Do We Need To Study The Brain To Understand The Mind?

September 01, 2006

The brain is the most complex object in the known universe. Some 100 billion neurons release hundreds of neurotransmitters and peptides in a dynamic spanning timescales from the microsecond to the lifetime. Given this complexity, neurobiologists can spend productive careers studying a single receptor. Might psychologists more productively understand the mind by ignoring the brain altogether?

Marr (1977) suggested that mental processes may be studied at three levels of analysis: computational (the goals of the process), algorithmic (the method), and implementation (the hardware). The separation implies that the same computational goals and algorithms may be accomplished by a human brain or a computer, and the physical medium—neuron or silicon—is irrelevant. This concept was fundamental to the cognitive science movement and has given its practitioners permission to comfortably ignore the brain. But it has been seriously challenged: A high-level computation (e.g., deciding the next move in a chess game) can be accomplished in a virtually infinite number of ways. Building a computer model that accomplishes the computational goal says little about whether it does so in the same way that a human would. The hardware provides critical constraints on the space of possible models.

The debate about whether we need to study the brain to understand the mind is now being conducted among a network of thousands of scientists and scholars worldwide. The emerging consensus appears to be that implementation is important. Interestingly, the inverse question is also being asked by neurobiologists—do we need consider the mind to understand the brain?—and answered largely and increasingly in the affirmative.

We can learn much about the mind without knowing a neuron from an astrocyte. As I often repeat to myself and occasionally to others, “If you want to understand human performance, study human performance.” But brain data provide information about the mind that cannot be gleaned from even the most careful studies of behavior. In short, brain data provide a physical grounding that constrains the myriad otherwise-plausible models of cognition. They give us a direct window into which mental processes involve similar and different neurobiological processes, allowing us to use biology to ‘carve nature at its joints’ and understand the structure of mental processes (Kosslyn, 1994). Brain function also provides a common language for directly comparing and contrasting processes that are otherwise ‘apples and oranges,’ such as attention and emotion. This common language is a basis for the integration of knowledge across different types of research—basic and clinical, human and nonhuman.

As the general uses of neuroimaging have been eloquently discussed elsewhere, I focus here on a few examples of how functional magnetic resonance imaging (fMRI) has been useful in my work (see Jonides, Nee, & Berman, 2006). Also, as every method has its limitations, I discuss some of the pitfalls of making psychological inferences from neuroimaging data.

One use for me has been in understanding the structure of emotion and executive control processes, and the ways in which cognitive control operates in emotional and nonemotional situations. My colleagues

and I have asked: Is pain different from negative emotions such as sadness and anger, or are they variants on a common theme? In meta-analyses we have found that pain and negative emotions activate distinct brain networks, but share features such as anterior cingulate and frontal cortex activity with a broader class of processes, including attention (Wager & Barrett, 2004; Wager, Reading & Jonides, 2004). In contrast, different varieties of negative emotion engage largely overlapping networks. Thus, pain appears to be distinct from negative emotion, but commonalities suggest ways in which they may share underlying processes such as heightened attention.

Questions about the similarity and distinctiveness of mental processes have been at the heart of psychology since its inception, but definitive answers have been elusive. Inferences have been based largely on correlations in performance across tasks (or in physiological responses, for emotion). But performance data are relatively information-poor: the fact that two tasks take about as long to complete says little about whether processes involved in selecting the response were the same. Physiological responses suffer from similar problems of specificity. Neuroimaging provides a much richer source of information: if two tasks activate the same brain regions to the same degree, they are likely to involve similar processes. This logic provides a way to assess the structure of mental processes based on the similarity of their brain activation patterns. In a study based on these principles, we asked whether diverse 'executive control' tasks involve a common brain substrate (Wager, et al., 2005). Substantial overlapping activation suggested a common network for controlled response selection.

Though questions about mechanism are more difficult to address, neuroimaging can be informative here as well. In an fMRI study of pain, my colleagues and I found that expectation of pain relief induced by a placebo engages the frontal cortex and midbrain pain-relieving mechanisms (Wager et al., 2004). Frontal activation suggests a common substrate for maintaining cognitive context that shapes both perceptual/motor and affective processes, and midbrain activation suggests engagement of opioid analgesic systems. Such direct evidence on the mechanisms by which expectations affect pain would be hard to come by without studying the brain.

The study also points to an additional benefit of neuroimaging: In cases where self-report may be inaccurate, imaging can provide converging direct measures of central processing of a stimulus. Whereas expectations might affect pain reports for uninteresting reasons related to cognitive reporting bias, the evidence that expectations affect ongoing pain processing provides converging evidence that they shape pain experience.

Yes, there are many ways in which neuroimaging data can be misused or misinterpreted. Gross levels of regional brain activity might in some cases be uninformative about the similarity of psychological tasks: Two dissimilar tasks may involve the same regions but use different populations of neurons or involve different patterns of connectivity between regions. Two similar tasks might involve different regions but involve the same type of computation. Neural activity may be missed, as observed imaging signal only indirectly reflects neural activity, and observed imaging activation may not be essential for the task.

One of the biggest pitfalls is the temptation to observe brain activity and make inferences about the psychological state—for example, to infer episodic memory retrieval from hippocampal activity, fear from amygdala activity, or visual processing from activity in the 'visual cortex' (Barrett & Wager, 2006; Poldrack, 2006; Wager et al., in press). These inferences ignore the scope of processes which may activate each of these areas and involve a fallacy in reasoning: "if memory then hippocampus" is not the

same thing as “if hippocampus then memory.” The fact that few brain areas, including the ‘visual cortex,’ are dedicated to one process means that self-report is still the gold standard for assessing emotional experience and the contents of thought (Shuler & Bear, 2006). This is a serious challenge for those who would like, for example, to assess your brand preferences or your political affiliation from a brain scan. (And isn’t it easier just to ask?)

These problems are significant, but there is no perfect method—an understanding of the mind must emerge from a coordinated effort using converging evidence from all the tools at our disposal. Many of the issues above are being addressed by advances in data acquisition and analysis methods, the accumulation of more data on the mapping between brain structure and psychological function, and more nuanced views of what kinds of inferences are plausible. I believe that as the field matures, the exuberance of youth will give way to a more level-headed view of when and how neuroimaging can inform us about the mind. What we have learned already is considerable, and the accelerated integration across fields is leading to ever more and sophisticated and veridical models of the mind.

References

Barrett, L.F. and Wager, T.D. (2006). The structure of emotion: Evidence from neuroimaging studies. *Current Directions in Psychological Science, 15*, 79-83.

Jonides, J., Nee, D.E., Berman, M.G. (2006). What has functional neuroimaging told us about the mind? So many examples, so little space. *Cortex, 42,* 414-427.

Kosslyn, S. M. (1994). Carving a system at its joints. *In image and brain: The resolution of the mental imagery debate.* Cambridge, MA: MIT Press.

Marr, D. and Poggio, T. (1977). From understanding computation to understanding neural circuitry. *Neurosciences Res Prog Bull, 15,* 470-488.

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

Shuler, M.G., Bear, M.F. (2006). Reward timing in the primary visual cortex. *Science, 311,* 1606-1609.

Wager, T.D. and Barrett, L.F. (2004). *From affect to control: Functional specialization of the insula in motivation and regulation*.

Wager, T.D., Reading S., Jonides, J. (2004). Neuroimaging studies of shifting attention: A meta-analysis. *Neuroimage, 22,* 1679-1693.

Wager, T.D., et al. (2005). Common and unique components of response inhibition revealed by fMRI. *Neuroimage, 27*, 323-340.

Wager, T.D. et al. (in press). Elements of functional neuroimaging. In J. Cacioppo and R.J. Davidson (Ed.), *Handbook of Psychophysiology*. Cambridge, MA: Cambridge University Press.

Wager, T.D., et al. (2004). Placebo-induce changes in fMRI in the anticipation and experience of pain. *Science*, 303, 1162-1167.