University of Pennsylvania
ScholarlyCommons

Departmental Papers (Psychology)

Department of Psychology

2017

Cognitive Neuroscience Methods: An Introductory Overview for Social Scientists

Trishala Parthasarathi
University of Pennsylvania

Joseph W. Kable
University of Pennsylvania, kable@psych.upenn.edu

Follow this and additional works at: https://repository.upenn.edu/psychology_papers

Part of the [Cognitive Neuroscience Commons](#), and the [Psychology Commons](#)

Recommended Citation

Parthasarathi, T., & Kable, J. W. (2017). Cognitive Neuroscience Methods: An Introductory Overview for Social Scientists. *Thinking About Bribery: Neuroscience, Moral Cognition and the Psychology of Bribery*, 33-66. <http://dx.doi.org/10.1017/9781316450765.002>

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/psychology_papers/32
For more information, please contact repository@pobox.upenn.edu.

Cognitive Neuroscience Methods: An Introductory Overview for Social Scientists

Abstract

A challenge to doing interdisciplinary research is developing a solid understanding of the methods used by different fields. This challenge is particularly acute when the methods may be technologically complex. This chapter provides an introductory overview of the methods of cognitive neuroscience. This overview is intended for the people who are new to the field and who have very little, if any, knowledge of cognitive neuroscience techniques. Though the primary goal is to provide social scientists who are interested in bribery the necessary foundation to understand subsequent chapters that use these techniques, a broader aim is to also spark ideas about how future research could use neuroscience to better inform our understanding of bribery.

Disciplines

Cognitive Neuroscience | Psychology

Cognitive Neuroscience Methods: An Introductory Overview for Social Scientists

TRISHALA PARTHASARATHI AND JOSEPH W. KABLE

A challenge to doing interdisciplinary research is developing a solid understanding of the methods used by different fields. This challenge is particularly acute when the methods may be technologically complex. This chapter provides an introductory overview of the methods of cognitive neuroscience. This overview is intended for people who are new to the field and who have very little, if any, knowledge of cognitive neuroscience techniques. Though the primary goal is to provide social scientists who are interested in bribery the necessary foundation to understand subsequent chapters that use these techniques, a broader aim is to also spark ideas about how future research could use neuroscience to better inform our understanding of bribery.

Given these goals, this chapter is meant as a starting point for beginning students and researchers, not as the last word. No one method is covered in great detail; rather, only the basics of each method are discussed, and interested readers are directed to key tutorial references for each technique to get more detailed information. In addition to these references, beginning researchers will also want to consult one of several good methodology textbooks (e.g., Senior, Russell, & Gazzaniga 2006; Toga & Mazziotta 2002).

This chapter also focuses exclusively on neuroscience methods that can be used in human subjects. This omits a large swath of neuroscience methods that can only be used in nonhuman animals. While these more invasive techniques provide unparalleled ability to answer questions regarding basic neural mechanisms, they are unlikely to be of as much interest to scholars

Adapted with permission from: Kable, J. W. (2011). The cognitive neuroscience toolkit for the neuroeconomist: A functional overview. *Journal of Neuroscience, Psychology and Economics*, 4: 63–84.

studying bribery or, more broadly, to those whose focus is human social behavior.

This introductory overview organizes human neuroscience methods functionally, according to the kinds of inferences each method provides about brain-behavior relationships. This organization differs from the canonical presentation of cognitive neuroscience methods, which emphasizes the spatial and temporal resolution of each technique (Churchland & Sejnowski 1988). The motivation for the current framework arises from answering the question, "What different kinds of evidence should we gather to allow the strongest inferences about brain-behavior relationships?" Neuroscience methods differ according to the type of inference they provide concerning the relationship between brain and behavior. Broadly, methods can test the *association* between brain activity and behavior, the *necessity* of brain activity for behavior, or the *sufficiency* of brain activity for behavior. The most solid claims about brain-behavior relationships are based on converging evidence from these three kinds of tests. Research programs would, therefore, do best to employ a combination of methods that can address questions of association, necessity, and sufficiency.

This discussion takes it as given that the goal of some social scientists using neuroscience is to build theories about the relationship between the brain and cognition or behavior. Of course, some social scientists may be interested in neuroscience findings and techniques only inasmuch as these can inform theories of cognition or of human economic or social behavior. The question of how neuroscience research can inform theories in the social sciences is an important one, which has been extensively debated (Camerer et al. 2005; Glimcher 2010; Gul & Pesendorfer 2008; Harrison 2008; McCabe 2008). However, the methodological framework and methods overview below stress the role of each technique in elucidating brain-behavior relationships. The concluding section then returns to the question of how neuroscience methods, and the knowledge of brain-behavior relationships gained through them, can inform theories in the social sciences.

DIFFERENT KINDS OF INFERENCE

So, what different kinds of evidence combine to make the strongest inferences regarding the mappings between brain systems and cognitive functions or behaviors? To answer this question, we can examine the evidence underlying neuroscience claims that have achieved the status of broad

agreement. As an example, take the statement that "cortical area V5 subserves visual motion perception." (Early visual cortical regions are named V1, V2, V3, etc.) This claim is often taught in introductory lectures on the visual system, and V5 is often used as a paradigmatic example of specialization in cortical organization. What is the evidence that V5 is important for motion perception?

Converging evidence from many different kinds of studies point to V5's role in motion perception. Noninvasive brain imaging (specifically functional magnetic resonance imaging – fMRI; see below for details) has been used to map the location of different visual areas based on their responses to stimuli presented in different parts of the visual field. Compared to other visual areas, neural activity in V5 is selective to moving compared to stationary stimuli, and this activity is further sensitive to different directions of motion (Tootell et al. 1995). Other techniques have been used to disrupt neural activity in V5 (specifically transcranial magnetic stimulation – TMS; see below for details). Disrupting V5 impairs performance in motion-direction discrimination tasks and in visual search tasks that require paying attention to motion (Laycock et al. 2007; Walsh et al. 1998). Yet other techniques can enhance V5 activity (specifically transcranial direct current stimulation – tDCS; see below for details), and such enhancement increases the accuracy of tracking the motion of stimuli (Antal et al. 2004). Parallel experiments have recorded, disrupted, or enhanced V5 activity using more invasive methods in the model system of the Rhesus macaque (*Macaca mulatta*) monkey and observed similar results (for a detailed review, see Newsome 1997; Parker & Newsome 1998).

From these studies, we see the combined use of three broad kinds of evidence (Sarter, Berntson, & Cacioppo 1996; Silva 2007).¹ First, there are tests of *association*. These involve observing or experimentally manipulating psychological states or behavior, simultaneously measuring neural activity, and examining the correlation between the two.² In probabilistic terms,

¹ The framework used here borrows from Silva (2007) and Sarter, Berntson, and Cacioppo (1996). Silva's (2007) "Convergent Four" hypothesis argues that the strongest way to test a hypothesis concerning the connection between two phenomena is to gather converging evidence from theoretical work and three different kinds of experiments: *nonintervention* (here called *association*), *negative alteration* (*necessity*), and *positive alteration* (*sufficiency*). In their comment on functional imaging techniques, Sarter, Berntson, and Cacioppo (1996) distinguish between the ability of neuroimaging to test $P(\phi | \psi)$ and the ability of other techniques to test $P(\psi | \phi)$.

² Despite the name, tests of association can in some cases provide causal evidence (Weber & Thompson-Schill, 2010), depending on the particulars of the experiment. For example, neuroimaging experiments typically control the stimuli presented to the subjects and measure the subject's responses and brain activity. At one end of the spectrum, when

tests of association focus on the probability of a neural state (φ) given a psychological state (ψ), $P(\varphi | \psi)$ (though the final section of the chapter discusses analyses that shift this focus to the probability of a psychological state given a neural state, $P(\psi | \varphi)$). An example of a test of association is the correlation between neural activity in V5 measured with fMRI and perception of motion by the subject.

Second, there are tests of *necessity*. These involve disrupting neural activity and showing that this manipulation impairs a specific behavioral or psychological function. In probabilistic terms, tests of necessity reduce the probability of a neural state and see how this changes the probability of a psychological state, $P(\Delta\psi | \downarrow\varphi)$.³ An example of a test of necessity is the impairment in motion-detection discrimination tasks caused by TMS to V5.

Third, there are tests of *sufficiency*. These involve enhancing neural activity and showing that this manipulation results in a specific behavioral or psychological state. In probabilistic terms, tests of sufficiency increase the probability of a neural state and see how this changes the probability of a psychological state, $P(\Delta\psi | \uparrow\varphi)$. An example of a test of sufficiency is improvement in performance induced by tDCS stimulation of V5.

Successful research programs in neuroscience employ a multimethod approach that combines tests of association, necessity, and sufficiency. This combination provides strong evidence for the causal influence of neural states on psychological states. In tests of association, neural activity is *measured*, while in tests of necessity and sufficiency, neural activity is *manipulated*. The first kind of test establishes an important precondition for causality, that occurrences of the respective neural and psychological states are correlated. The second and third tests then test causality directly, by establishing whether the neural states are necessary and sufficient conditions for the psychological states.

examining the relationship between stimuli and brain activity, causal inference is clearly warranted. At the other, when examining the neural activity linked to different behavioral responses to the same stimuli, causal inference is clearly not warranted. Note that the causal inference provided is only ever about how changes in the environment or psychological states cause changes in brain activity, while tests of necessity and sufficiency are interested in the opposite direction of causality.

³ Technically, the probabilistic framework is not concerned with causality, and it does not matter whether the change in neural state is due to an experimental manipulation or is simply observed. The only distinction is between $P(\varphi | \psi)$ and $P(\psi | \varphi)$, so $P(\psi | \varphi)$, $P(\Delta\psi | \downarrow\varphi)$, and $P(\Delta\psi | \uparrow\varphi)$ are formally the same. The important point to emphasize is that tests of association are often of $P(\varphi | \psi)$, while tests of necessity and sufficiency are of $P(\psi | \varphi)$.

WHAT NEUROSCIENCE METHODS PROVIDE WHICH KINDS OF INFERENCE?

This methodological framework – distinguishing between tests of association, necessity, and sufficiency – provides a way to functionally organize the methods of human neuroscience (see Tables 2.1 and 2.2). Table 2.1 uses this framework to categorize methods that link regional neural activity to mental function. Table 2.2 similarly categorizes methods that link neurochemical systems to mental function. The following sections provide a brief description of each technique listed in Tables 2.1 and 2.2, organized by the categories used in these tables. For each technique, tutorial references dedicated to that technique are listed, which the interested reader should consult for additional information. For each technique, a recent study demonstrating its use is also described. As bribery per se has yet to be widely studied with neuroscience methods, these examples draw on the wider but still highly relevant literature on the neuroscience of social bargaining and moral judgments.

Methods for Linking Regional Neural Activity to Mental Function

Tests of Association Between Regional Neural Activity and Mental Function

Functional Magnetic Resonance Imaging (fMRI). fMRI is the most widely used method in cognitive neuroscience (for more information on fMRI, see Bandettini 2009a, 2009b; Huettel, Song, & McCarthy 2004; Jezzard, Matthews, & Smith 2001). Like all MRI, fMRI involves measurements made with an MRI scanner. MRI scanners have a large static magnetic field, typically with a field strength of 3 Tesla, which is 60,000 times that of the earth's gravity. Signals based on the properties of nuclear magnetic resonance are then measured by making small, local modulations in the strength of this field. In fMRI, the dominant source of these signals comes from protons in water molecules. Different kinds of MR images have different *contrasts* – that is, they are sensitive to different properties of the measured tissue. The kind of contrast used in most fMRI is referred to as T2* contrast (“T-2-star,” so named in comparison to T1 or T2 contrast, both of which are used in anatomical imaging). T2* is sensitive to, among other things, the oxygenation level of the blood. This is because deoxygenated hemoglobin has magnetic properties and, therefore, causes local susceptibility effects that decrease T2* contrast. Oxygenated hemoglobin does not

TABLE 2.1 *Functional overview of human neuroscience methods, Part 1: Methods that link regional neural activity to mental processes*

	Method linking regional neural activity to mental function	Example(s) using this method
Tests of association	fMRI	Sanfey et al. (2003) demonstrated an association between brain activity and fairness during the Ultimatum Game.
	PET (^{15}O water, etc.)	de Quervain et al. (2004) showed that increased dorsal striatal activity was associated with the willingness to incur costs to punish defectors in an economic exchange.
	EEG, MEG	Sarlo et al. (2012) demonstrated that instrumental and incidental moral dilemmas had differential effects on ERP responses.
	NIRS	Strait et al. (2013) showed that hemodynamic activity in anterior prefrontal cortex was greater when interacting with human agents.
	Anatomical imaging	Prehn et al. (2015) found increased gray matter volume in the vmPFC in individuals who showed postconventional moral reasoning.
Tests of necessity	Lesion studies	Koenigs et al. (2007) and Ciaramelli et al. (2007) observed that individuals with vmPFC damage made more utilitarian judgments in moral dilemmas.
	TMS	Knoch et al. (2006) found that disruption of right dorsolateral prefrontal cortex resulted in increased acceptance of unfair offers in the Ultimatum Game.
	tDCS (cathodal)	Knoch et al. (2008) demonstrated that disruption of right lateral prefrontal cortex led to increased acceptance of unfair offers in the Ultimatum Game.

	Method linking regional neural activity to mental function	Example(s) using this method
Tests of sufficiency	tDCS (anodal)	Ruff et al. (2013) showed that stimulation of lateral prefrontal cortex increased monetary offers from the proposer in the Ultimatum Game when the responder could punish the proposer compared to when no punishment took place.

TABLE 2.2 *Functional overview of human neuroscience methods, Part 2: Methods that link neurochemical systems to mental processes*

	Method linking regional neural activity to mental function	Example(s) using this method
Tests of association	PET (^{11}C raclopride, etc.)	Buckholz et al. (2010) found that differences in dopamine release in the striatum is associated with trait impulsivity.
	Genetics	Zhong et al. (2010) observed a relationship between fairness and the 4-allele repeat of the DRD4 gene.
Tests of necessity	Pharmacology (antagonists)/neurotransmitter depletion	Crockett et al. (2008) found that serotonin depletion increased rejection of unfair offers in the Ultimatum Game.
Tests of sufficiency	Pharmacology (agonists)/neurotransmitter loading	Sáez et al. (2015) showed that administration of a dopamine agonist results in greater egalitarian tendencies in the Dictator Game.

have magnetic properties. Since local increases in neural activity are associated with increases in cerebral blood flow, and these increases in blood flow are associated with increases in blood oxygenation, local increases in neural activity are associated with increases in the fMRI signal. This fMRI signal is commonly referred to as the blood oxygenation level-dependent (BOLD) signal.

BOLD fMRI has several advantages that make it a preferred technique in human neuroscience. Although many methods measure correlates of neural activity noninvasively in humans, fMRI provides the best combination of spatiotemporal resolution and anatomical coverage. fMRI provides a simultaneous measurement of neural activity across almost the entire brain, including subcortical regions (the exceptions being a few regions hampered by technical signal dropout artifacts, such as orbitofrontal cortex and the temporal pole; Ojemann et al. 1997; Weiskopf et al. 2006). This whole-brain coverage is achieved at a high spatial resolution (typically 3 mm) and at a good temporal resolution (typically a measurement every 2 s). On the practical side, many research institutions have MRI scanners and the technical expertise necessary to run them, and several good software options for analyzing fMRI data are freely available, which lowers the barriers to starting an fMRI research project.

The basic analysis of fMRI data involves comparing BOLD activity under two different conditions, with this statistical comparison usually being performed independently at each spatial location. By designing the right set of comparisons between stimuli or tasks, one can ask questions about where regional brain activity is associated with different mental functions.

Because fMRI has good spatiotemporal resolution and the possibility for many repeated trials, more sophisticated experimental designs and analyses are also possible. Parametric designs look for BOLD activity that changes with parametric variations in stimuli or task variables, rather than categorical changes, which can strengthen inferences regarding the causes of neural activity (Aguirre & D'Esposito 1999; Buchel et al. 1996). Model-based fMRI (O'Doherty, Hampton, & Kim 2007) builds on parametric designs, by testing for activity that is correlated with parametrically varying hidden variables that are estimated in a statistical model of the subject's task behavior. Both parametric and model-based fMRI designs identify not just *where* neural activity changes but also *how* activity in these regions changes with experimental variables. Adaptation and repetition suppression designs (Grill-Spector & Malach 2001; Henson & Rugg 2003; Weigelt, Muckli, & Kohler 2008) capitalize on the finding that repeated stimuli are associated with reduced BOLD activity, using this effect to ask what stimuli are treated as the same in a given neural population. Multivariate techniques such as

pattern classification (Haynes & Rees 2006; Norman et al. 2006) ask how the spatial pattern of neural activity changes under different conditions. By testing what stimuli a region can and cannot distinguish between, both adaptation and pattern classification techniques show how information is encoded in a region. Finally, other statistical techniques have been used to test for interactions ("functional" or "effective" connectivity) between different brain regions, using the correlations or coherences between time series, and how those interactions change with experimental manipulations (Friston 1994; Friston et al. 1997; McIntosh et al. 1996; Sun, Miller, & D'Esposito 2004). All of these designs and analyses go beyond localization, or a simple mapping of the regional brain activity associated with specific stimuli or tasks, to ask more detailed and textured questions about how brain activity is related to mental activity.

fMRI does have both theoretical and practical limitations. Perhaps the most important theoretical limitation is the source of the fMRI signal. The BOLD signal reflects mass activity across the more than one million neurons in a voxel and may be more closely linked to changes in synaptic conductances than neuronal spiking (Logothetis 2008; Logothetis et al. 2001). This means that the signal cannot easily distinguish bottom-up versus top-down signals or neuromodulatory versus other inputs (Logothetis 2008). The BOLD signal does not provide information about neural dynamics that occur at a fast time scale and may be insensitive to other neural dynamics as well. Practically, the requirement of being in an MRI scanner constrains the kind of experiments that are possible. For example, experiments that require any head movement or large body movements are not possible. There are also restrictions on the kinds of subjects that can be tested. Subjects who are contraindicated for MRI (e.g., have metal in their body) or those who cannot lie motionless in a relatively small space (e.g., young children, claustrophobics) cannot be tested. fMRI is also more expensive than behavioral testing or several of the other techniques discussed below.

One of the earliest fMRI studies of decision making provides a good example of how fMRI can be used. Sanfey et al. (2003) measured BOLD activity while participants played the Ultimatum Game. In the Ultimatum Game, participants play the role of either the "proposer" or "responder." The proposer offers to split a monetary sum in a certain way, and the responder chooses whether to accept or reject the offer. If the responder accepts, the split is made accordingly. If the responder rejects, however, neither subject receives any money. Though there are cultural differences (Henrich et al. 2006), responders typically accept offers when the split is close to 50–50 and tend to reject offers when the division is unequal. When Sanfey et al.

(2003) scanned participants in the responder role, they found increased activation in bilateral anterior insula, anterior cingulate cortex, and the dorsolateral prefrontal cortex for unfair offers compared to fair offers. This study demonstrates the involvement of these specific brain regions in recognizing and responding to unfairness and provided a foundation for several studies discussed below to further delineate the functional role these regions play in that response.

A further study by Hollmann et al. (2011) demonstrates the use of more advanced analytic techniques. These authors used pattern classification to predict whether a responder accepted or rejected an unfair offer in the Ultimatum Game. They found that spatial pattern of activity in anterior insula and lateral prefrontal cortex, two of the regions that show increased activity to unfair offers, could successfully distinguish between acceptance and rejection.

Positron Emission Tomography (PET). PET predates fMRI as a technique for noninvasively mapping changes in cerebral blood flow (for more information on PET, see Raichle 1998). In PET, a radioactive tracer is injected into the blood stream. For measurements of cerebral blood flow, [^{15}O] water is used. The decay of this tracer in the brain is then measured. This radioactive decay creates positrons, which are then quickly annihilated when they react with nearby electrons, thereby creating two photons that leave the point of annihilation in exactly opposite directions. PET scanners are designed to detect the coincident arrival of these photon pairs. The location of their origin, and thus the point of radioactive decay, can then be estimated from these measurements. In the case of [^{15}O] water, differences in the amount of radioactive decay across regions indicate differences in cerebral blood flow.

With the development of fMRI, the use of PET in cognitive neuroscience has waned. fMRI offers several advantages over PET. fMRI does not involve radioactivity, which restricts the number of measurements you can make in a given subject. fMRI also offers better spatial (3 mm vs. 1 cm) and temporal (2 s vs. 1 min) resolution. A typical PET study involves one to two measurements per subject per condition, with each measurement reflecting neural activity averaged over 1 minute. This restriction limits the kind of designs that can be used in PET compared to fMRI; for example, “event-related” designs are not possible with PET. In addition, fMRI is less expensive and more widely available than PET.

There is still a role for PET, however. In particular, PET can measure blood flow in absolute terms (mL/g/min), while BOLD fMRI only provides

a relative measure. Absolute measures of blood flow permit comparisons across subjects, across sessions, and across brain regions. PET can also measure quantities other than blood flow. For example, other tracers can be used to measure cerebral metabolism (Fox & Raichle 1986) or, as discussed below, different components of specific neurotransmitter systems (Zald et al. 2004).

A study by de Quervain et al. (2004) illustrates the use of PET. These authors used H_2^{15}O PET to study the neural mechanisms involved in punishing unfair behavior. Participants first experienced another player behaving unfairly toward them in an exchange game called the Trust Game. Then they were scanned while considering whether to punish this other player, under conditions where punishment would be symbolic (no real consequences) or effective (real consequences) and when punishment accrued different costs to participant. They found greater activity in the dorsal striatum, a brain area implicated in goal-directed behaviors, for effective compared to symbolic punishment. In addition, players with greater dorsal striatal activity were willing to undergo greater costs in order to carry out the punishment, suggesting a role for this region in administering “satisfactory” punishment for violating social norms (de Quervain et al. 2004).

Electroencephalography (EEG) and Magnetoencephalography (MEG).

Aside from lesion studies, EEG is one of the oldest human neuroscience techniques (for more information on this technique, see Handy 2005; Luck 2005). In EEG, electrodes are placed on the scalp to measure electrical changes that result from brain activity. All neuronal signaling is electrochemical, and electrical signals within neurons are accompanied by simultaneous extracellular changes. These electrical signals can propagate through the brain and skull and be detected on the scalp. The size of these signals at the scalp is very small, 10–100 μV , an order of magnitude smaller than at their source.

Classically in EEG studies, the time-locked responses to many repetitions of a stimulus are averaged. The spatial and temporal profile of this average response, called the event-related potential – ERP, is then investigated. More recent statistical techniques have focused on decomposing the ERP so that varying responses on single trials can be investigated (Debener et al. 2006). In addition to the ERP, EEG can also be used to measure changes in oscillatory activity that accompany task performance.

MEG (for more information on MEG, see Lounasmaa et al. 1996) measures fluctuations in the magnetic field at the scalp that are caused by changes in neural activity. The source of this signal is similar to that of EEG

signals. Neuromagnetic fields have a strength of 50–500 fT, which is many orders of magnitude smaller than the Earth's magnetic field (0.5 mT) or fluctuations due to noise in the environment (1 nT–1 μ T). Detecting neuromagnetic signals, therefore, requires very sensitive detectors called SQUIDS (superconducting quantum interference devices), as well as sophisticated noise cancellation techniques. In a typical MEG setup, a fixed helmet containing more than 100 detectors is placed on top of the subject's head. A clear advantage of both EEG and MEG is that they provide great temporal resolution (<1 ms). This allows investigations of the timing and evolution of neural activity at the timescale at which cognition occurs. It also permits studies of neural signals that occur at a very fast time scale, such as synchronized or oscillatory activity. EEG is also less expensive and more widely available than fMRI. (MEG equipment is much more expensive than that used for EEG.)

An important disadvantage of both EEG and MEG is that these signals cannot be spatially localized in the manner that fMRI or PET can. This is because of the “inverse problem” (Michel et al. 2004). EEG and MEG measure signals at the scalp, and the question of what configuration of brain sources could have led to the observed data is an ill-posed problem – an infinite number of different configurations are possible in principle. Thus, calculating an inverse solution is only possible by making certain assumptions, and the validity of this inverse solution is complicated to test. Solutions to the inverse problem in MEG are somewhat more reliable and accurate, because the magnetic permeability of the head is more uniform than its electrical conductivity, which greatly simplifies the calculation of the MEG signal pattern produced by a given neural current source (Lounasmaa et al. 1996). Recent efforts to improve solutions to the inverse problem in EEG have focused on developing new statistical techniques for calculating the inverse solution, developing new hardware such as dense electrode nets that help further constrain this solution, and conditionalizing the solution on other data such as fMRI activations (Michel et al. 2004).

EEG has been used to study the temporal dynamics of moral decision making. Sarlo et al. (2012) posed subjects with two types of moral dilemmas – “instrumental” dilemmas involved killing someone intentionally to save others, while “incidental” dilemmas involved killing someone as an unintended consequence of saving others. ERPs were time-locked to the onset of the decision, and movement-related potentials were time-locked to the onset of the behavioral response. Instrumental dilemmas were associated with reduced utilitarian choices and increased negative affect, and a larger early ERP component (around 250 ms) was found over frontal

and frontopolar regions for instrumental dilemmas compared to incidental ones. In contrast, incidental dilemmas were associated with larger components of the movement potentials at a later point in time, suggesting facilitated action selection for these choices. This study illustrates how EEG can be used to look at distinct temporal stages of a decision.

Near-Infrared Spectroscopy (NIRS). NIRS (for more information on NIRS, see Gratton & Fabiani 2001; Hoshi 2005) is another noninvasive imaging method that can be used in humans. NIRS uses a light source and detector to measure the absorption and scatter of near-infrared light by the skull and underlying tissue. In its most widely used form, NIRS detects changes in absorption due to changes in oxygenated and deoxygenated hemoglobin and, therefore, is sensitive to the same kind of hemodynamic signals measured by BOLD fMRI. There are also reports that NIRS can be sensitive to faster signals arising because of changes in scatter associated with neuronal firing (Gratton & Fabiani 2001), but these signals are much weaker than the hemodynamic signals that are typically assessed (Obrig & Villringer 2003).

Despite measuring a similar signal, NIRS provides much poorer spatial localization than fMRI and is restricted to measuring the dorsal cortical surface directly underneath the skull. Recent technical developments are aimed at improving spatial localization, by using different types of measurement (time- or frequency-resolved in addition to continuous wave) or by using imaging techniques that combine multiple sources and detectors (Hoshi 2005; Obrig & Villringer 2003). NIRS does have some advantages relative to fMRI, including much higher temporal resolution, as well as specific measurements of both oxygenated and deoxygenated hemoglobin. One of the biggest advantages of NIRS is that the measurement devices are portable, unobtrusive, and less costly. NIRS can, therefore, be used in populations that cannot easily be scanned, such as infants and young children, claustrophobics and others contraindicated for fMRI, and specific patient populations. It is also possible to construct wearable NIRS systems that allow measurements while subjects move around (Hoshi 2005), allowing experimenters to conduct experiments in more realistic settings.

A recent study by Strait et al. (2013) illustrates the use of NIRS. They measured activity in anterior prefrontal cortex with NIRS while participants considered moral dilemmas and found significant increases in activation when making decisions about other human agents compared to robots or canines. In future work, NIRS could allow investigators to examine this

agency effect in social interactions that are more naturalistic than those that could be contrived for an MRI scanner.

Anatomical Imaging. The above methods all test for dynamic changes in regional neural activity associated with the engagement of particular mental functions. There are also techniques for testing of static differences in brain anatomy associated with differences in mental function. These techniques test for associations across (rather than within) subjects, using an individual differences approach. The relevant individual differences might be in the structure (e.g., size, cortical thickness, or gray/white matter composition) of brain regions as measured with anatomical MRI (Ashburner & Friston 2000) or in the structure of white matter tracts as measured with diffusion-tensor imaging (Johansen-Berg & Rushworth 2009; Mori & Zhang 2006).

For example, a study by Prehn et al. (2015) used voxel-based morphometry to assess structural brain differences between individuals with different levels of moral reasoning. Participants took the Defining Issues Test (DIT-2; Rest et al. 1999) to determine the type of schema they used while evaluating moral dilemmas. Based on Kohlberg's theory of the development of moral reasoning, there are three types of schemas: preconventional, which only involves thinking about the individual; conventional, which follows conventional rules and norms; and postconventional, which involves thinking about how these rules can impact society as a whole. The study found that participants who use postconventional reasoning when presented with moral dilemmas had greater gray matter volume in the ventromedial prefrontal cortex and subgenual anterior cingulate cortex compared to those who only use preconventional reasoning (Prehn et al. 2015).

Invasive Recordings. In addition to the above noninvasive techniques, there are also rare opportunities to measure neural activity more directly in the human brain (for more information on invasive recordings, see Engel et al. 2005). These opportunities are only available in specific patient populations where invasive neural recordings are already being performed as part of a clinical treatment. Examples include patients with Parkinson's disease undergoing surgery to implant a deep brain stimulator, and patients with epilepsy undergoing invasive monitoring and surgery to detect and remove epileptic tissue. In such patients, neural recordings might occur acutely during surgery or more chronically (over days) as part of monitoring. Obviously, since any research study is secondary to the patient's treatment, these studies have many practical limitations. Also, by definition, the

subjects in these studies all have some kind of brain pathology, which could limit generalizability of the data. Nevertheless, these studies provide a rare source of data from the human brain that is comparable to that obtained in nonhuman animal models.

Multiple neurophysiological methods are used in these studies. During surgery, recordings are usually performed with a single (or small number of) electrode(s). Presurgical monitoring can be performed with grids of multiple large electrodes (3–4 mm) placed on the cortical surface, a technique referred to as electrocorticography (Miller et al. 2007), or with depth electrodes that penetrate deep into the brain, a technique referred to as stereotactic-EEG (Lachaux, Rudrauf, & Kahane 2003). Both these techniques can record oscillatory activity, at a greatly increased spatial precision and signal-to-noise compared to noninvasive EEG/ERP. Depth electrodes can also record multiunit activity and can be adapted to include microwire extensions for recording single-unit activity or probes for measuring neurotransmitter concentrations.

Tests of the Necessity and Sufficiency of Brain Regions for Mental Function

Lesion Studies. Perhaps the oldest technique in human neuroscience is the neuropsychological approach, which examines how damage to the brain results in the breakdown of mental function (for more information on lesion studies, see Kolb & Whishaw 2009; Rorden & Karnath 2004; Shallice 1988). This was the method of inquiry in the classic studies of Paul Broca and Carl Wernicke, and reports of striking dissociations after brain damage date back to the ancients (Benton 1991).

Modern experiments might study brain damage arising from many different sources. Stroke (ischemic or hemorrhagic) is most often studied, since it results in a relatively focal lesion. Other sources of focal lesions include neurosurgery (i.e., tumor removal, temporal lobe resection, corpus callosotomy) and penetrating head injury (Phineas Gage being a prominent historical example). Anoxia can cause more diffuse but still localizable damage. Degenerative diseases that cause more diffuse damage, such as frontotemporal dementia or Alzheimer's disease, have also been studied (Grossman et al. 2010). Investigators have studied Parkinson's disease to examine the effects of degeneration of the dopaminergic system (Rutledge et al. 2009; Shohamy et al. 2005).

Often, the neuropsychological approach has not been used for brain mapping at all, but rather as a tool for dissecting the structure of the mind.

In the tradition of cognitive neuropsychology (Shallice 1988), for example, the focus is on how a case illuminates the organization of mental functions, rather than where the brain is damaged. Central to cognitive neuropsychology is the search for dissociations. A single dissociation occurs when an individual with brain damage exhibits a deficit in one task while performing at normal levels in all others. Such dissociations are informative, but alone could be explained by tasks being differentially sensitive to brain damage. A double dissociation occurs when one individual exhibits a deficit on task A but not task B, while a second individual exhibits the opposite – a deficit on task B but not task A. Double dissociations cannot be explained by differential task sensitivity and, therefore, suggest that the two tasks depend on partially isolable mental processes.

Of course, the neuropsychological approach can also be used to address the question of what brain structures are necessary for a particular mental function. Especially with the advent of modern imaging techniques, neuropsychological studies have increasingly turned in this direction. One approach is to perform group studies, where subjects are grouped according to the location of their lesion, and the performance of different groups is compared (e.g., Bechara et al. 1997; Fellows & Farah 2007; Hsu et al. 2005). In these types of studies, it is important to have a brain-damaged control group in addition to an intact control group (Fellows 2012). An alternative is to group subjects by their performance and to compare the overlap of lesions in subjects with and without deficits (e.g., Adolphs et al. 2000; Tranel, Damasio, & Damasio 1997). A potential drawback to both these methods is that they involve a binary classification of the brain or behavioral data that overlooks any heterogeneity in lesion location or performance (i.e., a subject is either in or out of a particular lesion group, a subject either does or does not have a deficit). Methods such as voxel-lesion symptom mapping overcome these drawbacks by treating both the brain and behavioral data as continuous measures (Kimberg, Coslett, & Schwartz 2007; Rorden, Karnath, & Bonilha 2007).

The strength of lesion studies is the ability to perform tests of necessity in humans, to examine whether loss of a brain region leads to loss of function. One aspect of lesion studies that could limit inference is that the damage is not under experimental control, so it is possible that the behavior in question is related to the person's premorbid personality or medical condition rather than to the brain damage. In many cases, this is not a plausible interpretation, but there are some examples – for one, an experimenter observing a deficit after temporal lobe resection for epilepsy might worry whether the deficit was caused by the epilepsy rather than the resection. Such

concerns can be addressed by constructing the appropriate comparison groups to rule out potential confounds. Other issues that limit inference in lesion studies are: (1) cortical reorganization during recovery from stroke, so that intact regions could be compensating for damaged ones in chronic patients; (2) great variability across subjects, which can make it difficult to relate structure-to-function with precision; and (3) the often severe nature of the damage, which can make it difficult to examine the more subtle functions of a brain region. Finally, on the practical side, conducting large-scale lesion studies requires an infrastructure for the recruitment, retention, and testing of subjects that, although probably less expensive than neuroimaging, is less widely available (Fellows et al. 2008).

Investigators have used lesion studies to demonstrate the necessary role of ventromedial prefrontal cortex (vmPFC) in bargaining games and moral judgment. Individuals with lesions to vmPFC are more likely to reject unfair offers in the Ultimatum Game (Koenigs & Tranel 2007). In high-conflict moral dilemmas that pit the violation of a norm against maximizing welfare, individuals with lesions to vmPFC are more likely to make the utilitarian than the deontological choice, while their judgments are comparable to controls in low-conflict moral or nonmoral dilemmas (Ciaramelli et al. 2007; Koenigs et al. 2007). These single dissociations suggest that the vmPFC plays a causal role in certain kinds of decisions; to our knowledge, no one has yet established a double dissociation between the role of vmPFC and another brain region in social or moral decision making.

Transcranial Magnetic Stimulation (TMS). TMS (for more information on TMS, see Pascual-Leone, Walsh, & Rothwell 2000; Walsh & Cowey 2000) is based on Faraday's principle of electromagnetic induction. A rapidly fluxing magnetic field is induced by running an electrical current through the closed loop of the TMS coil. This magnetic field traverses the scalp and skull and induces an electrical current in the underlying neural tissue. A typical setup uses a figure eight-shaped coil that induces a magnetic field of ~2 T. The depth and width of the cortical area stimulated depends on the characteristics of the coil and its placement, but a rough estimate is that the cortical area stimulated is ~3 cm² at a depth of <2 cm. TMS, therefore, has a relatively focal spatial effect, but can only be used to affect the dorsal cortical surface (although specialized coils that can reach deeper tissues are being developed; Wagner, Valero-Cabre, & Pascual-Leone 2007). In a typical setup, a stereotaxic system is used to track the location of the subject's head and the TMS coil and, in conjunction with an anatomical scan of the subject's brain, to guide the precise placement of the TMS coil with respect to cortical landmarks.

The electric current induced in neural tissue by TMS is strong enough to cause neural firing. A single pulse of TMS, therefore, disrupts neural processing by inducing a burst of nonspecific neural firing, which is then terminated by cortical inhibition (Siebner et al. 2009). Since the time course of this effect is ~200 ms, single-pulse TMS has relatively good temporal specificity. Single-pulse TMS can, therefore, be used to test the temporal window within which disruption of cortical processing has an effect on behavior (Pascual-Leone et al. 2000; Walsh & Cowey 2000).

Repeated pulses of TMS (rTMS) can also be administered. Here it is critical that users remain within the established safety guidelines, since adverse effects of rTMS have been reported outside this range (Wassermann 1998). Short (<1 s), high-frequency bursts can be used if a temporal effect slightly longer than a single pulse is desired. Alternatively, longer trains (~10 min) of pulses can be administered. In this case, rTMS results in a prolonged (15 min–1 h) modulation of cortical excitability. Depending on the intensity and frequency of stimulation, this modulation can either enhance or suppress cortical activity, although most studies use low-frequency (<1 Hz) parameters that result in suppression of cortical activity.

An experiment that followed up on the results of Sanfey et al., discussed above, provides a nice example of the use of rTMS to provide insight about the causal role of a brain region. This study used rTMS to disrupt the right dorsolateral prefrontal cortex in responders playing the Ultimatum Game. This brain region is more active during unfair offers, but an open question was whether activity in this region promotes the acceptance of unfair offers (over concerns about fairness norms) or the rejection of unfair offers (over concerns about personal gain). They found that rTMS to dorsolateral prefrontal cortex did not alter the recognition of unfairness, but did increase the acceptance of unfair offers, showing that this region promotes the enforcement of fairness norms (Knoch et al. 2006).

Transcranial Direct Current Stimulation (tDCS). Like TMS, tDCS is a method for noninvasive brain stimulation (for more information on tDCS, see Been et al. 2007; Wagner et al. 2007). Typically, a constant current (<2 mA) is applied to the scalp through patch electrodes (25–35 cm² area) for a time period of up to 20 minutes. Depending on whether the anode or cathode is placed over the cortical region of interest, tDCS either increases (anodal tDCS) or decreases (cathodal tDCS) neuronal excitability in a region of the cortical surface beneath the electrode. These effects can persist for 15 minutes to 1 hour after stimulation. In cognitive studies, a behavior

task is performed either during or directly after tDCS, and the effects of stimulation on behavior are examined. Sham stimulation, where the current is ramped up and immediately back down rather than applied continuously, can be used as a control condition. The precise neuronal effects of TMS and tDCS likely differ, so they should not be considered interchangeable techniques. Nevertheless, it is useful to compare the benefits and drawbacks of the two methods. Similar to TMS, tDCS enables reversible activation and inactivation studies in humans and is restricted to studies of the dorsal cortical surface directly below the skull. In comparison to TMS, the equipment for tDCS is less expensive, more portable, and more easily scalable to study a group of subjects at the same time. In addition, no adverse effects of tDCS have been reported, beyond minor headaches or itching beneath the electrode, while rTMS has been associated with potential side effects that could affect the interpretation of behavioral results. TMS, however, permits greater spatial and temporal control of stimulation, since it can be applied in single pulses and directed to more localized cortical regions.

The same investigators who used TMS to delineate the function of dorsolateral prefrontal cortex in the Ultimatum Game further extended these results with tDCS. Knoch et al. (2008) administered tDCS over the right prefrontal cortex of participants who acted as the responder in the Ultimatum Game. Similar to the results in the TMS experiment discussed above, they observed that tDCS to right prefrontal cortex resulted in increased acceptance of unfair offers, further verifying the role of this region in enforcing fairness norms (Knoch et al. 2008).

Methods for Linking Neurochemical Systems to Mental Function

Tests of Association Between Neurochemical Systems and Mental Function

PET. As discussed above, PET can be used to obtain measures of regional cerebral blood flow or metabolism. PET can also be used to probe the function of neurotransmitter systems, using radioactive tracers that bind specific neurotransmitter receptors, transporters, or enzymes. For example, [¹¹C] raclopride has been used to measure dopamine D₂/D₃ receptors (Zald et al. 2004), while [¹⁸F] fluorodopa has been used to measure dopamine synthesis (Cools et al. 2009). The main limitation on what neurochemical processes can be studied is in the development of appropriate radioactive ligands.

One way to employ these kinds of PET scans is to look for associations *between subjects*. For example, individual differences in dopamine

release in the striatum have been related to differences in trait impulsivity (Buckholtz et al. 2010). Another way to employ these kinds of PET scans is to look for associations *within subjects*, across different tasks. For example, [^{11}C] raclopride has been used to measure changes in dopamine release under different task conditions, by assuming that neurotransmitter released endogenously competes with the radioactive tracer (Zald et al. 2004, but see Egerton et al. 2009 for several important caveats to this technique).

Genetics. Many techniques are now available for linking genetic differences (genotypes) to differences in behavior or brain function (phenotypes). Because most cognitive neuroscience studies incorporating genetics have focused on polymorphisms affecting one or more neurotransmitter systems, genetics is labeled here as a method for linking neurochemical systems to mental function. It is important to remember, however, that this is only a small portion of the possible genetic differences that could be studied.

Most cognitive neuroscience studies have used a *candidate gene* approach, in which a small number of preidentified genetic polymorphisms are tested (for more information on this approach, see Green et al. 2008). Some widely studied polymorphisms are in genes that code for neurotransmitter receptors (e.g., D2 or D4 dopamine receptors, A2-adrenergic receptor), neurotransmitter transporters (e.g., dopamine, norepinephrine, serotonin, or acetylcholine transporters), enzymes involved in neurotransmitter synthesis (e.g., dopamine beta-hydroxylase, tryptophan hydroxylase), enzymes involved in the breakdown of neurotransmitters (e.g., catechol-O-methyltransferase, monoamine oxidase), and proteins involved in neurotransmitter receptor signaling cascades (e.g., DARPP32). Usually, but not always, there is already molecular biological evidence that the genetic difference studied results in changes in the function of the associated protein (i.e., the protein is more or less efficient or is expressed to a greater or lesser degree). In these studies, genetic material is obtained from saliva samples, and each subject's genotype for the candidate gene is determined. Then behavioral performance (or neural activity) is correlated with genotype. Since genotype is not manipulated by the experimenter, these should be considered tests of association. These are also *between subjects* rather than *within subject* associations.

Although most published studies in cognitive neuroscience have used the candidate gene approach, questions about this approach have been raised. More recently, geneticists have moved toward *genome-wide association studies*, which scan the entire genome for loci associated with a particular phenotype (Manolio 2010). These studies are typically done at

large scale, gathering data from thousands of subjects, which necessitates a greater infrastructure of people and technology.

There are several complexities to keep in mind regarding genetic studies (Green et al. 2008). Most (if not all) behaviors are affected by many genes (*polygenicity*), so the effect of any one gene on the phenotype of interest is likely to be very small. This is why large samples and multiple replications have become the standard in genetics. Also, a given polymorphism is likely to have effects on multiple phenotypes (*pleiotropy*). Further, there could be interactions between different polymorphisms, including interactions between different polymorphisms within the same gene. Finally, these genetic differences are present throughout the lifespan, so they may exert effects through developmental or compensatory changes, and interactions between the gene and the environment are likely to play an important role.

In the domain of social decision making, the dopamine D4 receptor (DRD4) gene has been studied as a candidate gene contributing to punishing unfairness in the Ultimatum Game. This gene is characterized by different variable-number tandem repeats. Zhong et al. (2010) found that individuals with a 4-repeat allele of this gene, compared to those with a 7-repeat allele, are more sensitive to fairness and more likely to reject unfair offers (Zhong et al. 2010). Previous studies have also found that individuals with a 4-repeat allele have greater trait altruism and are less aggressive than individuals with the 7-repeat allele (Bakermans-Kranenburg, & van IJzendoorn, 2006; Bachner-Melman et al. 2005).

Tests of Necessity and Sufficiency of Neurochemical Systems for Mental Function

Pharmacology. Pharmacological interventions can be used to test the necessity and sufficiency of neurochemical systems for mental processes (for more information on this approach, see Robbins & Arnsten 2009). These studies involve administering a drug that blocks (an antagonist) or stimulates (an agonist) a particular kind of neurotransmitter receptor and comparing behavioral performance on and off the drug.

A main limiting factor to these studies is the existence of a drug that modulates the system of interest and that can be safely administered to drug-naïve subjects. Another practical limitation is the possible need for a physician's oversight. Theoretical interpretation of these studies should take into account that: (1) drugs act at many locations in the brain (and body) simultaneously, and these effects may interact; (2) most drugs are

not perfectly specific and can have effects at multiple sites (e.g., multiple receptor subtypes); and (3) the same drug can have different and sometimes opposing effects at different doses.

In addition to manipulating neurotransmitter systems pharmacologically, there are also methods for increasing or decreasing the basal levels of serotonin and dopamine/norepinephrine (Mendelsohn, Riedel, & Sambeth 2009; Silber & Schmitt 2010). These methods involve dietary manipulations that overload or deplete tryptophan (the metabolic precursor of serotonin) or that overload or deplete phenylalanine and tyrosine (metabolic precursors of dopamine and norepinephrine).

An example of using pharmacological techniques to enhance the function of a neurotransmitter system can be found in a study by Sáez et al. (2015). They used the drug tolcapone to selectively enhance dopamine levels in brain areas where there are low levels of dopamine transporters (e.g., frontal cortex and hippocampus). Participants then played a version of the Dictator Game, in which they received an initial number of tokens and could choose how many to share with another person. Tokens had varying exchange rates with differing values for both the player and the recipient. Tolcapone enhanced inequity-averse behavior, that is, it promoted the tendency of the participant to ensure an equal division of money between the two players (Sáez et al., 2015).

Both pharmacological and depletion techniques have been used to demonstrate the importance of the serotonergic system in rejecting unfairness. Reducing serotonergic tone through serotonin depletion makes responders in the Ultimatum Game more likely to reject unfair offers (Crockett et al. 2008), while enhancing serotonergic tone by administering selective serotonin reuptake inhibitors makes responders less likely to reject unfair offers (Crockett et al. 2010).

Methods That Monitor Peripheral Nervous System Activity

So far, this review has focused on methods that link mental functions to regional brain activity or to the action of diffuse neurochemical systems. Several other methods in the cognitive neuroscientist's toolkit do not fit into these two categories, but can be broadly characterized as measurements of peripheral nervous system activity. This includes indirect indices of autonomic (sympathetic/parasympathetic) activity, such as measurements of heart rate, blood pressure, respiration rate, pupil diameter, and galvanic skin response. These measurements have shown associations with arousal, learning, and decision making (Nassar et al. 2012; Satterthwaite et al. 2007)

and are also particularly useful in studies of emotion (Olsson & Phelps 2007). Other measures, such as eye tracking or facial electromyography (EMG), index peripheral somatomotor activity. Eye tracking can provide information about attention and information processing (Krajchich, Armel, & Rangel 2010), while facial EMG can indicate the presence of different emotions (Halberstadt et al. 2009). As an example of how these techniques can be used, Van't Wout et al. (2006) measured the galvanic skin response in Ultimatum Game responders and showed that it increased when receiving unfair (versus fair) offers and when rejecting (versus accepting) these offers, demonstrating an association between emotional arousal and recognition and punishment of unfairness.

Combining Methods

The functional organization of this methods overview stresses how different methods provide different kinds of inference and, therefore, highlights what can be gained by studying the same behavior with multiple different methods. Many of these gains occur even when different methods are used in separate studies. In some cases, though, the simultaneous use of two different methods enables one to ask novel questions that could not be answered with one technique alone or with the two techniques used at separate times.

Some good examples involve the simultaneous combination of fMRI with another method. Simultaneous EEG and fMRI allows one to look for BOLD activity that varies on a trial-by-trial basis with the amplitude or timing of an ERP, providing a stronger way to link data from the two methods (Debener et al. 2006). Simultaneous TMS and fMRI can map the functional interactions between brain regions, by testing how brain regions distant from the target respond to TMS pulses, or by testing how distant regions compensate during task performance when the targeted region is depressed (Siebner et al. 2009). Combining fMRI with genetics can identify intermediate phenotypes – neural responses that vary with genotype and that might prove more reliable than behavioral phenotypes (Canli & Lesch 2007; Green et al. 2008; Hariri, Drabant, & Weinberger 2006). Combining fMRI with pharmacology can allow one to better pinpoint the location of any pharmacological effects (Leslie & James 2000; Shah & Marsden 2004). Finally, peripheral measures can be combined with fMRI to study the central correlates of peripheral effects (Gray et al. 2009).

As an example, Baumgartner et al. (2011) combined rTMS and fMRI to look at how rTMS to right dorsolateral prefrontal cortex led to reduced

rejections of unfair offers in the Ultimatum Game (see the Knoch et al. study discussed above). They administered rTMS to either the right or left dorsolateral prefrontal cortex of participants and then scanned them 15 minutes later while they acted as responders in the Ultimatum Game. rTMS to right dorsolateral prefrontal cortex led to increased acceptance of unfair offers as well as decreased activity in the region stimulated. In addition, however, there was also reduced activity in vmPFC. This reveals that there is connectivity between vmPFC and right dorsolateral prefrontal cortex that is important for deciding to reject unfair offers, illustrating how the combination of rTMS and fMRI can be used to uncover the broader network through which rTMS effects occur (Baumgartner et al. 2011).

HOW CAN NEUROSCIENCE METHODS BE USED TO INFORM THEORIES IN THE SOCIAL SCIENCES?

A recurring question anytime neuroscience methods are applied in a new field is “Why should we study the brain in order to understand the mind or behavior?” An example is the recent debate regarding neuroeconomics (Camerer 2007; Camerer et al. 2004, 2005), though similar issues were raised by cognitive psychologists about the explosion of cognitive neuroscience research in the 1990s (Coltheart 2004). This introductory overview has stressed how different neuroscience techniques contribute to our understanding of brain–behavior relationships. Many scholars believe that a multilevel investigation of brain–behavior relationships, aimed at building mechanistic theories of human behavior, will ultimately lead to a coherent integration at the border of neuroscience and the social sciences (Glimcher 2010; Kable & Glimcher 2009). Setting aside the prospects for a broad integration, however, there are still some specific ways that the neuroscience methodologies reviewed above can be used to inform research questions in social sciences. That is, even for scholars who are *only* interested in human behavior, studying the brain can sometimes be informative.

As introduced above, tests of association such as fMRI usually make inferences about the probability of a neural state (ϕ) given a psychological state (ψ), $P(\phi | \psi)$. This kind of reasoning has been called “forward inference,” and it involves making assumptions about the psychological or behavioral states in the task and then inferring what neural processes are linked with those psychological or behavioral states (Aguirre & D’Esposito 1999; Poldrack 2006). Sometimes, though, investigators are interested in using fMRI data (or other association techniques) to make inferences about the probability of a psychological state given a neural state, $P(\psi | \phi)$. This

kind of reasoning has been called “reverse inference,” and it involves making assumptions about the linkages between neural processes and psychological processes and then inferring based on neural activity what psychological processes are involved in a given task (Aguirre & D’Esposito 1999; Poldrack 2006). A particularly weak form of reverse inference is to assume the function of brain region based on past studies and then, based on activation in that region, infer that the task the subject performed involved that particular psychological function. For example, one might find that certain kinds of decisions are associated with activity in the amygdala, note that the amygdala has previously been implicated in the emotion of fear, and then conclude that these decisions involved fear. The central problem with this kind of reverse inference is that it assumes there is only one psychological process associated with activity in a given brain region (e.g., the amygdala is *only* active when you experience fear), and such a one-to-one mapping is unlikely. Poldrack works through an example in probabilistic terms to show that this kind of reverse inference is unlikely to provide very strong information, at least at the spatial scale of brain activity and coarseness of the cognitive ontology that are typically used (Poldrack 2006, but see Ariely & Berns 2010 for a more promising example).

Reverse inferences in fMRI can be strengthened, however, by using the spatial pattern of BOLD activity across different regions of the brain. Tools such as Neurosynth (www.neurosynth.org) can provide quantitative estimates of the likelihood of different psychological processes given a brain activity pattern, based on the associations previously reported in the literature. An even stronger method is to specify the psychological states of interest ahead of time and use multivariate pattern analysis (MVPA) techniques to build a predictor to distinguish the states. MVPA has been used in this manner to distinguish different emotions (Kassam et al. 2013) and to distinguish different motives for altruistic behavior (Hein et al. 2016).

Beyond formal reverse inference, tests of association such as fMRI can also use the logic of associations and dissociations to inform theories about cognition and behavior. Whether two different tasks use similar or different neural processes provides information about whether these tasks are likely to use similar or different cognitive processes. For example, the fact that neural activity differs depending on whether people are interacting with another person or a computer algorithm suggests that people use different cognitive processes when engaging with human agents (Carter et al. 2012; McCabe et al. 2001; Strait et al. 2013). Similarly, whether two *individuals* use similar or different neural processes provides information about whether these individuals are likely using similar or different cognitive processes.

Bhatt and colleagues provide one example of this with regard to identifying different types of players in a bargaining game (Bhatt et al. 2010), and Houser and colleagues discuss how the discovery and validation of different types of decision-makers could improve econometric predictions (Houser, Schunk, & Xiao 2007).

Of course, the logic of associations and dissociations can also be applied to tests of necessity and sufficiency. Indeed, for more than a hundred years before the advent of modern imaging techniques, human neuropsychologists studied the effects of brain damage in order to elucidate mental structure. The search for double dissociations in neuropsychology has been motivated as much by a desire to understand the organization of the mind and, in turn, the brain. Though this may not be how they are commonly used, TMS, tDCS, and pharmacology can uncover similar dissociations. Further, since behavior is the dependent variable in tests of necessity and sufficiency, these techniques might provide even more compelling evidence to social scientists for whom human behavior is a primary interest. Thus, in addition to providing evidence about the causal role of brain structures in behavior, tests of necessity and sufficiency may be particularly well suited to inform theories in the social sciences about the organization of human behavior.

REFERENCES

- Adolphs, R., Damasio, H., Tranel, D., Cooper, G., & Damasio, A. R. (2000). A role for somatosensory cortices in the visual recognition of emotion as revealed by three-dimensional lesion mapping. *Journal of Neuroscience*, 20(7), 2683–2690.
- Aguirre, G. K., & D'Esposito, M. (1999). Experimental design for brain fMRI. In C. T. W. Moonen & P. A. Bandettini, eds., *Functional MRI* (369–380). Berlin: Springer Verlag.
- Antal, A., Nitsche, M. A., Kincses, T. Z., Kruse, W., Hoffmann, K. P., & Paulus, W. (2004). Facilitation of visuo-motor learning by transcranial direct current stimulation of the motor and extrastriate visual areas in humans. *European Journal of Neuroscience*, 19(10), 2888–2892.
- Ariely, D., & Berns, G. S. (2010). Neuromarketing: The hope and hype of neuroimaging in business. *Nature Reviews Neuroscience*, 11(4), 284–292.
- Ashburner, J., & Friston, K. J. (2000). Voxel-based morphometry – the methods. *Neuroimage*, 11(6 Pt 1), 805–821.
- Bachner-Melman, R., Gritsenko, I., Nemanov, L., Zohar, A. H., Dina, C., & Ebstein, R. P. (2005). Dopaminergic polymorphisms associated with self-report measures of human altruism: a fresh phenotype for the dopamine D4 receptor. *Molecular Psychiatry*, 10(4), 333–335.
- Bakermans-Kranenburg, M. J., & Van IJzendoorn, M. H. (2006). Gene-environment interaction of the dopamine D4 receptor (DRD4) and observed maternal insensitivity predicting externalizing behavior in preschoolers. *Developmental Psychobiology*, 48(5), 406–409.
- Bandettini, P. A. (2009a). Seven topics in functional magnetic resonance imaging. *Journal of Integrative Neuroscience*, 8(3), 371–403.
- (2009b). What's new in neuroimaging methods? *Annals of the New York Academy of Sciences*, 1156, 260–293.
- Baumgartner, T., Knoch, D., Hotz, P., Eisenegger, C., & Fehr, E. (2011). Dorsolateral and ventromedial prefrontal cortex orchestrate normative choice. *Nature Neuroscience*, 14(11), 1468–1474.
- Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, 275(5304), 1293–1295.
- Been, G., Ngo, T. T., Miller, S. M., & Fitzgerald, P. B. (2007). The use of tDCS and CVS as methods of non-invasive brain stimulation. *Brain Research Reviews*, 56(2), 346–361.
- Benton, A. (1991). Aphasia: Historical perspectives. In M. T. Sarno, ed., *Acquired Aphasia*, 2nd ed. (1–26). New York, NY: Academic Press.
- Bhatt, M. A., Lohrenz, T., Camerer, C. F., & Montague, P. R. (2010). Neural signatures of strategic types in a two-person bargaining game. *Proceedings of the National Academy of Sciences*, 107(46), 19720–19725.
- Buchel, C., Wise, R. J., Mummary, C. J., Poline, J. B., & Friston, K. J. (1996). Nonlinear regression in parametric activation studies. *Neuroimage*, 4(1), 60–66.
- Buckholtz, J. W., Treadway, M. T., Cowan, R. L., Woodward, N. D., Li, R., Ansari, M. S., Baldwin, R. M., Schwartzman, A. N., Shelby, E. S., Smith, C. E., Kessler, R. M., & Zald, D. H. (2010). Dopaminergic network differences in human impulsivity. *Science*, 329(5991), 532–532.
- Camerer, C. F. (2007). Neuroeconomics: Using neuroscience to make economic predictions. *Economic Journal*, 117(519), C26–C42.
- (2008). The potential of neuroeconomics. *Economics and Philosophy*, 24(3), 369–379.
- Camerer, C. F., Loewenstein, G., & Prelec, D. (2004). Neuroeconomics: Why economics needs brains. *Scandinavian Journal of Economics*, 106(3), 555–579.
- (2005). Neuroeconomics: How neuroscience can inform economics. *Journal of Economic Literature*, 43(1), 9–64.
- Canli, T., & Lesch, K. P. (2007). Long story short: The serotonin transporter in emotion regulation and social cognition. *Nature Neuroscience*, 10(9), 1103–1109.
- Carter, R. M., Bowling, D. L., Reeck, C., & Huettel, S. A. (2012). A distinct role of the temporal-parietal junction in predicting socially guided decisions. *Science*, 337(6090), 109–111.
- Churchland, P. S., & Sejnowski, T. J. (1988). Perspectives on cognitive neuroscience. *Science*, 242(4879), 741–745.
- Ciaramelli, E., Muccioli, M., Ladavas, E., & di Pellegrino, G. (2007). Selective deficit in personal moral judgment following damage to ventromedial prefrontal cortex. *Social Cognitive and Affective Neuroscience*, 2(2), 84–92.
- Coltheart, M. (2004). Brain imaging, connectionism, and cognitive neuropsychology. *Cognitive Neuropsychology*, 21(1), 21–25.

- Cools, R., Frank, M. J., Gibbs, S. E., Miyakawa, A., Jagust, W., & D'Esposito, M. (2009). Striatal dopamine predicts outcome-specific reversal learning and its sensitivity to dopaminergic drug administration. *Journal of Neuroscience*, 29(5), 1538–1543.
- Crockett, M. J., Clark, L., Tabibnia, G., Lieberman, M. D., & Robbins, T. W. (2008). Serotonin modulates behavioral reactions to unfairness. *Science*, 320(5884), 1739.
- Crockett, M. J., Clark, L., Hauser, M. D., & Robbins, T. W. (2010). Serotonin selectively influences moral judgment and behavior through effects on harm aversion. *Proceedings of the National Academy of Sciences*, 107(40), 17433–17438.
- de Quervain, D. J., Fischbacher, U., Treyer, V., et al. (2004). The neural basis of altruistic punishment. *Science*, 305(5688), 1254–1258.
- Debener, S., Ullsperger, M., Siegel, M., & Engel, A. K. (2006). Single-trial EEG-fMRI reveals the dynamics of cognitive function. *Trends in Cognitive Sciences*, 10(12), 558–563.
- Egerton, A., Mehta, M. A., Montgomery, A. J., et al. (2009). The dopaminergic basis of human behaviors: a review of molecular imaging studies. *Neuroscience and Biobehavioral Reviews*, 33(7), 1109–1132.
- Engel, A. K., Moll, C. K., Fried, I., & Ojemann, G. A. (2005). Invasive recordings from the human brain: Clinical insights and beyond. *Nature Reviews Neuroscience*, 6(1), 35–47.
- Fellows, L. K., & Farah, M. J. (2007). The role of ventromedial prefrontal cortex in decision making: Judgment under uncertainty or judgment per se? *Cerebral Cortex*, 17(11), 2669–2674.
- Fellows, L. K., Stark, M., Berg, A., & Chatterjee, A. (2008). Patient registries in cognitive neuroscience research: Advantages, challenges, and practical advice. *Journal of Cognitive Neuroscience*, 20(6), 1107–1113.
- Fellows, L. K. (2012). Group studies in experimental neuropsychology. In *APA Handbook of Research Methods in Psychology, Vol. 2* (647–659). Washington, DC: American Psychological Association.
- Fox, P. T., & Raichle, M. E. (1986). Focal physiological uncoupling of cerebral blood flow and oxidative metabolism during somatosensory stimulation in human subjects. *Proceedings of the National Academy of Sciences*, 83(4), 1140–1144.
- Friston, K. J. (1994). Functional and effective connectivity in neuroimaging: A synthesis. *Human Brain Mapping*, 2(1–2), 56–78.
- Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage*, 6(3), 218–229.
- Glimcher, P. W. (2010). *Foundations of Neuroeconomic Analysis*. Oxford, New York: Oxford University Press.
- Gratton, G., & Fabiani, M. (2001). Shedding light on brain function: The event-related optical signal. *Trends in Cognitive Sciences*, 5(8), 357–363.
- Gray, M. A., Minati, L., Harrison, N. A., Gianaros, P. J., Napadow, V., & Critchley, H. D. (2009). Physiological recordings: Basic concepts and implementation during functional magnetic resonance imaging. *Neuroimage*, 47(3), 1105–1115.
- Green, A. E., Munafò, M. R., Deyoung, C. G., Fossella, J. A., Fan, J., & Gray, J. R. (2008). Using genetic data in cognitive neuroscience: From growing pains to genuine insights. *Nature Reviews Neuroscience*, 9, 710–720.

- Grill-Spector, K., & Malach, R. (2001). fMR-adaptation: A tool for studying the functional properties of human cortical neurons. *Acta Psychologica*, 107(1–3), 293–321.
- Grossman, M., Eslinger, P. J., Troiani, V., et al. (2010). The role of ventral medial prefrontal cortex in social decisions: Converging evidence from fMRI and frontotemporal lobar degeneration. *Neuropsychologia*, 48(12), 3505–3512.
- Gul, F., & Pesendorfer, W. (2008). The case for mindless economics. In A. Caplin & A. Schotter, eds., *The Foundations of Positive and Normative Economics: A Handbook*. New York, NY: Oxford University Press.
- Halberstadt, J., Winkelman, P., Niedenthal, P. M., & Dalle, N. (2009). Emotional conception: How embodied emotion concepts guide perception and facial action. *Psychological Science*, 20(10), 1254–1261.
- Handy, T. C. (2005). *Event-Related Potentials: A Methods Handbook*. Cambridge, MA: MIT Press.
- Hariri, A. R., Drabant, E. M., & Weinberger, D. R. (2006). Imaging genetics: Perspectives from studies of genetically driven variation in serotonin function and corticolimbic affective processing. *Biological Psychiatry*, 59(10), 888–897.
- Harrison, G. W. (2008). Neuroeconomics: A critical reconsideration. *Economics and Philosophy*, 24(3), 303–344.
- Haynes, J. D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, 7(7), 523–534.
- Hein, G., Morishima, Y., Leiberg, S., Sul, S., & Fehr, E. (2016). The brain's functional network architecture reveals human motives. *Science*, 351(6277), 1074–1078.
- Henrich, J., McElreath, R., Barr, A., et al. (2006). Costly punishment across human societies. *Science*, 312(5781), 1767–1770.
- Henson, R. N., & Rugg, M. D. (2003). Neural response suppression, haemodynamic repetition effects, and behavioural priming. *Neuropsychologia*, 41(3), 263–270.
- Hollmann, M., Rieger, J. W., Baecke, S., et al. (2011). Predicting decisions in human social interactions using real-time fMRI and pattern classification. *PloS One*, 6(10), e25304.
- Hoshi, Y. (2005). Functional near-infrared spectroscopy: Potential and limitations in neuroimaging studies. *International Review of Neurobiology*, 66, 237–266.
- Houser, D., Schunk, D., & Xiao, E. (2007). Combining brain and behavioral data to improve econometric policy analysis. *Analyse & Kritik*, 29, 86–96.
- Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., & Camerer, C. F. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science*, 310(5754), 1680–1683.
- Huettel, S. A., Song, A. W., & McCarthy, G. (2004). *Functional Magnetic Resonance Imaging*. Sunderland, MA: Sinauer Associates.
- Jezzard, P., Matthews, P. M., & Smith, S. M. (2001). *Functional MRI: An Introduction to Methods*. Oxford, New York: Oxford University Press.
- Johansen-Berg, H., & Rushworth, M. F. (2009). Using diffusion imaging to study human connective anatomy. *Annual Review of Neuroscience*, 32, 75–94.
- Kable, J. W., & Glimcher, P. W. (2009). The neurobiology of decision: Consensus and controversy. *Neuron*, 63(6), 733–745.
- Kassam, K. S., Markey, A. R., Cherkassky, V. L., Loewenstein, G., & Just, M. A. (2013). Identifying emotions on the basis of neural activation. *PloS One*, 8(6), e66032.

- Kimberg, D. Y., Coslett, H. B., & Schwartz, M. F. (2007). Power in voxel-based lesion-symptom mapping. *Journal of Cognitive Neuroscience*, 19(7), 1067–1080.
- Knoch, D., Nitsche, M. A., Fischbacher, U., Eisenegger, C., Pascual-Leone, A., & Fehr, E. (2008). Studying the neurobiology of social interaction with transcranial direct current stimulation – The example of punishing unfairness. *Cerebral Cortex*, 18(9), 1987–1990.
- Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., & Fehr, E. (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science*, 314(5800), 829–832.
- Koenigs, M., Young, L., Adolphs, R., et al. (2007). Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature*, 446(7138), 908–911.
- Koenigs, M., & Tranel, D. (2007). Irrational economic decision-making after ventromedial prefrontal damage: Evidence from the ultimatum game. *The Journal of Neuroscience*, 27(4), 951–956.
- Kolb, B., & Whishaw, I. Q. (2009). *Fundamentals of Human Neuropsychology*, 6th ed. New York, NY: Worth Publishers.
- Krajchich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, 13(10), 1292–1298.
- Lachaux, J. P., Rudrauf, D., & Kahane, P. (2003). Intracranial EEG and human brain mapping. *Journal of Physiology (Paris)*, 97(4–6), 613–628.
- Laycock, R., Crewther, D. P., Fitzgerald, P. B., & Crewther, S. G. (2007). Evidence for fast signals and later processing in human V1/V2 and V5/MT+: A TMS study of motion perception. *Journal of Neurophysiology*, 98(3), 1253–1262.
- Leslie, R. A., & James, M. F. (2000). Pharmacological magnetic resonance imaging: A new application for functional MRI. *Trends in Pharmacological Sciences*, 21(8), 314–318.
- Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature*, 453(7197), 869–878.
- Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, 412(6843), 150–157.
- Lounasmaa, O. V., Hamalainen, M., Hari, R., & Salmelin, R. (1996). Information processing in the human brain: magnetoencephalographic approach. *Proceedings of the National Academy of Sciences*, 93(17), 8809–8815.
- Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique*. Cambridge, MA: MIT Press.
- Manolio, T. A. (2010). Genome-wide association studies and assessment of the risk of disease. *New England Journal of Medicine*, 363(2), 166–176.
- McCabe, K. A. (2008). Neuroeconomics and the economic sciences. *Economics and Philosophy*, 24(3), 345–368.
- McCabe, K., Houser, D., Ryan, L., Smith, V., & Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of Sciences*, 98(20), 11832–11835.
- McIntosh, A. R., Bookstein, F. L., Haxby, J. V., & Grady, C. L. (1996). Spatial pattern analysis of functional brain images using partial least squares. *Neuroimage*, 3(3 Pt 1), 143–157.

- Mendelsohn, D., Riedel, W. J., & Sambeth, A. (2009). Effects of acute tryptophan depletion on memory, attention and executive functions: A systematic review. *Neuroscience and Biobehavioral Reviews*, 33(6), 926–952.
- Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., & Grave de Peralta, R. (2004). EEG source imaging. *Clinical Neurophysiology*, 115(10), 2195–2222.
- Miller, K. J., denNijs, M., Shenoy, P., Miller, J. W., Rao, R. P., & Ojemann, J. G. (2007). Real-time functional brain mapping using electrocorticography. *Neuroimage*, 37(2), 504–507.
- Mori, S., & Zhang, J. (2006). Principles of diffusion tensor imaging and its applications to basic neuroscience research. *Neuron*, 51(5), 527–539.
- Nassar, M. R., Rumsey, K. M., Wilson, R. C., Parikh, K., Heasley, B., & Gold, J. I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. *Nature Neuroscience*, 15(7), 1040–1046.
- Newsome, W. T. (1997). The King Solomon Lectures in Neuroethology. Deciding about motion: linking perception to action. *Journal of Comparative Physiology A – Sensory Neural & Behavioral Physiology*, 181(1), 5–12.
- Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, 10(9), 424–430.
- O'Doherty, J. P., Hampton, A., & Kim, H. (2007). Model-based fMRI and its application to reward learning and decision making. *Annals of the New York Academy of Sciences*, 1104, 35–53.
- Obrig, H., & Villringer, A. (2003). Beyond the visible – Imaging the human brain with light. *Journal of Cerebral Blood Flow and Metabolism*, 23(1), 1–18.
- Ojemann, J. G., Akbudak, E., Snyder, A. Z., McKinstry, R. C., Raichle, M. E., & Conturo, T. E. (1997). Anatomic localization and quantitative analysis of gradient refocused echo-planar fMRI susceptibility artifacts. *Neuroimage*, 6(3), 156–167.
- Olsson, A., & Phelps, E. A. (2007). Social learning of fear. *Nature Neuroscience*, 10(9), 1095–1102.
- Parker, A. J., & Newsome, W. T. (1998). Sense and the single neuron: Probing the physiology of perception. *Annual Review of Neuroscience*, 21, 227–277.
- Pascual-Leone, A., Walsh, V., & Rothwell, J. (2000). Transcranial magnetic stimulation in cognitive neuroscience – Virtual lesion, chronometry, and functional connectivity. *Current Opinion in Neurobiology*, 10(2), 232–237.
- Plassmann, H., O'Doherty, J., & Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *Journal of Neuroscience*, 27(37), 9984–9988.
- Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, 10(2), 59–63.
- Prehn, K., Kordzykowski, M., Rao, H., Fang, Z., Detre, J. A., & Robertson, D. C. (2015). Neural correlates of post-conventional moral reasoning: A voxel-based morphometry study. *PloS One*, 10(6), e0122914.
- Raichle, M. E. (1998). Behind the scenes of functional brain imaging: A historical and physiological perspective. *Proceedings of the National Academy of Sciences*, 95(3), 765–772.

- Rest, J., Narvaez, D., Bebeau, M., & Thoma, S. (1999). A neo-Kohlbergian approach: The DIT and schema theory. *Educational Psychology Review*, 11(4), 291-324.
- Robbins, T. W., & Arnsten, A. F. (2009). The neuropsychopharmacology of fronto-executive function: Monoaminergic modulation. *Annual Review of Neuroscience*, 32, 267-287.
- Rorden, C., & Karnath, H. O. (2004). Using human brain lesions to infer function: A relic from a past era in the fMRI age? *Nature Reviews Neuroscience*, 5(10), 813-819.
- Rorden, C., Karnath, H. O., & Bonilha, L. (2007). Improving lesion-symptom mapping. *Journal of Cognitive Neuroscience*, 19(7), 1081-1088.
- Ruff, C. C., Ugazio, G., & Fehr, E. (2013). Changing social norm compliance with noninvasive brain stimulation. *Science*, 342(6157), 482-484.
- Rutledge, R. B., Lazzaro, S. C., Lau, B., Myers, C. E., Gluck, M. A., & Glimcher, P. W. (2009). Dopaminergic drugs modulate learning rates and perseveration in Parkinson's patients in a dynamic foraging task. *Journal of Neuroscience*, 29(48), 15104-15114.
- Sáez, I., Zhu, L., Set, E., Kayser, A., & Hsu, M. (2015). Dopamine modulates egalitarian behavior in humans. *Current Biology*, 25(7), 912-919.
- Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science*, 300(5626), 1755-1758.
- Sarlo, M., Lotto, L., Manfrinati, A., Rumiat, R., Gallicchio, G., & Palomba, D. (2012). Temporal dynamics of cognitive-emotional interplay in moral decision-making. *Journal of Cognitive Neuroscience*, 24(4), 1018-1029.
- 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(1), 13-21.
- Satterthwaite, T. D., Green, L., Myerson, J., Parker, J., Ramaratnam, M., & Buckner, R. L. (2007). Dissociable but inter-related systems of cognitive control and reward during decision making: Evidence from pupillometry and event-related fMRI. *Neuroimage*, 37(3), 1017-1031.
- Senior, C., Russell, T., & Gazzaniga, M. S. (2006). *Methods in Mind*. Cambridge, MA: MIT Press.
- Shah, Y. B., & Marsden, C. A. (2004). The application of functional magnetic resonance imaging to neuropharmacology. *Current Opinion in Pharmacology*, 4(5), 517-521.
- Shallice, T. (1988). *From Neuropsychology to Mental Structure*. Cambridge, New York: Cambridge University Press.
- Shohamy, D., Myers, C. E., Grossman, S., Sage, J., & Gluck, M. A. (2005). The role of dopamine in cognitive sequence learning: evidence from Parkinson's disease. *Behavioural Brain Research*, 156(2), 191-199.
- Siebnner, H. R., Bergmann, T. O., Bestmann, S., et al. (2009). Consensus paper: Combining transcranial stimulation with neuroimaging. *Brain Stimulation*, 2(2), 58-80.

- Silber, B. Y., & Schmitt, J. A. (2010). Effects of tryptophan loading on human cognition, mood, and sleep. *Neuroscience & Biobehavioral Reviews*, 34(3), 387-407.
- Silva, A. J. (2007). The science of research: the principles underlying the discovery of cognitive and other biological mechanisms. *Journal of Physiology (Paris)*, 101(4-6), 203-213.
- Strait, M., Briggs, G., & Scheutz, M. (2013, September). Some Correlates of Agency Ascription and Emotional Value and Their Effects on Decision-making. 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII) (505-510). IEEE.
- Sun, F. T., Miller, L. M., & D'Esposito, M. (2004). Measuring interregional functional connectivity using coherence and partial coherence analyses of fMRI data. *Neuroimage*, 21(2), 647-658.
- Toga, A. W., & Mazziotta, J. C. (2002). *Brain Mapping: The Methods*, 2nd ed. Amsterdam, Boston: Academic Press.
- Tootell, R. B., Reppas, J. B., Dale, A. M., et al. (1995). Visual motion aftereffect in human cortical area MT revealed by functional magnetic resonance imaging. *Nature*, 375(6527), 139-141.
- Tootell, R. B., Reppas, J. B., Kwong, K. K., et al. (1995). Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *The Journal of Neuroscience*, 15(4), 3215-3230.
- Tranel, D., Damasio, H., & Damasio, A. R. (1997). A neural basis for the retrieval of conceptual knowledge. *Neuropsychologia*, 35(10), 1319-1327.
- Van't Wout, M., Kahn, R. S., Sanfey, A. G., & Aleman, A. (2006). Affective state and decision-making in the ultimatum game. *Experimental Brain Research*, 169(4), 564-568.
- Wagner, T., Valero-Cabre, A., & Pascual-Leone, A. (2007). Noninvasive human brain stimulation. *Annual Review of Biomedical Engineering*, 9, 527-565.
- Walsh, V., & Cowey, A. (2000). Transcranial magnetic stimulation and cognitive neuroscience. *Nature Reviews Neuroscience*, 1(1), 73-79.
- Walsh, V., Ellison, A., Battelli, L., & Cowey, A. (1998). Task-specific impairments and enhancements induced by magnetic stimulation of human visual area V5. *Proceedings of the Royal Society of London B: Biological Sciences*, 265(1395), 537-543.
- Wassermann, E. M. (1998). Risk and safety of repetitive transcranial magnetic stimulation: Report and suggested guidelines from the International Workshop on the Safety of Repetitive Transcranial Magnetic Stimulation, June 5-7, 1996. *Electroencephalography & Clinical Neurophysiology*, 108(1), 1-16.
- Weber, M. J., & Thompson-Schill, S. L. (2010). Functional neuroimaging can support causal claims about brain function. *Journal of Cognitive Neuroscience*, 22(11), 2415-2416.
- Weigelt, S., Muckli, L., & Kohler, A. (2008). Functional magnetic resonance adaptation in visual neuroscience. *Reviews in the Neurosciences*, 19(4-5), 363-380.
- Weiskopf, N., Hutton, C., Josephs, O., & Deichmann, R. (2006). Optimal EPI parameters for reduction of susceptibility-induced BOLD sensitivity losses: A whole-brain analysis at 3T and 1.5T. *Neuroimage*, 33(2), 493-504.

- Zald, D. H., Boileau, I., El-Dearedy, W., et al. (2004). Dopamine transmission in the human striatum during monetary reward tasks. *Journal of Neuroscience*, 24(17), 4105–4112.
- Zhong, S., Israel, S., Shalev, I., Xue, H., Ebstein, R. P., & Chew, S. H. (2010). Dopamine D4 receptor gene associated with fairness preference in ultimatum game. *PLoS One*, 5(11), e13765.

3

The Conforming Brain and Deontological Resolve

MELANIE PINCUS, LISA LAVIERS,
MICHAEL J. PRIETULA, AND GREGORY BERNS

INTRODUCTION

A person's set of values helps define one's individual and group identity. Of particular importance in defining one's identity are protected, or sacred, values – those defended resolutely and not compromised for material trade-offs (Baron & Spranca 1997). These include core religious beliefs and moral norms that constrain decision making across a person's lifetime. In many cultures, violating sacred values is tantamount to disavowing group membership (Atran 2010), underscoring the importance of sacred values to group identity.

Research demonstrates that individuals process sacred values deontologically, whereby individuals construct and represent these values according to a rule-based schema of rights and wrongs, irrespective of context or outcome (Baron & Spranca 1997; Ginges et al. 2011; Tetlock 2003). Evidence suggests that individuals differ in the degree to which they endorse deontological rules (Bartels 2008; Lombrozo 2009), and individuals with a deontological disposition may have a stronger neurobiological representation of deontological rules. Consequently, we hypothesize that they may be more unwilling to bend the rules of right/wrong in the face of social influence. This hypothesis is supported by the finding that the certainty with which an attitude is held is correlated with greater resistance to social influence (Babad et al. 1987; Bassili 1996; Tormala & Petty 2002). Accordingly, individuals with a strong deontological disposition who readily appeal to absolute rules that confer greater certainty should be expected to be less conformist generally. Because sacred values are categorical in nature, it would be difficult to test this hypothesis and discern individual differences in deontological

Reprinted with permission from *PLoS One*, August 2014, Volume 8, e106061.