Furthermore, activation in these regions has been observed in a wide range of imaging studies associated with positive and negative subjective feelings expressed by subjective ratings of facial trustworthiness (34), musical enjoyment (35), sensual touch (36), and distress resulting from social exclusion (37). More generally, these regions may play a critical role in second-order representations of bodily homeostatic states that underpin core representations of self (38, 39).

Our data suggest that empathizing with the pain of others does not involve the activation of the whole pain matrix, but is based on activation of those second-order re-representations containing the subjective affective dimension of pain. Accordingly, we propose that these cortical re-representations have a dual function. First, they form the basis for our ability to form subjective representation of feelings that allow us to predict the effects of emotional stimuli with respect to the self. Second, they serve as the neural basis for our ability to understand the emotional importance of a particular stimulus for another person and to predict its likely associated consequences. From a functional and evolutionary perspective, a detailed representation of the source and nature of a noxious stimulus (i.e., intensity, location) is of functional relevance when it concerns our own body, allowing effective engagement of executive action systems (i.e., removing the noxious source). In contrast, the understanding of someone else’s emotional reaction to pain does not necessitate such a detailed sensory-discriminative representation of the noxious stimulus but rather a representation of the subjective relevance of the stimulus as reflected in the subjective unpleasantness that the other person feels. Such decoupled representations—which are independent of the sensory inputs of the outside world—have been postulated to be necessary for our ability to mentalize, that is, to understand the thoughts, beliefs, and intentions of others (2). Our data suggest that we use similar decoupled representations to understand the feelings of others and that our ability to empathize has evolved from a system for representing our internal bodily states and subjective feeling states.

Fig. 4. Activation level (parameter estimates) observed within peaks of the ACC and the left insula during empathy-related conditions (pain–no pain in other) are significantly correlated with individual differences in empathy as measured by (A) the Empathic Concern Scale of Davis (20) and (B) the Balanced Emotional Empathy Scale of Mehrabian (18, 19). The lines represent the linear best fit; $r$ refers to the correlation coefficient. All correlations are significant on the $P < 0.05$ level. Peak activations lie within regions of ACC and left insula that were activated in the simple contrast pain–no pain in others. Coordinates refer to peak activations and are in mm.

References and Notes

16. Materials and methods are available as supporting material on Science Online.
The experience of pain arises from both physiological and psychological factors, including one’s beliefs and expectations. Thus, placebo treatments that have no intrinsic pharmacological effects may produce analgesia by altering expectations. However, controversy exists regarding whether placebos alter sensory transmission, pain affect, or simply produce compliance with the suggestions of investigators. In two functional magnetic resonance imaging (fMRI) experiments, we found that placebo analgesia was related to decreased brain activity in pain-sensitive brain regions, including the thalamus, insula, and anterior cingulate cortex, and was associated with increased activity during anticipation of pain in the prefrontal cortex, providing evidence that placebos alter the experience of pain.

The idea that sensory experience is shaped by one’s attitudes and beliefs has gained currency among psychologists, physicians, and the general public. Perhaps nowhere is this more apparent than in our ability to modulate pain perception. A special case of this phenomenon is placebo analgesia, in which the mere belief that one is receiving an effective analgesic treatment can reduce pain (1–5). Recently, some researchers have attributed placebo effects to response bias and/or to publication biases (6), which raises the issue of whether placebo treatments actually influence the sensory, affective, and cognitive processes that mediate the experience of pain.

One important piece of evidence that placebo effects are not simply due to response or publication bias is that such effects can be reversed by the mu-opioid antagonist naloxone (2, 3, 7), suggesting that some kinds of placebo effects may be mediated by the opioid system. However, naloxone has also been shown to produce hyperalgesia independent of placebo, in some cases offsetting rather than blocking the effects of placebo analgesia (8). Although pharmacological blockade provides suggestive evidence regarding the neurochemical mechanisms mediating placebo effects, such data do not illuminate the nature of the information-processing system that gives rise to such effects. Neuroimaging data can provide complementary evidence of how pain processing in the brain is affected by placebos and about the time course of pain processing. Identifying placebo-induced changes in brain activity in regions associated with sensory, affective, and cognitive pain processing (9) may provide insight into which components of pain processing are affected by placebo. In addition, identifying changes that occur at particular times—in anticipation of pain, early or late during pain processing—may shed light on how cognitive systems mediating expectancy interact with pain and opioid systems.

In two functional magnetic resonance imaging (fMRI) experiments (n = 24 and n = 23), we examined two hypotheses regarding the psychological and neural mechanisms that underlie placebo analgesia. Our first hypothesis was that if placebo manipulations reduce the experience of pain, pain-responsive regions of the brain should show a reduced fMRI blood oxygen level-dependent (BOLD) signal (a measure related to neural activity) during pain. Pain-responsive regions, or the “pain matrix,” include thalamus, somatosensory cortex, insula, and anterior cingulate cortex (10–14). Our second hypothesis was that placebo modulates activity of the pain matrix by creating expectations for pain relief, which in turn inhibit activity in pain-processing regions. Converging evidence suggests that the prefrontal cortex (PFC), the dorsolateral aspect (DLPFC) in particular, acts to maintain and appropriately update internal representations of goals and expectations, which modulate processing in other brain areas (15, 16). Thus, stronger PFC activation during the anticipation of pain should correlate with greater placebo-induced pain relief as reported by participants and greater placebo-induced reductions in neural activity within pain regions (17).

Placebo reduces reported pain and brain activity in Study 1 (shock pain). The design of Study 1 is illustrated in Fig. 1A (see the figure legend for a description) (18). First, to confirm that application of shock elicits a neural response in pain-related areas, we compared brain activity in the intense shock versus no shock conditions. This revealed activation of the classic pain matrix (11, 14, 19, 20), including thalamus, primary somatosensory cortex/primary motor cortex (S1/M1), secondary somatosensory cortex (SII), midbrain, anterior insula, anterior cingulate cortex (ACC), ventrolateral prefrontal cortex, and cerebellum (Fig. S1). As expected, activations in thalamus, S1, SII, and M1 were larger in the left hemisphere, contralateral to the wrist where shocks were applied, whereas cerebellar activation was ipsilateral, although some bilateral activation was observed in each of these areas. We also
The Prepared Mind: Neural Activity Prior to Problem Presentation Predicts Subsequent Solution by Sudden Insight
John Kounios, Jennifer L. Frymiare, Edward M. Bowden, Jessica I. Fleck, Karuna Subramaniam, Todd B. Parrish and Mark Jung-Beeman

Psychological Science 2006 17: 882
DOI: 10.1111/j.1467-9280.2006.01798.x

The online version of this article can be found at:
http://pss.sagepub.com/content/17/10/882
The Prepared Mind

Neural Activity Prior to Problem Presentation Predicts Subsequent Solution by Sudden Insight

John Kounios, Jennifer L. Frymiare, Edward M. Bowden, Jessica I. Fleck, Karuna Subramaniam, Todd B. Parrish, and Mark Jung-Beeman

ABSTRACT—Insight occurs when problem solutions arise suddenly and seem obviously correct, and is associated with an “Aha!” experience. Prior theorizing concerning preparation that facilitates insight focused on solvers’ problem-specific knowledge. We hypothesized that a distinct type of mental preparation, manifested in a distinct brain state, would facilitate insight problem solving independently of problem-specific knowledge. Consistent with this hypothesis, neural activity during a preparatory interval before subjects saw verbal problems predicted which problems they would subsequently solve with, versus without, self-reported insight. Specifically, electroencephalographic topography and frequency (Experiment 1) and functional magnetic resonance imaging signal (Experiment 2) both suggest that mental preparation leading to insight involves heightened activity in medial frontal areas associated with cognitive control and in temporal areas associated with semantic processing. The results for electroencephalographic topography and frequency (Experiment 1) and functional magnetic resonance imaging signal (Experiment 2) both suggest that mental preparation leading to insight involves heightened activity in medial frontal areas associated with cognitive control and in temporal areas associated with semantic processing. The results for electroencephalographic topography and frequency (Experiment 1) and functional magnetic resonance imaging signal (Experiment 2) both suggest that mental preparation leading to insight involves heightened activity in medial frontal areas associated with cognitive control and in temporal areas associated with semantic processing. The results for electroencephalographic topography and frequency (Experiment 1) suggest that noninsight preparation, in contrast, involves increased occipital activity consistent with an increase in externally directed visual attention. Thus, general preparatory mechanisms modulate problem-solving strategy.

When Louis Pasteur said, “Chance favors only the prepared mind,” he was likely referring to preparation for the sort of sudden illumination that enables one to solve a difficult problem or reinterpret a situation in a new light (Wallas, 1926). Psychologists later called this type of sudden comprehension insight (Smith & Kounios, 1996; Sternberg & Davidson, 1995), a phenomenon associated with performance on tests of intelligence and creativity (Ansbury & Hill, 2003; Davidson, 1995). Although insights pop into awareness unexpectedly, or even unbidden (Kvavilashvili & Mandler, 2004; Metcalfe & Wiebe, 1987; Smith & Kounios, 1996), Pasteur apparently believed that some form of preparation facilitates insight. One type of preparation involves studying a problem or relevant background information (Seifert, Meyer, Davidson, Patalano, & Yaniv, 1995; Wallas, 1926). Such study is obviously helpful, but probably facilitates problem solving by both insight and noninsight analytic processing.

We hypothesized another type of preparation, one that does not depend on information related to specific problems, but that biases a person toward processing that facilitates solution by insight. Here, we demonstrate that preparation for problem solving can be associated with distinct brain states, one biasing toward solution with insight, the other biasing toward solution without insight. We examined neural activity associated with subjects’ preparation immediately prior to the presentation of each problem and found that the spatial distribution and oscillatory frequency of this activity predicts whether the problem that follows will be solved with insight or noninsight processing, as marked by the presence or absence of an “Aha!” experience. Insight has typically been studied by comparing performance on insight problems, which are often solved with an “Aha!” experience, with performance on noninsight problems, which are usually solved without an “Aha!” (Mayer, 1995; Weisberg, 1995). Unfortunately, such classification is not definitive, because any particular problem could be solved with or without insight (Bowden, Jung-Beeman, Fleck, & Kounios, 2005). Instead, in the present study, we used each subject’s trial-by-trial judgments of whether each solution became available incrementally or as a sudden insight to classify solutions to individual problems as either insight or noninsight solutions.
problems as resulting from insight or noninsight processing. Using this approach, previous studies have demonstrated unique patterns of behavioral results (Bowden & Jung-Beeman, 2003a) and neural activity (Jung-Beeman et al., 2004) associated with insight versus noninsight solutions.

For several reasons, we have inferred that these self-reported “Aha!” experiences reflect the sudden conscious availability of a solution rather than some ancillary process. For example, (a) the associated neural activity, a sudden burst of gamma-band oscillatory activity in the right anterior superior temporal gyrus, does not reflect subjects’ affective or surprise reactions following solutions, because the onset of this activity coincides with, rather than follows, the conscious availability of the solution; (b) task-related activity occurs in the same region when people first start processing a problem, before they experience any solution-related emotional response (Jung-Beeman et al., 2004); and (c) this region is a polymodal association area that has not been implicated in affective or novelty processing (Jung-Beeman, 2005).

In two experiments, using electroencephalography (EEG) and functional magnetic resonance imaging (fMRI), we assessed neural activity as people prepared to solve each problem in a series. We examined activity prior to the presentation of each problem in order to assess patterns independent of specific problems and their difficulty. Experiment 1 focused on EEG power within the alpha (8–13 Hz) frequency band. Alpha, the brain’s dominant rhythm (Shaw, 2003), reflects cortical deactivation and is inversely related to hemodynamic and metabolic measures of neural activity (Cook, O’Hara, Uijtdehaage, Mandelker, & Leuchter, 1998; Cooper, Croft, Dominey, Burgess, & Gruzelier, 2003; Goldman, Stern, Engel, & Cohen, 2002; Laufs et al., 2003; Ray & Cole, 1985; Worden, Foxe, Wang, & Simpson, 2000). The topographic distribution of alpha across the scalp therefore mirrors the spatial distribution of neural activity (Pfurtscheller & Lopes da Silva, 1999). We focused on the low-alpha frequency band (8–10 Hz), because high alpha (10–13 Hz) is typically dominated by an occipital alpha rhythm reflecting gating of visual information (Gevins & Smith, 2000). Effects were also found in the gamma band (> 30 Hz), but could not be reliably distinguished from electromyograph artifact, so they are not discussed in this report.

**EXPERIMENT 1**

**Method**

After giving informed consent, 19 subjects attempted to solve 186 problems. On each trial (Fig. 1), they indicated by bimanual button press that they were prepared to begin working on a problem, thereby initiating the display of a visual fixation mark. This button press was the midpoint of the 2-s epoch selected as the **preparation interval** of interest. After 1 s, this fixation mark was replaced by the three words of a **compound remote-associates problem**. For each problem (e.g., pine, crab, sauce), subjects attempted to produce a solution word (e.g., apple) that could be combined with each of the three problem words to form a common compound or phrase (pineapple, crabapple, applesauce). These problems were used because a large number are available, they can be solved relatively quickly, and similar problems have been used successfully in numerous studies of insight and creativity (for review, see Bowden & Jung-Beeman, 2003b). Also, they can be solved either with insight or with noninsight analytic processes (Bowden & Jung-Beeman, 2003a; Jung-Beeman et al., 2004), so we could compare these two general strategies while holding task and problem type constant. Previous studies have shown that subjects solving such problems tend to use each of these strategies about half the time (Bowden et al., 2005).

If subjects achieved solution, they made an immediate bimanual button press indicating that they had solved the problem, verbalized the solution when prompted, and then, when prompted, pressed one of two buttons to indicate whether or not the solution had been achieved by insight. (Prior instructions to subjects explained the notion of insight as sudden awareness of the solution—Jung-Beeman et al., 2004). After a 2-s intertrial

---

**Fig. 1.** Time line of events on a trial in Experiment 1. A “Ready?” message was displayed on a monitor, and the subject made a bimanual button press (“S”) when he or she was prepared to start the trial and view the three words constituting a problem. The problem was displayed after a 1-s visual fixation mark. The subject responded with another button press (“R”) as soon as he or she had solved the problem. This initiated a prompt to verbalize the solution (in this case, goose), followed by another prompt to make a button press to indicate whether the verbalized solution was accompanied by an experience of insight. This report presents electroencephalography (EEG) results for the bracketed preparatory interval consisting of the 2 s prior to the display of the problem.
interval, a “Ready?” prompt appeared; when ready, subjects initiated the next trial with a bimanual button press. High-density (128-channel) EEG was continuously recorded (referenced to digitally linked mastoids) at 250 Hz (0.02–100 Hz). Eyeblink artifacts were removed using EMSE 5.0 (Source Signal Imaging, Inc., www.sourcesignal.com). Trials containing other artifacts were identified by visual inspection and deleted. EEG segments corresponding to the preparation intervals were extracted and, on the basis of subjects’ performance on the following problem, were sorted into three types of preparation: preparation leading to solution with insight, preparation leading to solution without insight, or time-out (no response within 30 s). (Errors were rare, and trials with errors were deleted.) EEG power was estimated by computing power spectra on these segments. The University of Pennsylvania’s institutional review board approved the study.

Results
Participants solved 46.2% (SD = 8.2) of the problems correctly within the 30-s time limit. Subjects labeled 56.2% (SD = 8.3) of solutions as insight solutions (average median response time = 7.70 s, SD = 2.60) and 42.5% (SD = 8.9) as noninsight solutions (7.48 s, SD = 3.08). Of all responses, 11.5% (SD = 11.3) were errors.

As predicted, distinct alpha topographies were associated with preparation for solving problems with insight versus non-insight processing. Moreover, visual inspection of power spectra suggested that insight preparation and noninsight preparation were associated with different peak oscillatory frequencies within the low-alpha frequency band. These observations were substantiated by two sets of analyses.

The first set contrasted insight preparation against preparation that led to time-outs, and separately contrasted noninsight preparation against time-out preparation. These analyses also allowed us to assess whether preparation preceding successful problem solving differed from preparation preceding unsuccessful problem processing. Two repeated measures analyses of variance (ANOVAs) were performed on log-transformed EEG power to examine anterior-posterior and hemispheric differences in two low-alpha subbands (the frequency factor: 8–9 Hz vs. 9–10 Hz, determined by visual inspection of the power spectra) for insight, noninsight, and time-out trials (the trial-type factor). The first ANOVA examined power at left and right anterior-frontal (AF1/2) and left and right occipital (O9/10) electrodes (allowing comparison of scalp topographies). These electrodes were chosen a priori in order to maximize distance between electrodes, thereby minimizing the influence of EEG volume conduction. Topographic factors with two levels were utilized in the initial ANOVA to avoid reduction in statistical power associated with correction for nonsphericity (Dien & Santuzzi, 2005). Relevant significant effects included a Frequency × Trial Type × Anterior-Posterior interaction, $F(2, 36) = 3.95$, $p_{rep} = .93, \tilde{\nu}^2 = .18$. A second ANOVA utilized electrodes with more lateral placements (left and right frontal, F7/8; parietal, P7/8; and temporal, T7/8) in order to assess possible hemisphere effects; this analysis yielded a significant Trial Type × Hemisphere interaction, $F(2, 36) = 4.07$, $p_{rep} = .93, \tilde{\nu}^2 = .18$.

To specify the topographic differences driving the interactions in the omnibus ANOVAs, we computed the statistical parametric maps shown in Figure 2a as follow-up tests. The use of these follow-up tests to specify effects driving the interaction is justified by the presence of the interaction in the overall

![Fig. 2. Results from Experiment 1: alpha-band electroencephalographic (EEG) topography during the 2-s preparatory interval before the problem was displayed. Plotted values are t scores of electrode-by-electrode comparisons. The maps in (a) show results for comparisons between unsolved problems (time-outs, or TOs) and problems solved with insight processing (I) or with noninsight processing (NI); the left column shows maps for comparisons in the 8- to 9-Hz frequency band, and the right column shows maps for comparisons in the 9- to 10-Hz band. Red and orange regions indicate electrode sites at which I or NI trials exhibited less alpha power (i.e., more neural activity) than TO trials; the middle 66% of the color scale is grayed out. The maps in (b) show comparisons between insight preparation at peak power (9–10 Hz) and noninsight preparation at peak power (8–9 Hz; insight preparation minus noninsight preparation). Yellow regions show electrode sites at which insight preparation exhibited less alpha power than noninsight preparation. Blue regions show electrode sites at which noninsight preparation exhibited less alpha power than insight preparation. The middle 33% of the color scale is grayed out.](image-url)
ANOVA. Compared with time-out preparation, insight preparation was associated with greater neural activity (i.e., less alpha power) peaking over midfrontal cortex (9-10 Hz; Fig. 2a, top right) and left anterior-temporal cortex (8-9 Hz and 9-10 Hz; Fig. 2a, top row). In contrast, noninsight preparation, compared with time-out preparation, was associated with greater neural activity (decreased 8- to 9-Hz alpha; Fig. 2a, bottom left) peaking over occipital cortex.

Notably, insight and noninsight preparation (i.e., the two forms of successful preparation) showed no common differences from time-out (i.e., unsuccessful) preparation. This suggests that subjects did not fail to prepare on some trials (e.g., by not attending); rather, they failed on some trials because they could not solve those particular problems (quickly enough). They likely engaged in more insight processing than noninsight processing on some unsuccessful trials and in more noninsight processing than insight processing on others, though we have no way of sorting those trials.

We explored the data further in a set of more focused analyses. An ANOVA compared power values for insight trials (in the 9- to 10-Hz range) and noninsight trials (in the 8- to 9-Hz range) at left and right anterior-frontal (AF1/2) and occipital (O9/10) electrodes. Direct comparison of the topography corresponding to the peak low-alpha frequency associated with insight preparation against the topography corresponding to the peak low-alpha frequency associated with noninsight preparation optimized the spatial resolution of the comparison. Relevant findings included a significant Insight × Anterior-Posterior interaction, $F(1, 18) = 8.46, p_{rep} = .97, e^2 = .32$. There were also marginally significant Insight × Hemisphere, $F(1, 18) = 4.04, p_{rep} = .91, e^2 = .18$, and Insight × Anterior-Posterior × Hemisphere, $F(1, 18) = 3.04, p_{rep} = .88, e^2 = .14$, interactions. To specify the topographic differences underlying these interactions, we computed a statistical parametric map as a follow-up test (Fig. 2b). Direct comparison of peak low-alpha topographies associated with insight preparation (9-10 Hz) versus noninsight preparation (8-9 Hz) showed greater neural activity (i.e., less alpha) associated with insight preparation peaking over midfrontal, left temporal, right temporal, right inferior frontal, and bilateral somatosensory cortex and greater activity associated with noninsight preparation over a broad region of posterior cortex (Fig. 2b). (The somatosensory activation likely reflects activity associated with the bimanual button press subjects made to initiate a trial.)

It is possible that a subject’s brain state changes slowly over the course of an experimental session, because of practice or fatigue. However, the proportion of problems solved by insight did not change from the first to the second half of the session ($F < 1$). Another possibility is that a subject’s brain state varies on a shorter time scale, perhaps over the course of a few trials, entering an insight or a noninsight mode for some time before changing state. This would result in temporal clustering of trial types. However, temporal clustering of trial types in our data did not differ from that of simulated random data (all $F$s < 1.0). These results suggest that activity associated with insight versus noninsight processing changed within the few seconds between trials. Given that subjects initiated the presentation of each problem when they felt prepared, these trial-by-trial changes in neural activity likely reflect differential preparation prior to the presentation of individual problems.

**EXPERIMENT 2**

EEG allowed us to isolate the preparatory interval with high precision, but provided less precise spatial information about the relevant brain regions. In Experiment 2, fMRI confirmed and specified the brain regions involved in preparation, though with less precise isolation of the preparatory interval.

**Method**

After giving informed consent, 25 subjects attempted to solve 135 compound remote-associate problems during fMRI scanning; the time limit for each problem was 15 s. Five subjects were replaced: Four showed excessive movement or poor MRI signal, and one responded “noninsight” on only two trials.

The paradigm described for Experiment 1 was modified slightly to optimize fMRI data acquisition. The preparation interval preceding each problem comprised a rest period (with a fixation cross) of varying length (2, 4, 6, or 8 s, randomly chosen) that followed the insight judgment (after a solved problem) or the time-out event (after an unsolved problem). No button press was required to initiate a trial, and subjects could not predict the length of the rest period, so the preparation was likely somewhat more passive than in Experiment 1.

Scanning was performed at Northwestern University’s Center for Advanced MRI using a 3-T Siemens Trio scanner with standard head coil. Head motion was restricted with plastic calipers built into the coil and a vacuum pillow. Anatomical high-resolution T1-weighted images were acquired in the axial plane at the end of every session. In-plane functional images were acquired using a gradient echo-planar sequence (time to repetition, TR = 2 s for 38 slices that were 3 mm thick, time to echo = 20 ms, matrix size of 64 × 64 in a 220-mm field of view). Each of five runs began with an 8-s saturation period; then participants solved problems for up to 10 min, 20 s (the final run was truncated when subjects finished solving problems).

Functional and anatomical images were co-registered through time, spatially smoothed with a 7.5-mm Gaussian kernel, and fit to a common template. The data were analyzed using general...
linear model analysis that extracted average estimated response to each trial type, correcting for linear drift and removing signal changes correlated with head motion. For each participant, an event-related analysis contrasted fMRI signal for insight versus noninsight preparation intervals (for 4 TRs reflecting 4.1–12.0 s following the onset of the preparation period). Between-subjects consistency in difference scores (insight preparation minus noninsight preparation) was analyzed in a second-stage random-effects analysis. The significance threshold combined cluster size and t values for every voxel within a cluster: Clusters exceeded 1,000 mm$^3$ in volume, with each voxel reliably different across participants, $t(24) = 3.374, p < .0025$, uncorrected. This combination threshold yields very low false-positive rates in simulations.

Signal acquisition and initial data analyses for this experiment are described in detail elsewhere (Virtue, Haberman, Clancy, Parrish, & Jung-Beeman, in press). Northwestern University’s institutional review board approved the study.

Results

Participants solved 47% ($SD = 9.9$) of the problems correctly within the 15-s time limit and labeled 55.8% ($SD = 13.9$) of solutions as insight solutions (median response time = 5.84 s, $SD = 1.24$) and 43.4% ($SD = 13.6$) as noninsight solutions (median response time = 7.81 s, $SD = 1.79$). Of all responses, 6.6% ($SD = 5.4$) were errors.

The fMRI signal changed in several brain areas during the preparation period. Most areas showed decreasing signal during preparation, as neural activity returned to baseline. A few areas showed increased signal, indicating increasing neural activity during “rest.” This increase was strongest in anterior cingulate cortex (ACC). Small regions within posterior cingulate cortex (PCC) and bilateral posterior middle and superior temporal gyri (M/STG) also showed slightly increasing or sustained activity during preparation.

Of greater interest is that signal change during the preparation interval varied systematically according to whether the subsequent problem was solved with insight versus noninsight (Fig. 3, Table 1). The fMRI results essentially replicate those observed with EEG in Experiment 1. In ACC, preparation that preceded solutions with insight increased fMRI signal more than did preparation that preceded solutions without insight. In PCC and bilateral posterior M/STG, signal was stronger for insight than noninsight preparation, mostly because signal decreased more for noninsight than insight preparation. Both significant temporal clusters appeared to comprise several smaller clusters (a small subregion in each hemisphere showed greater signal increase for insight preparation; the remainder of each cluster showed greater signal decrease for noninsight preparation). There were also small, nonsignificant clusters in the anterior-temporal region. No clusters showing the opposite effect exceeded our combined significance threshold. However, at lower t thresholds, the largest cluster (375 mm$^3$ at $p < .01$, 1,344 mm$^3$ at $p < .05$) showing stronger signal for noninsight than insight preparation was in left middle and inferior occipital cortex (Table 1).

GENERAL DISCUSSION

The observed effects on EEG topographies, peak low-alpha frequencies, and fMRI signal all demonstrate that the neuronal populations active prior to the presentation of problems subsequently solved with insight are different from the neuronal populations active prior to the presentation of problems subsequently solved with noninsight processing. Note that differences in brain activity during preparation were not influenced by the specific content of the problems, which were not displayed until after the epoch examined. Thus, subjects engaged distinct patterns of mental preparation.

Although the mere demonstration of such preparatory states is important, the anatomical pattern evident across the two experiments also suggests specific preparatory mechanisms facilitating insight. The convergence across methods is critical, as the fMRI data, providing specific anatomical information, replicate the effects of insight preparation observed with EEG.$^2$

We begin with the ACC region identified with fMRI. This is the likely source of the insight-related midfrontal activity observed with EEG.$^3$ Beyond showing an insight-noninsight difference, this region showed the most robust pattern of increasing fMRI signal during the preparation period, that is, during “rest.” ACC has been associated with monitoring for competition among potential responses or processes. Such conflict monitoring can signal the need for top-down cognitive control mechanisms facilitating the maintenance or switching of attentional focus, or selection from competing responses (Badre & Wagner, 2004; Botvinick, Cohen, & Carter, 2004; Kerns et al., 2004; Miller & Cohen, 2001). According to this interpretation, increased ACC activity is followed by increased top-down control. Notably, ACC activity increased during the preparation period, when no obvious response conflict existed. Recent studies suggest a possible reason: ACC may be involved in suppressing irrelevant thoughts (Anderson et al., 2004; Wyland, Kelley, Macrae, Gordon, & Heatherton, 2003), such as daydreams or thoughts related to a preceding event (in this case, the preceding trial). Thus, ACC activity may have allowed participants to attack the next problem with a “clean slate.” This explanation assumes that insight processing is more susceptible to internal interference.

$^2$We did not contrast insight or noninsight preparation against time-out preparation in Experiment 2 (as we did in Experiment 1), because the contrast would have been confounded by some temporally proximal solutions, which would have been absent in the time-out trials, and which would have affected fMRI signal.

$^3$Most prior studies have associated ACC activity with changes in theta-band EEG power (Luu, Tucker, & Makeig, 2004), but some studies have associated it with alpha-band oscillations (Goldman et al., 2002).
than is noninsight processing (Schooler, Ohlsson, & Brooks, 1993), thereby necessitating greater suppression of extraneous thoughts.

Thus, in the context of problem solving, the activity observed in ACC prior to insight may reflect increased readiness to monitor for competing responses, and to apply cognitive control mechanisms as needed to (a) suppress extraneous thoughts, (b) initially select prepotent solution spaces or strategies, and, if these prove ineffective, (c) subsequently shift attention to a nonprepotent solution or strategy. Such shifts are characteristic of insight.

We speculate that during insight preparation, cognitive control mechanisms were modulating activity in brain areas related to semantic processing. Greater neural activity was observed for insight than for noninsight preparation in bilateral temporal cortex (with this activity being more extensive on the left than on the right, in both experiments). We propose that this temporal-lobe activity reflects preparation for semantic activation and
results from top-down processes such as those associated with the ACC. Within a theoretical framework of bilateral semantic processing (Jung-Beeman, 2005; also Faust & Lavidor, 2003) partly based on cytoarchitectonic studies (Hutsler & Galuske, 2003), the bilateral (or slightly leftward) distribution of this activity indicates that subjects were prepared to retrieve both prepotent associations (predominantly in the left posterior M/STG) and weaker associations (in the right posterior M/STG).

This hypothesis is also consistent with a recently proposed model of the neural basis of insight in verbal problem solving (Bowden et al., 2003; Jung-Beeman et al., 2004) according to which solvers initially focus on prepotent associations, possibly reaching an impasse if this does not lead to a solution. However, solvers may maintain weak activation for the solution (Beeman & Bowden, 2000) due to coarser semantic coding in the right hemisphere (Bowden & Beeman, 1998; Bowden & Jung-Beeman, 2003a). Solvers may overcome impasse by switching attention to this weakly activated representation, suddenly increasing its strength. This shift of attention to a nonprepotent solution (or solving process) involves cognitive control mechanisms such as those associated with the ACC. This shift of attention may be less likely to occur when people use a less controlled mode of processing based on passive accrual of evidence.

This interpretation is consistent with a sudden increase in high-frequency EEG power (i.e., gamma-band oscillations peaking at 40 Hz) observed to occur about 0.3 s before insight (compared with noninsight) solutions (Jung-Beeman et al., 2004). This burst of gamma-band activity was focused at right anterior superior-temporal electrodes and corresponded to activity detected by fMRI in the underlying cortex (with no significant insight effect in the left anterior temporal lobe). This right anterior temporal activity may reflect the sudden emergence into consciousness of the correct solution (Jung-Beeman et al., 2004).

The current results extend this prior work by demonstrating that this insight response is the culmination of mechanisms that begin even before a problem is presented. Insight preparation modulates processing and biases toward insight solutions by increasing readiness in (a) left posterior temporal cortex, whose activation is hypothesized to reflect readiness to initially pursue prepotent associations, and (b) ACC, whose activation is hypothesized to reflect cognitive control mechanisms that facilitate initially attending to prepotent associations, then discretely shifting attention to nonprepotent associations— which may include solution-related information activated in the right anterior temporal region.

One additional area, PCC, showed stronger fMRI signal for insight than for noninsight preparation, perhaps reflecting attentional differences (Small et al., 2003). No corresponding effect was evident in the EEG data. It is possible, however, that an effect was present in this deep area, but was canceled out by the opposite effect measured over posterior cortex, which was likely generated in more superficial brain areas. Specifically, EEG revealed greater occipital-parietal activity during noninsight than during insight preparation; there was a similar, though not significant, effect in the fMRI results. It is possible that this effect was stronger in the EEG experiment because the preparation period was more active and predictable; subjects pressed a button to indicate readiness in the EEG experiment, but not in the fMRI experiment. This posterior cortical activity during

<table>
<thead>
<tr>
<th>Gyrus or structure</th>
<th>Brodmann Area</th>
<th>Volume (mm³)</th>
<th>Center coordinates</th>
<th>Signal change (%)</th>
<th>Effect size (d)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>y</td>
<td>z</td>
</tr>
<tr>
<td>Left posterior M/STG</td>
<td>39, 37, 22</td>
<td>2,031</td>
<td>−49</td>
<td>−62</td>
<td>15</td>
</tr>
<tr>
<td>Anterior cingulate</td>
<td>32, 24</td>
<td>1,922</td>
<td>−4</td>
<td>41</td>
<td>8</td>
</tr>
<tr>
<td>Posterior cingulate</td>
<td>31</td>
<td>1,594</td>
<td>−3</td>
<td>−47</td>
<td>32</td>
</tr>
<tr>
<td>Right posterior M/STG</td>
<td>37, 39, 22</td>
<td>1,266</td>
<td>52</td>
<td>−62</td>
<td>9</td>
</tr>
<tr>
<td>Left amygdala</td>
<td>—</td>
<td>438</td>
<td>−25</td>
<td>−7</td>
<td>−9</td>
</tr>
<tr>
<td>Left middle temporal gyrus</td>
<td>21</td>
<td>438</td>
<td>−62</td>
<td>−32</td>
<td>−6</td>
</tr>
<tr>
<td>Noninsight preparation &gt; insight preparationa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left middle and inferior occipital gyrus</td>
<td>18</td>
<td>375</td>
<td>−30</td>
<td>−98</td>
<td>−2</td>
</tr>
</tbody>
</table>

Note. We used a strict threshold for significance, requiring a cluster size of 1,000 mm³ and requiring that all voxels show a consistent effect across subjects, t(24) = 3.374, p < .0025. For thoroughness, all clusters down to 300 mm³ are listed. For each cluster listed, we report the signal difference between insight and noninsight preparation as a percentage of the average signal within the cluster, as well as the average and maximum t score for all voxels within the cluster. For significant clusters, we include the cluster-wise effect size, d. M/STG = middle and superior temporal gyri.

aNo clusters showed significantly greater signal for noninsight preparation than for insight preparation, so we list the largest area at a lower threshold, 375 mm³ at t(24) = 2.795, p < .05.
noninsight preparation may suggest readiness to pursue a less controlled, bottom-up mode of processing, for example, by increasing visual attention just before the problem is displayed.

CONCLUSIONS

The present study demonstrates that a person’s preparatory brain state even prior to seeing a problem influences whether the person will solve that problem with insight or noninsight processing. Insight preparation could be characterized as preparing to strongly activate prepotent candidate solutions while also preparing to switch attention to nonprepotent candidates, thereby enabling retrieval of weakly activated solutions characterized by remote associations among problem elements. In contrast, noninsight preparation could be characterized as external attentional focus on the source of the imminent problem. The fact that subjects use both of these forms of preparation suggests either that they spontaneously alternate strategies or that one form of preparation, presumably insight preparation with its top-down component, is perhaps too cognitively demanding to use for every problem in a series.

Future studies will further specify these preparation-related brain states and their determinants. Ideally, this line of research could lead to the development of techniques for facilitating or suppressing insight in order to optimize performance for different types of problems and contexts. In sum, this work may show how to lessen the impact of chance on efforts to reach insightful solutions, a goal that Pasteur would likely have endorsed.

Acknowledgments—Grants DC-04818 (to J.K.) and DC-04052 (to M.-J.-B.) from the National Institutes of Deafness and Other Communication Disorders supported this research. We thank D. Stephen Lindsay, James Cutting, and an anonymous reviewer for helpful comments.

REFERENCES


Davidson (Eds.), *The nature of insight* (pp. 3–32). Cambridge, MA: MIT Press.


(RECEIVED 8/1/05; REVISION ACCEPTED 11/29/05; FINAL MATERIALS RECEIVED 2/1/06)
Inhibitory Control is Slowed in Patients with Right Superior Medial Frontal Damage

Darlene Floden and Donald T. Stuss

Abstract

Inhibitory control is an essential part of behavior. Comprehensive knowledge of the neural underpinnings will shed light on complex behavior, its breakdown in neurological and psychological disorders, and current and future techniques for the pharmacological or structural remediation of disinhibition. This study investigated the neural mechanisms involved in rapid response inhibition. The stop signal task was used to estimate inhibitory speed in a group of neurologically normal control subjects and patients with discrete frontal lobe lesions.

INTRODUCTION

Response inhibition is a cardinal element of efficient behavioral control. A range of tasks, populations, and methods has been used to investigate response inhibition. Impaired inhibitory control is associated with the functional disturbances found in disorders involving known or suspected frontal lobe dysfunction such as attention deficit/hyperactivity disorder (Schachar, Mota, Logan, Tannock, & Klim, 2000) and traumatic brain injury (Robertson, Manly, Andrade, Baddeley, & Yiend, 1997). Early neuropsychological studies of go/no-go tasks demonstrated deficits in patients with frontal lobe damage (i.e., Decary & Richer, 1995; Leimkuhler & Mesulam, 1985; Drewe, 1975), but the precise neuroanatomical correlates have not been extensively studied in patients. Several recent functional neuroimaging studies suggest that sites on the medial and lateral aspects of the right frontal lobe participate in networks underlying response inhibition (i.e., Garavan, Ross, Murphy, Roche, & Stein, 2002; Menon, Adleman, White, Glover, & Reiss, 2001; Rubia et al., 2001). Converging evidence from lesion studies, controlling for response speed and assessing patients with focal and defined frontal lesions, is needed to verify the necessary role of these regions and the nature of their contributions to inhibitory processing. To address these issues, we administered a carefully controlled response inhibition procedure known as the stop signal task to patients with well-documented focal frontal lesions.

The stop signal task is unique among response inhibition tasks in that it provides an estimate of inhibitory speed (Logan & Cowan, 1984). The task is a choice reaction time (RT) procedure where, on a fraction of trials, a stop stimulus signals the subject to inhibit an already initiated response (see Figure 1). The onset asynchrony of the stop signal (stimulus onset asynchrony [SOA]) is varied to increase or decrease the likelihood of successfully stopping, and the observed inhibition success rate at each SOA is used to estimate RT to the stop signal (or stop signal RT [SSRT]). The task has demonstrated sensitivity to inhibitory control problems in children diagnosed with attention deficit/hyperactivity disorder (Schachar et al., 2000) and its treatment with methylphenidate (Tannock, Schachar, & Logan, 1995). We reasoned that the task would also be sensitive to inhibitory deficits in patients with focal frontal lobe lesions. We specifically hypothesized that in the experimental condition that controls for strategic changes in response speed, reduced inhibitory control would be related to damage at right medial and/or lateral sites but not other frontal lesion locations.

There is some limited evidence that inhibitory speed on the stop signal task is impaired after frontal lobe damage (although see Dimitrov et al., 2003). Rieger, Gauggel, and Burmeister (2003) found that right or bilateral frontal lobe lesions or basal ganglia damage produced slower SSRTs, but they did not attempt to identify more specific anatomical correlations. Another study in patients with lesions of the right frontal lobe (Aron, Fletcher, Bullmore, Sahakian, & Robbins, 2003)
reported that slow SSRTs were related to lesions in the inferior frontal gyrus, which has been implicated repeatedly in functional neuroimaging studies of go/no-go and stop signal tasks (Brass, Derrfuss, & von Cramon, 2005; Menon et al., 2001; Rubia et al., 2001). These same imaging studies also demonstrate consistent activations in superior medial regions, including anterior cingulate, supplementary motor area (SMA), and pre-SMA. In fact, Aron et al. (2003) found a correlation between anterior cingulate lesions and slow inhibitory speed. However, they also reported that when inferior frontal lesions were factored into the analysis, the influence of anterior cingulate lesions was not significant.

Recent work by Garavan et al. (2002) offers a possible explanation for the conflicting imaging and lesion findings regarding the role of superior medial regions. Using functional magnetic resonance imaging and electroencephalography during performance of go/no-go tasks, they showed that right lateral frontal and parietal activation were related to successful stopping on relatively “easy” inhibition conditions (i.e., relatively slower RTs, or go RTs, possibly reflecting weaker response set). In contrast, a superior medial region involving the posterior portion of the anterior cingulate cortex and pre-SMA was more active for successful stop trials when go RTs were fast leading up to the stop trial, suggesting a rapid “kill-switch” type of inhibitory control. From this, they proposed two networks that participate in inhibition under different conditions. Correlations with absentmindedness scores on the Cognitive Failures Questionnaire (Broadbent, Cooper, Fitzgerald, & Parkes, 1982) suggested that high absentminded subjects tended to rely on the fast medial inhibitory system, whereas low absentminded subjects relied more on a strategic lateral inhibitory system.

In the stop signal task study of Aron et al. (2003), the frontal lobe group had significantly longer go RTs than controls and it is possible that patients deliberately slowed their response times to avoid making errors on stop trials. The observations of Garavan et al. (2002) would suggest that inhibition under these conditions (strategically slowed response speed) would rely on more lateral frontal regions rather than medial regions. A control condition measuring go RT without additional stop trials is necessary to evaluate baseline response speed to confirm that subjects are performing the choice RT task as quickly as possible. Prior work on the psychometric properties of the stop signal task has shown that when subjects delay responses in anticipation of the rare stop signals, estimates of inhibitory speed are artificially inflated (van den Wildenberg, van der Molen, & Logan, 2002). To circumvent this potential obstacle, the current study included a control condition to verify that subjects continued to perform the task at baseline speeds.

METHODS

Participants

Participants were 23 patients (17 men, 19 right-handed; age, 50.6 years, SD = 12.7 years; education = 14.3 years, SD = 2.4 years) with single focal frontal lobe lesions (7 vascular, 10 tumor/epilepsy resections, 6 traumatic focal contusions without evidence of diffuse injury; mean chronicity = 32, SD = 22 months) (see Stuss, Alexander, et al., 2002, for rationale on inclusion of these patient groups) and 19 age- and education-matched control subjects (11 men, 14 right-handed; age, 49.5 years, SD = 11.6 years; education = 14.6 years, SD = 1.9 years). One patient with a left lateral lesion arising from a vascular...
insult was excluded because he failed to show any stopping behavior throughout the task, which precludes estimation of inhibitory speed. Exclusion criteria included impaired/uncorrected vision/hearing, history of psychiatric disorder, substance abuse, or neurological disorder unrelated to lesion, estimated full-scale IQ <90, and diet-controlled diabetes.

Lesion Analysis

Lesions were documented from anatomical scans performed in the course of clinical care and depicted on a standard anatomical template based on the cytoarchitectonic divisions of Petrides and Pandya (1994). The full procedure is described in detail elsewhere (Stuss, Alexander, et al., 2002). Lesion locations were coded as present or absent in seven regions for each hemisphere: dorsolateral (4 [lateral], 6A [lateral], 8Ad, 8Av, 9 [lateral], 46, 9/46D, 9/46V), ventrolateral (4 [ventral], 6B, 44, 45A, 45B, 47/12 [lateral]), polar (10), superior medial (6A [medial], 8B, 9 [medial]), inferior medial (14 [medial], inferior 24, 25, 32), dorsal anterior cingulate (superior 24, 32), and orbitofrontal (11, 13, 14, 47/12 [orbital]). Lesion location, etiology, chronicity, and size are displayed in Table 1.

Stop Signal Task

Stimuli and instructions were programmed by using Mel v 2.01 for DOS and presented on an IBM-compatible

### Table 1. Lesion Characteristics

<table>
<thead>
<tr>
<th>Subject No.</th>
<th>Laterality</th>
<th>Pol</th>
<th>Orb</th>
<th>IM</th>
<th>AC</th>
<th>SM</th>
<th>DL</th>
<th>VL</th>
<th>NF</th>
<th>Etiology</th>
<th>Chronicity (months)</th>
<th>Lesion Size (% Whole Brain)</th>
</tr>
</thead>
<tbody>
<tr>
<td>517</td>
<td>Bilateral</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>0</td>
<td>0</td>
<td>Tumor</td>
<td>33</td>
<td>9.09</td>
</tr>
<tr>
<td>500</td>
<td>Bilateral</td>
<td>B</td>
<td>R</td>
<td>B</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>CVA</td>
<td>51</td>
<td>6.47</td>
</tr>
<tr>
<td>520</td>
<td>Bilateral</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>R</td>
<td>B</td>
<td>R</td>
<td>B</td>
<td>0</td>
<td>Trauma</td>
<td>59</td>
<td>4.33</td>
</tr>
<tr>
<td>526</td>
<td>Bilateral</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>L</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Trauma</td>
<td>5</td>
<td>3.32</td>
</tr>
<tr>
<td>528</td>
<td>Bilateral</td>
<td>B</td>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>Trauma</td>
<td>12</td>
<td>2.37 0.40</td>
</tr>
<tr>
<td>529</td>
<td>Bilateral</td>
<td>B</td>
<td>B</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>L</td>
<td>0</td>
<td>Trauma</td>
<td>24</td>
<td>1.42</td>
</tr>
<tr>
<td>515</td>
<td>Bilateral</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Tumor</td>
<td>82</td>
<td>1.36</td>
</tr>
<tr>
<td>531</td>
<td>Bilateral</td>
<td>B</td>
<td>B</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>R</td>
<td>0</td>
<td>Trauma</td>
<td>20</td>
<td>1.31 0.21</td>
</tr>
<tr>
<td>522</td>
<td>Bilateral</td>
<td>0</td>
<td>0</td>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>CVA</td>
<td>41</td>
<td>0.33</td>
</tr>
<tr>
<td>523</td>
<td>Right</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>CVA</td>
<td>48</td>
<td>12.67 0.22</td>
</tr>
<tr>
<td>527</td>
<td>Right</td>
<td>R</td>
<td>0</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>Trauma</td>
<td>10</td>
<td>2.05</td>
</tr>
<tr>
<td>507</td>
<td>Right</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>Epilepsy</td>
<td>34</td>
<td>1.19</td>
</tr>
<tr>
<td>518</td>
<td>Right</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>R</td>
<td>Tumor</td>
<td>59</td>
<td>0.68</td>
</tr>
<tr>
<td>535</td>
<td>Right</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>R</td>
<td>CVA</td>
<td>17</td>
<td>0.24 0.06</td>
</tr>
<tr>
<td>508</td>
<td>Right</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Tumor</td>
<td>19</td>
<td>0.20</td>
</tr>
<tr>
<td>506</td>
<td>Left</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>L</td>
<td>L</td>
<td>0</td>
<td>Epilepsy</td>
<td>39</td>
<td>3.07</td>
</tr>
<tr>
<td>504</td>
<td>Left</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>L</td>
<td>L</td>
<td>0</td>
<td>0</td>
<td>Tumor</td>
<td>72</td>
<td>2.08</td>
</tr>
<tr>
<td>505</td>
<td>Left</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Epilepsy</td>
<td>20</td>
<td>1.66</td>
</tr>
<tr>
<td>509</td>
<td>Left</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>0</td>
<td>Tumor</td>
<td>32</td>
<td>1.44</td>
</tr>
<tr>
<td>525</td>
<td>Left</td>
<td>0</td>
<td>L</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>L</td>
<td>0</td>
<td>Trauma</td>
<td>8</td>
<td>0.91</td>
</tr>
<tr>
<td>502</td>
<td>Left</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>L</td>
<td>L</td>
<td>0</td>
<td>0</td>
<td>Epilepsy</td>
<td>12</td>
<td>0.57</td>
</tr>
<tr>
<td>533</td>
<td>Left</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>L</td>
<td>L</td>
<td>0</td>
<td>L</td>
<td>CVA</td>
<td>40</td>
<td>0.21 0.02</td>
</tr>
<tr>
<td>534</td>
<td>Left</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>L</td>
<td>0</td>
<td>0</td>
<td>CVA</td>
<td>10</td>
<td>NA</td>
</tr>
</tbody>
</table>

Pol = polar (10); ORB = orbitofrontal (11, 13, 14, 47/12 [orbital]); IM = inferior medial (14 [medial], inferior 24, 25, 32); AC = dorsal anterior cingulate (superior 24, 32); SM = superior medial (6A [medial], 8B, 9 [medial]); DL = dorsolateral (4 [lateral], 6A [lateral], 8Ad, 8Av, 9 [lateral], 46, 9/46D, 9/46V); VL = ventrolateral (4 [ventral], 6B, 44, 45A, 45B, 47/12 [lateral]); NF = nonfrontal; B = bilateral; R = right; L = left; CVA = cerebrovascular accident.
Subjects were seated directly in front of the monitor at a distance of approximately 0.5 m. In the initial control condition, subjects received 40 choice RT trials requiring keypress responses to the letters X and O to evaluate baseline go RT to control for strategic slowing in subsequent blocks. Subjects were then informed that the task would be repeated but that an auditory cue would occasionally occur after a letter and that responses to that letter should be withheld if possible. Subjects performed four experimental blocks of 55 trials, each involving 15 (27%) random stop trials. Ten stop signals occurred at each of six fixed SOAs: 75, 150, 225, 300, 375, and 450 msec poststimulus. We adopted a fixed SOA procedure in light of computational studies of the psychometric properties of the stop signal task (Band, van der Molen, & Logan, 2003), which indicate that titrating SOA procedures are more susceptible to artifact introduced when subjects fail to initiate inhibitory processing (i.e., no stopping behavior is attempted).

SSRTs were calculated from the observed data at each SOA according to the method described in Logan and Cowan (1984). Briefly, the stop signal task is thought to reflect the race between a “go” process and a “stop” process. On a stop trial, a response is executed if the total time to receive and respond to the stop signal exceeds the go RT. Therefore, for each subject, the probability of responding given a stop signal was calculated for each SOA ($p_r|\text{SOA}$). This was multiplied by the total number of correct go trials to give the percentage of trials ($n$) where the stop process is slower than the go process. Go RT values were then rank ordered and the $n$th fastest go RT selected. For example, a $0.4$ probability of responding given SOA$_1$ means that $40\%$ of go RTs are faster than the total stopping time. The $n$th go RT minus the SOA equals the estimated SSRT at that SOA. Therefore, $\text{go RT}_n - \text{SOA}_1 = \text{SSRT}_1$. SSRT estimates were averaged for all SOAs (where $0 < p_r|\text{SOA} < 1$) for each subject.

Reaction time feedback for correct go trials was presented after each block to discourage subjects from delaying responses in the attempt to avoid responding on stop trials. Comparison of go RTs in the experimental and control conditions verified that all subjects maintained consistent go RTs or “response readiness” during stop blocks.

**RESULTS**

**Inhibitory Speed**

When ranked according to SSRT, seven patients showed SSRTs longer than 1.5 $\text{SD}$ above the control mean. Five of the seven had damage to the right superior medial region ($\chi^2 = 4.0$, $p < .05$). The other two patients had either left dorsolateral damage or left orbital and polar damage. Maximal lesion overlap in slow-SSRT patients occurred in the region that likely corresponds to the SMA and pre-SMA (see Figure 2). Mean SSRT (see Table 2) for all nine patients with right superior medial damage was significantly longer than control subjects, $t(26) = 2.1, p < .05$. For all other frontal regions, slowed SSRTs were present in less than half the subjects with...
lesions in that region. For example, five patients had
damage to the right inferior frontal gyrus but only one
demonstrated slowed SSRTs. SSRT was unrelated to le-
sion etiology, chronicity, and all demographic variables.

Choice RT Performance

Patients and control subjects showed equivalent go RT,
t(39) < 1, in the control condition. The encouragement
to maintain baseline speeds during experimental trials
was effective; go RT during stop blocks did not differ
from baseline levels for patients or controls (both t < 1).
Moreover, the subgroup of patients with right superior
medial damage did not differ from controls in baseline
go RT, t(26) < 1, and go RT did not differ from baseline
levels during stop blocks, t(8) < 1. This rules out the
possibility that a delayed responding strategy contribut-
ed to observed increases in SSRT. Go RT correlated with
age (r = .57, p < .001) but not lesion size (r = .12, p = ns),
or SSRT (r = .13, p = ns). Rate of response
errors on trials without a stop signal did not differ
between groups (F < 1).

DISCUSSION

The current study is the first demonstration that damage
to right superior medial frontal regions (specifically, SMA
and pre-SMA) impairs inhibitory control in the stop
signal task. In contrast to prior work (Aron et al.,
2003), lesions of the right lateral region did not contrib-
ute to inhibitory speed (SSRT). On the surface, these
two studies completely contradict each other. We would
argue, however, that the current results are complemen-
tary and provide a key differentiation in the processes
that underlie performance in response inhibition tasks.
The design of our task was unique in that subjects were
not able to strategically slow go RTs, and our results
conform to the findings of Garavan et al. (2002), which
suggested that inhibition in the context of speeded go
RTs relies more on a rapid, kill-switch inhibitory system
that involves superior medial regions. The anatomical
connectivity of this region provides the anatomical sub-
strates to accomplish such processing; major cortico-
spinal projections originate in this region, providing a
route for rapid control over motor effectors (Dum &

In contrast, when the task does permit a slowed go RT
strategy (as the procedure of Aron et al., 2003, might
have done), a more controlled response readiness pro-
cess related to right lateral may dictate inhibitory speed.
In a recent review, Aron, Robbins, and Poldrack (2004)
discussed a number of attention tasks, including their
stop signal task, that recruit the right inferior frontal
gyrus in a more strategic type of inhibitory control, and
our own work has suggested a right lateral involve-
m ent in response readiness over a number of task con-
texts (i.e., Stuss et al., 2005; Stuss, Binns, Murphy, &
Alexander, 2002). The anatomical connections of this
right lateral region with the SMA (Dum & Strick, 2002)
and the basal ganglia (Alexander, DeLong, & Strick, 1986)
also suggest a response biasing role rather than a direct
access to motor effectors capable of interrupting an
initiated response. Thus, we would argue that, together,
the results support the suggestion that both superior
medial and lateral regions of the right frontal lobe
appear to participate in response inhibition depending
on task requirements.

The probability of observing slow inhibitory control
on experimental tasks following damage to these re-
gions is also dependent on the method used to analyze
lesion locations. Volumetric analysis is an excellent
technique for detecting relationships between the se-
verity of impairment and the size of the lesion within a
set of predefined regions. The trade-off is that anatom-
ical regions must be relatively large in order to keep the
number of statistical comparisons (independent corre-
lations for each region) small to avoid problems with
statistical power. We have found in our earlier work that
volume measures do not allow identification of critical
areas within large regions and, in fact, result in lower
sensitivity to brain–behavior correlations. This is partic-
ularly true in the case of small sample sizes such as ours
where volume analysis may confound lesion size with
lesion location. Our approach is to use a priori criteria
for identifying poor performance in individual patients
(i.e., cutoff scores) and only then investigate whether
meaningful relationships exist with more focused lesion
locations. Using this technique, we were able to identify
a role for SMA and pre-SMA, separate from other

| Group               | Simple Go RT | | Go RT | | SSRT | | % Go Errors |
|---------------------|--------------|-----|-------|------|------|----------|
|                     | Mean         | SD  | Mean  | SD   | Mean | SD       | Mean | SD |
| Frontal RSM         | 532.8        | 124.8 | 549.6 | 116.9 | 308.7 | 81.2 | 3.1  | 2.7 |
| Frontal non-RSM     | 503.1        | 76.1  | 509.6 | 109.7 | 274.0 | 99.2 | 2.4  | 3.9 |
| Control             | 502.6        | 105.2 | 493.4 | 74.9  | 246.9 | 67.9 | 1.9  | 1.8 |

RSM = right superior medial.

Table 2. Reaction Time (RT) and Errors for Stop Signal Task

Floden and Stuss 1847
structures on the medial and lateral surface. In the study of Aron et al. (2003), these medial regions were incorporated into a much larger area involving substantial real estate on the lateral surface. It is therefore possible that the SMA and pre-SMA were in fact contributing to performance on their task, but that this was obscured by the inclusion of irrelevant areas within the same region of interest.

Although the SMA is intimately involved in general planning of movement, slowed inhibition of already initiated responses does not appear to be a result of a general motor slowing. Patients with SMA lesions demonstrated go RTs equivalent to control subjects. Moreover, correlational analyses support the dissociation of go RT and SSRT; although age correlated with go RT, neither showed any relationship to SSRT in the patients or controls, suggesting that these are independent processes. The lack of association between go RT and SSRT is one of the fundamental assumptions of the SSRT calculation. Aron et al. (2003) argued that go RTs probably had no impact on long stopping times in right inferior frontal patients because go RTs did not correlate with inferior frontal lesion volumes. However, a baseline measure of go RT speed would be necessary to validate this suggestion. Choice go RT can be slowed after damage to a number of frontal sites and under different conditions (Stuss, Binns, et al., 2002), and, without a baseline measure, it is difficult to determine whether slow go RTs in any particular subject reflected a strategic delay of go RTs or a genuine RT impairment.

At the same time, the absence of slowed go RTs in the current sample of patients was somewhat surprising. Our previous choice RT studies in patients with frontal lobe damage show differences in go RTs (Stuss, Binns, et al., 2002). However, others have demonstrated that go RT slowing in frontal patients is related to the complexity of stimulus–response mapping (Godefroy, Cabaret, Petit-Chenal, Pruvo, & Rousseaux, 1999). It may be the case that the simple two-choice task was too simple to reveal differences in go RTs.

Go RTs for both patients and controls were longer in the current study than that of Aron et al. (2003). This is most likely a function of differences in the choice RT tasks. Subjects in their study responded to arrows that pointed in the direction of the correct response (i.e., left arrow for left button press). The current study used consistent mappings of nondirectional stimuli (i.e., X for left button press). Decades of Stroop and Simon tasks (for a review, see MacLeod, 1991) demonstrate facilitation (faster RT) when stimulus and response are inherently related rather than arbitrary. What is more difficult to determine is whether the relationship between stimuli and responses also influenced estimates of SSRTs in the study by Aron et al. The inhibitory speeds reported in that study are much faster than those reported here or in prior work using the stop signal task (i.e., Dimitrov et al., 2003; Rieger et al., 2003; Williams, Ponesse, Schacher, Logan, & Tannock, 1999). The reason for this remains unclear.

As a final note, an interesting, although speculative, alternative explanation for the conflicting findings in our study and those of Aron et al. (2003) lies in the fact that long SSRTs can arise from problems other than simple inhibitory slowing. For example, one might fail to initiate a “stop” response for one of several reasons: impaired stimulus discrimination, difficulty in processing two stimuli occurring close in time, or poor switching between task sets or responses, all of which have been associated with right lateral frontal function (Aron, Monsell, Sahakian, & Robbins, 2004; Marcantoni, Lepage, Beaudoin, Bourgouin, & Richer, 2003; Stuss, Binns, et al., 2002). Although it is impossible to differentiate slowed inhibition from failure to initiate inhibition on any single trial where a response is made following a stop signal, computational studies that simulate these processes separately have shown that a fixed SOA procedure is less sensitive to failures to trigger inhibitory processing than titration procedures (Band et al., 2003). If right lateral lesions prevent the initiation of inhibition for some noninhibitory reason, it may be that our stop signal procedure, selected to minimize the influence of failures to initiate inhibitory processing, thereby reduced the relevance of right lateral frontal regions. In contrast, stop procedures using a titration procedure are more sensitive to these other noninhibitory processes and therefore the contribution of right lateral regions is more salient. However, this remains speculative on the basis of the current findings. To begin to assess this possibility, it would be necessary to directly compare stop procedures while manipulating other task parameters, such as stimulus discriminability or response cueing. This was beyond the scope of the present study but may be a fruitful avenue of research to pursue.

Summary

This study demonstrated that lesions of the right superior medial frontal lobe, particularly the pre-SMA and SMA, impair estimated inhibitory control speed on the stop signal task. Procedural comparisons with prior work suggest that this region is involved in contexts where strategic response modulation cannot be implemented to bolster inhibitory control. This study constitutes an important step toward understanding the anatomical basis of different components underlying behavioral control. It demonstrates how lesion studies with careful manipulation of task parameters can validate and clarify the nature of regional contributions to processing networks, offer vital insight into the necessary functional role of particular anatomical regions, and aid the interpretation of neuroimaging findings. Identifying the anatomical substrates underlying inhibitory control will also help to elucidate the neural basis of clinical syndromes that involve disinhibition.
References


