

IN DEFENSE OF BRAIN MAPPING IN SOCIAL AND AFFECTIVE NEUROSCIENCE

William A. Cunningham
Ohio State University

The primary question asked for this special issue concerns how neuroscience data might contribute to and generate new theories in social psychology. Interestingly, in attempts to address this question, social neuroscientists have rushed to make sometimes overly bold claims about their data to demonstrate the utility of social neuroscience. The implicit logic in this response is a fear that to be useful, social neuroscience data needs quickly to move beyond brain mapping—the localization of specific social processes such as morality, the self, attitude, or prejudice to a particular brain region—and contribute in a more meaningful way. The promise of social neuroscience has always been to link multiple levels of analysis, constrain interpretation across these levels, and derive new predictions; yet there is now some question as to whether the field can deliver on this promise. Although this is an important question to ask, I suggest that requiring specific answers to this concern may be counterproductive at this stage of the field and that shifting the focus from brain mapping to process inference may be too early. Reverse inferences from the brain to social theory need to be based on accurate models of brain function and to the extent there are errors in our understanding of observed brain activation patterns our inferences will be necessarily invalid. What is required in social neuroscience research is an iterative cycle of brain mapping with one goal to understand how social and affective processes are computed in the brain, and with each iteration a second goal to determine whether (a) current social psychological models can account for these data, and (b) if not, to derive new theories. Thus, brain mapping as a research goal cannot be maligned, as it is a critical element of the social neuroscience enterprise. Indeed, at this stage of this newly emerging field, brain mapping should perhaps even be our dominant focus.

I thank Evan Thompson, Matt Lieberman, Andy Jahn, Samantha Mowrer, Nathan Arbuckle, and Jacqueline Matthews for useful discussions and comments on this manuscript.

Correspondence concerning this article should be addressed to William A. Cunningham, Department of Psychology, The Ohio State University, 140H Lazenby Hall, 1827 Neil Avenue, Columbus, OH 43210. E-mail: cunningham.417@osu.edu.

In the last decade, there has been an exponential growth in research in social neuroscience. Although researchers have sought to link brain, body, and mind to understand social cognition and emotion for some time (see Cacioppo & Berntson, 1992), a facilitating factor in this expansion has been the development of functional magnetic resonance imaging (fMRI) and its ability to associate specific brain regions with various psychological processes (Ochsner & Lieberman, 2001). Although research using fMRI to study cognitive processes in humans may have an obvious influence on our understanding of the brain, new questions have arisen about whether this research can contribute back to theory in social psychology. In other words, can understanding where a particular process is localized in the brain “constrain” or modify theories in social psychology? Given the large amount of resources being directed to fMRI research (e.g., running a single participant can cost hundreds of dollars), people in the social psychological community want to know whether they will receive any return on their investment.

To adequately scrutinize this variant of social neuroscience research, one must consider the two interrelated goals of the field. First, social neuroscience seeks to understand how neurons in the brain can interact to give rise to complex cognitive processes such as understanding another person’s mind, or generating novel solutions to economic challenges. Although we spend many of our conscious and unconscious resources struggling to understand our social world, how our brain accomplishes this is relatively unknown. For example, although from our initial research we know the dorsal medial prefrontal cortex and the right temporo-parietal junction appear to be involved in generating person representations (Mitchell, Banaji, & Macrae, 2005; Saxe & Wexler, 2005), why these brain regions are used and what exactly they are doing still escape us. Second, social neuroscience seeks to take the findings from neuroimaging studies to reinterpret behavioral findings from social psychology, specifically with the goal of developing new theories. The question here is: Can knowing how the brain accomplishes a particular task (typically defined in terms of localization) inform our understanding of the computational process? Whereas the second goal is relatively uncontroversial, it is the first goal that inspires some debate. That is, the first goal is sometimes attacked as being simply “brain mapping,” relevant only perhaps to the neurosciences, and the second goal of conducting direct tests of social psychological theory using fMRI is considered legitimate.

In support of fMRI data, researchers sometimes point to cases when an inference from brain data can point toward one hypothesis or another. For instance, it is believed that one can determine whether performing a task requires emotional processing or not by observing activity in the amygdala (an area that is often activated when participants are engaged in emotional processing; Phelps, 2006) while participants perform a task. For example, a study by Cunningham et al. (2004) is thought to provide support for the suggestion that implicit attitudes operate through unconscious emotion because the presentation of subliminal Black faces activated the amygdala more so than White faces. Similarly, Greene, Sommerville, Nystrom, Darley, and Cohen (2001) tested whether moral judgments were relatively more emotional than cognitive by examining whether brain regions assumed to be involved in emotion were activated when participants made moral

judgments. In each of these two studies, assumptions about what a particular region performed were used to make theoretical inferences.

Yet, some critics suggest that fMRI data has not provided any new theoretical advances that a well-controlled behavioral study would not have also provided. More radically, some worry that it may never do so. With so many brain areas being investigated, and more critically, our currently limited understanding of the specific functions for most of these regions, it is possible to interpret nearly any pattern of results to be consistent with a priori hypotheses. If we believe that two processes are dissociated, we can point to two regions that are different between tasks to confirm our prediction. If we already believe that there is a single process, we can point to the similarities in brain activation between the tasks. Researchers might simply be reading tea leaves and finding evidence for what they already thought was true. Moreover, even in cases where two competing hypotheses are tested, the theories existed prior to running the fMRI study—the data did not lead to new theory development.

In the majority of social neuroscience studies, a two stage approach is used to link brain and cognition. We first manipulate social psychological variables during scanning and observe changes in the brain. Assuming that the study was run well with few or no confounds, this is the easy part. We now know that brain areas A, B, and C were relatively more recruited for task X than task Y. The second step, with its use of *reverse inference*, or the attribution of the recruitment of a psychological process or presence of a particular psychological state following the observation of a pattern of responses in a dependent measure (see Cacioppo & Tassinary, 1990; Poldrack, 2006), is the more controversial step. After cataloging all the differences, we go to the literature and use reverse inference in an attempt to determine what psychological states these brain areas represent. Finally, we work to spin a coherent story that allows us to make new theoretical claims about our construct of interest. For example, if we find that the amygdala responds to pictures of fathers more so than mothers in children with conduct disorder, we can look to the large literature on fear conditioning (which always activates the amygdala), and make the inference that these children are more afraid of their fathers than their mothers (despite what they may have said on their self-report measures!). Depending on our theoretical perspective, this interpretation can lead to new interventions that can aid the suffering of both these children and their families.

Regardless of whether current data has yet provided information that will point to a new theoretical perspective, theoretical advances are limited by how much is known about the brain and particularly the accuracy surrounding the specific inferences that can be drawn about the functions and processes of activated brain regions. For instance, if the regions identified by Greene et al. (2001) and Cunningham et al. (2004) are not uniquely involved in emotional processing or negative affect, respectively, have the studies contributed back to social psychological theory as much as previously thought? For this reason, understanding the neuroscience question of what the brain does and how it does it is necessary for any credible application to social psychological theory. Thus, there are not two social neuroscience questions, one relevant to neuroscience and another relevant to social psychology, but one question: How does the brain/mind process information?

The issue that needs to be considered is that our *reverse inference* is only as good as the original inference that established the agreed upon psychological process that the brain region is supposed to be involved in.¹ As Cacioppo et al. (2003) noted, there is rarely a 1:1 mapping between a particular psychological state (Ψ) and an accompanying physiological state (Φ). For example, it is true that research using fMRI in humans has shown that the amygdala is involved in the detection of threat in all stimulus modalities, including the perception of visual facial expressions of fear (Adolphs, Tranel, & Damasio, 1998; Morris et al., 1996), cognitive representations of fear (Phelps et al., 2001), threat-related words (Isenberg et al., 1999), and aversive odors (Zald & Pardo, 1997). Thus, if one is processing something threatening, there is a high probability that the amygdala will be active. To make the opposite claim—that amygdala activation means a participant is feeling fear—is more problematic. Several studies have now shown that the amygdala is not only sensitive to fearful or negative information, but also positive information (Hamann, Ely, Hoffman, & Kilts, 2002; Hamann & Mao, 2002; Garavan, Pendergrass, Ross, Stein, & Risinger, 2001; Liberzon, Phan, Decker, & Taylor, 2003). Further, studies independently manipulating valence and intensity (Anderson et al., 2003; Small et al., 2003), or using statistical methods to separate the contributions of the two (Cunningham, Raye, & Johnson, 2004), have demonstrated that amygdala activity appears to be more associated with processing affective intensity than with processing any particular valence and may actually be modified to respond differentially to positive or negative information depending on the goals of the perceiver (Cunningham, Raye, & Johnson, 2005; Cunningham, Van Bavel, & Johnsen, 2008).²

Some conclusions can be drawn from this line of work. Specifically, a more refined understanding of the operating characteristics of each brain region is necessary. With the example above, it is likely that there is some process that is accomplished by the amygdala that is necessary for fear conditioning, but that is also useful for making general negativity judgments, processing positive rewards, and perhaps even tracking motivationally significant stimuli. Thus, rather than considering this region “the fear center,” it is necessary to first better understand what the region does, when it is used, and specifically, what types of inputs, transformations, and outputs are associated with amygdala activity. In others words, what

1. Yet, it should be noted that these conceptual challenges are not unique to neuroscience research. Each time that we use the Rosenberg Self-Esteem Scale (Rosenberg, 1965), for example, we make the inference that our observed responses reflect self-esteem (and thus, our manipulation affects self-esteem, not pencil placement). A critical assumption here is that a stable construct of self-esteem exists, and that it is measurable with a self-report measure. Evidence for self-esteem is bolstered when the scale works the way that the researcher expects, and when divergent findings emerge, one needs to change the theory of the construct or of the operationalization of the measure. Importantly, when this happens, it is often not clear whether the theory or the measure needs to be modified. That is, just as our understanding of brain data is incomplete, so too are most measures of psychological constructs.

2. It is important to note that not all reverse inference is bad. First, once we understand what a region does (and what it does not do), the process of reverse inference will provide powerful insights into psychological states. When we see a region active, we will know why. Secondly, reverse inference has provided counterintuitive findings. By finding similar regions of the brain active to social and physical pain, Eisenberger, Lieberman, and Williams (2003) made the claim that social pain uses the same neural architecture as physical pain. In subsequent years, this inference has been validated in genetic studies and following drug administrations (Way, Taylor, & Eisenberger, 2009; DeWall et al., 2010).

this line of research indicates is that we need to develop a computational model of brain function before we can make strong inferences about how activation in this region can inform social psychological theory. More generally, the steps associated with our initial inference are more critical for theory development than the later steps of reverse inference. For this reason, social neuroscience research will only be useful to social psychology to the extent that the inferences made reflect our best knowledge of the computational properties of the brain regions.

REFERENCES

- Adolphs, R., Tranel, D., & Damasio, A. R. (1998). The human amygdala in social judgment. *Nature*, 393, 470-474.
- Anderson, A. K., Christoff, K., Stappen, I., Panitz, D., Ghahremani, D. G., Glover, G., et al. (2003). Dissociated neural representations of intensity and valence in human olfaction. *Nature Neuroscience*, 6, 196-202.
- Cacioppo, J. T., & Berntson, G. G. (1992). Social psychological contributions to the decade of the brain: Doctrine of multi-level analysis. *American Psychologist*, 47, 1019-1028.
- 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., & Tassinary, L. G. (1990). Inferring psychological significance from physiological signals. *American Psychologist*, 45, 16-28.
- Cunningham, W. A., Johnson, M. K., Raye, C. L., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2004). Separable neural components in the processing of Black and White Faces. *Psychological Science*, 15, 806-813.
- Cunningham, W. A., Raye, C. L., & Johnson, M. K. (2004). Implicit and explicit evaluation: fMRI correlates of valence, emotional intensity, and control in the processing of attitudes. *Journal of Cognitive Neuroscience*, 16, 1717-1729.
- Cunningham, W. A., Raye, C. L., & Johnson, M. K. (2005). Neural correlates of evaluation associated with promotion and prevention regulatory focus. *Cognitive, Affective, & Behavioral Neuroscience*, 5, 202-211.
- Cunningham, W. A., Van Bavel, J. J., & Johnsen, I. R. (2008). Affective flexibility: Evaluative processing goals shape amygdala activity. *Psychological Science*, 19, 152-160.
- Cunningham, W. A., & Zelazo, P. D. (2007). Attitudes and evaluations: A social cognitive neuroscience perspective. *TRENDS in Cognitive Sciences*, 11, 97-104.
- DeWall, C. N., MacDonald, G., Webster, G. D., Masten, C., Baumeister, R. F., Powell, C., et al. (2010). Acetaminophen reduces social pain: Behavioral and neural evidence. *Psychological Science*, 21, 931-937.
- Eisenberger, N. I., Lieberman, M. D., & Williams, K. D. (2003). Does rejection hurt? An fMRI study of social exclusion. *Science*, 302, 290-292.
- Garavan, H., Pendergrass, J. C., Ross, T. J., Stein, E. A., & Risinger, R. C. (2001). Amygdala response to both positively and negatively valenced stimuli. *NeuroReport*, 12, 2779-2783.
- Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, 293, 2105-2108.
- Hamann, S. B., Ely, T. D., Hoffman, J. M., & Kilts, C. D. (2002). Ecstasy and agony: Activation of the human amygdala in positive and negative emotion. *Psychological Science*, 13, 135-141.
- Hamann, S., & Mao, H. (2002). Positive and negative emotional verbal stimuli elicit activity in the left amygdala. *NeuroReport*, 13, 15-19.

- Isenberg, N., Silbersweig, D., Engelen, A., Emmerich, S., Malavade, K., Beattie, B., et al. (1999). Linguistic threat activates the human amygdala. *Proceedings of the National Academy of Sciences, USA*, 96, 10456–10459.
- LeDoux, J. E. (1996). *The emotional brain*. New York: Simon & Schuster.
- Liberzon, I., Phan, K. L., Decker, L. R., & Taylor, S. F. (2003). Extended amygdala and emotional salience: A PET activation study of positive and negative affect. *Neuropsychopharmacology*, 28, 726–733.
- Mitchell, J. P., Banaji, M. R., & Macrae, C. N. (2005). General and specific contributions of the medial prefrontal cortex to knowledge about mental states. *NeuroImage*, 28, 757–762.
- Morris, J. S., Frith, C. D., Perrett, D. I., Rowland, D., Young, A. W., Calder, A. J., & Dolan, R. J. (1996). A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature*, 383, 812–815.
- Ochsner, K. N., & Gross, J. J. (2008). Cognitive emotion regulation: Insights from social cognitive and affective neuroscience. *Currents Directions in Psychological Science*, 17, 153–158.
- Ochsner, K. N. & Lieberman, M. D. (2001). The emergence of social cognitive neuroscience. *American Psychologist*, 56, 717–734.
- Phelps, E.A. (2006) Emotion and cognition: Insights from studies of the human amygdala. *Annual Review of Psychology*, 24, 27–53.
- Phelps, E. A., O'Connor, K. J., Gatenby, J. C., Grillon, C., Gore, J. C., & Davis, M. (2001). Activation of the human amygdala to a cognitive representation of fear. *Nature Neuroscience*, 4, 437–441.
- Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *TRENDS in Cognitive Sciences*, 10, 59–63.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Saxe, R., & Wexler, A. (2005). Making sense of another mind- the role of the right temporo-parietal junction. *Neuropsychologia*, 43, 1391–1399.
- Small, D. M., Gregory, M. D., Mak, Y. E., Gitelman, D., Mesulam, M. M., & Parrish, T. (2003). Dissociation of neural representation of intensity and affective valuation in human gustation. *Neuron*, 39, 701–711.
- Way, B.M., Taylor, S.E., & Eisenberger, N.I. (2009). Variation in the mu-opioid receptor gene (OPRM1) is associated with dispositional and neural sensitivity to social rejection. *Proceedings of the National Academy of Sciences*, 106, 15079–15084.
- Zald, D. H., & Pardo, J. V. (1997). Emotion, olfaction, and the human amygdala: Amygdala activation during aversive olfactory stimulation. *Proceedings of the National Academy of Sciences, USA*, 94, 4119–4124.