Great Expectations:

What Can fMRI Research Tell Us About Psychological Phenomena?

Tatjana Aue[1,2], Leah A. Lavelle[1], and John T. Cacioppo[1]

[1]Center for Cognitive and Social Neuroscience, University of Chicago

[2]Swiss Center for Affective Sciences, University of Geneva, Geneva, Switzerland

Address correspondence for all authors to:

Tatjana Aue

Swiss Center for Affective Sciences

7, Rue des Battoirs

1205 Geneva

Switzerland

tatjana.aue@pse.unige.ch

Abstract

Expectations for what functional magnetic resonance imaging (fMRI) can offer psychophysiology vary greatly. Overreaching enthusiasm such as the idea that fMRI can reveal lies and political attitudes are as common as the opinion that fMRI, in its current form, is useless for the advancement of psychological theories. Errors in the inferences being drawn from fMRI data may be contributing to each of these extreme positions, so the present paper addresses these several common inferential errors and describes some of the potential of fMRI for psychophysiological theory and research.

Great Expectations:

What Can fMRI Research Tell Us About Psychological Phenomena?

Functional magnetic resonance imaging (fMRI) is a tool used by a growing number of scientists who seek to investigate the brain mechanisms underlying psychological phenomena (Cacioppo & Decety, in press). To determine the extent of this growth, we performed a literature search on the terms “fMRI” or “functional magnetic resonance imaging” on Medline, ISI, and PsycInfo. To provide a relevant point of comparison in these searches, we also performed a search on the terms “electroencephalography” and “EEG”. The results are displayed in Figure 1. From this figure it is clear that the number of papers published in fMRI has increased steadily since systematic fMRI research began in 1991. As would be expected, there have been increases in the number of studies published using EEG, as well, but the rate of growth has been far greater for fMRI, especially in the psychological literature (see Figure 1, bottom panel).

Why is fMRI becoming the measure of choice in psychological studies of brain activity? What can fMRI tell us about ongoing psychological processes? The rise in the use of fMRI techniques in psychological research, and the diversion of resources from traditional methodologies to fMRI, has led many to ask these questions. Different authors have come to widely different answers, ranging from a conviction that fMRI provides a veritable readout of mental contents to the equally extreme conviction that fMRI is a waste of time and resources in psychological science (see Page, 2006). In this paper, we attempt to place these extremes into a broader context with the aim of encouraging the effective use of fMRI to advance psychological theory and research.

Divergent Opinions on the Utility of fMRI for the Study of Psychological Phenomena

There are several reasons for the increasing frequency of application of fMRI to psychological questions. First, there have been considerable advances in the fMRI technique during the last decade (e.g., improved spatial and temporal resolution, improved signal-to-noise ratio) that have made fMRI data more precise and reliable. Second, fMRI is noninvasive and relatively safe in contrast to some other neuroimaging tools such as positron emission tomography. Third, fMRI allows for the continuous collection of data, which is an advantage when tracking ongoing processes. Lastly, because verbal reports can be distorted and are difficult to collect effectively online (e.g., Redelmeier & Kahneman, 1996), there is a strong demand for more objective measures of psychological processes. fMRI is appealing because it can be used to investigate psychological operations to which people have little or no verbal access.

For some investigators, fMRI holds a special promise for the objective study of phenomena that people would not be willing or capable of providing in verbal reports (e.g., lies and deception, repression, nonconscious information processing operations), leading thus to the overly optimistic opinion that fMRI can be used to simply read people's knowledge, beliefs, intentions, and political attitudes. While these phenomena can certainly be informed by fMRI data, the opinion that fMRI provides a readout of mental contents is overly optimistic because it is based on the denial that other antecedent conditions exist for the observed pattern of brain activation.

Iacoboni, Freedman, and Kaplan (2007), for example, used fMRI to infer political attitudes from changes arising during the presentation of pictures and video excerpts of presidential candidates. Iacoboni and colleagues reported that people who had rated Hillary Clinton unfavorably displayed increased activity in the anterior cingulate cortex (ACC) upon viewing pictures and videos of Hillary Clinton. The authors attributed this activity to “battling unacknowledged impulses,” given that the ACC had been associated with conflict monitoring and response conflict in earlier studies (see Bush, Luu, & Posner, 2000, for an overview). They also interpreted differences in activity in the amygdala and the insula as indexing anxiety and disgust for political parties, respectively.

Research on lie detection represents another context in which fMRI data have been interpreted as a read-out of mental processes. Spence and collaborators (2008) studied a single patient suffering from Munchhausen’s syndrome who had been accused of poisoning her child. In an MRI scanner, the patient was told to respond in two different ways to questions about the alleged poisoning, expressing either her innocence (consistent with her own version of events) or her guilt (consistent with the accuser’s version of events). Spence et al. observed greater activity in the ventrolateral prefrontal cortex (vlPFC) and the ACC when the patient expressed guilt than when she expressed innocence. Based on the assumption that deception incurs a cognitive processing cost, this differential brain activity was interpreted as being “consistent with (though they [the data] do not prove) the thesis that she [the patient] is telling the truth when she denies having harmed the child in her care” (p. 312).

Both the Iacoboni et al. (2007) and the Spence et al. (2008) interpretations may be consistent with their observations but there are alternative interpretations that are also plausible. For instance, increased activity in the amygdala and in the insula has been linked to various aversive states besides anxiety or disgust (and even to appetitive states;

see Feldman Barrett & Wager, 2006, for an overview), and increased activity in the ACC is not limited to conflict much less to conflict between unacknowledged impulses (Bush et al., 2000). Thus, the interpretation of the fMRI data offered by Iacoboni et al. (2007) is but one of several that are plausible. Similarly, the fMRI data reported by Spence et al. (2008) is just as consistent with the examinee being guilty as with the examinee being innocent. As Spence and colleagues acknowledged, as a result of repeated questioning on prior occasions, it is conceivable that the examinee's own version of the events were so well-rehearsed as to have automatic components. The endorsement of the accuser's version of the events, even if true, was not as rehearsed a response and may have engendered conflict. Moreover, admission of a culturally reprehensible behavior such as harming one's child may trigger strong control processes in an attempt to avoid detection.

The inference that regional brain activation indexes a specific psychological state without considering the possibility that other antecedent conditions produce the same regional brain activation is an example of affirming the consequent. Inferential errors of this type are not unique to fMRI research (Cacioppo & Tassinary, 1990), of course, but the high cost of fMRI research and the concomitant reduction in resources available to support more traditional forms of behavioral research have led some researchers to be skeptical of using fMRI as a tool for the study of psychological phenomena.

One criticism of fMRI comes from researchers who see current fMRI research as being limited to localization of function, which alone would not help to advance psychological theory (e.g., Coltheart, 2006a,b). They argue that brain imaging data do not further scientific theory because they are often simply interpreted in line with prior psychological theories. This sentiment is clearly expressed by Kihlstrom (2006) who claims that "There does not appear to be any instance where neuroscientific findings have constrained social-psychological theory…To the contrary, it appears that precisely the reverse is true: psychological theory constrains the interpretation of neuropsychological and neuroscientific data" (p.16). A similar argument was made by Coltheart (2006b), who states, that "no functional neuroimaging research to date has yielded data that can be used to distinguish between competing psychological theories" (p. 323).

From the above, it is evident that fMRI research has elicited a variety of responses, ranging from exaggerated claims to complete skepticism. This divergence of opinions may ultimately contribute to a more realistic assessment of new scientific tools. Some of the same teething troubles were seen when somatovisceral methods were applied to psychological issues some decades ago (see Landis, 1930; Cacioppo, Petty, & Tassinary, 1989). On the one hand, physiological measures ranging from EEG, heart rate, and skin conductance to event-related brain potentials, impedance cardiography, and facial electromyography were embraced by many psychologists because they were seen as having less bias than traditional psychological measures. On the other hand, other researchers saw psychophysiology as a failure because the data did not support invariant inferences about ongoing psychological processes. Whereas enthusiasm ensures the application and continuing use of a scientific tool in a variety of scientific contexts, skepticism may help to reveal shortcomings and stimulate fundamental work to address these shortcomings. Many of these shortcomings revolve around a misunderstanding of which inferences are appropriate to draw from the data in light of the experimental design and the extant literature. We, therefore, focus here on inferential limitations and possibilities in fMRI research.

Reducing Errors in Inference

Well defined theories allow for the use of hypothetico-deductive logic, which entails formulating hypotheses that can be falsified using fMRI data (Cacioppo, Berntson, & Nusbaum, 2008). The steps of this process are familiar: devising alternative hypotheses; devising an experiment with possible outcomes each of which could exclude one or more of the proposed hypotheses; conducting the experiment to obtain a clean result; and repeating the process in such a way as to refine the remaining hypotheses. Platt (1964) has argued that research based on hypothetico-deductive logic forms the basis for strong inferences and promotes scientific progress.

When experiments do not follow hypothetico-deductive logic, Platt (1964) argued that scientific progress is slowed by experiments and interpretations that suffer from a confirmatory bias. If a theory posits that a specific brain response (Φ) causes a psychological event (Ψ), the manipulation of the psychological event and the observation of activation over the specified brain response may provide only limited support for the experimental hypothesis. Hempel (1945) referred to inferences of this form as the paradox of confirmation (also called Hempel’s Raven Paradox). According to this famous paradox, the statements ‘all ravens are black' and ‘all non-black things are non-ravens' are logically equivalent. The hypothesis that 'all ravens are black' would be directly corroborated by observations of only black ravens. However, based on the logical structure of the argument, the statement 'all non-black things are non-ravens' also corroborates the hypothesis; for instance, a statement such as 'this is a pink flower.’ Continuing to corroborate a hypothesis does not imply that there is not another better explanation. Actively attempting to falsify a theorem by specifying situations where different theorems would suggest different outcomes minimizes this problem.

An approach to this problem has been suggested (Cacioppo & Tassinary, 1990; Sarter, Berntson, & Cacioppo, 1996). In cases in which one wishes to infer psychological events (ψ) based on a physiological event (Φ), one must know not only the rate at which ψ and Φ occur together ($P(\psi, \Phi)$), but also the rate at which Φ occurs overall ($P(\Phi)$). The appropriate measure of the probability of ψ given Φ is then $P(\psi|\Phi) = P(\psi, \Phi)/P(\Phi)$. Only if ψ systematically increases or decreases the likelihood of occurrence of Φ (or vice versa) can a real covariation of psychological and physiological events be assumed.

Inferential problems can occur if interpretations are made assuming an invariant relationship between ψ and Φ. Cacioppo and Tassinary (1990) describe four psychophysiological relationships which vary in specificity and generality. Specificity refers to whether an element on one (e.g., the psychological) level relates to a single element or several elements on another (e.g., the physiological) level. Generality refers to whether the relationship is bound by context or generalizes across contexts. Invariant relationships are those, which meet the criteria for both specificity and generality, allowing for the inference that a change in Φ implies the change of ψ across contexts. For simplification, throughout the rest of the article, we will simply refer to many-to-one and many-to-many relationships (low specificity) as compared to one-to-one relationships (high specificity).

Unless inferences are based on hypothetico-deductive logic, one cannot infer ψ from Φ without risking an inferential error when the type of relationship under investigation is not known. For instance, if a given brain area (Φ) shows a hemodynamic response as a function of an experimentally manipulated emotional state (ψ), it does not logically follow that the activation in that brain area in another condition or study marks the occurrence of that emotional state. Unless a relationship between ψ and Φ can be determined to be one-to-one, at least within a specific measurement context, changes in activation in a given brain area or neural network following a certain psychological phenomenon do not allow for strong inferences that those same activations signal changes in the psychological phenomenon (Sarter, Berntson, & Cacioppo, 1996). This is because the observed activations could arise from a number of quite distinct psychological phenomena. The following sections describe how research questions are informed by knowledge of psychophysiological relations and their corresponding inferences.

Examples of Research Questions That Can Be Addressed By fMRI

Contrasting Distinct Psychological Theories

fMRI can be a useful tool for comparing psychological theories that predict the same behavioral outcomes, but differ in the hypothesized mechanism explaining the outcome (e.g., Dickhaut et al., 2003). Although it is always desirable to have prior knowledge of the psychophysiological relation under study, hypothetico-deductive logic, as outlined above, allows for falsification of one or more hypotheses even in the absence of prior knowledge regarding the specificity or generality of the relationship between ψ and Φ. Henson (2006a), for instance, used fMRI to compare single process or dual process theories of recognition memory. According to single process theories, items that are presented and later explicitly remembered (Remember) as compared to those that just seem familiar (Know) differ only in memory strength. Dual process theories, in contrast, postulate that Remember and Know judgments are qualitatively different because they rely on two distinct memory processes. Henson (2006a), therefore, argued that any qualitative difference in fMRI data between Remember and Know judgments would favor dual process theories over single process theories.

According to Henson (2006a), two important characteristics are required to infer qualitative differences between two experimental conditions from fMRI data. First, analogous to techniques used in lesion studies, there should be a significant interaction between the two conditions in at least two brain regions (crossover interaction), increasing the likelihood that the observed differences are qualitative rather than quantitative in nature. Additionally, for at least one of the experimental conditions, the two brain areas should both be activated or both be deactivated in comparison to a baseline condition. Otherwise, the crossover pattern could simply result from reciprocal connections between the two brain areas (i.e., the increase in the first region leads to a decrease in the second region or vice versa) and not reflect a true qualitative difference.

Henson (2006a) observed such a true qualitative difference in his research on memory recognition, revealed by a crossover interaction between the experimental conditions (Remember versus Know) and brain activations. The posterior cingulate cortex was more strongly activated for Remember than Know judgments and the right lateral frontal cortex was more strongly activated for Know than Remember judgments. This result favors the idea of Remember and Know responses relying on distinct memory processes, thus supporting dual process theories (see also Henson, 2006b; for an alternative interpretation of the data, still favoring the existence of two different processes, see, Henson, Rugg, Shallice, & Dolan, 2000).

In many instances investigators have some prior knowledge of the psychophysiological relationship under investigation and have specific hypotheses about which brain regions are associated with a psychological process. De Lange et al. (2008), for example, contrasted different theories about how other people's action intentions are recognized. According to a first view, motor simulation is central to the understanding of third-person action intentions (Gallese & Goldman, 1998). Motor simulation is believed to occur in the putative mirror neuron system, which spans regions in the premotor and parietal cortex. Neurons in these regions fire both when people perform an action and when they observe another person performing the same action. A second view postulates instead that other people's action intentions are inferred by mentalizing their desires and goals (Frith & Frith, 2003). The neural network linked to such mentalizing activities includes the medial prefrontal cortex, the posterior cingulate cortex, the superior temporal cortex, and the temporoparietal junction. Finally, a third view sees motor simulation and mentalizing as complementary processes for the recognition of other people's intentions (Keysers & Gazzola, 2007), with the former being automatic and intuitive and the latter being more reflective in nature. Thus, hypothetico-deductive reasoning permitted a strong test of these hypotheses.

De Lange et al. (2008) tested the predictions of the three theories by instructing participants to either judge the extent to which an actor's *means* of performing an action were unusual (e.g., coffee cup that is brought to the mouth is held in a power grip; no conscious processing of intentions) or the extent to which the actor's action *intentions* were unusual (e.g., coffee cup is brought to the ear; conscious processing of intentions). The authors found that, irrespective of whether or not intentions were processed consciously, unusual intentions involved increased activity in the inferior frontal gyrus, a structure belonging to the mirror neuron system. In contrast, only conscious processing of the actor's intentions activated the medial prefrontal cortex, the posterior cingulate cortex, and the right posterior superior temporal sulcus, all of them being structures belonging to the mentalizing brain network. These observations, therefore, are inconsistent with predictions from the first two views.

The described fMRI studies illustrate that theoretical tests based on hypothetico-deductive logic can and have been performed to distinguish between theories. What is more, these studies helped to evaluate competing theories in a way that would have been difficult, if not impossible, using behavioral measures alone. Not unique to fMRI research, studies designed to falsify one or more competing hypotheses only provide evidence for which of two theories better accounts for the obtained data. While the above studies provided evidence that falsified one or more specific theories, it is also possible that fMRI data do not allow for the rejection of either hypothesis or are consistent with both hypotheses. A researcher may therefore come up with an alternative theory and corresponding hypotheses, which are critically tested in subsequent studies, as suggested in the process of strong inference (cf. Henson, 2006a).

Hypotheses Generation

fMRI data have the potential to inform research questions by generating new hypotheses for psychological phenomena. New hypotheses about psychological phenomena can be formed when fMRI data violate expectations derived from behavioral data, provide data that extend behavioral observations and thus alter the understanding of a current theory, and even when fMRI studies fail to yield data that unambiguously favor one theory over opposing psychological theories.

An example demonstrating fMRI's utility for hypothesis generation comes from a debate within research on visual short-term memory (VSTM). Behavioral data yielded ambiguous information about the number of objects that can be held in VSTM, which led to a debate about whether object representation in VSTM was limited to a fixed number or was affected by object complexity. Whereas some studies reported a fixed number of 4 objects could be stored in VSTM (Luck & Vogel, 1997), others suggested that the number of objects stored in VSTM decreases with increasing object complexity ($n \leq 4$; Alvarez & Cavanagh, 2004).

In a series of four studies, Xu and Chun (2006) demonstrated that each theory was supported by fMRI data in specific regions of the parietal and occipital cortices. Activity in the inferior intraparietal sulcus (IPS) increased with increasing numbers of objects, reaching a plateau at set size 3 or 4, regardless of object complexity. This region also encoded the spatial location of the displayed objects. Activity in the superior IPS and the lateral occipital complex, in contrast, was influenced by object complexity. For simple objects, activity in these regions also reached a plateau at set size 3 or 4. With increasing complexity of the objects presented, however, the plateau was reached with smaller object set sizes already, proposing that the superior IPS and the lateral occipital complex encode object features instead of whole objects.

These data permitted the generation of the hypothesis that VSTM capacity relies on two or more interacting psychological processes. The data suggest a first process, associated with activity in the inferior IPS, which encodes the spatial locations of the objects and determines which objects in the environment receive closer attention.

A second process, associated with activity in the superior IPS and the lateral occipital complex, then more deeply processes and encodes the objects that were preselected in the early stage. However, due to limited processing capacities, the number of complex objects that can receive closer inspection in the second stage is further reduced. This hypothesis helped explain inconsistent findings in the VSTM research and unified the two competing theories in a way that had proven difficult to achieve based on analyses of behavioral data alone.

Hypotheses derived from fMRI research can, in this way, act as a starting point for subsequent systematic investigations of psychological phenomena and provide insight about the mechanisms underlying these phenomena. For instance, psychological theorizing can be further advanced by directly and reliably relating psychological processes to specific brain activations.

Strategies That Increase the Interpretability of fMRI data

As we have seen thus far, fMRI data can address questions about how neural activation might be related to certain mental states or behavior. However, inferences about a specific psychological process based on the activation of a particular brain region should be considered tentative given the current limited knowledge about the functions served by individual brain regions and the possibility of more than one psychological antecedent for the activation of a particular brain region.

The possibility of many-to-many and many-to-one relationships makes it important to search for alternative, testable hypotheses that might fit the data (cf. Poldrack, 2006). A researcher therefore may want to decrease the possible number of alternative interpretations for specific brain activations when planning an experiment.

The following actions foster the accomplishment of this goal.

1. Refinement of Psychological Concepts

In some studies, possible interpretations of fMRI data can be limited by refining concepts on the psychological level (ψ). For example, interference in different versions of the Stroop task can be cognitive (color Stroop) or affective (emotional Stroop) in nature. Different subsections of the ACC have been linked to these cognitive and affective Stroop tasks (Bush, Luu, & Posner, 2000), and these differences on the psychological level have been shown to account for a considerable part of the spatial distribution of activations within the ACC in Stroop tasks. Breaking down psychological concepts in subcomponents, in this instance, increased the interpretability of brain data by equating these specific subcomponents with specific brain activations.

2. Consideration of an Intervening Computational Level

Better inferences may also be obtained by considering the psychological concept under investigation within a computational framework. Inference is improved when psychological processes are decomposed into their constituent information processing operations, instead of assuming a one-to-one relationship between psychological concepts and neural responses. These basic information processing operations may be more easily related to activity in specific brain regions. Additionally, identification of information processing components allows for a better understanding of more complex psychological phenomena by elucidating when and how these neural processes are integrated.

Such a strategy has been pursued in recent research on the visual system, which has demonstrated that decomposing psychological entities into simpler computational features can enhance the accuracy of inferences drawn from brain responses (Kay, Naselaris, Prenger, & Gallant, 2008). The neural activity associated with processing of simple images was better understood by decomposing images into visual processing components such as space, orientation, and spatial frequency. In the study by Kay et al. (2008), participants looked at pictures while in an MRI scanner. From the acquired brain data, a qualitative receptive-field model was estimated for each voxel, accounting for space, orientation, and spatial frequency characteristics of the presented images. Based on this model, which defined psychological concepts by its related basic computational units in the visual system, the authors improved their ability to predict brain responses to a new set of pictures and were able to correctly identify which picture had been displayed in most cases (110 out of 120 pictures for a first participant; 86 out of 120 pictures for a second participant).

3. Consideration of Smaller, Circumscribed, Brain Regions

Another way to decrease errors of inference is by considering well-defined, circumscribed, brain regions instead of gross anatomical landmarks. Considering smaller regions where activation is particularly strongly associated with a specific psychological event may simultaneously reduce the probability of activation in the exact same region due to other psychological events. Because, as mentioned above, $P(\psi|\Phi) = P(\psi ,\Phi)/P(\Phi)$, choosing an area that has a strong relationship could increase $P(\psi ,\Phi)$, while excluding parts of the area that are associated with other mental processes simultaneously reduces $P(\Phi)$ by decreasing the probability $P(\Phi$, not $\psi)$ in the concerned region. Both effects should increase $P(\psi|\Phi)$, and thus the validity of an inference deriving mental processes from brain activities.

For instance, Poldrack (2006) found that smaller brain regions allowed for more valid inferences about language processing. A brain region is characterized by high selectivity if it is activated by the phenomenon of interest and only few or no other phenomena. High selectivity of a region, in turn, allows for stronger inferences. On the basis of the results of a representative study in the field (McDermott et al., 2003), Poldrack (2006) defined a region of interest (ROI) in the dorsal left inferior frontal gyrus that had been shown to be active during both phonological and semantic processing. To assess the ROI's selectivity for language processing, Poldrack (2006) compared the number of language-related studies with the number of not language-related studies showing activity in this region. While the center for this ROI was comparable for different analyses, cluster width varied from 4 mm to 40 mm. The author demonstrated that smaller region sizes were characterized by greater selectivity (i.e., fewer errors of inference) than were larger region sizes.

4. Consideration of neural networks instead of isolated brain areas

Brain circuits rather than single brain regions may be a better marker of psychological phenomena or processes, especially when these phenomena or processes are complex. Poldrack (2008) posits that "to the degree that neuroimaging data are predictive of mental states, this prediction will come both from fine-grained patterns of activity (which are not evident to the human eye) and from the coordinated activity across many brain regions, not from localized activity in specific brain regions."

Using coordinated patterns of activity strengthens the association $P(\psi, \Phi)$ in much the same way that choosing small, distinct brain regions strengthens associations. Physiological responses to a given psychological phenomenon may share some of the same regions of brain activation with other psychological phenomena, but it is less likely that the exact same pattern of activation will happen in the absence of the psychological phenomenon. P(Φ) is, therefore, reduced by reducing P(Φ, not ψ). Defining Φ as activity of a neural network instead of activity of an isolated brain area thus may increase the interpretability of the observed fMRI data.

5. Testing fMRI Data for Homogeneity

In efforts to increase the signal to noise ratio in fMRI studies, patterns of activation are often aggregated across trials and across individuals. For the results of aggregations to be informative, similar trials within a subject must be homogenous and individuals must also be homogenous on a particular trial type. The existence of heterogeneity within an individual or between individuals may lead to erroneous interpretations. For example, if the same behavioral outcome can result from different strategies, reflected by different brain activations, composite activation may be an average of the two or more different activation patterns. This can occur if a single person uses different strategies on different trials, if different individuals use different processing strategies, or if both of these types of heterogeneity are present. The aggregate pattern of activation that is obtained would reflect none of the individual patterns of activation accurately. In order to reduce problems of aggregation within individuals, tests of heterogeneity are typically conducted. Some examples include comparing activation patterns on even versus odd trials, or early trials versus later trials (e.g., McKeown et al., 2002). Across individuals, such tests of heterogeneity are still relatively rare in fMRI studies (e.g., Kherif et al., 2003).

The aggregate activation pattern, in cases of heterogeneity, reflects the average activation not the typical activation. In fact, in some cases, the aggregate activation pattern may not be representative of any individual's pattern of activation. Inferring that the average activation pattern is the typical activation pattern can lead to spurious inference. One example of aggregate data not representing a typical pattern of activation is in lie detection. Monteleone et al. (in press) attempted to predict whether each individual in a group was lying from the aggregate activation of the group. They found that even though the group analysis showed greater activation in a number of regions when participants were lying, none of these regions predicted lying across all individuals.

6. Experimentally Controlling for Alternative Antecedents and Use of Convergent Measures

When planning a study, the researcher may want to control a priori for some alternative interpretations of expected brain activations. Controlling for alternative causes of brain activation can also reduce P(Φ, not ψ). This can be achieved by holding these alternative antecedents constant across experimental conditions. Additionally, selecting specific experimental tasks or participants may reduce the number of possible alternative explanations.

Alternative explanations may also be controlled for through the use of convergent methods of data collection. For example, in research on conflict, registration of reaction times and eye tracking can be used to exclude visual attention as an alternative interpretation for obtained fMRI data. Likewise, a study of emotional state mirroring used facial EMG to exclude an alternative hypothesis that activity in premotor cortices reflected emotional expression rather than automatic preparatory responses (Warren et al., 2006). Eye tracking or facial EMG used in other types of studies, however, might not provide information that can differentiate between theories.

Behavioral measures and fMRI therefore, act as complementary sources of information, but only improve inference when they are selected to provide information about hypotheses that could not be appropriately or easily tested with fMRI data or complementary methods alone.

In sum, the increased use of fMRI research for the investigation of psychological questions has been a cause for both enthusiasm and skepticism. The aim of the present article is to draw a realistic picture of what psychophysiology currently offers to fMRI and may offer in the future by discussing problems in inference occasionally found in this area of research. An investigator's aim may be to uncover one-to-one relationships between psychological phenomena or processes and activity in specific brain networks. If such invariant relationships are established, the number of potential antecedents is minimized and affirming the consequent would be correctly implicated by the results. However, this goal will perhaps be reached more quickly if investigators do not assume any invariant relationships but instead expect to have to marshal empirical evidence to support the claim.

References

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science, 15,* 106-111.

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences, 4,* 215-222.

Cacioppo, J. T., Berntson, G. G., Lorig, T. S., Norris, C. J., Rickett, E., & Nusbaum, H. (2003). Just because you're imaging the brain doesn't mean you can stop using your head: A primer and set of first principles. *Journal of Personality and Social Psychology, 85,* 650-661.

Cacioppo, J. T., Berntson, G. G., & Nusbaum, H. C. (2008). Neuroimaging as a new tool in the toolbox of psychological science. *Current Directions in Psychological Science, 17,* 62-67.

Cacioppo, J. T., & Decety, J. (in press). What are the brain mechanisms on which psychological processes are based? *Perspectives on Psychological Science*.

Cacioppo, J. T., Petty, R. E., & Tassinary, L. G. (1989). Social psychophysiology: A new look. *Advances in Experimental Social Psychology, 22*, 39-91.

Cacioppo, J. T., & Tassinary, L. G. (1990). Inferring psychological significance from physiological signals. *American Psychologist, 45,* 16-28.

Coltheart, M. (2006a). What has functional neuroimaging told us about the mind (so far)? *Cortex, 42,* 323-331.

Coltheart, M. (2006b). Perhaps functional neuroimaging has not told us anything about the mind (so far). *Cortex, 42,* 422-427.

De Lange, F. P., Spronk, M., Willems, R. M., Toni, I., Bekkering, H. (2008). Complementary systems for understanding action intentions. *Current Biology, 18,* 454-457.

Dickhaut, J., McCabe, K., Nagode, J. C., Rustichini, A., Smith, K., & Pardo, J. V. (2003). The impact of the certainty context on the process of choice. *Proceedings of the National Academy of Sciences of the United States of America, 100,* 3536-3541.

Feldman Barrett, L., & Wager, T. (2006). The structure of emotion: Evidence from the neuroimaging of emotion. *Current Directions in Psychological Science, 15,* 79-85.

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London B, 358,* 459-473.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences, 2,* 493-501.

Hempel, C. G. (1945). Studies in the logic of confirmation. *Mind, 5, 1-26, 97-121.*

Henson, R. (2006a). Forward inference using functional neuroimaging: Dissociations versus associations. *Trends in Cognitive Sciences, 10,* 64-69.

Henson, R. (2006b). What has (neuro)psychology told us about the mind (so far)? A reply to Coltheart (2006). *Cortex, 42*, 387-392.

Henson, R. N. A., Rugg, M. D., Shallice, T., & Dolan, R. J. (2000). Confidence in recognition memory for words: Dissociating right prefrontal roles in episodic retrieval. *Journal of Cognitive Neuroscience, 12,* 913-923.

Iacoboni, M., Freedman, J., Kaplan, J., Jamieson, K. H., Freedman, T., Knapp, B., et al. (2007, November 11). This is your brain on politics. *New York Times.* Retrieved December 19, 2007, from http://www.nytimes.com

Kay, K. N., Naselaris, T., Prenger, R. J., & Gallant, J. L. (2008). Indentifying natural images from human brain activity. *Nature*, *452*, 352-356.

Keysers, C., & Gazzola, V. (2007). Integrating simulation and theory of mind: From self to social cognition. *Trends in Cognitive Sciences*, *11,* 194-196.

Kherif, F., Poline, J. B., Meriaux, S., Benali, H., Flandin, G., & Brett, M. (2003). Group analysis in functional neuroimaging: Selecting subjects using similarity measures. *Neuroimage*, *20,* 2197-2208.

Kihlstrom, J.F. (2006). Does neuroscience constrain social-psychological theory? *Dialogue, 21*, 16-17.

Landis, C. (1930). Psychology and the psychogalvanic reflex. *Psychological Review, 37*, 381-398.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390,* 279-281.

McDermott, K. B., Petersen, S. E., Watson, J. M., & Ojemann, J. G. (2003). A procedure for identifying regions preferentially activated by attention to semantic and phonological relations using functional magnetic resonance imaging. *Neuropsychologia, 41*, 293-303.

McKeown, M. J., Varadarajan, V., Huettel, S., & McCarthy, G. (2002). Deterministic and stochastic features of fMRI data: Implications for analysis of event-related experiments.

Monteleone, G. T., Phan, K. L., Nusbaum, H. C., Fitzgerald, D., Irick, J. S., Fienberg, S. E., & Cacioppo, J. T. (in press). Detection of deception using functional magnetic imaging: Better than chance, but well below perfection. *Social Neuroscience.*

Page, M. P. A. (2006). What can't functional neuroimaging tell the cognitive psychologist? *Cortex, 42,* 428-443.

Platt, J. R. (1964). Strong inference. *Science, 146,* 347-353.

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences, 10,* 59-63.

Poldrack, R. A. (2008, June 3). Poldrack Replies to Iacoboni Neuropolitics Discussion. *Neuroethics and Law Blog.* Retrieved June 26, 2008, from http://kolber.typepad.com

Redelmeier, D. A., & Kahneman, D. (1996). Patients' memories of painful medical treatments: Real-time and retrospective evaluations of two minimally invasive procedures. *Pain, 66,* 3-8.

Sarter, M., Berntson, G. G., & Cacioppo, J. T. (1996). Brain imaging and cognitive neuroscience: Toward strong inference in attributing function to structure. *American Psychologist, 51,* 13-21.

Spence, S. A., Kaylor-Hughes, C. J., Brook, M. L., Lankappa, S. T., & Wilkinson, I. D. (2008). 'Munchausen's syndrome by proxy' or a 'miscarriage of justice'? An initial application of functional neuroimaging to the question of guilt versus innocence. *European Psychiatry, 23,* 309-314.

Warren, J. E., Sauter, D. A., Eisner, F., Wiland, J., Dresner, M. A., Wise, R. J. S., et al. (2006). Positive emotions preferentially engage an auditory-motor "mirror" system. *Journal of Neuroscience, 26,* 13067-13075.

Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature, 440,* 91-95.

Author Note

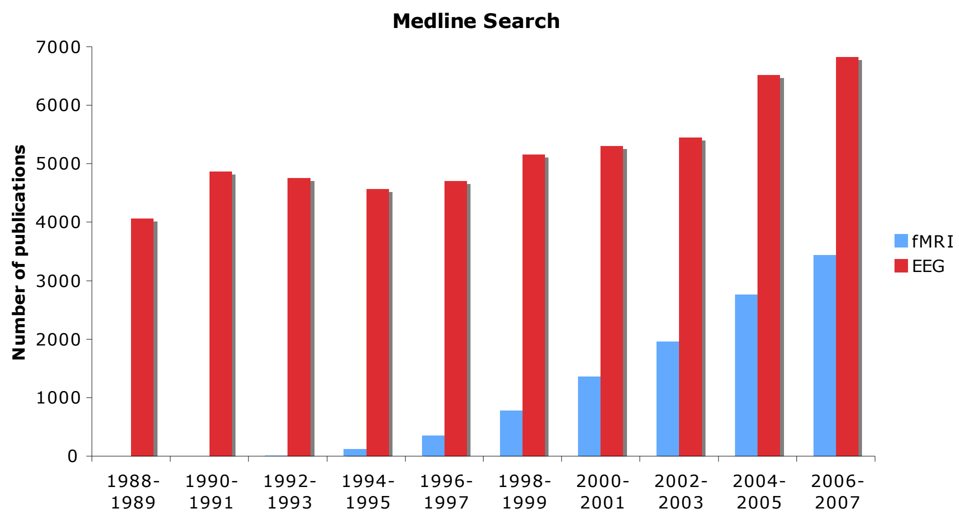
Tatjana Aue, Center for Cognitive and Social Neuroscience, University of Chicago, and Swiss Center for Affective Sciences, University of Geneva, Geneva, Switzerland, Leah A. Lavelle and John T. Cacioppo, Center for Cognitive and Social Neuroscience, University of Chicago.

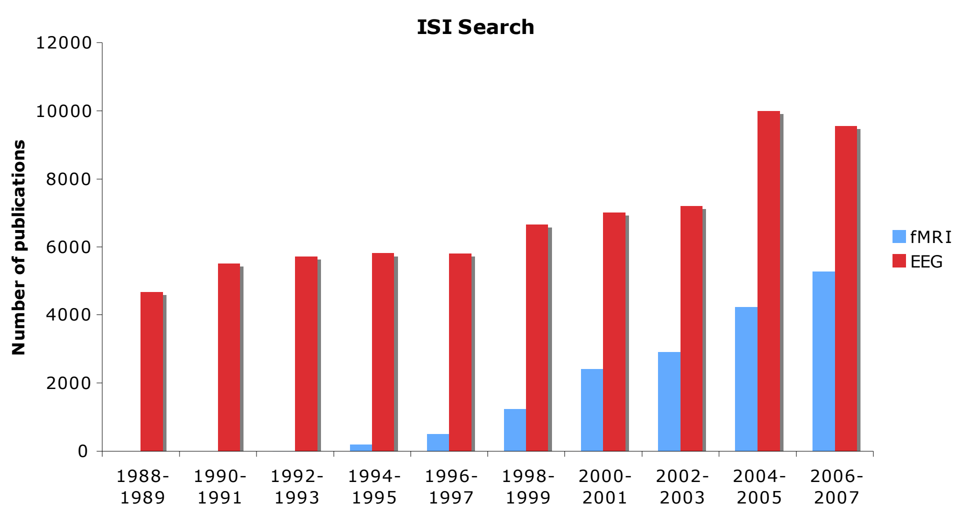
This research was supported by National Institute of Mental Health Grant No. P50 MH72850.

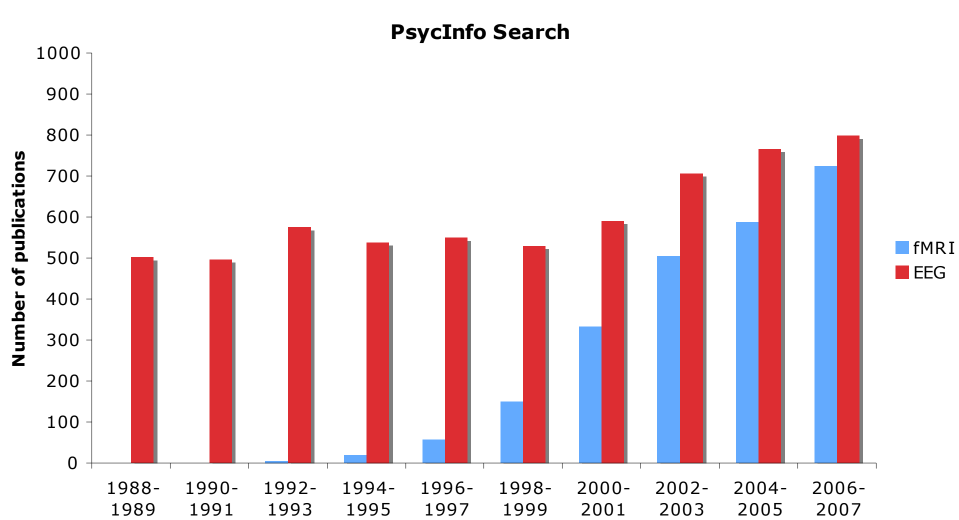
Correspondence concerning this article should be addressed to Tatjana Aue, Swiss Center for Affective Sciences, University of Geneva, Rue des Battoirs 7, 1205 Geneva, Switzerland. E-mail: tatjana.aue@unige.ch

Figure Caption

Figure 1. Number of fMRI and EEG publications within the last 20 years as revealed by literature searches performed on Medline (top panel), ISI (middle panel), and PsycInfo (bottom panel).


Medline Search
Number of publications
7000
6000
5000
4000
3000
2000
1000
0
1988-1989
1990-1991
1992-1993
1994-1995
1996-1997
1998-1999
2000-2001
2002-2003
2004-2005
2006-2007
fMRI
EEG


ISI Search
Number of publications
12000
10000
8000
6000
4000
2000
0
1988-1989
1990-1991
1992-1993
1994-1995
1996-1997
1998-1999
2000-2001
2002-2003
2004-2005
2006-2007
fMRI
EEG


PsycInfo Search
Number of publications
1000
900
800
700
600
500
400
300
200
100
0
1988-1989
1990-1991
1992-1993
1994-1995
1996-1997
1998-1999
2000-2001
2002-2003
2004-2005
2006-2007
fMRI
EEG