

Using Neuroimaging to Resolve the Psi Debate

Samuel T. Moulton and Stephen M. Kosslyn

Abstract

■ Parapsychology is the scientific investigation of apparently paranormal mental phenomena (such as telepathy, i.e., “mind reading”), also known as psi. Despite widespread public belief in such phenomena and over 75 years of experimentation, there is no compelling evidence that psi exists. In the present study, functional magnetic resonance imaging (fMRI) was used in an effort to document the existence of psi. If psi exists, it occurs in the brain, and hence, assessing the brain directly should be more sensitive than using indirect behavioral methods (as have been used previously). To increase sensitivity, this experiment was designed to produce positive results if telepathy, clairvoy-

ance (i.e., direct sensing of remote events), or precognition (i.e., knowing future events) exist. Moreover, the study included biologically or emotionally related participants (e.g., twins) and emotional stimuli in an effort to maximize experimental conditions that are purportedly conducive to psi. In spite of these characteristics of the study, psi stimuli and non-psi stimuli evoked indistinguishable neuronal responses—although differences in stimulus arousal values of the same stimuli had the expected effects on patterns of brain activation. These findings are the strongest evidence yet obtained against the existence of paranormal mental phenomena. ■

INTRODUCTION

“But it is a miserable thing for a question of truth to be confined to mere presumption and counter-presumption, with no decisive thunderbolt of fact to clear the baffling darkness.”—William James (1896, p. 884)

If psi exists, sciences from physics to psychology may require fundamental revision. If psi does not exist, roughly half of the general population (Moore, 2005) should be disabused of their fallacious beliefs. In theory, science has the capacity to settle this debate, yet in practice empirical investigations into psi have produced much heat and scarce light. With the advent of sophisticated neuroimaging techniques, however, psychologists are in a position to advance the debate over psi beyond presumption and counterpresumption. In this article, we describe a method that has the potential ultimately to resolve the psi debate, and present results from implementing this method.

Many people give credence to the existence of psi because they have had compelling personal experiences or heard descriptions of such experiences. For example, people sometimes claim to know who is about to call them on the phone, and then moments later have this hunch confirmed after answering a ringing phone. And people sometimes seem to “know” about a loved one’s injury or death, without any good reason to have this

knowledge. For instance, consider the following report from Rhine (1981):

One Thursday morning about 4 a.m., I jumped out of bed, feeling as if I was dying. I felt as if blood or something was pouring down from my head choking me and I was trying desperately to get my breath. My husband got up to help me. He tried to get me to the bathroom for some water to drink to stop the terrible choking spasms I seemed to be having. They soon diminished and I grew very weak. I thought I must be really dying. My husband put me down on the bed where I rested but felt so “all gone.” Then I thought my son had called, saying “Oh, Mama help me,” in such anguish.

Later in the day I went to the doctor for an X-ray of my chest. I thought with such acute pain that something must be wrong. But the doctor could find nothing. That was February 10th and on the 12th we received a telegram saying our son was killed by gunshot in the head at one o’clock on February 10th. There is a nine-hour difference in time. I feel he called me as it happened, and I heard his groan and felt his dying. (p. 138)

Although such experiences compel many—most commonly those who experience them—to accept the existence of psi, psychologists remain skeptical of such anecdotal evidence, and for good reason: Cognitive biases such as the clustering illusion (Gilovich, Vallone, & Tversky, 1985), availability error (Tversky & Kahneman, 1973), confirmation bias (Wason, 1960), illusion of

control (Langer, 1975), and the bias blind spot (Pronin, Lin, & Ross, 2002) may explain many apparently paranormal events that people report.

To move beyond the inherent uninterpretability of anecdotal evidence, researchers have employed a broad range of experimental methodologies to investigate psi. The most common experimental psi paradigm in recent years is the Ganzfeld (Honorton & Harper, 1974), named after a sensory deprivation technique of the same name. In these experiments, a sensorially deprived participant (the “receiver”) free associates while another participant (the “sender”), located in a properly isolated room, views a randomly selected stimulus and tries telepathically to send this information to the receiver. Following this period, which typically lasts 20 min, the receiver is presented with four stimuli and asked to identify which stimulus was viewed by the sender. Any significant deviation from chance guessing rate (25%) represents a psi effect.

Past experimental research, such as that using the Ganzfeld method, has not empirically resolved the psi debate for three fundamental reasons. First, psi phenomena are notoriously and inexplicably unreliable. No effects consistently replicate on an experiment-by-experiment basis, and meta-analytic effects suffer from similar instability (e.g., Milton & Wiseman, 1999; Bem & Honorton, 1994). Second, the positive evidence that has been reported is merely “anomalous.” As many have noted previously (e.g., Mumford & Rose, 1995), the absence of a normal explanation does not justify the presence of a paranormal explanation. Third, experiments with negative (i.e., null) results have always been vulnerable to the reasonable criticism that they assessed the wrong behavior. With no clear consensus on how psi might manifest itself behaviorally and a near-infinite set of possible behaviors to test, behavioral psi research is unusually susceptible to the argument that “absence of proof is not proof of absence.”

In the present experiment, we operationalize the psi hypothesis by asking the following question: Does the brain respond selectively to psi stimuli? By “psi stimuli” we mean stimuli that not only are presented through the usual senses (e.g., visually), but also are presented telepathically (mind to mind), clairvoyantly (world to mind), and precognitively (future to present); by “non-psi stimuli,” we mean identical stimuli that are only presented through normal sensory channels. Under the null hypothesis, these psi and non-psi stimuli are one and the same (because the additional aspects of psi stimuli do not in fact exist) and thus should evoke indistinguishable neuronal responses. Under the psi hypotheses, the stimuli are categorically different and should evoke different neuronal responses.

Research in cognitive neuroscience suggests two general effects that psi might have on the brain. On the one hand, psi might provide participants with specific, implicit knowledge of stimuli. In this case, we would

expect a suppressed brain response to psi stimuli compared to non-psi stimuli. This expectation is based on extensive research that shows that the brain’s response to a stimulus is suppressed by prior exposure to that stimulus—even when the prior exposure is subliminal (Naccache & Dehaene, 2001), conceptual (Wheatley, Weisberg, Beauchamp, & Martin, 2005), milliseconds earlier (Kourtzi & Kanwisher, 2000), days earlier (van Turennout, Ellmore, & Martin, 2000), a different size, position (Grill-Spector et al., 1999), orientation (Vuilleumier, Schwartz, Duhoux, Dolan, & Driver, 2005), or format (Badgaiyan, Schacter, & Alpert, 1999). On the other hand, psi might increase participants’ attention to stimuli without providing them with stimulus-specific content. In this case, we would expect an enhanced brain response to psi stimuli, given the evidence that attention enhances brain activity (Corbetta, Shulman, Miezin, & Petersen, 1995). In the present experiment, we tested for suppression effects, enhancement effects, or some combination thereof (with different effects in different brain areas). By hypothesizing merely a difference in activation—without specifying a direction or neuroanatomical locus—we make minimal assumptions about psi and offer the broadest possible test of the psi hypothesis.

The present approach holds promise of circumventing the problems that typically plague psi research. First, by turning to the possible effects of psi on the brain, one cannot obtain positive evidence for an anomaly without simultaneously obtaining evidence about its nature. Any differences in activation will have a neuroanatomical location, time course, and relationship with stimulus, participant, or experimental variables that can, when combined with knowledge from the field of cognitive neuroscience, elucidate the underlying mechanism. Second, given the instability of behavioral findings, one way to increase sensitivity is to go directly to the source—and assess the effects of psi on the brain itself. As the source of behavior, the brain may offer a more stable gauge of psi-mediated knowledge. Finally, the current approach promises to decrease the ambiguity of negative evidence. Although one can never affirm the null hypothesis, not all null results are epistemologically equal. Because this paradigm uniquely minimizes assumptions about the source of knowledge, the kind of processing, or the nature of mental content responsible for psi, any ensuing null results will be qualitatively more informative than those from behavioral methods. Moreover, we can compare any null results with positive results that reflect other aspects of the same stimuli; thus, conceptually, such null results can be considered part of an interaction, where one variable has effects but another does not.

In our experiment, participants played one of two roles: “sender” or “receiver.” On each trial, sender participants viewed a randomly selected target stimulus from outside the scanner (see Figure 1), and tried to

send this information to the receiver participant by mental means alone. While the senders were doing this, receiver participants completed a simple binary guessing task, and functional magnetic resonance imaging (fMRI) was used to monitor their brain activity. On each trial of the guessing task, the receivers sequentially viewed two stimuli, guessed which one was the stimulus being “sent” (i.e., the psi stimulus), and then saw the psi stimulus a second time. This paradigm allowed us simultaneously to test all three hypothesized mechanisms of psi: telepathy (i.e., “mind reading”), clairvoyance (i.e., direct sensing of remote events), and precognition (i.e., knowing future events). The sender served as the potential telepathic source, the sender’s computer monitor served as the potential clairvoyance source, and the second presentation of the psi stimulus served as the potential precognition source.

We adapted our design from a series of low-density event-related potential (ERP) experiments that first investigated the possibility of differential psychophysiological responses to psi versus non-psi stimuli (McDonough, Don, & Warren, 2002; Don, McDonough, & Warren, 1998; Warren, McDonough, & Don, 1992a, 1992b). Our proposal relies on logic similar to that used in the earlier experiments, while at the same time taking advantage of a dependent measure (blood oxygenation level dependent [BOLD] hemodynamic response) that is substantially more informative than low-density ERPs; fMRI provides far more information about the location of activation than does low-density ERP, and in so doing affords us the opportunity to detect more possible distinct “sources” for an effect. In an additional effort to document a psi effect, we recruited biologically or emotionally related participants (e.g., identical twins) because of anecdotal and experimental evidence suggesting that they are more prone to psi (Sheldrake & Smart, 2003; Taylor, 2003; Playfair, 1999). We also used emotional stimuli, which have similarly been implicated in psi (Sherwood, Dalton, Steinkamp, & Watt, 2000; Cornell, 1999; Moss & Gengerelli, 1967, 1970; Myers, 1903).

The behavioral task we used served two purposes: It motivated participants to detect the psi stimuli, and it provided the structure necessary for the BOLD contrasts. We did not include the behavioral task because we expected to find an explicit (guessing rate) or implicit (response time) effect. Unfortunately, there are no behavioral tasks that can be used as “standard metrics” of psi; it is precisely this failure of behavioral research that has motivated our neuroimaging work. This observation highlights a critical asymmetry in cognitive neuroscience: All behavior requires neural activity, but not all neural activity yields behavior. Thus, in the current experiment, although the unexpected presence of a behavioral effect mandates a neural correlate, the expected absence of a behavioral effect does not necessarily imply null neuroimaging results.

METHODS

Participants

We recruited 19 pairs of individuals (6 couples, 5 emotionally close roommate/friend pairs, 3 identical twin pairs, 1 mother–son pair, and 1 pair of sisters) from the local community to participate in this experiment. Data from 3 pairs (1 twin pair, 1 friend pair, and 1 couple) were eliminated because of scanner spiking, equipment malfunction, or excessive head motion. Of the 32 remaining participants, 14 were men (mean age 22.8 years, range 19–47 years) and 18 were women (mean age 23.4 years, range 19–58 years). Each session was approximately 2.5 hr, with 1.5 hr devoted to the experimental task and the rest reserved for paperwork, instructions, training, and debriefing. In addition to travel expenses, we compensated receivers and senders with \$77–134 (depending upon task performance, see subsequent description) and \$50, respectively. The receiver in each pair was selected based on the participants’ preference, right-handedness, as well as magnetic resonance (MR) safety guidelines.

Materials

Stimuli

Using the International Affective Picture System (IAPS) stimulus set (Lang, Bradley, & Cuthbert, 1995), we selected 240 pairs of photographs as test stimuli based on content uniqueness and normative emotionality ratings. Negative high-arousal pictures (e.g., a snake) were paired with either neutral or positive low-arousal pictures (e.g., a tissue box, giraffe); positive high-arousal pictures (e.g., an erotic couple) were paired with either neutral or negative low-arousal pictures (e.g., a neutral face, cemetery). By maximally separating pictures within each stimulus pair on dimensions of valence and arousal, we sought to increase the likelihood of a psi effect. Figure 2 presents four examples of the sorts of stimulus pairs used (although the actual pairs were drawn from the IAPS).

Stimulus Lists

We first randomly assigned each picture pair to a *stimulus category* (psi, non-psi). Next, we randomly assigned half of the psi stimuli to each *position* (first, second), to ensure that psi assignment was not confounded with stimulus presentation. To ensure the adequacy of this randomization, we tested for differences in arousal and valence across Stimulus Category × Position conditions (paired *t* tests, all $p > .1$, two-tailed). We then created four stimulus lists to counterbalance completely each pair on these variables. Therefore, across lists, any given picture was assigned once to each of the following conditions: psi/first, psi/second, non-psi/first, non-psi/second. We also randomly

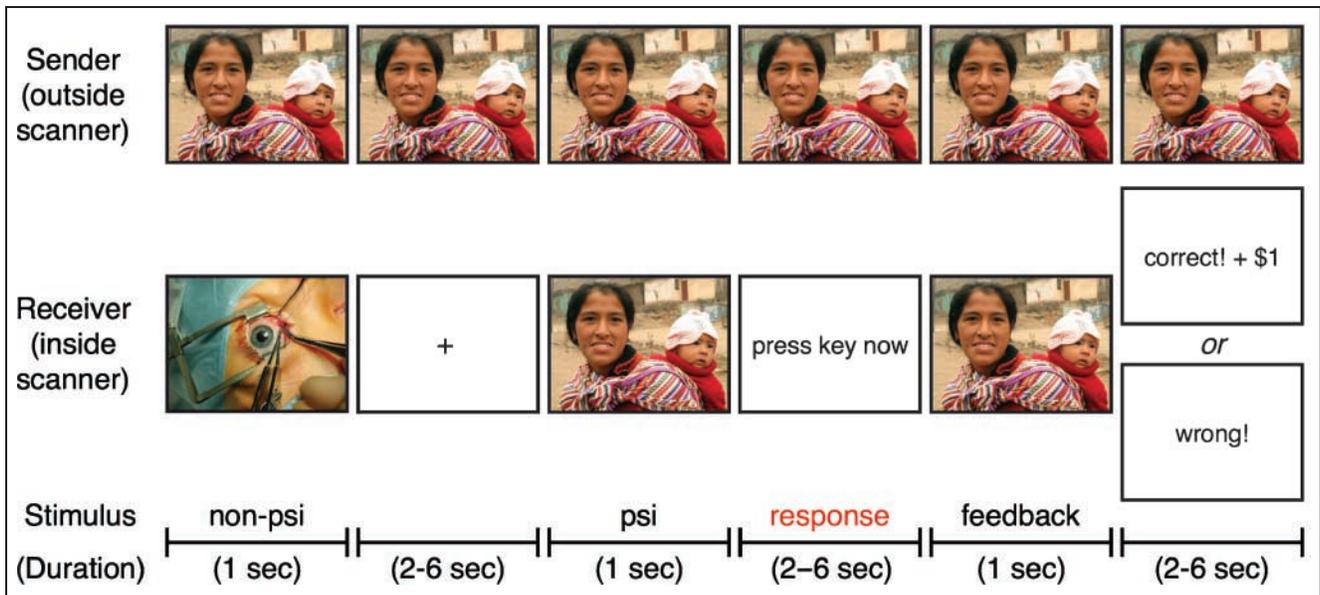


Figure 1. A schematic of one trial. In this trial for the receiver, the non-psi stimulus appears first and the psi stimulus second. The third stimulus presentation (feedback) in each trial is always the same as the psi stimulus. The sender sees only the psi stimulus for each trial.

assigned pairs to trial positions within each list and created an additional counterbalancing variable of *trial order* (forwards, backwards). We crossed trial order with stimulus category and position to create eight total between-participants stimulus lists.



Apparatus

All stimuli presented during MR scanning were generated using a Macintosh G4 computer and PsyScope software (Macwhinney, Cohen, & Provost, 1997). Visual stimuli were backprojected via an LCD projector and a mirror attached to the head coil. We collected behavioral responses by using an MR-compatible button box connected to the Macintosh via a custom USB interface.

Tasks and Procedure

Receiver

After obtaining informed consent, while still outside the scanner, participants completed a practice version of the experimental task for approximately 10 min to

Figure 2. Examples of the sorts of stimulus pairs used in the study. From top to bottom: negative high-arousal paired with positive low-arousal; positive high-arousal paired with negative low-arousal; negative high-arousal paired with neutral low-arousal; and positive high-arousal paired with neutral low-arousal. The actual stimuli were selected from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 1995) set. Because copyright restrictions prevent us from reprinting IAPS pictures, the images in this figure are meant merely to illustrate the types of stimuli used in this experiment.

familiarize themselves with the stimuli, procedure, and minimum response window (2 sec). None of the experimental stimuli were used in the practice task. Once positioned in the scanner, we first obtained anatomical scans from the receiver participants; this required approximately 25 min, during which time they passively viewed a screensaver display designed to entertain and relax them. Following this, we administered five series of functional scans while the participants completed five sets of the experimental task. Each set consisted of 48 trials, and each trial contained three sequential exposures to two unique IAPS stimuli. The first and second exposures, lasting 1 sec each, presented participants with the assigned psi and non-psi stimuli from each stimulus pair (with the presentation order varying across trials), and were separated by a randomly jittered period of fixation (2, 4, or 6 sec). From the participants' perspective, each of the first two stimuli was a potential psi stimulus. After the second exposure, participants saw "press key now" on the screen and guessed which item was the psi stimulus by pressing one of two buttons (Button "1" for the first item, Button "2" for the second item). Participants had between 2 and 6 sec to respond depending upon the trial, and their responses did not advance the script. Following this forced-choice response, participants were presented with the psi stimulus (1-sec duration) and feedback ("correct! + \$1" or "wrong!," 2–6 sec). Receivers were told that they would receive an additional dollar for each correct response.

Sender

After providing informed consent, sender participants observed their partner practicing the receiver's task and were concurrently given instructions about the details of their task. We then took the senders to another room, where they could not see or hear the scanner, and they waited for the investigator's cue to begin their task. At the start of each functional series, the investigator signaled the sender to begin, at which point the sender started sequentially viewing the psi images. The sender and receiver scripts were temporally synchronized such that during each entire trial of the receiver's task, the sender viewed that trial's psi stimulus. Senders viewed each stimulus for 10–22 sec, depending upon timing of the receivers' stimulus presentations. The investigator asked senders to "influence the receiver" with the psi stimulus, adopt a "playful" attitude (for a discussion of attitudes that are purportedly conducive to psi, see Targ, Braud, Stanford, Schlitz, & Honorton, 1991), maintain an active interest in the stimuli, and use whatever "sending" tactics they deemed appropriate. The investigator did not provide senders with real-time, trial-by-trial feedback; both receivers and senders were informed of the overall hit rate at the conclusion of the experiment.

fMRI Acquisition

MR scanning was performed using a 3T Siemens Allegra head-only scanner with high-speed imaging gradients and a quadrature head coil. Structural images included a high-resolution MP-RAGE scan (128 sagittal slices, 1.33-mm thickness, $T_R = 2530$ msec, $T_E = 3.25$ msec). Whole-brain functional imaging was performed using a gradient-echo echo-planar imaging (EPI) pulse sequence (33 interleaved slices oriented along the AC–PC line, 3 mm thick with 1-mm interslice gap, $T_R = 2$ sec, $T_E = 30$ msec, flip angle = 90°). Each functional series lasted 13 min 8 sec and consisted of 394 sequential whole-brain acquisitions. Four additional volumes were acquired at the start of each functional series to account for T1 saturation effects.

fMRI Analysis

We analyzed and preprocessed the functional data using the software package FSL (Smith et al., 2004). Brain and non-brain image data were segmented using FSL's Brain Extraction Tool (Smith, 2002a). Each extraction was manually inspected and conducted again with new parameters (center of gravity, fractional intensity threshold) if necessary to ensure maximal segmentation. We applied slice timing correction to account for the sampling offsets inherent in slicewise EPI acquisition sequences, and rigid-body transformation of 3-D functional volumes to correct for spatial misalignment due to head motion. Mean changes in MR signal between functional series were eliminated with 4-D global intensity normalization (Smith, 2002b). Data were convolved in 3-D space with a 5-mm (full width half maximum) Gaussian kernel to increase signal-to-noise ratio and allow later GRF-based statistical thresholding. To inform decisions about temporal filtering, we conducted a Fourier analysis of our unfiltered design matrix. This revealed significant spectral power in high-frequency domains and little power in frequencies lower than 25 Hz. Based on this spectral analysis, we decided against low-pass filtering and set a high-pass filter cutoff of 25 sec to eliminate low-frequency noise (e.g., scanner drift). Using the tool FLIRT (Jenkinson, Bannister, Brady, & Smith, 2002; Jenkinson & Smith, 2001), functional data were registered to the brain-extracted structural images using a 6 degrees of freedom (DOF) rigid-body transformation and to standard Montreal Neurological Institute (MNI; Evans et al., 1993) space using a 12-DOF affine transformation. We visually inspected the accuracy of each registration and manually checked each brain volume for evidence of scanner spiking or other image anomalies.

Functional series were analyzed using FMRIB's Improved Linear Model (FILM; Woolrich, Ripley, Brady, & Smith, 2001), which removes nonparametrically estimated temporal autocorrelation in each voxel's time series

before applying the general linear model (GLM). For each series we modeled neural responses to the psi stimulus, non-psi stimulus, and feedback stimulus, as well as their temporal derivatives by convolving the basic waveforms (based on onset time and duration) of these variables with a double-gamma canonical hemodynamic response function (Glover, 1999). The same temporal filtering that was applied to the data was also applied to the model. For every functional volume, the following linear contrasts were employed to create statistical parametric maps (SPMs): non-psi > psi, and psi > non-psi.

We analyzed the data both from the group as a whole and also for each individual participant. To increase power, we analyzed the fMRI data using a fixed-effects model (i.e., only within-series variance was modeled) instead of the typical mixed-effects model, which was justified because the hypotheses were focused exclusively on the tested group of participants rather than the general population. Moreover, we wanted to err on the side of finding an effect, whenever justifiable. Using parametric statistics based on Gaussian random-field theory (Forman et al., 1995; Friston, Worsley, Frackowiak, Mazziotta, & Evans, 1993; Worsley, Evans, Marrett, & Neelin, 1992), the group-level SPMs for our two psi contrasts were thresholded on the voxel level at $z = 3.1$ (uncorrected $p = .001$) and on the cluster level at corrected $p = .025$ (the alpha level was halved to correct for the number of contrasts). To ensure that the results from the primary analysis did not critically depend on an arbitrary chosen thresholding parameter, we tested whether the number, location, or significance of clusters qualitatively changed as a function of the voxel height threshold (tested range: $2.7 < z < 3.4$).

For each participant, we tested the same two contrasts described above but with two appropriate revisions. First, because there was much less data, we changed the primary voxel z threshold to 2.3 and the exploratory range to $1.9 < z < 2.7$. Second, because we now had 32 independent tests (16 participants \times 2 contrasts per participant), we lowered the alpha value to .0016.

Three functional series, all from separate participants, were removed from the analysis. We eliminated two series because of excessive head motion (absolute displacement > 1 mm), and one because of scanner malfunction.

RESULTS

Behavioral Results

Considering first the behavioral data, the participants performed almost exactly at chance on the guessing task. Out of 3687 recorded responses, they correctly guessed the psi stimulus 1842 times (50.0%). None of the results from any individual participant deviated from what would be expected on the basis of chance variation alone. On average, participants responded 19 msec fast-

er when they correctly selected the psi stimulus (mean response time [RT] = 939 msec) compared to when they incorrectly selected the non-psi stimulus (mean RT = 958 msec), an insignificant difference (paired t test, $p = .109$, two-tailed).

fMRI Results

The key results are the comparisons of brain activation for psi stimuli versus for non-psi stimuli. We sought anatomical loci across the entire brain that responded differently to these stimuli for the entire group and for each individual participant. Analyses of group data revealed no evidence whatsoever of psi. Psi and non-psi stimuli evoked widespread but indistinguishable neural responses (see Figure 3). However, data from a single participant (Participant 14, a member of a romantic couple) revealed several brain regions that appeared to respond less strongly to psi stimuli than to non-psi stimuli. All other analyses of individual participants revealed no significant differences in activation.

For Participant 14, the non-psi > psi contrast revealed multiple clusters of significant activation, as seen in Figure 4. The first functional series from this participant was excluded from analysis because of excessive head motion, leaving only Series 2–5 for analysis. The clusters were centered bilaterally in the superior temporal gyrus (Brodmann's area [BA] 21, 22, 41), with the largest, most significant, and most robust activity in the left hemisphere. Using our a priori voxel correction, activation in both hemispheres reached significance (after correcting for multiple analyses). In the left hemisphere, the cluster contained 592 voxels (corrected $p = 2.48 \times 10^{-5}$), and remained significantly large across all explored voxel thresholds ($z = 1.9$ – 2.7 , $p = .03$ – $.003$). In the right hemisphere, the cluster contained 342 voxels (corrected $p = .0107$), and remained significantly large across most of the explored voxel thresholds ($z = 1.9$ – 2.3 , $p = .03$ – $.01$).

However, the apparently positive result from this participant could simply reflect uninteresting artifacts. One possibility is that stimulus-correlated head motion spuriously produced the activation. Although we accounted for head motion artifacts by (a) eliminating series with excessive head motion, (b) spatially aligning all 3-D functional images in the preprocessing stage, and (c) scrutinizing the anatomy of any significant activation, stimuli-correlated motion might still have produced spurious activation. As an additional control, we reduced the six motion parameters (three translations, three rotational) into two factors through principal components analysis, and added these two factors as covariates in the GLM analysis of Participant 14's data. The addition of these covariates had a negligible effect on both the location and size of significant clusters.

Because motion induces not just spatial misalignment but complex, unmeasured changes in the MR signal

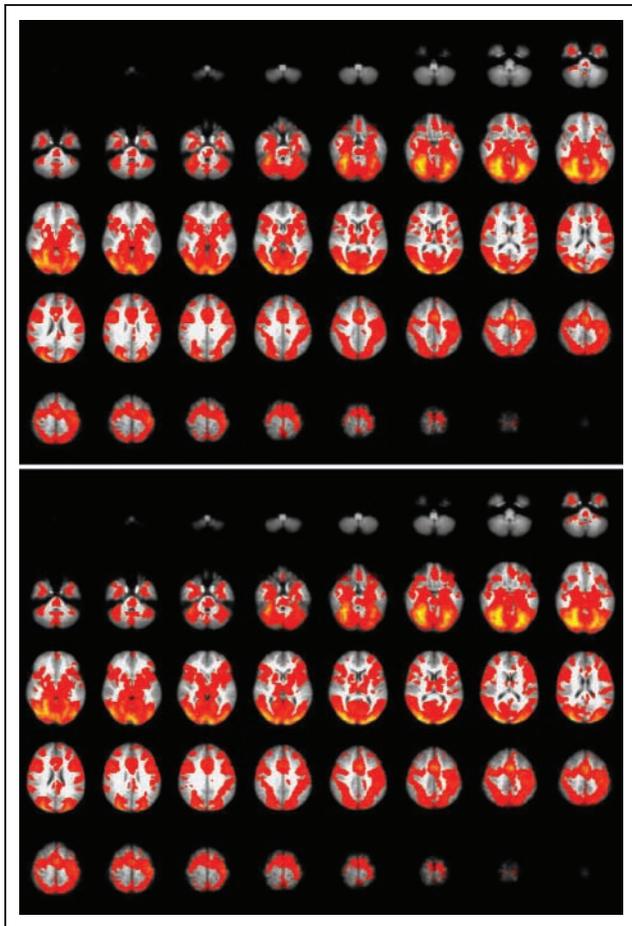
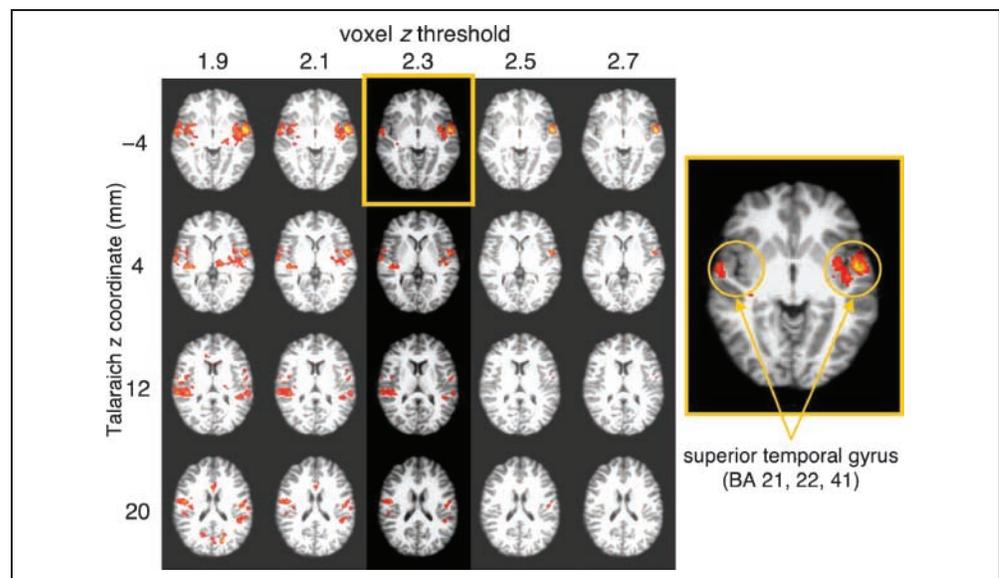


Figure 3. Activation maps for psi > baseline (top) and non-psi > baseline (bottom) contrasts on a group level using a fixed-effects analysis, voxel thresholding at $z < 3.1$, and cluster thresholding at $p < .025$. In all figures, we adopt the radiological convention of displaying the left and right hemispheres towards the right and left of the page, respectively.

Figure 4. Significant activation for non-psi > psi contrast in Participant 14 as a function of voxel z threshold. Clusters were centered in the superior temporal gyrus, with the most significant activity occurring in the left hemisphere. Functional activation is overlaid on the participant's T1 structural image.



(e.g., interpolation and magnetic susceptibility effects), all the motion “correction” methods listed above do not, in fact, completely eliminate the effect of head motion. We can, however, correlate motion parameters with the design matrix to evaluate directly the possibility that stimulus-correlated motion produced spurious activation. For each of Participant 14’s four functional series, we correlated measures of absolute and relative head motion with the convolved psi and non-psi variables (as well as the difference between them). We found no significant correlations, and thus, no evidence that stimulus-correlated motion explains our results.

A second possibility is that the activation in question represents processes that underlie a later decision (a “gleam in the eye” effect; McDonough, Warren, & Don, 1992) rather than implicit knowledge of psi stimuli. For example, because Participant 14 did not select the psi stimuli exactly 50% of the time on an individual series level or overall (success rate = 40.0%, 52.1%, 56.3%, and 59.6% for individual series and 52.0% overall), there exists a confound between the psi conditions and whether or not stimuli were later selected. This confound could produce spurious psi results by merely revealing brain regions that predict future decision making. To test for this possibility, we divided the first two stimulus presentations into selected and unselected stimuli instead of psi and non-psi stimuli, and tested the following two contrasts: selected > unselected, unselected > selected. These contrasts revealed no significant clusters of activation. Thus, the possibility of a gleam in the eye confound appears unlikely.

A third counterexplanation is that the observed activation difference reflects the participant’s idiosyncratic reactions to perceptual, conceptual, or affective differences between the psi and non-psi stimuli. Although no such content differences existed on a group level be-

cause of counterbalancing, they did exist for individual participants. This account of the data is consistent with the neural correlates of emotion, as found in other fMRI research that used the IAPS stimulus database: All but one of the regions that were activated in the one participant with apparently positive results have previously been shown to be sensitive to emotional valence (Kensinger & Schacter, 2006). To test this content-differences hypothesis directly, we conducted 100 simulations on Participant 14's data (excluding the first series). In each simulation, we randomly permuted condition assignment (psi, non-psi) within each stimulus pair and reanalyzed the data (using the same voxel thresholding: $z < 2.3$). We applied the same constraints to the permutations that were applied to the original stimulus lists: Psi and non-psi stimuli were balanced for onset times and stimulus position, and were not significantly different in normative ratings of valence or arousal. In no simulations were the condition reassignments significantly correlated ($r < .1$) with the real condition assignments. Out of 100 simulations, psi versus non-psi contrasts revealed 10 with at least one cluster that was equal to or larger than the largest cluster (592 voxels) found in the real data. Moreover, this simulation activation occurred in all of the same areas as observed in the actual data. The probability that the results from the real series would pass significance ($p < .0016$) after an exhaustive permutation analysis and that the present simulation results (with 100 permutations) are due to chance is $(100 \times 0.0016)^{10} = 1.1 \times 10^{-8}$. Thus, the confounding between stimulus characteristics and condition for Participant 14 can in fact account for the observed patterns of activation.

Null findings are always difficult to defend, if only because they could simply indicate that the technique lacks adequate sensitivity. To examine the statistical power of our approach, we conducted two "control" analyses of psychological (rather than parapsychological) mental processes on the same data set using identical analysis parameters. First, we contrasted activation elicited by the psi stimuli with that elicited by the feedback stimuli. Although these two sets of stimuli were visually identical, there is good reason to suspect that they were processed in subtle but systematically different ways. On the one hand, the psi stimuli were novel and likely provoked response preparation; on the other hand, the feedback stimuli were recognizable and probably provoked reward anticipation. A combination of repetition suppression, fMRI adaptation, and attentional enhancement effects might also underlie any differential activation. This analysis revealed numerous regions of the brain that discriminated between psi and feedback stimuli (see Figure 5). We will not offer ad hoc interpretations for these results, but simply note that their number and strength indicate that the technique was in fact sensitive to subtle changes in information processing.

Nevertheless, this first control analysis is not ideal because it confounds trial sequence with stimulus condition (psi stimuli always preceded feedback stimuli) and also reflects the effects of several different psychological processes (making the results difficult to interpret). We therefore conducted a second control analysis that does not suffer from these limitations. Here we investigated the neural correlates of emotional arousal by dividing the first two stimuli of each trial into categories of high and low arousal; in the previous psi analysis, we segregated these same stimuli into psi and non-psi categories. As in the psi analysis, trial sequence was counterbalanced across conditions: On half the trials, low-arousal pictures preceded high-arousal pictures and, on the other half, high-arousal pictures preceded low-arousal pictures. We contrasted activation evoked by high-arousal stimuli with that evoked by low-arousal stimuli and discovered three significant clusters of activation in the following brain regions (see Figure 6): right occipital-temporal regions (BA 18, 19, 37, 39; 2081 voxels, corrected $p = 9.35 \times 10^{-21}$), left occipital-temporal regions (BA 18, 19, 37, 39; 1589 voxels, corrected $p = 3.16 \times 10^{-17}$), and the right precuneus (BA 7, 144 voxels, corrected $p = .0063$). These regions have been implicated in emotional processing in prior studies (e.g., Bermpohl et al., 2006; Kensinger & Schacter, 2006; Paradiso et al., 1999). For example, similar arousal-related activation in occipital-temporal regions and the precuneus were reported by Kensinger and Schacter (2006), who also had participants passively view IAPS stimuli—despite the fact that we relied solely on normative ratings (instead of participants' ratings) and presented stimuli for 1 sec (instead of 2.5 sec).¹

DISCUSSION

The results support the null hypothesis that psi does not exist. The brains of our participants—as a group and individually—reacted to psi and non-psi stimuli in a statistically indistinguishable manner. Given the relatively large number of participants, the use of fixed-effects statistics, the extensive activation elicited separately by both types of stimuli, the subtle psychological effects revealed in the much smaller data set from a single participant, and the non-psi effects we documented on a group level using identical statistical criteria, a lack of statistical power does not reasonably explain our results. Even if the psi effect were very transient, as are many mental events, it should have left a footprint that could be detected by fMRI—as did the other subtle effects we detected. In particular, the large and massively significant activation revealed by our arousal contrast shows that the psi effect, if it exists, must be substantially smaller than the effect of arousal on brain activity.

But what of the truism that one cannot affirm the null hypothesis? We note that some null results should be

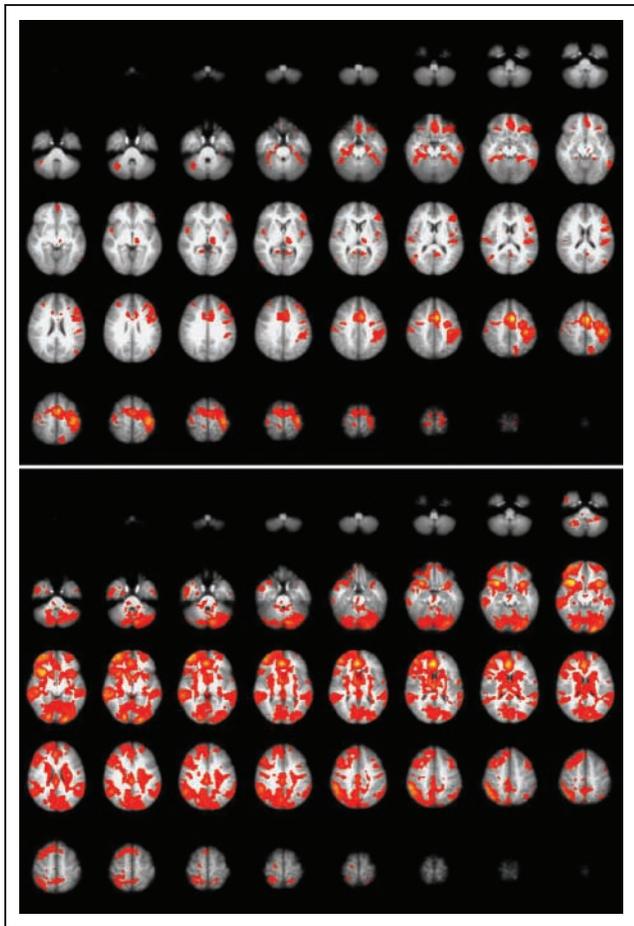
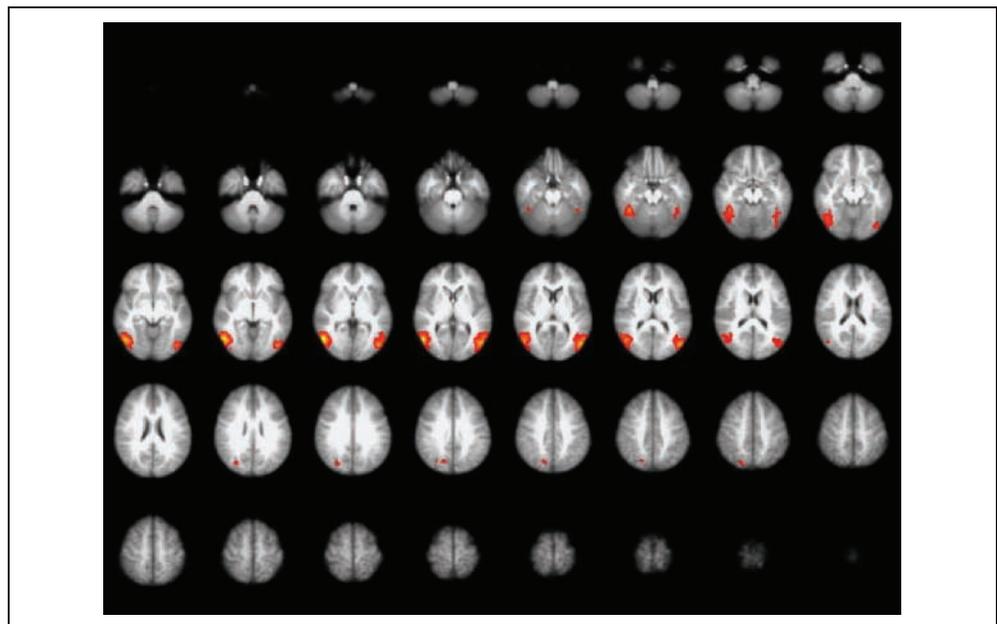


Figure 5. Activation maps for psi > feedback (top) and feedback > psi (bottom) contrasts on a group level using a fixed-effects analysis, voxel thresholding at $z < 3.1$, and cluster thresholding at $p < .025$.

Figure 6. Activation map for high-arousal > low-arousal contrast on a group level using a fixed-effects analysis, voxel thresholding at $z < 3.1$, and cluster thresholding at $p < .025$.



taken more seriously than others. For example, take the famous Michelson and Morley (1887) experiment, which showed no difference in the speed of light moving in different directions; these results had a profound influence on theories in physics. Or consider the possibility of water on Mars. If a set of close-up images of its surface failed to capture frozen lakes, few would accept the nonexistence of Martian water. Yet if a planetwide analysis of its subsurface soil content failed to show telltale signs of water, most would accept the null hypothesis of a Martian desert. Past null results from parapsychology are comparable to scattered snapshots of the surface in that they measure a small sample of outwardly observable variables. The current neuroimaging approach, however, seeks anomalous knowledge at its source, inside the brain, using methods validated by cognitive neuroscience. It is also exhaustive: With the exception of spinal reflexes (Creed, Denny-Brown, Eccles, Liddell, & Sherrington, 1932), all behaviors stem from brain events. Finally, the study incorporated methodological variables (e.g., biological and emotional relatedness of participants, evocative stimuli) widely considered to facilitate psi by parapsychologists. As such, the current null results do not simply fail to support the psi hypothesis: They offer strong evidence against it. If these results are replicated over a range of participants and situational contexts, the case will become increasingly strong, with as much certainty as is allowed in science, that psi does not exist.

Finally, one goal of this study was to develop, implement, and test a new way to address the psi hypothesis, one that incorporates the intellectual and technological gains of contemporary cognitive neuroscience. By using neuroimaging to compare the brain's response to psi and non-psi stimuli, we argue that one can effectively

circumvent three typical limitations of psi research: the ambiguity of positive results, the ambiguity of negative results, and the lack of experimental replicability. Thus, this method has much to offer should researchers wish to investigate further the possible circumstances in which psi might exist.

Acknowledgments

We acknowledge the generous support of the Bial Foundation (Grant 118/02), and thank William Thompson, Jennifer Shephard, Giorgio Ganis, Rebecca Wright, Patricia Zadnik, Elissa Aminoff, Elizabeth Chua, and Penny Hoorenbeek for their assistance and commentary.

Reprint requests should be sent to Samuel T. Moulton, 844 William James Hall, 33 Kirkland Street, Cambridge, MA 02138, or via e-mail: moulton@wjh.harvard.edu.

Note

1. Although this analysis did not reveal arousal-related activation in the amygdala, we did find such activation with less stringent thresholding parameters. It is also worth noting that amygdalar activity is particularly difficult to detect because of nearby sinuses that interfere with the BOLD signal (see Johnstone et al., 2005).

REFERENCES

- Badgaiyan, R. D., Schacter, D. L., & Alpert, N. M. (1999). Auditory priming within and across modalities: Evidence from positron emission tomography. *Journal of Cognitive Neuroscience, 11*, 337–348.
- Bem, D. J., & Honorton, C. (1994). Does psi exist? Replicable evidence for an anomalous process of information transfer. *Psychological Bulletin, 115*, 4–18.
- Bermpohl, F., Pascual-Leone, A., Amedi, A., Merabet, L. B., Fregni, F., Gaab, N., et al. (2006). Dissociable networks for the expectancy and perception of emotional stimuli in the human brain. *Neuroimage, 30*, 588–600.
- Corbetta, M., Shulman, G. L., Miezin, F. M., & Petersen, S. E. (1995). Superior parietal cortex activation during spatial attention shifts and visual feature conjunction. *Science, 270*, 802–805.
- Cornell, T. (1999). *Investigating the paranormal: Spontaneous cases and other relevant phenomena*. London: Athlone.
- Creed, R. S., Denny-Brown, D., Eccles, J. C., Liddell, E. G. T., & Sherrington, C. S. (1932). *Reflex activity of the spinal cord*. Oxford, UK: Clarendon Press.
- Don, N. S., McDonough, B. E., & Warren, C. A. (1998). Event-related potential (ERP) indicators of unconscious psi: A replication using subjects unselected for psi. *Journal of Parapsychology, 62*, 127–145.
- Evans, A. C., Collins, D. L., Mills, S. R., Brown, E. D., Kelly, R. L., & Peters, T. M. (1993). *3D statistical neuroanatomical models from 305 MRI volumes*. Paper presented at the IEEE Nuclear Science Symposium and Medical Imaging Conference.
- Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., & Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. *Magnetic Resonance in Medicine, 33*, 636–647.
- Friston, K. J., Worsley, K. J., Frackowiak, R. S. J., Mazziotta, J. C., & Evans, A. C. (1993). Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping, 1*, 210–220.
- Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology, 17*, 295–314.
- Glover, G. H. (1999). Deconvolution of impulse response in event-related BOLD fMRI. *Neuroimage, 9*, 416–429.
- Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzhak, Y., & Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron, 24*, 187–203.
- Honorton, C., & Harper, S. (1974). Psi-mediated imagery and ideation in an experimental procedure for regulating perceptual input. *Journal of the American Society for Psychical Research, 68*, 156–168.
- James, W. (1896). Address of the president before the Society for Psychical Research. *Science, New Series, 3*, 881–888.
- Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage, 17*, 825–841.
- Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis, 5*, 143–156.
- Johnstone, T., Somerville, L. H., Alexander, A. L., Oakes, T. R., Davidson, R. J., Kalin, N. H., et al. (2005). Stability of amygdala BOLD response to fearful faces over multiple scan sessions. *Neuroimage, 25*, 1112–1123.
- Kensinger, E. A., & Schacter, D. L. (2006). Processing emotional pictures and words: Effects of valence and arousal. *Cognitive, Affective, and Behavioral Neuroscience, 6*, 110–126.
- Kourtzi, Z., & Kanwisher, N. (2000). Activation in human MT/MST by static images with implied motion. *Journal of Cognitive Neuroscience, 12*, 48–55.
- Lang, P., Bradley, M., & Cuthbert, B. (1995). *International Affective Picture System (IAPS): Technical manual and affective ratings*. Gainesville, FL: The Center for Research in Psychophysiology, University of Florida.
- Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology, 32*, 311–328.
- Macwhinney, B., Cohen, J., & Provost, J. (1997). The PsyScope experiment-building system. *Spatial Vision, 11*, 99–101.
- McDonough, B. E., Don, N. S., & Warren, C. A. (2002). Differential event-related potential to targets and decoys in a guessing task. *Journal of Scientific Exploration, 16*, 187–206.
- McDonough, B. E., Warren, C. A., & Don, N. S. (1992). Event-related potentials in a guessing task: The gleam in the eye effect. *International Journal of Neuroscience, 65*, 209–219.
- Michelson, A., & Morley, E. (1887). On the relative motion of the Earth and the luminiferous ether. *The American Journal of Science, 203*, 333–345.
- Milton, J., & Wiseman, R. (1999). Does psi exist? Lack of replication of an anomalous process of information transfer. *Psychological Bulletin, 125*, 387–391.
- Moore, D. W. (2005). *Three in four Americans believe in paranormal*. Retrieved September 26, 2006, from www.galluppoll.com/content/?ci=16915.
- Moss, T., & Gengerelli, J. A. (1967). Telepathy and emotional stimuli: A controlled experiment. *Journal of Abnormal Psychology, 72*, 341–348.
- Moss, T., & Gengerelli, J. A. (1970). ESP effects generated by affective states. *Behavioral Neuropsychiatry, 2*, 8–12.

- Mumford, M. D., & Rose, A. M. (1995). *An evaluation of remote viewing: Research and applications*. American Institutes for Research.
- Myers, F. W. H. (1903). *Human personality and its survival of bodily death*. Oxford, UK: Longmans, Green.
- Naccache, L., & Dehaene, S. (2001). The priming method: Imaging unconscious repetition priming reveals an abstract representation of number in the parietal lobes. *Cerebral Cortex*, *11*, 966–974.
- Paradiso, S., Johnson, D. L., Andreasen, N. C., O’Leary, D. S., Watkins, G. L., Ponto, L. L., et al. (1999). Cerebral blood flow changes associated with attribution of emotional valence to pleasant, unpleasant, and neutral visual stimuli in a PET study of normal subjects. *The American Journal of Psychiatry*, *156*, 1618–1629.
- Playfair, G. L. (1999). Identical twins and telepathy. *Journal of the Society for Psychical Research*, *63*, 86–98.
- Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin*, *28*, 369–381.
- Rhine, L. E. (1981). *The invisible picture: A study of psychic experiences*. Jefferson, NC: McFarland.
- Sheldrake, R., & Smart, P. (2003). Experimental tests for telephone telepathy. *Journal of the Society for Psychical Research*, *67*, 184–199.
- Sherwood, S. J., Dalton, K., Steinkamp, F., & Watt, C. (2000). Dream clairvoyance study II using dynamic video-clips: Investigation of consensus voting judging procedures and target emotionality. *Dreaming*, *10*, 221–236.
- Smith, S. M. (2002a). Fast robust automated brain extraction. *Human Brain Mapping*, *17*, 143–155.
- Smith, S. M. (2002b). Preparing fMRI data for statistical analysis. In P. Jezzard, P. M. Matthews, & S. M. Smith (Eds.), *Functional MRI: An introduction to methods* (pp. 129–141). Oxford: Oxford University Press.
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, *23*, S208–S219.
- Targ, R., Braud, W., Stanford, R., Schlitz, M., & Honorton, C. (1991). Increasing psychic reliability—A panel discussion presented at the 33rd Annual Conference of the Parapsychological Association, Chevy Chase, MD, August 16–20, 1990. *Journal of Parapsychology*, *55*, 59–83.
- Taylor, R. (2003). Evolutionary theory and psi: Reviewing and revising some need-serving models in psychic functioning. *Journal of the Society for Psychical Research*, *67*, 1–17.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*, 207–232.
- van Turennout, M., Ellmore, T., & Martin, A. (2000). Long-lasting cortical plasticity in the object naming system. *Nature Neuroscience*, *3*, 1329–1334.
- Vuilleumier, P., Schwartz, S., Duhaux, S., Dolan, R. J., & Driver, J. (2005). Selective attention modulates neural substrates of repetition priming and “implicit” visual memory: Suppressions and enhancements revealed by fMRI. *Journal of Cognitive Neuroscience*, *17*, 1245–1260.
- Warren, C. A., McDonough, B. E., & Don, N. S. (1992a). Event-related potential changes in a psi task. *Journal of Parapsychology*, *56*, 1–30.
- Warren, C. A., McDonough, B. E., & Don, N. S. (1992b). Partial replication of single subject event-related potential effect in a psi task. In E. Cook (Ed.), *The Parapsychological Association 35th Annual Convention: Proceedings of Presented papers* (pp. 169–181). New York: Parapsychological Association, Inc.
- Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology*, *12*, 129–140.
- Wheatley, T., Weisberg, J., Beauchamp, M. S., & Martin, A. (2005). Automatic priming of semantically related words reduces activity in the fusiform gyrus. *Journal of Cognitive Neuroscience*, *17*, 1871–1885.
- Woolrich, M. W., Ripley, B. D., Brady, J. M., & Smith, S. M. (2001). Temporal autocorrelation in univariate linear modelling of fMRI data. *Neuroimage*, *14*, 1370–1386.
- Worsley, K. J., Evans, A. C., Marrett, S., & Neelin, P. (1992). A three-dimensional statistical analysis for CBF activation studies in human brain. *Journal of Cerebral Blood Flow and Metabolism*, *12*, 900–918.