Running head: SOCIAL NEURO METHODS

Research Methods in Social and Affective Neuroscience

Elliot T. Berkman

University of Oregon

William A. Cunningham

University of Toronto

Matthew D. Lieberman

University of California, Los Angeles

Main text word count: 19,380

Address correspondence to:
Elliot Berkman
Department of Psychology
University of Oregon
1227 University of Oregon
Eugene, OR, 97403-1227

Phone: (541) 346-4909
Fax: (541) 346-4911
e-mail: berkman@uoregon.edu

In press at *Handbook of Research Methods in Personality and Social Psychology*

H.T. Reis & C.M. Judd, Eds.

Introduction

In the two decades since the term *social neuroscience* was first used (Cacioppo & Berntson, 1992), social and personality psychologists have increasingly and enthusiastically adopted neuroscience methods including neuroimaging, endocrinology, and peripheral physiological measurement to address social psychological questions. The number of Google hits for the phrase "social neuroscience" rose from 393 in 2001 to 290,000 in 2011, and from 6 to 946,000 for the phrase "social cognitive neuroscience" across the same period. There are now two societies for social neuroscience, each with its own journal. As of 2012, more than 60 labs around the world identify themselves with social neuroscience, social cognitive neuroscience, or social-affective neuroscience, using neuroimaging to inform social psychological theories on such topics as self-knowledge, theory of mind, mentalizing, emotion regulation, empathy, and implicit attitudes (Lieberman, 2012).

Despite its increasing popularity, there remains a gap in the level of methodological knowledge between those who practice social neuroscience and those who comprise a large part of the intended audience but are not active researchers in the field, namely social / personality psychologists. This may be in part because the rapid and recent development of the field has outpaced changes in doctoral curricula and faculty hiring in psychology departments. In our experience, doctoral students, postdocs, and faculty alike in social and personality psychology are interested in social neuroscience and optimistic about its use to further psychological theory. They just may not have had the chance to learn about its methods in sufficient detail to be informed consumers or to integrate neuroscience into their own research.

Our aim here is to provide a concise and up-to-date description of neuroimaging methods used by social neuroscientists—particularly functional magnetic resonance imaging (fMRI) and

electroencephalography (EEG)—for an audience of social and personality psychologists. This chapter is not meant to be a comprehensive "how to" guide for practitioners, but rather a "what, how, and why" guide for consumers of social neuroscience research, particularly graduate students and faculty in psychology and related fields who are curious about but unfamiliar with the methods. For more comprehensive guides, we recommend *Methods in Social Neuroscience* (Harmon-Jones & Beer, 2009), for an overview of various methods, as well as *Functional Magnetic Resonance Imaging* (Huettel, Song, & McCarthy, 2009) and *Handbook of Functional MRI Data Analysis* (Poldrack, Mumford, & Nichols, 2011) for the details of fMRI acquisition and analysis, respectively. For an in depth discussion of EEG methods, we recommend *An Introduction to the Event-Related Potential Technique* (Luck, 2005). We focus on fMRI and EEG because they are the main neuroimaging modalities for social and affective neuroscience. Though we don't explicitly cover other modalities that are also commonly used such as structural MRI (sMRI; Mechelli, Price, Friston, & Ashburner, 2005) and transcranial magnetic stimulation (TMS; Hallett, 2007), we note that many of the conceptual issues that we discuss with respect to fMRI and EEG are also applicable to these and other imaging modalities (e.g., positron emission tomography).

This chapter reviews two methods in the context of eight conceptual questions in social neuroscience. First, we briefly discuss what kinds of questions can be answered using social neuroscience methods. Next, we describe how fMRI studies are designed and how their data are collected, analyzed, and reported. A section on EEG and event-related potential (ERP) studies follows. Finally, we turn to a discussion of recent debates and controversies in social neuroscience and related fields. Each section is written to be independent of the others, so readers can skip sections according to their content interest or desired level of detail.

Types of questions that social neuroscience methods can answer

Social/personality psychologists have been drawn to neuroimaging methods to answer at least three classes of questions that are relevant to social and personality theory (Lieberman, 2010). First, what are the neural mechanisms of social/personality processes? Second, do the neural systems involved in these processes overlap with the neural systems of other processes in ways that might not be easy to discover with other methods? And third, are there neural systems that are uniquely involved in social/personality processes? We consider each briefly below.

Brain mapping: What are the neural mechanisms of social/personality processes?

The ability to make inferences about which mental processes are associated with a given pattern of neural activation depends critically on the quality of our understanding of the function of various brain regions. The more detailed understanding we have of what a region does, the stronger our inference can be about which mental process is occurring when we observe activation in that region. Because of this fact, the enterprise of social neuroscience actually consists of two steps that happen in an iterative loop: using brain mapping to understand which brain regions are involved in a given psychological process, then testing psychological theory based on this new information and updating it when necessary (Cunningham, 2010). For example, if we are highly confident that the left ventrolateral prefrontal cortex and the medial temporal lobes are involved in processing depth during memory encoding, and observe activation in these regions during non-social information encoding but in a distinct set of brain regions during encoding of social information, then we can infer that social information is encoded in a qualitatively different way than non-social information (Mitchell, Macrae, & Banaji, 2004).

The key point is that high quality brain mapping is critical for social neuroscience to be successful. Considerable progress has been made in this area in recent decades, but even more work remains. For instance, we know that the right inferior frontal gyrus is involved in self-control (Aron, Robins, & Poldrack, 2004), but it is also involved in related processes such as task switching (Cools, Clark, Owen, & Robbins, 2002), conditional rule use (Bunge & Zelazo, 2006), and competitive biasing of abstract information (Munakata, Herd, Chatham, Depue, Banich, & O'Reilly, 2011). These kinds of brain mapping studies are and will continue to be essential in triangulating a more precise understanding of what computations this region is performing, and their results are directly informative to social psychological theory on self-control. In general, these refinements in brain mapping knowledge include: identifying the relationship between mental process and function or connectivity of a region; specifying boundary conditions or contextual moderators; and charting functional co-activations among two or more regions that are characteristic of a specific process. Each contributes to our understanding of the mapping between mind and brain, and thus each is a critical part of social neuroscience.

Convergences: How do social/personality neural systems overlap with other systems?

Identifying convergences in the brain systems of mental processes that subjectively feel distinct or are otherwise predicted to be different is another way that neuroscience can contribute to psychological theory. Examples of convergences include strong overlaps in the brain systems involved in social and non-social rewards (Izuma, Saito, & Sadato, 2008), and between social and physical pain (Eisenberger, 2012). This later set of findings suggested a novel intervention for social pain—Tylenol administration—that would have been difficult to justify based on behavior and self-report data alone (DeWall, et al., 2010). Convergences in neural function between different mental processes have also successfully been used to compare among theories

when each offers a competing process model. For example, psychologists have debated whether we understand the emotions of others by mirroring their experience directly, or by first simulating their experience in our own mind and then projecting our emotions onto them (Wheatley, Milleville, & Martin, 2007). It turns out that these accounts—mirroring and self-projection—have distinct neural systems, and that only the system for self-projection is consistently active when subjects attempt to empathize with the emotions of another (Waytz & Mitchell, 2011). This overlap between the brain systems for directly experiencing an emotion and for understanding it in others has provided support for the self-projection theory of empathy.

We note again that the strength of inference that can be made based on neural convergence is limited by the specificity of brain mapping. A critical open question is: what level of precision on a neural level is required to infer that two processes are the same? Clearly, it is finer than the entire brain (i.e., "any two processes that both occur in the brain are the same" makes no sense), and probably finer than a hemisphere or a lobe, too. Are two processes that both activate the same 1 cm^3 chunk of tissue the same? How about $1mm^3$? It is unlikely that two instances of even the *same* mental process share that level of specificity (see Points #4 and #5 on replication below), but the exact threshold for ontological "sameness" is currently unknown, and may be beyond the resolution of our current neuroimaging technologies. This issue precludes us from making definitive claims about shared mechanism for now, but these will improve along with advances in methods and knowledge.

Divergences: Are there brain systems unique to social/personality processes?

Neuroscience methods can also be used to find surprising divergences in the mechanisms of various processes that would otherwise seem similar. An emerging theme in social neuroscience is that the brain networks for processing social stimuli may be distinct from those

that process non-social stimuli. The study by Mitchell and colleagues (2004) cited above is an example: social and non-social memory encoding did not merely differ quantitatively on depth of processing, but also qualitatively in that they used distinct neural regions. Memory retrieval was correlated with one set of regions for non-social stimuli, and with an entirely different set of regions for social stimuli.

We note that the social / non-social distinction is just one potential divergence uncovered thus far by social neuroscience. It happens to have considerable support, but there are others. One example comes from the emotion literature, where LeDoux (1988) famously identified two distinct pathways, the "low road" and the "high road" to achieve a fear response. Another comes from social cognition, in which Foerde and colleagues (2006) demonstrated that implicit and explicit learning (differentiated with a cognitive load manipulation) produced equivalent behavioral results but used distinct neural regions to do so. Social neuroscience is still young; surely more surprising divergences are yet to be discovered.

fMRI methods

Study design

In addition to the usual considerations that one must take into account when designing a social/personality psychology study, there are several special design considerations that are unique to fMRI. These can be classified into three categories that we call, "statistical," "technological," and "human factors" considerations.

The main statistical considerations derive from the facts that all fMRI studies contain at least one within-subjects factor (because, as discussed below, the raw data are in arbitrary units that vary considerably from subject to subject), and that the data are arranged in a time-series with a potentially strong autocorrelation (i.e., later time points cannot be assumed to be

independent of earlier time points). Another factor is that the variable that is measured by fMRI, the blood oxygenation level-dependent (BOLD) signal, is slowed and elongated in time relative to the stimuli used in our experiments. For example, even a very brief (e.g., 100ms) burst of white noise will generate a BOLD signal in auditory cortex that begins 3-5 seconds following onset and peaks around 6-8 seconds following onset, with a total duration of approximately 6-8 seconds (Glover, 1999; Figure 1). Longer stimuli will produce even longer BOLD responses, and identical stimuli presented in rapid succession will sum in their BOLD response (approximating the rules of a linear system within a certain range).

Together, these facts imply that stimuli presented too close to one another in time or in regular succession will be confounded in the BOLD response and difficult statistically to differentiate from one another. There are two main design solutions to this problem: blocked and event-related designs (Figure 2). In a blocked design, trials of the same type are grouped together into relatively long "blocks" that are usually 20 to 50 seconds each, and separated by either a fixation cross "resting" baseline, various other conditions, or both. Blocked designs maximize the efficiency to detect differences between conditions throughout the brain, but provide no information about the shape of the curve of the BOLD response over time, called the hemodynamic response curve (Buckner, Bandettini, O'Craven, Savoy, Petersen, Raichle, & Rosen, 1996). The other solution is to use a "rapid event-related" design in which brief events (typically 1-5 seconds each) are presented in close succession. (A third type of design, "slow event-related," which can be thought of as a hybrid of blocked and rapid event-related designs, has fallen out of favor because of its low detection power.) Rapid event-related designs are popular because, when used appropriately, they allow for a reasonable balance between signal detection and hemodynamic response estimation efficiency (Birn, Cox, & Bandettini, 2002), and

allow stimuli in various conditions to intermingle in ways that may be advantageous psychologically (e.g., in studies on emotion where habituation to negatively valenced stimuli can be a problem). Fortunately, researchers can either use "jitter" (insert a systematically varying amount of time between successive stimuli of the same trial type, Dale & Buckner, 1997) or optimize the order of the trials (Wager & Nichols, 2003) to both remove the temporal correlation between conditions and also tease apart the BOLD signal corresponding to each. In other words, by varying the order of trials or the length of the gap between trials, or even between parts of trials, statistical procedures are better able to recover the contrast between particular trials (or trial components).

One last important statistical consideration is power. What sample size is sufficient and how many trials per condition are required? These questions are difficult to answer analytically because the variance of fMRI data and the effect size ranges widely from study to study and from brain region to brain region. However, researchers are beginning to develop methods to conduct power analyses based on formal statistical models (Mumford & Nichols, 2008) and data simulations (Desmond & Glover, 2002; Murphy & Garavan, 2004). We anticipate especially rapid accumulation of knowledge in this area in the coming years because of the high demand relative to the current information available. For the time being, we are reluctant to provide informal rules of thumb because, just as with behavioral studies, the sample size required to achieve a given amount of power can vary considerably from study to study depending on the amount of noise, and the covariance between within-subjects conditions, among other factors. We provide further information about sample size in the section below on between-subjects correlations (Point #2, below).

The technological design factors concern the limitations of the fMRI machine and its environment. First, fMRI scanning is expensive: an hour of scanner time typically costs between $400 and $800 in the U.S. In that time, a researcher must leave 10-15 minutes for subject setup inside the scanner, 10 minutes for various structural scans and other scans that are not directly relevant to the task, and several minutes for instructions. Efficient researchers can maximize the amount of scanner time spent on experimental tasks, but even in this case the tasks are shorter than they otherwise would be (e.g., in a behavioral laboratory study). This can present a problem for behavioral experimental paradigms that didn't originate in neuroscience and that require a large number of trials and/or subjects (either of which increases the total cost of an experiment). Generally speaking, it is more cost-effective to run fewer subjects each with more trials (vs. more subjects each with fewer trials), so social neuroscientists rely almost exclusively on within-subjects designs—even for experimental paradigms that are usually between-subjects—and try to employ manipulations that maximize the difference between conditions with as few trials as possible.

Second, fMRI scanning is loud. Tasks that rely upon auditory cues (e.g., the auditory stop-signal task) must be adjusted accordingly and tested to ensure that task-relevant cues are not presented in the same frequency range as the scanner noise that accompanies functional MRI. And third, fMRI scanning is awkward. The most common type of study involves a single participant, alone and supine in the bore of the scanner, wearing goggles or a mirror for visual stimuli and headphones for auditory stimuli, holding a customized response device such as a scanner-safe button box, mouse, or joystick, and attempting to hold as still as possible. Because MRI relies on a very strong magnetic field (usually 1.5 to 4 Tesla), no ferromagnetic objects (e.g., most computers, cameras, or input devices typically used in behavioral studies) can be near the

scanner. Some experiments can be run within these constraints, but many cannot. However, neuroscientists have recently made great strides in overcoming some of the limitations of this environment, for example by having another person in the scanner room to study physical contact (Coan et al., 2006), using eye-tracking to simulate gaze-contingent social interaction (Wilms, Schilbach, Pfeiffer, Bente, Fink, & Vogeley, 2010) and studying gustatory responses to food with a milkshake pump system (Stice et al., 2011) or an elaborate olfactory cue delivery device (Small, Gerber, Mak, & Hummel, 2005).

Finally, to add to the technological limitations listed above, researchers seeking to use fMRI for social neuroscience must put human beings into the scanner environment. Many humans are uncomfortable in dark and narrow tubes, including young children and individuals with ADHD or claustrophobia, making those challenging populations to study (Pierce, 2011; Redcay, Kennedy, & Courchesne, 2007) and precluding true random sampling of individuals. Further, from an embodied cognition perspective, lying motionless and flat on one's back may dampen some of the psychological processes that are of greatest interest to social and personality psychologists, such as approach motivation (Harmon-Jones & Peterson, 2009), and inhibiting head motion (which is typically desired to be at 2mm or less) may lead to reductions in emotion experience (Davis, Senghas, Brandt, & Ochsner, 2010). Researchers could pre-test designs in a realistic way (e.g., using a mock scanner setup) to ensure their manipulations are successful.

Data acquisition

Participants are situated in the scanner in a supine position. Stimuli are presented via special scanner-safe goggles or a rear-projection system and headphones. Many scanning centers require participants to wear earplugs to dampen the noise, even if they are also wearing headphones. Depending on the task, participants may also hold a response device in one or both

hands such as a button box or joystick. Scanning centers frequently take a variety of precautions to reduce head motion during the scan, ranging from foam padding and inflatable cushions to bite bars and custom face masks that strap to the scanner bed. During the scan, the experimenter remains immediately outside of the scanning room in a control room, and can communicate with the participant through a microphone connected to the participant's headphones. Direct auditory communication with participants is typically limited during the scan because of scanner noise, so participants are given an emergency call exit button in case they need to discontinue the scan for any reason.

Once situated in the scanner, participants' brain activation during the experimental tasks is measured indirectly by the fMRI scanner. Neural activity increases local blood flow, leading to increases in the oxygenated-to-deoxygenated hemoglobin ratio that perturbs the local magnetic field, and it is these small fluctuations in the local magnetic field that are detected in fMRI. Though the BOLD signal is an indirect measure of neural activity, there is now considerable evidence using a combination of optical imaging (Kennerley, Berwick, Martindale, Johnston, Papadakis, & Mayhew, 2005) and direct electrical recording (Viswanathan & Freeman, 2007) that the BOLD response is tightly coupled to synaptic activity and particularly to electrical local field potentials (Goense & Logothetis, 2008).

Though there are now dozens of ways to use MRI scanners to examine changes in blood flow and volume, nearly every social neuroscience study employs whole-brain echo-planar imaging, which refers to the particular sequence of spatially-varying magnetic fields and radio frequency pulses used by the fMRI scanner to collect the data. The considerable advantage of this technique is that it can be used to acquire an entire two-dimensional brain image in as little as 50 milliseconds, enabling whole-brain scanning (i.e., three-dimensional images) in under two

seconds. Each two dimensional image, also known as a "slice," typically has a resolution of 64x64, containing a total of 4,096 volume elements, or "voxels." As shown in Figure 3, whole-brain images are comprised of a series of slices that are acquired either sequentially (i.e., from bottom-up or top-down) or interleaved (i.e., slices 1, 3, 5, etc., then slices 2, 4, 6., etc), and then stitched together to form a single three-dimensional image (or "volume"). A typical volume contains 30-35 slices, totaling 120,000 to 150,000 voxels. In the language of MRI, the time between subsequent volumes is known as the repetition time, or TR. Ultimately, a raw fMRI data set consists of a time-series of whole-brain images acquired at each TR (usually once every two seconds or so) for the duration of the task (usually 5-10 minutes, or 150-300 scans).

The gain of such rapid scanning comes at a cost. Echo-planar imaging is extremely sensitive to distortions, or inhomogeneities, in the magnetic field. These distortions are particularly strong at boundaries between tissue and air-filled cavities, and cause susceptibility artifacts in the images – regions where the signal is poor or unusable. Unfortunately, some of the regions that are of greatest interest to social neuroscientists such as orbitofrontal cortex (OFC; Beer, John, Scabini, & Knight, 2006), the temporal poles, and the medial temporal lobe including the amygdala (Olson, Plotzker, & Ezzyat, 2007), are prone to susceptibility artifacts because of their proximity to the nasal cavity and the auditory canal, respectively. Fortunately, several methods have been developed recently to minimize signal loss in these regions (e.g., using an oblique axial-coronal acquisition angle; Weiskopf, Hutton, Josephs, Turner, & Deichmann, 2007), and are increasingly adopted in social neuroscience (e.g., Hare, Camerer, & Rangel, 2009; Xu, Monterosso, Kober, Balodis, & Potenza, 2011).

Functional imaging data are nearly always collected in tandem with at least one structural image that has relatively higher resolution (e.g., voxel dimensions of $1mm^3$ for structural vs.

$3mm^3$ for functional images). These images serve as a reference map to help localize activations identified using the functional images and for presentation purposes, which is why they are often described as "anatomical" images. For example, the "blobs" of activation found using the EPI images are often displayed overlaid on a group average structural image (see Figure 4a), which helps readers identify the brain regions involved. Using a group average overlay for presentation is preferred to using a template (Figure 4b), which is often more aesthetically pleasing but misleading regarding the true resolution obtained in the study. Though there are dozens if not hundreds of varieties of structural images available, the magnetization-prepared rapid gradient echo, or MP-RAGE, is most commonly used in social neuroscience studies because of its high resolution, excellent gray-to-white matter contrast, and fast acquisition time (Brant-Zawadzki, Gillan, & Nitz, 1992).

Data cleaning and preprocessing

Following acquisition, fMRI data must go through a number of cleaning, shuffling, and normalizing steps collectively known as "preprocessing" before they are ready for analysis (Figure 5). Preprocessing is computationally elaborate, memory intensive, and still under active mathematical and theoretical development. Because of these challenges, the number of software packages available to preprocess and analyze fMRI data is ever increasing. The most common are Statistical Parametric Mapping (SPM), Functional MRI of the Brain Centre Software Library (FSL), Analysis of Functional NeuroImages (AFNI), and FreeSurfer, which are all free, and BrainVoyager, which is proprietary. Most labs use SPM, FSL, or a combination of those two (See Poldrack, Mumford, & Nichols, 2011, for a comparison). There are many others. To complicate matters further, the consensus within the research community about how best to preprocess and analyze data is still in flux, and each of these packages is constantly updated to

reflect these changing opinions. Also, each software package preprocesses raw data in a slightly different order. Thus, instead of describing any one piece of software or processing algorithm in detail, we will focus on the basic steps of analysis that are implemented in most or all of the packages (though not necessarily in the order described here). The specific software and computational methods will surely change in the coming years, so our aim is to explain the purpose of each step more than to describe how that step is calculated.

Raw data conversion. Data arrive off of the scanner in a raw file format known as DICOM. There is one DICOM file for each whole-brain volume and several for the anatomical scan. Each software package contains its own tool for converting the DICOMS to a more usable format, and there are several excellent stand-alone utilities freely available (e.g., MRIConvert). Most packages now use the a format developed by the Neuroinformatics Technology Initiative (NifTI) that allows for a time-series of whole-brain volumes from a task to be grouped into a single 4-dimensional file (i.e., three spatial dimensions plus sequences of volumes over time).

After conversion, it is prudent to make a quality check of some kind, either by scanning through manually or using an automated tool. The main sources of contamination at this stage are participant motion and data spikes from ambient electromagnetic noise. Both of these may be correctable, but doing so requires extra steps beyond the standard preprocessing stream. This is important to note, as many of the tools allow start-to-finish batching of an entire preprocessing and analysis stream without forcing any data quality inspection. Without visual inspection, it is easy for noisy or distorted raw data to sneak by unexamined.

Fieldmap correction. It is now common to begin preprocessing by correcting for distortions in the magnetic field. This can be done using a set of “field map” scans that contain a vector field of local magnetic inhomogeneities at each voxel. As noted above, attempting to

correct for these inhomogeneities is particularly important if the hypotheses involve brain regions that are highly susceptible to dropout such as the OFC. Though the correction is not perfect, it can restore a considerable amount of signal, particularly if participant motion is low and a separate field map is obtained immediately before or after the functional run of interest (Hutton, Bork, Josephs, Deichmann, Ashburner, & Turner, 2002).

Realignment. Next, researchers typically correct for motion between functional volumes and changes in timing within them. These steps are known as "realignment" and "slice-timing correction," respectively. Realignment uses a six-parameter rigid body transformation (i.e., motion along the *x*, *y*, and *z* axes and rotation of pitch, roll, and yaw without warping or stretching) to quantify and subsequently correct for motion from volume to volume, similar to stacking a deck of cards into a neat pile after shuffling them. Slice-timing correction adjusts the slices within each volume to account for changes in signal that may have occurred within the TR, and is generally only necessary in studies where precise timing is critical such as event-related designs or those measuring functional connectivity. There is some debate regarding the best order to compute realignment and slice-timing correction as they are mutually dependent. For this reason, the field is moving toward consensus that an algorithm that computes them simultaneously is needed (Smith, Jenkinson, Woolrich, Beckmann, Behrens, Johansen-Berg, et al., 2004).

Coregistration. The purpose of the next step, "coregistration," is to move the anatomical and/or the (now realigned) functional scans in space so they are perfectly overlapping with one another. (Note that the term "registration" is sometimes used to refer collectively to what we call "coregistration" and "realignment". However, we will maintain a distinction between them throughout.) If realignment is akin to stacking a deck of cards, then coregistration is like placing

the cut card on top; the cut card is qualitatively different than the other cards in the deck, but it has the same dimensions and must be oriented in exactly the same way as the other cards to be of any use. The algorithm used for coregistration is the same as the one used for realignment—the purpose is still to correct for motion—but coregistration adjusts for movement that occurs between the functional and structural scans. Considering this, it makes sense to compute coregistration separately from realignment because there can be a considerable amount of time between the functional and structural scans, depending on the experimental protocol. For example, the task of interest might take place at the very beginning of the scanning session, followed by several other tasks lasting 10-15 minutes each, with the anatomical scan at the end. Coregistration must account for the cumulative motion within each scan and also between scans (when participants are often encouraged to move in order to prevent within-task movement). Thus, acquiring the functional and anatomical scans in close proximity or minimizing between-task movement can facilitate high-quality coregistration.

Reorientation. Realignment and coregistration aim to put all of the acquired scans into the same space as one another; the next step, "reorientation," is the first step toward placing the scans into a common space that can be used to generalize results and communicate them to other scientists. After conversion from DICOM files, the coordinate structure of the files is arbitrary. After successful realignment and coregistration, a given *x,y,z* coordinate on one image within the experiment will correspond to the same brain location on all the other images within the experiment but will not necessarily be the same as any other data set. And after reorientation, the brain images will all be angled and shifted in the same, standardized way. Specifically, the origin of the coordinate space (i.e., $x = 0$, $y = 0$, $z = 0$) will have been moved to a piece of white matter called the anterior commissure, the *x-y* plane will be parallel to the anterior commissure –

posterior commissure (AC-PC) line, the *y-z* plane will delineate the interhemispheric fissure, and the *x-z* plane will be perpendicular to the other two at the AC (Figure 6). For now, reorientation must be done manually because it involves mapping points in space to specific anatomical structures, and thus is highly labor intensive. Nonetheless, it is worth the effort because it can greatly improve normalization (the next step), which assumes that the images have been reoriented to the AC-PC line with the origin at the AC.

Normalization. Normalization adjusts for the simple fact that different people's brains are shaped differently. Group inference requires a way of averaging across people that preserves specificity of neural regions. For example, even though one subject's anterior cingulate might not be located at the same coordinates as another, we still want to average the activation in the anterior cingulate across those subjects without mixing in nearby regions. One solution to this problem is to specify *a priori* a canonical brain with a known mapping between each point in space and anatomical landmarks (e.g., everyone agrees that the coordinates [18, -4, -20] on this brain correspond to the right amygdala). Then, find a way to map each participant's brain to this canonical brain. Once such a map is found, it becomes possible to generate a new version of the participant's brain that is in the canonical space. "Normalization" refers to this process of mapping then re-drawing a brain into a common space.

Normalization is typically done into either Talairach space (Talairach & Tournoux, 1988), or, more commonly, into Montreal Neurological Institute (MNI; Evans, Kamber, Collins, & MacDonald, 1994) space. The MNI atlas is generally preferred because it provides a probability atlas based on 152 brains (versus just one brain described in detail in the Talairach atlas), and there are several versions available for developmental populations (e.g., Fonov, Evans, Botteron, Almli, McKinstry, Louis Collins, et al., 2010). Other advantages are that most software packages

contain several high resolution template brain images in MNI space, and there are an increasing number of databases available for labeling brain structures based on MNI coordinates (e.g., Fischl, Salat, Busa, Albert, Dieterich, Haselgrove, et al., 2002; Shattuck, Mirza, Adisetiyo, Hojatkashani, Salamon, Narr, et al., 2008).

Spatial smoothing. The last step of preprocessing is spatial smoothing, or blurring the brain images based on pre-set radius around each voxel. Smoothing accomplishes several things, including reducing outlier voxels (because they are blurred with nearby voxels), accounting for small differences in normalization between subjects, improving detection of activations larger than the smoothing kernel, and introducing a spatial correlation into the image (which is a prerequisite for some statistical methods). Smoothing is conceptually akin to averaging several related items on a personality questionnaire to get a better estimate of the true effect because the errors cancel each other out. The radius and the intensity of the smoothing is determined by a "smoothing kernel," which typically follows a Gaussian distribution with a full width at half maximum (FWHM, or the width of the curve at half of its maximum value) of 4-8mm. It is recommended to make the smoothing kernel no larger than the smallest hypothesized activation, as larger kernels will decrease power to detect activations smaller than the kernel.

Data analysis

Fully preprocessed images are entered into statistical analysis first on a subject-by-subject basis ("first-level" models) and then as a group for population inference ("second-level" models). This approach can be thought of as a proto-multilevel model where two levels (within- and between-subjects) are estimated separately (Schoemann & Little, ch x, this volume). In this section, we will describe each of the two levels, the standard analyses that are computed within

this framework, and some alternative models that have been adopted to address specific research questions.

First-level models. The subject-level models capture the entire time-course of the task by breaking it down into component events (i.e., trials, rest periods, etc.). For example, the blocked design in the top of Figure 2 could be described as: condition "A" beginning at seconds 10 and 130 and lasting 30 seconds each, condition "B" beginning at second 70 and lasting 30 seconds, and fixation baseline all other times. Even more complicated designs (e.g., the rapid event-related design at the bottom of Figure 2) can be described using lists of onsets and durations for each condition (called "vectors" in matrix algebra). In most software packages, any period of time that is not explicitly modeled is automatically included in a so-called "implicit baseline" condition, which is often used as a low-level comparison condition. For this reason, it is important (and sometimes forgotten!) to explicitly model every non-baseline event, even those of no interest such as instruction periods. The on-off time-course of each condition is then convolved with a canonical hemodynamic response function to create a predicted BOLD time-course of a voxel that responds to only that condition, and those time-courses are entered as regressors in a multiple regression model predicting the observed BOLD response (Figure 7). First-level models frequently also include several "covariates of no interest" that control for potentially confounding factors such as subject motion or peripherial physiological measures.

The solution to this regression model, which includes predictors for each condition, run, and covariate, is then estimated accounting for the auto-correlated nature of the errors, producing one beta weight for each condition. This approach is known as a "massively univariate general linear model" (Bowman, Guo, & Derado, 2007; Monti, 2011) because a regression model is created and solved at each voxel independently. The resulting beta weights (and corresponding t-

values) are stored in three-dimensional maps representing the association between the predicted and observed BOLD response at each point in the brain.

Task effects are assessed via "contrasts" between conditions. This subtractive approach is necessary for neuroimaging based on the BOLD response because the raw data values are on an arbitrary scale, and thus only interpretable relative to one another. Contrasts in fMRI are exactly like those used in the context of ANOVA by social and personality psychologists in that they are weighted combinations of means from various conditions, only with beta values as the dependent measure in each condition. For example, in a simple experiment with two conditions, A and B, if the beta weight at one voxel were 0.8 for A and 0.3 for B, the parameter estimate of the contrast, [1 -1], which compares the model fit between the two conditions in that voxel, would be 0.5. A pairwise *t*-test can be used in this way to test for reliable differences between conditions. The vast majority of social neuroscience studies that examine the main effect of task conditions (or other comparisons that are based on within-subjects ANOVA such as interactions or custom linear combinations of conditions) use this *t*-test approach. For example, this strategy has been used to identify the neural systems involved in interpreting a series of cartoon panels in terms of mental state inference or physical motion (Brunet, Sarfati, Hardy-Bayle, & Decety, 2000).

"Conjunction" analysis is based on *t*-tests between conditions, as it involves examining the overlap between two or more pairwise *t*-tests. In a study with 4 conditions (e.g., a classic 2x2), conjunction identifies regions that are significantly more active during condition A vs. B *and also* significantly more active during condition C vs. D. This complements interaction analysis, which asks whether a region is differentially active during condition A vs. B *to a greater extent than* during condition C vs. D. Thus, conjunction is often used together with interaction in the same set of analyses. For example, McRae and colleagues used both

conjunction and interaction analyses to identify the overlapping and distinct neural systems involved in two distinct forms of emotion regulation, distraction and reappraisal (McRae, Hughes, Chopra, Gabrieli, Gross, & Ochsner, 2010). First, distraction and reappraisal were each contrasted in a pair-wise *t*-test to a "view negative" condition, generating Distraction vs. View and Reappraisal vs. View contrasts. A conjunction analysis between these two contrasts identified brain regions that were involved in both forms of emotion regulation, and an interaction analysis, (Distraction vs. View) vs. (Reappraisal vs. View), identified brain regions that were differentially involved in one form greater than the other.

Another common analytic strategy is known as "parametric modulation," and is used to compare varying degrees of treatment effect within one condition. This is qualitatively different than the *t*-test approaches, because instead of comparing two or more conditions to each other, parametric modulation weighs each trial *within a single condition* according to some parameter (e.g., reaction time or Likert rating), hence the name, and compares trials that are higher on that parameter to trials that are lower on that parameter. This analysis is independent from between-condition contrasts, and can be used to complement them or to replace them entirely depending on the design. A number of studies on reward anticipation and responsivity use this approach to identify regions that are sensitive to increasing amounts of reward (e.g., where \$3 > \$2 > \$1, etc.) in addition to contrasting reward to loss (Eisenberger, Berkman, Inagaki, Rameson, Mashal, & Irwin, 2010; Knutson, Adams, Fong, & Hommer, 2001; Tom, Fox, Trepel, & Poldrack, 2007).

The last major class of within-subject models is used to pinpoint connectivity between regions, either at rest or as a function of condition. These models are alternately called "functional connectivity" or "effective connectivity," and there are some subtle differences between them (Friston, 2009) that are not relevant for the present chapter. In resting state

connectivity, the observed BOLD response at rest from a "seed" region is used as the predictor in each subject's first-level design file (instead of a set of condition predictors). The resulting beta map estimates the strength of the linear relationship between the seed region and all other regions in the brain. This approach has been instrumental in identifying the so-called "default mode network" of regions that is preferentially active at rest in humans (Uddin, Kelly, Biswal, Castellanos, & Milham, 2009). In functional (or effective) connectivity, the BOLD response from a seed region is combined with task condition predictors to create a seed-by-condition interaction term (i.e., a psycho-physiological interaction, or PPI; Friston, Buechel, Fink, Morris, Rolls, & Dolan, 1997). This interaction term is the predicted BOLD response of a region that is more correlated with the seed region during some conditions than others. PPI is ideal to assess functional connectivity because it controls for resting state and anatomical connectivity to identify only regions whose correlation with the seed region changes as a function of task. This kind of analysis has been used in social neuroscience to find regions that are inversely related to emotion response areas (e.g., the amygdala) during emotion regulation but not other conditions (e.g., Berkman, Burklund, & Lieberman, 2009).

Second-level models. Once first-level models have been specified and estimated and contrasts have been generated for each subject in the sample, summary statistics for each subject (e.g., mean contrast values or mean connectivity estimates) are brought to a separate, group-level, model for a random effects analysis. These summary statistics often take the form of a whole-brain map of contrast values at each voxel (i.e., one map per subject per contrast), but can also be subsets of the whole-brain map known as "regions-of-interest," or ROIs. Besides representing more spatially targeted hypotheses, the main advantage of analyzing ROIs at the second level is

that they require fewer statistical comparisons compared to a whole-brain map. (See the section on statistical thresholding below).

The main inferential tests at the group level are *t*-tests and correlations. Even though fMRI analysis always involves comparing means of different trial types by subtracting one from the other (i.e., in a pair-wise *t*-test) , this subtraction is always done at the subject level; at the group-level, there is usually only one data point per participant (per voxel). Thus, one-sample *t*-tests are used to compare the group mean contrast value (e.g., condition A beta - condition B beta) to the null hypothesis value of 0. Independent-samples *t*-tests are used when comparing different groups such as smokers to non-smokers (Galvan, Poldrack, Baker, McGlennen, & London, 2011) or children to adolescents (Pfeifer, Masten, Borofsky, Dapretto, Fuligni, & Lieberman, 2009). Correlation/regression is used when there is a subject-level individual difference or behavioral measure that is hypothesized to relate to brain activity. In these types of second-level models, the correlation across subjects is estimated between the contrast value and the measure at each voxel, producing a map of where in the brain the difference between conditions is related to the behavioral measure. Though often called a "brain-behavior correlation," this approach is more accurately described as a moderation or interaction because it tests for regions where the difference between conditions varies as a function of the behavioral measure. For example, some of the authors recently found that a region of the frontoparietal attention network was active overall during mindfulness meditation vs. control (i.e., a main effect), and that trait mindfulness moderated activity in a separate region such that this region was more active during mindfulness meditation vs. control only for highly mindful people (i.e., a task-by-trait interaction; Dickenson, Berkman, Arch, & Lieberman, in press).

Statistical thresholding. Each of the statistics described above is computed at every voxel in the brain (or within the hypothesized ROI), and some of these voxels are likely to evince a large test value by chance. For example, a typical map of contrast values contains 64x64x34 = 139,264 voxels that are each 3x3x3mm, though many of them are outside the brain or contain white matter or cerebrospinal fluid which are not expected to show blood flow differences between condition. (Only about 53,500 are likely inside gray matter.) Even if there is no difference between the conditions in the contrast, without eliminating these extraneous voxels and using an α-level of .05, this analysis is expected by chance to identify 6,963 voxels as "significant" (i.e., commit Type 1 error in 6,963 voxels). As the main problem is the large number of statistical comparisons, one approach would be to use a traditional adjustment such as a Bonferroni correction, but these are overly conservative (e.g., Bonferroni would suggest a voxel-wise alpha threshold of 3.59×10^{-7} in this case, which would eliminate even the most reliable activations known to neuroimaging such as visual cortex activity during visual stimulation). Another approach is to use ROI analyses to reduce the total number of comparisons, but this denies researchers the opportunities to make discoveries that were not hypothesized. The challenge in statistical thresholding of whole-brain fMRI data is to find a way to account for multiple comparisons in a way that balances Type I and Type II error rate (Lieberman & Cunningham, 2009).

There are now at least a dozen types of statistical thresholding ranging from more conservative to more liberal, and more surely will be developed in the coming years. For example, familywise error rate (FWE) restricts the probability of finding *any* false positives, is relatively conservative, and includes Bonferroni correction and random field theory (Nichols & Hayasaka, 2003), whereas false discovery rate (FDR) controls the fraction of detected voxels that

are false positives (Benjamini & Hochberg, 1995; Genovese, Lazar, & Nichols, 2002), is less conservative, and includes some simulation methods (Forman, Cohen, Fitzgerald, Eddy, Mintun, & Noll, 1995) and cluster-level correction approaches (Chumbley & Frison, 2009). There is even now a "threshold-free" method of selecting significant clusters that appears to be highly sensitive without inflating the Type I error rate (Smith & Nichols, 2009), but it is not yet implemented in most software packages. All of these methods are actually more conservative than the typical approaches in social psychology publications (Lieberman & Cunningham, 2009). There is consensus in the field that some correction for multiple comparisons is needed—particularly for whole-brain approaches—but there is not yet an optimal solution. (Using ROIs reduces this problem, but there is simply insufficient data on many topics at this early point in the field to make strong *a priori* predictions about where activation is likely to be for a given psychological process.) For now, researchers and reviewers must use their judgment whether the level of thresholding is too liberal, or too conservative, based on the potential value of false positives relative to missed true effects.

Reporting standards

Poldrack and colleagues have written the definitive standard for reporting a cognitive neuroscience fMRI study (Poldrack, Fletcher, Henson, Worsley, Brett, & Nichols, 2008). As this applies just as well to social neuroscience, we will not reiterate it here but instead highlight a few of the authors' recommendations and also mention some others that are particularly relevant to social and affective neuroscience. The two main types of standards that are relevant across all the cognitive neurosciences are those that ensure unbiased reporting of results and those that facilitate meta-analysis (see Johnson & Eagly, this volume).

No paper can report the entirety of the large corpus of data generated in an fMRI study. How does one select which analyses to report, and at which threshold? The best solution is to be as forthcoming as possible in describing all conditions in the design, and all comparisons among them that are relevant to the hypotheses. If a comparison is relevant to the hypothesis, then it should be reported regardless of the results. For example, if a study on emotional reactivity included conditions with positive and negative emotional stimuli, as well as neutral stimuli, then the authors should report a set of orthogonal contrasts that completely describes the variability between the conditions. Furthermore, it is not appropriate to omit conditions that were in the design from the report because all epochs from a task must be included in the statistical model. In the emotion reactivity example, for instance, it would be inappropriate to omit the positive emotion condition from the report even if (maybe especially if) there were no brain regions that showed a significant difference between it and either of the other two conditions. Not all comparisons must be featured in a figure, but all suprathreshold clusters from all hypothesis-relevant comparisons must be reported in a table. Statistical thresholds for whole-brain analyses should be decided in advance of the study (and hence described in the Methods section) and applied equally to all comparisons. If the authors have specific regional hypotheses, then ROI analyses should be used instead with standard thresholds (e.g., $\alpha < .05$ for one ROI, with appropriate corrections for multiple comparisons for >1 ROI), and supplemented with whole-brain analyses using a different threshold. We note that the lower threshold for ROIs is only appropriate if (1) the exact size, shape, and location of the ROI is specified in advance and (2) the "selective averaging" ROI approach is used in which all voxels within the ROI are included equally in the analyses (i.e., no within-ROI search for activation).

Meta-analysis is particularly important in social neuroscience given the relatively young age of the field and the unknown Type I error rate (see Johnson & Eagly, this volume). At this stage in the science, researchers have been understandably more interested to make new discoveries than to protect against potentially spurious ones, with the explicit understanding that true effects will emerge with replication and in meta-analysis. It is thus especially important to ensure that all activations are reported in a way that makes them accessible to meta-analysts. In addition to clearly describing each condition and reporting all contrasts among them as described above, this also requires reporting the coordinate system and normalization template, locating each cluster with coordinates in that system, providing the size of each cluster (in mm^3 or in voxels and per-voxel size in mm), labeling each cluster with an anatomical region, and describing the anatomical localization procedure (e.g., which atlas was used). Many researchers also report putative Brodmann's areas (BAs), with the caveat that this atlas is based on cytoarchitectural structure and cannot be directly inferred from fMRI data. (For example, there are boundaries between BAs within macroanatomical gyrii that cannot be identified using even the most high resolution anatomical scans available.) A good compromise is to report Brodmann's areas and complement them with descriptive labels generated by "tedious neuroanatomy" (Devlin & Poldrack, 2007), the invaluable process of sitting down and manually identifying each locus of activation.

Finally, there are a few reporting issues that are specific to social and affective neuroscience. Because scan time is so expensive, researchers often include several different tasks within a scanning protocol. The other tasks in a protocol might be irrelevant in some fields (e.g., vision), but can be critical in social neuroscience because our psychological processes of interest can be highly sensitive to context and recent events (e.g., ego depletion effects; Baumeister &

Heatherton, 1996, or priming; Tulving, Schacter, & Stark, 1982). For this reason, social neuroscientists should consider reporting the other tasks completed in the same scanning session that might affect performance on the task of interest. A related issue is participant expectations, which can also be influenced by contextual effects. For example, attributions about the cause of social rejection can alter its neural response (Masten, Telzer, & Eisenberger, 2011), and attributions can be influenced by subtle contextual information such as the likelihood of future interactions or whether the participant believes the interaction partner saw him/her. Hence, as in other types of social psychology research, these small but potentially powerful details must be included in the research report.

Electroencephalogram (EEG) / Event Related Potentials (ERP) Methods

As noted earlier, the primary weakness of fMRI methods are their inability to track the time course of neural activity, and many theories in social and personality psychology are not explicit about where psychological processes occur, but about when they occur and in which order. For example, several important models of prejudice regulation propose that automatic stereotypes and affect are activated automatically, and that controlled processes may inhibit or control this impulse (Devine, 1989). Although the processes of automatic activation and later control may be localized in different brain regions, models such as these are inherently more about time than they are about space. In contrast to fMRI, electroencephalogram (EEG) methods, allow for millisecond-by-millisecond recording of the electrical activity caused by neural activity. The temporal accuracy of EEG is possible because changes in postsynaptic potentials (when neurotransmitters bind to receptors) following neural activity create immediate changes in measureable electrical signals on the scalp (Buzsaki, Traub, & Pedley, 2003). The electrical signals, unlike the BOLD signal, are a direct measure of neural activity *at the time that activity*

occurred. Thus, EEG recordings contain the exact information that is lost in fMRI recordings – when the processes occur, and in what order. As models of social cognition begin to more fully embrace dynamical systems models of cognition rather than rigid dual-process models (see Cunningham & Zelazo, 2007), such information will be essential for providing more nuanced information regarding the interactions among multiple component processes.

Yet, this temporal precision gained with EEG comes at a cost. Whereas fMRI can tell us where a process is likely to have occurred, but not when, EEG methods can tell us when a process occurred but not where. Multiple factors contribute to this loss of information. First, neural activity is summed across multiple foci (or "generators") inside the brain by the time that this information reaches the scalp. The signal coming from a right anterior electrode is not necessarily, or even primarily, coming from right prefrontal cortex. A left parietal area may be primarily responsible for this effect, if the electrical activity is oriented toward the right frontal regions. Furthermore, if multiple regions are involved in a process, these processes are summed across all the electrodes. Second, the scalp itself blurs the signal in space such that the electrodes, even when using a high-density system (128 or more channels), do not provide independent signals. Although new methods for determining the spatial location of EEG signals are being developed (Pascual-Marqui, Michel, & Lehmann, 1994; Scherg & Ebersole, 1993), they suffer from the fact that a near infinite number of solutions can be generated when attempting to isolate the location of the EEG signal – there is no unique solution (Plonsey, 1963). Thus, EEG provides quantitatively different information about neural activity than fMRI. Because we must choose to sacrifice knowing where for when, or when for where (leading to something analogous to the Heisenberg Uncertainty Principle in physics), the choice to use fMRI or EEG should come from

theoretical questions, rather than ease or cost.[1]

Study Design

Whereas the electroencephalogram (EEG) is the raw time series of electrical activity recorded (see Figure 8A), an event related potential (ERP) is the averaged neural activity associated with a particular trial type. Specifically, an ERP reflects the averaged time locked signal following neural activity. Given the oscillatory nature of EEG recording, ERPs typically have multiple deflections, some with a more positive voltage than baseline and some with a more negative voltage than baseline. Unlike fMRI, where a positive signal reflects activation, and a negative signal reflects deactivation, the direction of the signal (positive or negative) is relatively meaningless for EEG. Rather, the strength of the absolute effect, positive or negative, is meaningful. To calculate an ERP, one creates *epochs* (time windows that correspond to specific trials) from the raw time series data to particular time points (e.g., when the stimulus was presented, or when a button press was recorded) and lines up all the EEG data in time for each of the predefined trial types. These time points are then averaged to create a *grand average* for each trial type for each participant that can be compared across condition (See Figure 8B for an example). Although this averaging can distort the individual trial-to-trial variability, or the participant-to-participant variability, this is important because it increases the signal to noise ratio and allows for meaningful comparisons to be made. Deflections, or ERP components or waveforms, are typically labeled with a letter and a number. The letter corresponds to whether the deflection was positive or negative, and the number corresponds to when the component occurs. The numbers can reflect either the number of deflections since baseline (e.g., a P1 would

[1] Of course, it is possible and necessary at times to run a study twice. Once as an fMRI study and once as an EEG study. Further, new technologies are being developed to collect EEG and fMRI data simultaneously, and statistical methods are being developed to fuse the two types of data together (see Moosman, Eichele, Nordby, Hugdahl, & Calhoun, 2008).

reflect the first positive deflection and an N3 would reflect the third negative deflection) or the time in milliseconds from the baseline (e.g., a P100 would be a positive deflection around 100ms after baseline, and an N170 would be a negative deflection 170ms following baseline). Other ERPs have specific names, such as the ERN (error related negativity) that follows behavioral mistakes (Gehring & Fencsik, 2001), or the LPC (late positive component associated with explicit recognition memory (Rugg et al., 1996).

In one of the first social psychological studies using EEG, Cacioppo, Crites, and Gardner (1996) identified an ERP associated with of valenced stimuli presented in an emotionally incongruous context. A series of valenced stimuli were presented before a critical stimulus that was of the same or different valence, and EEG signals that differentiated the stimuli presented in congruous versus incongruous contexts were examined. Cacioppo and colleagues identified a particular type of wave form termed a *late positive potential* (LPP) when participants saw a stimulus that was incongruous with a context; in these studies, a negative stimulus in the context of positive stimuli, or a positive stimulus in the context of negative stimuli. The amplitude of the LPP wave in these studies was shown to vary as a function of the degree of difference between the valence of the stimulus and the valence of the context in which it occurs. For example, when presented in the context of positive stimuli, a strongly negative stimulus will result in a larger LPP than a mildly negative stimulus (Cacioppo et al., 1996; Cacioppo, Crites, Gardner, & Berntson, 1994). The LPP associated with evaluative incongruity is widely distributed across scalp electrodes but is more pronounced over posterior (parietal) scalp regions than over frontal sites. There is also evidence that the amplitude of this posterior LPP is greater over the right hemisphere than over the left for both positive and negative stimuli presented in an incongruous evaluative context (Cacioppo et al., 1996). Researchers using this paradigm have shown that the

posterior LPP is evident when participants are making both evaluative and nonevaluative judgments, suggesting that evaluative incongruity may be detected automatically (Cacioppo et al., 1996; Ito & Cacioppo, 2000; see also Crites & Cacioppo, 1996). LPPs show a negativity bias in that they are typically larger for negative stimuli in a positive context than positive stimuli in a negative context (Ito, Larsen, Smith, & Cacioppo, 1998), and the degree of hemispheric asymmetry (right greater than left) is greater for negative stimuli as well (Cacioppo et al., 1996).

Like fMRI studies, EEG/ERP studies need to examine changes in neural activity as a function of a presumed change in perceptual or cognitive processing. The brain is always "active," in the sense that neurons throughout our brains fire spontaneously even when we are resting, and so it is only possible to quantify changes in brain activity. To examine these changes, it is necessary to have at least one within-subject condition that varies as a function of some psychological state. This variation can come from the manipulation of a categorical variable (pleasant vs. unpleasant photographs) or by manipulating a variable continuously (degrees of valenced photographs). As with all neuroscience methods and as discussed previously, it is important to ensure that either only one variable is being manipulated, or that appropriate control conditions have been added to statistically control for potential confounds. Further, because of the relatively low signal to noise ratio for neuroscience methods, it is necessary to have many trials of data. Because of these similarities, rather than repeat those design issues here, we highlight in this section the unique challenges that emerge when collecting ERP data and provide suggestions about how to optimally design studies to minimize these concerns (for a more detailed and exhaustive discussion of EEG methods, see Luck, 2005).

The goal of ERP data analysis is to characterize a neural response in time. As noted above, the advantage of the EEG method is its remarkable precision in detecting when a

particular neural response occurs. This precision, however, means that additional assumptions must be made during the study design and analysis, and it is possible that some questions simply cannot be answered using ERP methodology. The fundamental issue is that to average multiple trials into a single ERP waveform, we must assume that the processes occur at the same time point. For example, an early visual component in the occipital electrodes will be found if, for most trials and for most participants, there is a positive deflection approximately 100ms following stimulus presentation (see Figure 9A). However, note what happens if for some trials, the effect is offset by intervals of 40ms (Figure 9B). When these trials are averaged, the positive and negative deflections of the underlying neural process cancel each other out, leaving no effect whatsoever (see Figure 9C).

What these examples illustrate is that not only must it be assumed that the brain regions that produce the effect are the same for all trials and all participants, but that the timing of the brain activity must also be consistent. Because of this, it should not be surprising that many, if not most, ERP studies have examined low level perceptual process or motor responses. In both cases, it is more likely that the neural responses will "line up" in exactly the same way. For a perceptual process, there is a ballistic sequence of events that needs to automatically decode the stimulus properties to create a representation of an object. The ERPs following the stimulus presentation are likely very consistent from trial-to-trial for the first 100s of milliseconds. For a motor response, we know when the response occurred, so we can look at the ERPs that occur just before or just after the response. Again, we can assume that these responses are likely consistent for the 100s of milliseconds before or after the response.

Although the time locking to either a stimulus or response allows for an ability to understand the components of many automatic processes, this presents challenges for the study

of some constructs within social and personality psychology. For processes such as Theory of Mind, social rejection, or intertemporal discounting, it is unlikely that there are automatic sequences of processes events that (a) unfold in exactly the same way for all trials within subject (not even considering between subject difficulties), and (b) that we can a priori identify the temporal onset and offset of processing. For example, when deciding whether someone knows where an object is hidden, there is no reason that a person cannot consider another person's intention at 500ms, or 1200ms, or even 3000ms after the information is presented. Further, for a complicated decision, people can consider the various components of the information in different orders – in a gambling task, a person on one trial can consider the amount of money that can be earned ($50 vs. $1) before the probability of winning (60% vs. 20%), and on another trial can consider the probability before the magnitude. In this example, if amount of time spent considering the magnitude may vary as a function of the size of the potential win (i.e., if a participant spent more time thinking about $50 than $1), then these timings would further become blurred in time. Thus, as cognitive flexibility and reflective processing increases, we should see an exponential increase in the blurring of ERPs. Considering the importance of time locking (to the millisecond!) neural processes, it should not be surprising that most ERP effects discussed in the literature occur within the first hundreds of milliseconds before or after the time locked trial.

Data collection

Among the most important aspects of EEG/ERP data collection is to collect as clean data as possible. EEG signals from neural activity are small (less than 1/100,000th of a volt), and have a very low signal to noise ratio, so maximizing the signal and minimizing the noise components are important. Before all else, it is important to prepare the room for data collection and remove

as much as possible of the extraneous sources of electrical activity. Many people build rooms specifically for EEG recording that isolate electrical signals – removing computers from the rooms, using DC instead of AC lighting, and so on. The fewer sources of electrical activity in the testing room, the better the EEG recordings of neural activity will be. Next, it is important to get the best signal from the EEG cap. To do this, either a gel or a saline solution is placed under the electrode to decrease electrical resistance (impedance) between the electrode and the scalp. Reducing the impendence of the electrical recording allows for better recording of the neural activity and reduces the recording of environmental noise. Gel solutions allow for lower impedances (better signal), but are messy and take time to apply. Saline-based solutions typically cannot get as low impedance, but are less messy and faster to apply. If time is not an issue, gel tends to be preferable to get a better signal, but because application can take over 30 minutes for some high density systems, a saline solution can be preferred if working with certain populations, such as children, where there may be limited patience or time to apply the gel.

After preparing the electrodes for recording, it is necessary to determine a reference electrode. EEG is not an absolute signal, but the difference in voltage between two recordings. Because the activity at the reference electrode will be subtracted from all recordings, it is desirable to select a reference that is as isolated from the signals of interest as possible. Two of the most common reference locations are behind the ears (mastoid reference; recording from both the right and left and averaging these together to create a reference), or on the nose. The selection of reference location is important because it will partially determine the shape and location of observed ERP signals. In other words, a P1 recorded in an occipital electrode using a mastoid reference is *not* the same as a P1 recorded from the same electrode using a nose reference because they examine the voltage differential across different locations and distances.

When comparing across studies, or attempting to replicate previous work, it is critical to note the location of the reference (see Joyce & Rossion, 2005, for an example).

Because precise stimulus timing is critical for ERP data analysis, the precision of the stimulus presentation package and the calibration of the stimulus monitor and response box is far more important than in fMRI data collection. For example, software for experimental presentation varies in the quality of the precision of timing, which can interact with the particular hardware that it is being run on. Further, computer hardware matters far more for ERP studies than fMRI. One has far less control of the presentation of visual stimuli on an LCD monitor than a CRT monitor, and most keyboards have an error of response in the 50ms range. When time locking to a neural response that occurs in the order of 10s of milliseconds, these errors can add substantial noise and variability to the recording.

Data averaging, cleaning and preprocessing

To calculate the grand average ERP for each electrode, a marker file must be obtained that precisely (to the millisecond) codes when events occurred in the continuous EEG. These markers can represent when a stimulus was presented (and/or when it was removed from the screen), when a response was made, or any other event that has psychological meaning. It is always better to include too many than too few markers. When computing the grand average ERPs, it is necessary to determine which of these markers will be used to align the individual trial data in time. If interested in perceptual processing, the stimulus marker will most often be used. If interested in motor responses or decision processes, the response marker will most often be used. It is important to note that the resulting grand average ERPs time locking to the stimulus or the response will be completely different, as the stimulus marker does not perfectly predict the timing of the decision processes and motor responses, and the response marker does not perfectly

predict the timing of the stimulus presentation. Indeed, because of the blurring of ERPs discussed in the previous section, very little motor or decision related activity will be found when time locking to the stimulus, and very little perceptual processing will be found when time locking to the response. Because the temporal dynamics are critical in ERP studies, decisions about time locking should be made relative to how theory predicts consistency in temporal responses across trials and participants.

Before averaging, it is important to clean the data to ensure that most of the variability is due to psychological processing rather than extraneous sources. This is done through two primary means: data filtering and trial rejection. EEG data is typically bandpass filtered prior to calculating the individual trial epochs and the grand averages. Typically, a 0.1 to 0.30Hz bandpass filter (only frequencies between 0.1 and 0.30Hz are retained) is applied to the data to remove most extraneous sources of noise. Tighter filters can be used for particularly messy data, such as when uncooperative participants or children are run. Ideally, though, you would want to keep as much data as possible, or at least be confident that nothing you exclude could be meaningful, task-related neural activity. Following bandpass filtering, trials that have too much artifact are typically removed from the analysis. The most common artifacts are eye blinks, eye movements, and head motions. These artifacts create huge changes in voltage that dwarf the voltage changes that come from neural activity. Luckily, these artifacts are well characterized and can be deleted manually by examining each individual epoch, or can be detected using automated algorithms. More recent developments have allowed for the estimation of these artifacts (e.g., epoching ERPs to participant blinks) that may allow for artifact correction rather than removal. Eye blinks and motion can be reduced by allowing participants' periods of time when they can blink frequently throughout the task and by designing tasks that require limited

visual search.

Data Analysis

Assuming a sampling rate of 500Hz, and a 128-channel EEG system, 64,000 data points are collected per second. Thus, if an epoch is defined as the 1000ms following a marker, we are left with tens of thousands of data points per condition – an amount of data that leads to a serious problem concerning the appropriate ways to correct for multiple comparisons. Assuming only two conditions, if a researcher tested each time point at each electrode for these two conditions, 3200 false positives would be found using a $p < .05$ alpha level.

Luckily, EEG data is not independent in either the temporal or spatial domains, and several conventions allow for a relatively straightforward way to reduce the number of multiple comparisons. Specifically, researchers often examine only one or two meaningful ERP components, and average the data across multiple time points to get a single estimate of the ERP effect. For example, if one is interested in the P1 visual component, because this effect typically occurs around 100ms following stimulus presentation, it can be quantified as the mean signal averaging from 90 to 110ms. Similarly, the N170 visual component can be isolated as the mean signal from 160 to 180ms. By focusing on only these two components (in this case known to be involved in face processing), a researcher can already reduce the data from 64000 time points to 256. Not only does this reduce the number of comparisons, but because the EEG signal itself is not independent, it also helps increase the signal to noise ratio by averaging data that presumably capture the same information.

In addition to collapsing in the time domain, it may be beneficial to collapse in the spatial domain. In the faces processing example above, only some electrodes may be meaningful to examine. For example, given previous research, posterior electrodes are more likely to provide

more meaningful signal than more anterior electrodes when examining visual processing. Because of this, a researcher can focus on a subset of electrodes (perhaps 20 in the left hemisphere and 20 in the right hemisphere), thereby even further reducing the number of comparisons. Given a strong enough a priori hypothesis, there is no reason to go beyond the examination of a single electrode. For example, Amodio and colleagues (2004) focused on how people respond when making errors associating prejudicial objects such as guns with African-American faces. Specifically, participants were more likely to classify a tool as a gun following an African-American face than a Caucasian Face. Because they were interested in error processing, they focused on one particular ERP, the event-related negativity (ERN; Gehring & Fencsik, 2001; van Veen & Carter, 2002), which was associated with cortical activity after detecting errors. Because the ERN was well characterized in previous years, they were able to focus on a single frontocentral electrode (the Fcz) for all analyses and avoid problems of multiple comparisons. Similarly, researchers interested in face processing will tend to focus on the P100 and the N170 components time locked to the stimulus presentation, as these components have been shown repeatedly in the literature to be associated with face processing, and they may focus more on posterior lateralized electrodes (Benton et al., 1996). Thus, as in fMRI, the ways in which one reduces the data are dependent on the specific question being asked and the previous literature in this domain.

Once the data reduction strategy has been determined, ERPs are extracted from the single subject grand averages. There are two primary methods to quantify the size of the ERP component. One can determine the peak amplitude by identifying the highest (or lowest) point, and labeling that the size of the ERP effect. Although this approach has intuitive appeal, it can underestimate the size of an ERP for subjects (or conditions) where there is greater variability in

temporal dynamics of the ERP component. For example, if in one condition, the peak occurs at 100ms, ± 10ms, and in another condition the peak was at 100ms, ± 20ms, the fact that the peaks are more spread out in the second condition would reduce the observed average effect size even if the trial by trial peak amplitudes were identical. To circumvent this problem, it is preferred to calculate the mean amplitude across the entire ERP time window (the mean of time points 90-110ms). This method captures the average amplitude, and therefore the contributions of each of the independent trial amplitudes to the grand average.

If multiple electrodes are investigated simultaneously, it is standard to include electrode site and each of the ERP components as factors in an ANOVA design. So, if someone were interested in the P100 and the N170 mentioned earlier to be sensitive to faces, and wanted to compare upright vs. inverted faces, they may have to compare 20 left hemisphere and 20 right hemisphere posterior electrodes on a high density EEG system. To analyze, this would result in a a 2 (upright/inverted) X 2 (P100/N170) x 40 (electrode) MANOVA. Because the electrode sites are not independent from one another (here, the right electrodes are likely to be artificially more correlated with one another than they are with the left electrodes), it is necessary to perform a Greenhouse-Geisser correction any time that more than two electrodes are entered into the model. Because of this, if one believes that the right and left effects are for the most part isomorphic, one can simply average these effects to create a 2 (upright/inverted) X 2 (P100/N170) x 2 (left/right) MANOVA.

Analysis of Continuous EEG

Up to now, we have focused on the examination of single ERPs following stimulus presentation or a response. Although these ERPs are useful for investigating automatic processes, they have limited utility for examining cognitive processes that may not follow a stereotyped

temporal pattern. Although this is a limitation of ERP data collection process and analysis, the brain dynamics associated with more complex cognitive processes are recorded in the continuous EEG signal and can be recovered using different methods. One such technique is to compute the power in different frequency bands across several seconds (and often minutes) of EEG data. Power simply reflects the degree to which there is a large amount of activity within a particular frequency band. Greater power in these different bands at different electrode sites reflect different degrees of processing, *in general*, across tasks or people.

One of the best examples of using continuous EEG to examine affective processing comes from the work on frontal asymmetries for approach or avoidance motivation. Following the literature on emotional processing following brain injury, Davidson and his colleagues predicted that the right hemisphere may be more associated with negative affect and the left hemisphere with positive affect. To test this idea, Tomarken, Davidson, Wheeler, and Doss (1992) recorded EEG while participants were at rest, with the hypothesis being that greater activation in the right compared to left hemisphere would be found *across the entire resting epoch* for people who were predisposed to depression. To analyze the continuous EEG across the entire time window, the data were decomposed into different frequency ranges and the power of these ranges were estimated. By comparing the power estimates from right and left hemisphere, they found that differences in the alpha band (8 to 12 Hz) predicted affective style. Specifically, greater right-sided power is associated with greater negative symptoms. More recent research has suggested that these frontal asymmetries may be more associated with approach (left) vs. avoidance (right) motivation (Harmon-Jones, Lueck, Fearn, & Harmon-Jones, 2006). Regardless of the specific mechanism, frontal EEG power asymmetries at rest have been shown to predict depression, emotion regulation ability, and general well-being (Davidson, 1988; Jackson,

Mueller, Dolski, Dalton, Nitschke, Urry, et al., 2003; Urry, Nitschke, Dolski, Jackson, Dalton, Mueller, et al., 2004).

Eight conceptual issues in social neuroscience (and how to think clearly about them)

1. Forward and reverse inference

In neuroimaging, typically psychological processes are the independent variables and neural activity is the dependent variable. The *forward* direction of logical inference, therefore, is from the psychological to the neural: if psychological Process A is engaged and Region X is activated, then we can conclude that Region X is involved in Process A in some way (though not necessarily causally). Neuroimaging works this way because the methods were optimized to answer questions about localization of function in the cognitive neurosciences, for example, to identify regions in the brain that are responsive to motion perception. The vast majority of studies in social neuroscience still take the approach of manipulating a psychological process to the end of localizing the neural systems that are recruited during that process, and the methods described thus far in the chapter are valid inferential tools for this purpose.

However, it is tempting to do the reverse by inferring the presence of a psychological process based on neural evidence (e.g., concluding that motion perception must have been engaged based on activation in a particular area). This is known as *reverse inference*, and is often, though not always, an instance of the logical fallacy of affirming the consequent (Poldrack, 2006). The fallacy is that "B --> A" does not logically follow from "A --> B" when the relationship from psychological process A to neural activation B is not unique (i.e., when process A is not the only one that is associated with activation in B). And nearly all of the time, it is not. Indeed, the mapping from mental processes to neural activations seems to be *many-to-many* in that most brain regions are involved in many psychological processes, and vice versa. Thus, the

only logical conclusion that can be made based on observing activation B is that one of the many psychological processes that have been observed to be associated with B (or one that is not yet known to be) were engaged. We assume you agree that this is not a satisfying answer.

What is needed is a comprehensive database of social neuroscience data that tracks the mapping between cognitive processes and brain activation in order to estimate the probability of a psychological process given activation in a region (Poldrack, 2010). Bayes' Theorem provides a way to do this, but requires some base rate frequency information. Bayes' Theorem states that in order to make the reverse inference that a mental process was engaged based on a neural activation, then it is necessary to know the probability of activation in that region during the mental process, the probability of the mental process being elicited in the given task (which should be close to 1 if the task is valid), and the probability of activation of that region across all mental processes. In cases where the mental process is highly specific to a given region—that is, the probability of activation is high in the process of interest but low in others—then reverse inference is more likely to be valid. However, in cases where the mental process is not specific to a given region—that is, the probability of the activation is high regardless of the mental process—then reverse inference is less likely to be valid. The insular cortex is an example of a brain region falling into this later category, as activation there is observed across a broad variety of tasks (Kurth, Zilles, Fox, Laird, & Eickhoff, 2010). Thus, it is not valid to infer a specific mental process based on observed activation in the insula alone—a mistake that has been made at least once in the New York Times (Poldrack, 2011).

Fortunately, researchers increasingly recognize the need for tools to enable them to gauge how specifically a region is involved in a task. There are several such databases including NeuroSynth.org (Yarkoni, Poldrack, Nichols, Van Essen, & Wager, 2011), the Cognitive Atlas

(Poldrack, Kittur, Kalar, Miller, Seppa, Gil, et al., 2011), and the NeuroLex neuroscience lexicon (Bug, Ascoli, Grethe, Gupta, Fennema-Notestine, Laird, et al., 2008). Though the use of these systems is not yet standard in the field and many of them are still in beta, we anticipate that they will be commonly used within the next few years to support reverse inference claims with Bayesian data.

A second way to deal with reverse inference is to focus on networks of regions rather than isolated regions. Although a given region may show activity in response to multiple psychological demands, a set of brain regions may co-activate under a more selective set of psychological conditions. For instance, ventromedial PFC is a commonly activated region in a number of qualitatively different tasks. When seen alone it can be difficult to interpret. However, when seen in conjunction with ventral striatum and ventral tegmental area, it is more likely that this three-region network indicates some form of reward processing or valuation compared to when any one of those regions is seen alone.

2. Spuriously high correlations?

Social neuroscience has recently received criticism for reporting strong correlations between brain activation and individual differences (e.g., behavioral or self-report measures) that seemed to some to be "likely entirely spurious" (Vul, Harris, Winkielman, & Pashler, 2009). The correlations in question are indeed high, but the critics are incorrect that this implies the correlations must be meaningless. The flaw in their argument was misidentification of the source of the high correlations to be circular analysis (Lieberman, Berkman, & Wager, 2009); the high correlations are actually a result of the stringent thresholding procedures used to protect against Type I error. (We note that this criticism applies to cognitive neuroscience in general, which uses

identical methods as those described here, but the original critique focused on social neuroscience.)

The main assertion of the critics is that many correlations in social neuroscience are the product of circular statistics, or "double-dipping," and are therefore not valid (Kriegeskorte, Simmons, Bellgowan, & Baker, 2009). The claim is that some scientists were analyzing their data by first searching the whole brain for voxels that correlate with a measure, and then were selecting only those voxels that correlated highly with the measure for entry into a *second* analysis. We agree with the critics that such a second round of analysis would be circular and would not produce meaningful results. However, the problem with their critique is that nobody (or perhaps very few) researchers actually analyze their data in this way. Correlations in social neuroscience are generally conducted as described in the section above on second-level models: a whole-brain search is conducted (i.e., at every voxel) for correlations between contrast estimates and an individual differences measure, and those voxels surviving the multiple-comparisons threshold are reported. There is exactly one inferential step—the whole-brain search—and the reporting following that step is merely descriptive, regardless of whether a single voxel or an average across a cluster of voxels is reported. We note that the procedure for whole-brain correlations is exactly the same as for whole-brain condition effects (e.g., comparisons between two conditions), except the statistic computed at each and every voxel is a correlation instead of a *t*-test.

However, the correlations remain strikingly high—one study reported a correlation of 0.96 between activation in the middle frontal gyrus and a reaction time measure (Sareen, Campbell, Leslie, Malisza, Stein, Paulus, et al., 2007). What can explain this if not double dipping? The cause is small sample sizes combined with multiple comparisons procedures that

rely on a voxel-wise threshold. Supposing the common voxel-wise threshold of $p < .001$ and a sample size of 15, the corresponding correlation threshold (i.e., the lowest *r* value that could be considered significant) is 0.73. For a threshold of $p < .0001$, the threshold *r* is 0.82. And those correlations are the *minimum* values that can be reported with those thresholds—the tails of the distribution of values can be much higher in the presence of a true effect. The correlation values in the voxels that contain a true effect will be distributed around the true correlation according to sampling error, and one or more voxels in that population that emerge with a very high correlation (e.g., if the true correlation is 0.7 in two voxels, one might be observed at r=0.5 and another at r=0.9). Thus, no misapplication of statistical procedures is necessary to explain the high correlations—they are a direct result of the thresholding procedures that explicitly limit researchers to report only high correlation values.

The benefit that has resulted from this critique and the responses to it is an increased awareness of the limitations of whole-brain statistics, or at least to the limitations of the thresholding procedures that are necessitated by whole-brain statistics. The first lesson is to use larger samples. The relationship between a *p*-value threshold (e.g., <.001) and a statistic threshold (usually a *t* or *r* value) changes as a function of the degrees of freedom at the group level (Figure 10a). For a given statistical threshold, the Type I error rate decreases dramatically as the sample size (and thus the degrees of freedom) increases, particularly until N=20 (Figure 10b). The second lesson is that our raw statistics values are not unbiased measures of effect size (nor were they intended to be) because of the thresholding procedures (Lieberman et al., 2009). Only after an effect is detected at all (which *is* the intention of the procedures) can an effect size be estimated for an entire functional or structural brain region by examining the correlation across that entire area. Another lesson is to use *a priori* ROIs to compute correlations when

possible (e.g., Berkman & Lieberman, 2010), which can eliminate the need for multiple comparisons corrections in the first place and simultaneously provides effect *detection* (i.e., present or absent) and effect size *estimation*. Finally, as is always the case with empirical findings, only results that have been independently replicated should be considered reliable.

3. Experimental vs. ecological validity

In any psychology experiment, there is a tradeoff between experimental control of extraneous third variables or outside influences and ecological validity in realistically modeling complex human thoughts, feelings, and behaviors (Brewer & Crano, ch. 2, this volume). This tension is present in social neuroscience as much or even more so than in social psychology because the cost of null effects is quite high in terms of money and other resources. On the one hand, a key advantage of neuroimaging relative to other dependent measures is the ability to examine putative neural mechanisms of psychological processes. The best way to leverage this advantage is to examine as "pure" a process as possible by carefully controlling everything else—internal validity. On the other hand, a core aim of social neuroscience is to understand the brain systems that support real life thoughts, feelings, and behaviors, and these cannot always be broken down into a single mental process that can be easily manipulated in isolation—ecological validity. To be realistic, we must also relax some experimental control. What should the balance be in social neuroscience?

In our view, the field of social neuroscience to date has been biased toward ecological validity, in part because one advantage of whole-brain neuroimaging is that it can assess multiple processes in disparate brain regions simultaneously, and in part as a way of differentiating itself from cognitive neuroscience. There are several good reasons that justify this bias. One reason is that social neuroscience is not concerned merely with process but also with outcome—beliefs,

emotions, and behaviors that occur outside of the scanner (Berkman & Lieberman, 2011), such as relationship quality or addiction relapse. Outcomes like these cannot generally be explained by one process alone, so complex scanner tasks are needed to begin to get a handle on which brain systems contribute to them and how. Another reason is simply the relative youth of the field compared to other neuroscience fields. Social neuroscience is still in the discovery phase—the very first neuroimaging experiments in a number of areas (e.g., empathy, cognitive dissonance, emotion regulation) were conducted only within the last few years. Zeroing in on processes first requires some knowledge about which processes to examine. For example, researchers needed to conduct preliminary studies on the broad network involved in thinking about other people and their mental states (Frith & Frith, 2001; Gallagher & Frith, 2003) before breaking that network into its component pieces (e.g., Saxe & Powell, 2006; Spunt & Lieberman, 2012). We believe that, at this point, social neuroscience can make its most important contributions to psychological science by bringing real psychological phenomena into a neuroimaging environment where their corresponding brain systems can begin to be explored.

To the extent that experimental control is sacrificed for ecological validity, then the degree of ecological validity should be as high as possible. It can be tempting to simplify an experimental paradigm to a pallid and repetitive response time task, and some believe that psychology in general has been guilty of yielding to this temptation too often (Baumeister, Vohs, & Funder, 2007). Social neuroscientists have embraced the importance of creating an engaging, realistic scanner experience, which partly explains the popularity of certain tasks that reliably evoke strong feelings of, for example, rejection (Williams, Cheung, & Choi, 2000), fairness/unfairness (Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003), and disgust (Ochsner, Bunge, Gross, & Gabrieli, 2002). This same desire for realistic neuroimaging experiences has

motivated researchers to think outside of the task itself and manipulate the imaging environment by bringing in interaction partners either physically (Coan et al., 2006) or digitally (Redcay, Dodell-Feder, Pearrow, Mavros, Kleiner, Gabrieli, et al., 2010), or manipulating the temperature of tactile stimuli during the scan (Kang, Williams, Clark, Gray, & Bargh, 2011). We greatly anticipate seeing clever new ways for researchers to continue to bring the world into the scanning environment in the coming decades.

4. What counts as replication in fMRI and EEG?

Though test-retest reliability can be quite high both within- (Friedman, Stern, Brown, Mathalon, Turner, Glover, et al., 2008) and between-scanners (Casey, Cohen, O'Craven, Davidson, Irwin, Nelson, et al., 1998), there can be considerable variability in contrast-to-noise ratio, percent signal change, and spatial normalization across scanners and individuals for a variety of reasons that are beyond the scope of this chapter but are discussed extensively elsewhere (e.g., Bennett & Miller, 2010; Jovicich, Czanner, Greve, Haley, van de Kouwe, Gollub, et al., 2006). As a result, it can be hard to determine whether a particular voxel in one subject is located in the exact same brain region as the same voxel in another subject. And even if they are located in the same brain region, two subjects may activate the same region to differing extents for idiosyncratic reasons. By extension, the presence of activation in one voxel in a sample and the absence of activation in that same voxel in another sample doesn't necessarily imply a lack of replication across the two samples.

Achieving the scientific ideal of replication would seem hopeless based on this lack of voxels-wise concordance across studies. Fortunately, most social neuroscientists are not concerned about voxel-level replication. Instead, our interest is in understanding the brain at the level of functional regions—the chunks of gray matter that are consistently and coherently

involved in a given psychological process. Neighboring voxels, particularly those within the same anatomical structures, tend to perform a similar function, so exact voxel-wise replication is not as important as functional region-wise replication (which is another reason to praise "tedious neuroanatomy"). But this just punts the problem up a level of analysis: if exact voxel-wise replication is not possible, then exact region-wise replication might be challenging as well. How can researchers know which task-related activations replicate and which do not, even at the level of the cluster instead of the voxel?

The answer to this question is highly dependent on how "activation" is defined across studies. As we have described above, researchers make a large number of decisions in the course of preprocessing and analyzing their data that could potentially influence the precise size and location of their observed clusters. There are several existing tools to measure the amount of overlap in activation across studies ranging simple voxel counting methods (Cohen & DuBois, 1999) to more formal statistical methods (Nichols, Brett, Andersson, Wager, & Poline, 2005), and all of these must be interpreted with the following idea in mind. Even in the ideal case in which the *mental processes* are replicated exactly in two studies, the amount of replication in the *observed data* will vary as a function of fit in design, acquisition, and analysis. These features include, but are by no means limited to, design factors such as the level of complexity and abstraction of the task (see Point #5 below; Plichta, Schwarz, Grimm, Morgen, Mier, Sauer, et al., in press), whether the design is event-related or blocked (Bennett & Miller, 2010), and the subject population (Bosnell, Wegner, Kincses, Kortewek, Agosta, Ciccarelli, et al., 2008); acquisition factors including the field strength (Hoenig, Kuhl, & Scheef, 2005), scanning parameters (Bandettini, Wong, Jesmanowicz, Hinks, & Hyde, 1994), and thermal noise (Bodurka, Ye, Petridou, Murphy, & Bandettini, 2007); and analysis factors, mainly the thresholding

procedure (Bennett, Wolford, & Miller, 2009), but also the hemodynamic basis functions (Lindquist & Wager, 2007) and the smoothing kernel (Mikl, Marecek, Hlustik, Pavlicova, Drastich, Chlebus, et al., 2008).

Given the complexity of the dependent measure and the chain of decisions that must be made between when participants complete the task and when the data are reported, it seems amazing that fMRI results ever replicate at all. Nonetheless, using the intraclass correlation coefficient as the metric of replication (Caceres, Hall, Zelaya, Williams, & Mehta, 2009), several studies have observed reliabilities of 0.9 or greater (Aron, Gluck, & Poldrack, 2006; Raemaekers, Vink, Zandbelt, van Wezel, Kahn, & Ramsey, 2007). Intraindividual reliability tends to be quite high (Miller, Donovan, Van Horn, German, Sokol-Hessner, & Wolford, 2009), and the majority of variance in fMRI studies has been shown to be from between-subjects variance (Costafreda, Brammer, Vencio, Mourao, Portela, de Castro, et al., 2007). It follows that if variation in between-subjects factors (e.g., early life stress; Taylor, Eisenberger, Saxbe, Lehman, & Lieberman, 2006) is minimized then neural activations may replicate quite faithfully at the level of functional regions. Indeed, there are now several meta-analyses in social and affective neuroscience that point to areas reliably identified for processes of interest such as emotion (Kober, Feldman Barrett, Joseph, Bliss-Moreau, Lindquist, & Wager, 2008), mentalizing (Spreng, Mar, & Kim, 2009; Van Overwalle, 2009), and social cognition more generally (Amodio & Frith, 2006). Further approaches to accounting for between-subject variability are described below in Point #5.

5. Why is there greater fundamental variability in social neuroscience data?

It has been noted in the reviews cited above (e.g., Bennett & Miller, 2010) that some of the tasks used in social neuroscience tend to have high within- and between-subject variability.

In general, we agree, and would even conjecture that the source of this variability is that the mental processes studied in social neuroscience—and not just the tasks—are inherently more variable than those studied in other cognitive neurosciences. One reason for this is that social neuroscience tasks often involve manipulations of abstract concepts (e.g., one's own long-term goals or the emotional reactions of other individuals), so people are likely to vary from one another and across time in the strategies they take to engage in the tasks. We refer to these ideas as the *anteriorization-abstraction hypothesis*: that the axis that runs from the posterior to the anterior parts of the brain (i.e., from the back of the brain to the front) reflects increasing levels of abstraction of mental representation, and that activation changes from focal to diffuse in reliability along that axis. This diffuse activation for abstract representation reflects the high level of cognitive flexibility involved in that process. For example, viewing a picture of a given person relative to considering the mental state of that person can be expected to produce a more posterior/caudal pattern of activation that is more reliable over time and across people.

The anteriorization-abstraction hypothesis suggests some challenges for social neuroscience. The main challenge is that statistics that involve central tendency (e.g., means) are less powerful toward the abstract/anterior end of the gradient. And naturally, our standard statistical tools that are based on the GLM depend on high consistency within-subjects (at the first-level) and also between-subjects (at the second-level) to detect effects. Thus, more trials and more subjects are often needed to obtain effects compared to more "reliable" processes. Also, the aim of localization of functions (even abstract ones) can be challenging to meet when, by definition, some functions are more diffuse in their localization than others. What does it mean that top-down control is instantiated in the prefrontal cortex, for instance, if every person uses a slightly different part of the cortex to engage in control and a given person might use a different

part at different times? The answer to this question is related to how to define replication in social neuroscience studies (see Point #4 above), and also, in part, begins to undermine the granularization of function in more anterior parts of the brain. It may be the case that the prefrontal cortex is organized into larger functional units than more posterior regions, and that any small piece of it has no consistent, specific role beyond top-down control of any or all bottom-up systems (e.g., Heatherton & Wagner, 2011; Miller & Cohen, 2001; Munakata, Herd, Chatham, Depue, Banich, & O'Reilly, 2011).

So as not to sound too grim, there is also a large upshot that follows from the anteriorization-abstraction hypothesis: more variance means more variance to be explained in a statistical sense. At the within-subjects level, this means explaining how neural processes change across time, for example as a function of changing cognitive strategies over time (Kross, Davidson, Weber, & Ochsner, 2009) or engaging in a process for a sustained versus a brief period (e.g., Somerville, Wagner, Wig, Moran, Whalen, & Kelley, in press). At the between-subjects level, this means identifying factors that moderate the variation in brain activity across people. This is one reason why subject-level correlations are so appealing: they open a window to explain neural activation using idiographic differences as an alternative to mean tendency. In the future, we anticipate that more sophisticated multilevel modeling of fMRI data will enhance our ability to explain within- and between-subjects variance, and even the cross-level interactions between the two.

An issue related to increased variance in social neuroscience processes is the notion that the tasks used in social neuroscience are often not *process pure*, meaning that they are comprised of a number of mental processes that cannot be uniquely isolated from one another in terms of their neural underpinnings. For instance, some studies analyze epochs of six or more seconds of

thinking in a focused way (e.g., emotion regulation or mindfulness meditation), and experimenters simply have little control over exactly what subjects do and when they do it. Even with the most careful instructions and manipulation checks, we can never know for sure exactly what was happening inside our participants heads or whether they were doing the same thing as one another. (Additionally, even if they honestly told us what they believed they were doing, we know that many mental processes occur outside of awareness.) Experimenters can use *process analysis* (or *cognitive ontology*; Bilder, Sabb, Parker, Kalar, Chu, Fox, et al., 2009) to carefully specify the component operations that are required by a task during the study design phase, and then leverage existing databases (e.g., The Cognitive Atlas; Poldrack et al., 2011) during the analysis phase to triangulate the mental process – neural activation mappings in their data.

6. The rhetorical power of neuroimaging data

We hope that the main reason for excitement about neuroimaging among lay readers and members of the media is the same as it is for members of the scientific community: that neuroimaging data can provide unique information about the brain systems involved in mental processes and provide insight into the nature of the mental processes themselves. But casual observers have posited, and data now support, that another reason is the compelling nature of the visuals themselves. For example, one study found that undergraduates rated the scientific reasoning of articles to be more sound when the data were overlaid on a high-resolution brain image compared to the *exact same data* overlaid on a topographical map (McCabe & Castel, 2008). Another study found that even undergraduates with training in neuroscience (but not neuroscience experts) found scientific results more satisfying when they were accompanied by irrelevant neuroscience explanations (Skolnick Weisberg, Keil, Goodstein, Rawson, & Gray, 2008). This result also held for logically flawed arguments, suggesting that the presence of

irrelevant neuroscience information can trump basic reasoning in the minds of lay readers. The authors of this latter study conjectured that the "seductive allure" of neuroscience studies is attributable in part to a cognitive bias toward reductive arguments of mental phenomena. Other factors could include a "technical language" bias (Shafir, Smith, & Osherson, 1990), a "seductive details" effect (Harp & Mayer, 1998), and a "placebic information" heuristic (Langer, Blank, & Chanowitz, 1978).

Whatever their cause, the seemingly mind-numbing effects of neuroscience data among non-experts should be concerning to those in the field. We highlight three main lessons for researchers that follow from the "seductive allure" effect. First, we need to do a better job educating our students about when neuroimaging data do and do not inform theory. The fact that Ivy League undergraduates who completed an intermediate-level neuroscience class could not differentiate when neuroimaging data were relevant or not is a dramatic illustration of this need (Skolnick Weisberg et al., 2008). Second, we must be exceedingly clear when speaking to members of the media about what the neuroimaging data show. It is difficult enough to get media to report any scientific results with fidelity, and we have good evidence now that this is even more the case with neuroimaging data. And third, in light of the finding that good scientific explanations with irrelevant neuroimaging data are perceived as *less* satisfying by experts (Skolnick Weisberg et al., 2008), we should be careful not to oversell these data in professional venues.

7. Brain as predictor: Correlation versus prediction

Traditional functional neuroimaging studies are designed and analyzed with mental process as the independent variable and brain activation as the dependent variable. In the language of regression, the brain data are the criterion (Y vector) and the task conditions are the

predictors (X matrix). This statistical model (i.e., forward inference) has proven to be enormously useful over the past decades for developing and refining an extensive body of knowledge on the mapping from mental process to neural activation. Now, scientists across disciplines are eager to build upon this corpus to use neuroimaging data to predict real-world outcomes beyond the laboratory, but the traditional design and analysis tools are insufficient for this purpose. This necessitates a model that flips the traditional one on its head by treating neural activation as the *independent* variable predicting outcomes beyond the scanner in a *brain-as-predictor* approach (Figure 11; Berkman & Falk, under review).

Traditional correlation approaches in social neuroscience (e.g., those described in Point #2 above) are limited because they model the brain as the dependent measure. That is, they answer the question, which parts of the brain are correlated with an outcome? The design of these studies also commonly measures the outcome concurrently or even before the brain data, which rules out prospective prediction. For example, Mehta and Beer (2010) found that activation in medial orbitofrontal cortex while receiving unfair (versus fair) offers in an ultimatum game was correlated with subsequent aggressive responses. Even though brain activation was measured before the behavioral outcome, the statistical model used was designed to predict neural activity based on the behavior, and not the other way around. Though this model is highly useful for identifying linear relationships (which is most often its intended use), it is not a valid tool for prediction when the brain data are the criterion (Gelman & Hill, 2007).

As its name implies, the brain-as-predictor approach models brain activation as the predictor and allows any outcome that occurs after measurement of the activation to be the criterion. For example, we used this approach to test whether activation in regions hypothesized to be involved in inhibitory control were predictive of self-control in daily life (Berkman, Falk,

& Lieberman, 2011). First, in a baseline session, we measured activation in three such regions during a classic inhibitory control task. These were *a priori* regions selected based on their past involvement in self-control studies. Then, we measured participants' self-control abilities during a cigarette smoking cessation attempt using a daily diary. With all of these data in hand, we entered them simultaneously into a multilevel model which revealed that increased activation in each of the three brain regions predicted a reduced relationship between cigarette cravings and smoking over the course of three weeks. In another example of this approach, Masten and colleagues used activation in the subgenual anterior cingulate cortex during social rejection to predict subsequent increases in depression among adolescents (Masten, Eisenberger, Borofsky, McNealy, Pfeifer, & Dapretto, 2011).

The brain-as-predictor approach is a useful tool for both theory testing and translational science. In terms of theory, this approach allows for a strong test of the hypothesized function of various brain regions, and can measure not just the presence of predictive validity but also strength. An example of where this might be useful is in testing the specific function of the medial prefrontal cortex, a region thought to be involved in self-processing and comprised of various sub-divisions with different hypothesized functions (e.g., Northoff, Qin, & Feinberg, 2011). A researcher interested in a stringent test of these functions could measure the activation in each of the regions and then measure real-world effects of various types of self-processing, for example how new college students select a major (academic self-concept) and how they make new friends (social functioning). The brain-as-predictor approach is also useful in translational settings because, once validated, brain activation could be used to predict clinical outcomes such as response to treatment or substance use relapse. As the cost of imaging goes down and the body of knowledge about brain function goes up, using functional neural activation to predict

individual clinical outcomes could be particularly helpful for disorders in which the current diagnostic tools are unavailable, limited, or highly costly.

The largest drawback of this approach is that the hypothesized predicted brain regions must be known in advance. In that sense, brain-as-predictor is not primarily a brain mapping approach; it doesn't search for an answer to "where in the brain?" questions (though it can help to confirm existing answers to these). Instead, it more directly addresses "does the brain?" questions, ones that hypothesize a particular function for a region and test whether activation in that region is a valid predictor of outcomes dependent on that function. To do so, it depends on whole-brain analyses that *do* answer "where in the brain?" questions for the purpose of hypothesis generation. The brain-as-predictor approach is complementary to traditional forward inference approaches (and distinct from reverse inference). Forward inference and brain-as-predictor are both needed to first identify candidate brain regions involved in a process, and then test the predictive validity of those regions outside the scanner (Figure 11). We hope that in the future scientists will conduct series of studies that capitalize on the strengths of both methods.

8. Mind reading?

One of the popular fantasies about neuroimaging is that it can be used to read people's thoughts. Put someone in the scanner or an EEG, the claim goes, and you can know the contents of his mind, including whether he is lying, what his specific memories look like, and even what he will do in the future, like in the movie *Minority Report*. Even though there are several for-profit companies that make these claims, we want to be as clear as possible that this kind of mind reading, which we'll call *strong mind reading*, does not yet exist, nor is it likely to exist in the foreseeable future. However, there are several kinds of reliable pattern classification that do exist, which we call *weak mind reading*.

Why is strong mind reading so far fetched? There are dozens of reasons, but we'll focus on just a few here. First, strong mind reading, if possible, could only occur with a willing participant. Being able to infer a specific mental state requires knowledge of other examples of that state for comparison, often very many of them, and very little useful data would come from someone unwilling to provide them (e.g., a criminal defendant). Second, there is too much variability in how a given mental state (e.g., a feeling of guilt) is represented in the brain both within and between individuals (see Point #5). Even with many examples of a specific mental state to work from (e.g., a memory), the reliability of matching a single new example of that state to previous ones would be quite low. And third, in the case of predicting future actions from current brain states, future behavior is multiply determined, and the extent to which it is determined by information currently in the brain is probably beyond the resolution of current scanners. Experience does change the brain, but these experiences take years to manifest in changes that are visible to our current technology (e.g., Luby, Barch, Belden, Gaffrey, Tillman, Babb, et al., 2012). Even if any single thought, feeling, or memory were visible, we wouldn't even know where to look at this point.

However, there are a few forms of weak mind reading that may be valid and useful. For instance, we know that decisions are sometimes driven not by consciously accessible thoughts but rather by implicit associations or learning that are not directly accessible to consciousness (Nisbett & Wilson, 1977). Neuroimaging might help explain some variability in decisions or behaviors that would not otherwise be easy to measure. For example, Falk and colleagues found that brain activation in a "neural focus group" predicted changes in population-level behavior (statewide increases in calls to a smoking cessation quitline) in response to health messages (anti-smoking advertisements) above and beyond self-reports about the messages (Falk,

Berkman, & Lieberman, 2012). This result suggests that factors that people could not or would not report directly, but were coded in their neural activation patterns, were driving their actual responses to the messages.

Another kind of weak mind reading involves measuring brain activation during many instances of a small number of thoughts (e.g., 100 repetitions each of the nouns "carrot" and "airplane") and then using pattern classification algorithms to categorize new thoughts (e.g., "helicopter") into one of the trained categories. A recent study used this technique to demonstrate a 75% accuracy rate on trained words, and about a 60% accuracy on untrained, but semantically related words (Mitchell, Shinkareva, Carlson, Chang, Malave, Mason, et al., 2008), and another used it to communicate with minimally conscious patients using a yes/no response (Monti, Vanhaudenhuyse, Coleman, Boly, Pickard, Tshibanda, et al., 2010). One group of researchers combined pattern classification with real-time analysis to successfully predict decisions in real time during an ultimatum game (Hollmann, Rieger, Baecke, Lutzkendorf, Muller, Adolf, et al., 2011). As elegant as they are, note that all of these studies require either many exemplars from the same individual, or make prediction from one individual to a large group of others (i.e., on average), but none can yield the content of a given thought without at least some prior information.

Conclusion

We have provided a survey of the neuroimaging methods used in social and affective neuroscience, and a discussion of eight of the current controversies and open questions in the field. Our aim was to make this chapter simple but comprehensive so that by its end a reader with no previous neuroimaging experience would understand the purpose of each of the steps from start to finish of a social neuroscience study, and even have a basic idea of how to begin

designing his or her own study. Another aim was to provide thorough explanations of the major ongoing controversies and questions in the field that would be satisfying for those outside it, and viewed as even-minded by those inside it. In sum, we hope that this chapter might serve as a comprehensive guide to social neuroscience for our colleagues in social and personality psychology, and also as healthy introduction to the field for students and faculty seeking to enter it.

In a word, we see the theme of the future of the field as "expansion": expansion in theoretical scope, expansion in methodological and statistical depth, expansion in breadth of impact across fields, and expansion in size. A core issue is how to inform social and personality psychology theory using neuroimaging data, and scientists in the field are making encouraging progress in exploring a number of avenues to do so (e.g., brain-as-predictor design and Bayesian meta-analysis). The appeal and accessibility of these methods will increase even more as social neuroscience becomes a standard part of graduate education in social and personality psychology programs. Social and affective neuroscience is also gaining popularity among related fields such as psychopathology, child development, and addiction, and will continue to do so as our methods advance. Finally, developing more powerful statistical and computational methods such as multilevel modeling and structural equation modeling for neuroimaging data will further expand the toolkit available to researchers to test hypotheses with more nuance than null hypothesis significance testing. We hope that this chapter will be a first step into the field for many of those who will contribute to its progress.

Acknowledgements

We are grateful to the UO Social Affective Neuroscience, Developmental Social Neuroscience, and Social Endocrinology Laboratories for helpful feedback on a previous version.

Support for this work was provided by University of Oregon startup funds to ETB and National Science Foundation (BCS-1122352 & BCS-0819250) to WAC.

Correspondence should be addressed to ETB, Department of Psychology, 1227 University of Oregon, Eugene, OR 97403.

References

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience, 7*, 268-277.

Amodio, D. M., Harmon-Jones, E., Devine, P. G., Curtin, J. J., Hartley, S. L., & Covert, A. E. (2004). Neural signals for the detection of unintentional race bias. *Psychological Science, 15,* 88-93.

Aron, A. R., Gluck, M. A., & Poldrack, R. A. (2006). Long-term test–retest reliability of functional MRI in a classification learning task. *NeuroImage, 29*(3), 1000-1006.

Bandettini, P. A., Wong, E. C., Jesmanowicz, A., Hinks, R. S., & Hyde, J. S. (1994). Spin-echo and gradient-echo EPI of human brain activation using BOLD contrast: A comparative study at 1.5. *Nuclear Magnetic Resonance in Biomedicine, 7*, 12-20.

Baumeister, R. F., & Heatherton, T. F. (1996). Self-regulation failure: An overview. *Psychological Inquiry, 7*(1), 1-15.

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior? *Perspectives on Psychological Science, 2*(4), 396-403.

Beer, J. S., John, O. P., Scabini, D., & Knight, R. T. (2006). Orbitofrontal cortex and social behavior: Integrating self-monitoring and emotion–cognition interactions. *Journal of Cognitive Neuroscience, 18*(6), 871-879.

Benjamini, Y., & Hochberg, Y. (2008). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological), 57*(1), 289-300.

Bennett, C. M., & Miller, M. B. (2010). How reliable are the results from functional magnetic

resonance imaging? *Annals of the New York Academy of Sciences, 1191*(1), 133-155.

Bennett, C. M., Wolford, G. L., & Miller, M. B. (2009). The principled control of false positives in neuroimaging. *Social Cognitive and Affective Neuroscience, 4*(4), 417-422.

Berkman, E. T., Burklund, L., & Lieberman, M. D. (2009). Inhibitory spillover: intentional motor inhibition produces incidental limbic inhibition via right inferior frontal cortex. *NeuroImage, 47*(2), 705-712.

Berkman, E. T., & Falk, E. B. (under review). Beyond brain mapping: Using the brain to predict real-world outcomes.

Berkman, E. T., Falk, E. B., & Lieberman, M. D. (2011). In the trenches of real-world self-control: Neural correlates of breaking the link between craving and smoking. *Psychological Science, 22*(4), 498-506.

Berkman, E. T., & Lieberman, M. D. (2010). Approaching the bad and avoiding the good: Lateral prefrontal cortical asymmetry distinguishes between action and valence. *Journal of Cognitive Neuroscience, 22*(9), 1970-1979.

Berkman, E. T., & Lieberman, M. D. (2011). What's outside the black box?: The status of behavioral outcomes in neuroscience research. *Psychological Inquiry, 22*(2), 100-107.

Bilder, R. M., Sabb, F. W., Parker, D. S., Kalar, D., Chu, W. W., Fox, J., et al. (2009). Cognitive ontologies for neuropsychiatric phenomics research. *Cognitive Neuropsychiatry, 14*(4-5), 419-450.

Birn, R. M., Cox, R. W., & Bandettini, P. A. (2002). Detection versus estimation in event-related fMRI: Choosing the optimal stimulus timing. *NeuroImage, 15*, 252-264.

Bodurka, J., Ye, F., Petridou, N., Murphy, K., & Bandettini, P. A. (2007). Mapping the MRI voxel volume in which thermal noise matches physiological noise—Implications for

fMRI. *NeuroImage, 34*(2), 542-549.

Bosnell, R., Wegner, C., Kincses, Z. T., Korteweg, T., Agosta, F., Ciccarelli, O., et al. (2008). Reproducibility of fMRI in the clinical setting: Implications for trial designs. *NeuroImage, 42*(2), 603-610.

Bowman, F. D. B., Guo, Y., & Derado, G. (2007). Statistical approaches to functional neuroimaging data. *Neuroimaging Clinics of North America, 17*(4), 441-458.

Brant-Zawadzki, M., Gillan, G., & Nitz, W. (1992). MP RAGE: A three-dimensional, T1-weighted, gradient-echo sequence--initial experience in the brain. *Radiology, 182*(3), 769-775.

Brunet, E., Sarfati, Y., Hardy-Bayle, M. C., & Decety, J. (2000). A PET investigation of attribution of intentions to others with a non-verbal task. *NeuroImage*, 11, 157-166.

Buckner, R. L., Bandettini, P. A., O'Craven, K. M., Savoy, R. L., Petersen, S. E., Raichle, M. E., et al. (1996). Detection of cortical activation during averaged single trials of a cognitive task using functional magnetic resonance imaging. *Proceedings of the National Academy of Sciences of the United States of America, 93*(25), 14878.

Bug, W. J., Ascoli, G. A., Grethe, J. S., Gupta, A., Fennema-Notestine, C., Laird, A. R., et al. (2008). The NIFSTD and BIRNLex vocabularies: Building comprehensive ontologies for neuroscience. *Neuroinformatics, 6*(3), 175-194.

Buzsaki, G., Traub, R.D., Pedley, T.A. (2003). The cellular basis of EEG activity. In J.S. Ebersole & T.A. Pedley (Eds.), *Current Practice of Clinical Electroencephalography, 3rd edition* (pp. 1-11). Philadelphia: Lippincott Williams and Wilkins.

Caceres, A., Hall, D. L., Zelaya, F. O., Williams, S. C. R., & Mehta, M. A. (2009). Measuring fMRI reliability with the intra-class correlation coefficient. *NeuroImage, 45*(3), 758-768.

Cacioppo, J. T., & Berntson, G. G. (1992). Social psychological contributions to the decade of the brain: Doctrine of multilevel analysis. *The American Psychologist, 47*(8), 1019-1028.

Cacioppo, J. T., Crites, S. L., Jr., & Gardner, W. L. (1996). Attitudes to the right: Evaluative processing is associated with lateralized late positive event related brain potentials. *Personality and Social Psychology Bulletin, 22*, 1205-1219.

Cacioppo, J. T., Crites, S. L., Jr., Gardner, W. L., & Berntson, G. G. (1994). Bioelectrical echoes from evaluative categorizations: I. A late positive brain potential that varies as a function of trait negativity and extremity. *Journal of Personality and Social Psychology, 67*, 115-125.

Casey, B. J., Cohen, J. D., O'Craven, K., Davidson, R. J., Irwin, W., Nelson, C. A., et al. (1998). Reproducibility of fMRI results across four institutions using a spatial working memory task. *NeuroImage, 8*(3), 249-261.

Chumbley, J. R., & Friston, K. J. (2009). False discovery rate revisited: FDR and topological inference using Gaussian random fields. *NeuroImage, 44*(1), 62-70.

Coan, J. A., Schaefer, H. S., & Davidson, R. J. (2006). Lending a hand: Social regulation of the neural response to threat. *Psychological Science, 17*(12), 1032-1039.

Cohen, M. S., & DuBois, R. M. (1999). Stability, repeatability, and the expression of signal magnitude in functional magnetic resonance imaging. *Journal of Magnetic Resonance Imaging, 10*, 33-40.

Costafreda, S. G., Brammer, M. J., Vêncio, R. Z. N., Mourão, M. L., Portela, L. A. P., de Castro, C. C., et al. (2007). Multisite fMRI reproducibility of a motor task using identical MR systems. *Journal of Magnetic Resonance Imaging, 26*(4), 1122-1126.

Crites, S. L., Jr., & Cacioppo, J. T. (1996). Electrocortical differentiation of evaluative and

nonevaluative categorizations. *Psychological Science, 7*, 318-321.

Cunningham, W.A. (2010). In defense of brain mapping in social and affective neuroscience. *Social Cognition, 28*(6), 717-722.

Dale, A. M., & Buckner, R. L. (1997). Selective averaging of rapidly presented individual trials using fMRI. *Human Brain Mapping, 5*, 329-340.

Davis, J. I., Senghas, A., Brandt, F., & Ochsner, K. N. (2010). The effects of BOTOX injections on emotional experience. *Emotion, 10*(3), 433-440.

Davidson, R. J. (1988). EEG measures of cerebral asymmetry: Conceptual and methodological issues. *International Journal of Neuroscience, 29*, 71-89.

Desmond, J. E., & Glover, G. H. (2002). Estimating sample size in functional MRI (fMRI) neuroimaging studies: statistical power analyses. *Journal of Neuroscience Methods, 118*(2), 115-128.

Devine, P.G. (1989). Prejudice and stereotypes: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56*, 5–18.

Devlin, J. T., & Poldrack, R. A. (2007). In praise of tedious anatomy. *NeuroImage, 37*, 1033-1041.

Dickenson, J., Berkman, E. T., Arch, J., & Lieberman, M. D. (in press). Neural correlates of focused attention during a brief mindfulness induction. *Social Cognitive and Affective Neuroscience*.

Eisenberger, N. I., Berkman, E. T., Inagaki, T. K., Rameson, L. T., Mashal, N. M., & Irwin, M. R. (2010). Inflammation-Induced Anhedonia: Endotoxin Reduces Ventral Striatum Responses to Reward. *Biological Psychiatry, 68*, 748-754.

Evans, A., Kamber, M., Collins, D., & MacDonald, D. (1994). An MRI-based probabilistic atlas

of neuroanatomy. In S. Shorvon, D. Fish, F. Andermann, G. Bydder & H. Stefan (Eds.), *Magnetic Resonance Scanning and Epilepsy* (NATO ASI Series A, Life Sciences ed., Vol. 264, pp. 263-274). New York: Plenum.

Falk, E. B., Berkman, E. T., & Lieberman, M. D. (2012). From neural responses to population behavior: Neural focus group predicts population level media effects. *Psychological Science*, *23*, 439-445.

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., et al. (2002). Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron, 33*(3), 341-355.

Fonov, V., Evans, A. C., Botteron, K., Almli, C. R., Mckinstry, R. C., Collins, D. L., & the Brain Development Cooperative Group. (2011). Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage*, *54*, 313–327.

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., & Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. *Magnetic Resonance in Medicine, 33*(5), 636-647.

Friedman, L., Stern, H., Brown, G. G., Mathalon, D. H., Turner, J., Glover, G. H., et al. (2008). Test–retest and between-site reliability in a multicenter fMRI study. *Human Brain Mapping, 29*, 958-972.

Frith, U., & Frith, C. (2001). The biological basis of social interaction. *Current Directions in Psychological Science, 10*(5), 151-155.

Friston, K. (2009). Causal modelling and brain connectivity in functional magnetic resonance imaging. *PLoS Biology, 7*(2), e1000033.

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.

Gallagher, H.L., & Frith, C.D. (2003). Functional imaging of 'theory of mind'. *Trends in Cognitive Sciences, 7*(2), 77-83.

Galván, A., Poldrack, R. A., Baker, C. M., Mcglennen, K. M., & London, E. D. (2011). Neural correlates of response inhibition and cigarette smoking in late adolescence. *Neuropsychopharmacology, 36*(5), 970-978.

Gehring, W.J., & Fencsik, D.E. (2001). Functions of the medial frontal cortex in the processing of conflict and errors. *The Journal of Neuroscience, 21*(23), 9430-9437.

Gelman, A., & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York, NY: Cambridge University Press.

Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage, 15*(4), 870-878.

Glover, G. H. (1999). Deconvolution of impulse response in event-related BOLD fMRI. *NeuroImage, 9*(4), 416-429.

Goense, J. B. M., & Logothetis, N. K. (2008). Neurophysiology of the BOLD fMRI signal in awake monkeys. *Current Biology, 18*(9), 631-640.

Grezes, J., & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: A meta-analysis. *Human Brain Mapping, 12*(1), 1-19.

Hallett, M. (2007). Transcranial magnetic stimulation: A primer. *Neuron*, *55*(2), 187–199.

Hare, T. A., Camerer, C. F., & Rangel, A. (2009). Self-control in decision-making involves modulation of the vmPFC valuation system. *Science, 324*(5927), 646-648.

Harmon-Jones, E., & Beer, J. S. (2009). *Methods in Social Neuroscience*. New York, NY: The

Guilford Press.

Harmon-Jones, E., Lueck, L., Fearn, M., & Harmon-Jones, C. (2006). The effect of personal relevance and approach-related action expectation on relative left frontal cortical activity. *Psychological Science, 17*, 434-440.

Harmon-Jones, E., & Peterson, C. K. (2009). Supine body position reduces neural response to anger evocation. *Psychological Science, 20*(10), 1209.

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A cognitive theory of interest in science learning. *Journal of Educational Psychology, 90*(3), 414-434.

Heatherton, T. F., & Wagner, D. D. (2011). Cognitive neuroscience of self-regulation failure. *Trends in Cognitive Sciences, 15*(3), 132-139.

Hoenig, K., Kuhl, C. K., & Scheef, L. (2005). Functional 3.0-T MR assessment of higher cognitive function: Are there advantages over 1.5-T Imaging? *Radiology, 234*(3), 860-868.

Hollmann, M., Rieger, J. W., Baecke, S., Lützkendorf, R., Müller, C., Adolf, D., & Bernarding, J. (2011). Predicting decisions in human social interactions using real-time fMRI and pattern classification. *PLoS ONE*, *6*(10), e25304.

Huettel, S. A., Song, A. W., & McCarthy, G. (2009). *Functional Magnetic Resonance Imaging* (2nd ed.). Sunderland, MA: Sinauer Associates, Inc.

Hutton, C., Bork, A., Josephs, O., Deichmann, R., Ashburner, J., & Turner, R. (2002). Image distortion correction in fMRI: A quantitative evaluation. *NeuroImage, 16*(1), 217-240.

Ito, T. A., & Cacioppo, J, T. (2000). Electrophysiological evidence of implicit and explicit categorization processes. *Journal of Experimental Social Psychology, 36*, 660-676.

Ito, T. A., Larsen, J. T., Smith, N. K., & Cacioppo, J. T. (1998). Negative information weighs

more heavily on the brain: The negativity bias in evaluative categorizations. *Journal of Personality and Social Psychology, 75*, 887-900.

Jackson, D. C., Mueller, C. J., Dolski, I., Dalton, K. M., Nitschke, J. B., Urry, H. L. et al. (2003). Now you feel it, now you don't: Frontal EEG asymmetry and individual differences in emotion regulation. *Psychological Science, 75*, 887-900.

Joyce, C.A. & Rossion, B. (2005). The face-sensitive N170 and VPP components manifest the same brain processes: The effect of reference electrode site. *Clinical Neurophysiology, 116*, 2613-2631.

Jovicich, J., Czanner, S., Greve, D., Haley, E., van der Kouwe, A., Gollub, R., et al. (2006). Reliability in multi-site structural MRI studies: Effects of gradient non-linearity correction on phantom and human data. *NeuroImage, 30*(2), 436-443.

Kang, Y., Williams, L. E., Clark, M., Gray, J. R., & Bargh, J. A. (2011). Physical temperature effects on trust behavior: The role of insula. *Social Cognitive and Affective Neuroscience, 6*(4), 507-515.

Kennerley, A. J., Berwick, J., Martindale, J., Johnston, D., Papadakis, N., & Mayhew, J. E. (2005). Concurrent fMRI and optical measures for the investigation of the hemodynamic response function. *Magnetic Resonance in Medicine, 54*(2), 354-365.

Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *Journal of Neuroscience, 21*(16), RC159.

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S. F., & Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nature Neuroscience, 12*(5), 535-540.

Kross, E., Davidson, M., Weber, J., & Ochsner, K. (2009). Coping with emotions past: the neural bases of regulating affect associated with negative autobiographical memories. *Biological Psychiatry, 65*(5), 361-366.

Kurth, F., Zilles, K., Fox, P. T., Laird, A. R., & Eickhoff, S. B. (2010). A link between the systems: functional differentiation and integration within the human insula revealed by meta-analysis. *Brain Structure and Function, 214*(5-6), 519-534.

Langer, E., Blank, A., & Chanowitz, B. (1978). The Mindlessness of ostensibly thoughtful action: The role of "placebic" information in interpersonal interaction. *Journal of Personality and Social Psychology, 36*(6), 635-642.

Lieberman, M.D. (2012). A geographical history of social cognitive neuroscience. *NeuroImage, 61*, 432-436.

Lieberman, M. D. (2010). Social cognitive neuroscience. S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds). *Handbook of Social Psychology (*5th ed.), pp. 143-193. New York, NY: McGraw-Hill.

Lieberman, M.D., Berkman, E. T., & Wager, T. (2009). Correlations in social neuroscience aren't voodoo: Commentary on Vul et al. (2009). *Perspectives on Psychological Science, 4*(3), 299-307.

Lieberman, M. D., & Cunningham, W. A. (2009). Type I and Type II error concerns in fMRI research: Re-balancing the scale. *Social Cognitive and Affective Neuroscience, 4*(4), 423-428.

Lindquist, M. A., & Wager, T. D. (2007). Validity and power in hemodynamic response modeling: a comparison study and a new approach. *Human Brain Mapping, 28*(8), 764-784.

Luby, J. L., Barch, D. M., Belden, A., Gaffrey, M. S., Tillman, R., Babb, C., et al. (2012). Maternal support in early childhood predicts larger hippocampal volumes at school age. *Proceedings of the National Academy of Sciences, 109*(8), 2854-2859.

Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique.* Cambridge, MA: MIT Press.

Masten, C. L., Eisenberger, N. I., Borofsky, L. A., Mcnealy, K., Pfeifer, J. H., & Dapretto, M. (2011). Subgenual anterior cingulate responses to peer rejection: A marker of adolescents' risk for depression. *Development and Psychopathology, 23*(1), 283-292.

Masten, C. L., Telzer, E. H., & Eisenberger, N. I. (2011). An fMRI investigation of attributing negative social treatment to racial discrimination. *Journal of Cognitive Neuroscience, 23*(5), 1042-1051.

McCabe, D. P., & Castel, A. D. (2008). Seeing is believing: The effect of brain images on judgments of scientific reasoning. *Cognition, 107*(1), 343-352.

McRae, K., Hughes, B., Chopra, S., Gabrieli, J. D. E., Gross, J. J., & Ochsner, K. N. (2010). The neural bases of distraction and reappraisal. *Journal of Cognitive Neuroscience, 22*(2), 248-262.

Mechelli, A., Price, C. J., Friston, K. J., & Ashburner, J. (2005). Voxel-based morphometry of the human brain: methods and applications. *Current Medical Imaging Reviews*, *1*(2), 105–113.

Mehta, P. H., & Beer, J. (2010). Neural mechanisms of the testosterone-aggression relation: The role of orbitofrontal cortex. *Journal of Cognitive Neuroscience, 22*(10), 2357-2368.

Mikl, M., Mareček, R., Hluštík, P., Pavlicová, M., Drastich, A., Chlebus, P., et al. (2008). Effects

of spatial smoothing on fMRI group inferences. *Magnetic Resonance Imaging, 26*(4), 490-503.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*, 167-202.

Miller, M. B., Donovan, C.-L., Van Horn, J. D., German, E., Sokol-Hessner, P., & Wolford, G. L. (2009). Unique and persistent individual patterns of brain activity across different memory retrieval tasks. *NeuroImage, 48*(3), 625-635.

Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2004). Encoding-specific effects of social cognition on the neural correlates of subsequent memory. *Journal of Neuroscience, 24,* 4912–4917.

Mitchell, T. M., Shinkareva, S. V., Carlson, A., Chang, K.-M., Malave, V. L., Mason, R. A., et al. (2008). Predicting human brain activity associated with the meanings of nouns. *Science, 320*(5880), 1191-1195.

Monti, M. M. (2011). Statistical analysis of fMRI time-series: a critical review of the GLM approach. *Frontiers in Human Neuroscience, 5*(28).

Monti, M. M., Vanhaudenhuyse, A., Coleman, M. R., Boly, M., Pickard, J. D., Tshibanda, L., et al. (2010). Willful modulation of brain activity in disorders of consciousness. *The New England Journal of Medicine, 362*(7), 579-589.

Moosman, M., Eichele, T., Nordby, H., Hugdahl, K., & Calhoun, V. D. (2008). Joint independent component analysis for simultaneous EEG‐fMRI: Principle and simulation. *International Journal of Psychophysiology, 67*, 212-221.

Mumford, J., & Nichols, T. E. (2008). Power calculation for group fMRI studies accounting for arbitrary design and temporal autocorrelation. *NeuroImage, 39*, 261-268.

Munakata, Y., Herd, S. A., Chatham, C. H., Depue, B. E., Banich, M. T., & O'Reilly, R. C. (2011). A unified framework for inhibitory control. *Trends in Cognitive Sciences, 15*(10), 453-459.

Murphy, K., & Garavan, H. (2004). An empirical investigation into the number of subjects required for an event-related fMRI study. *NeuroImage, 22*, 879-885.

Nichols, T., Brett, M., Andersson, J., Wager, T., & Poline, J.-B. (2005). Valid conjunction inference with the minimum statistic. *NeuroImage, 25*(3), 653-660.

Nichols, T., & Hayasaka, S. (2003). Controlling the familywise error rate in functional neuroimaging: a comparative review. *Statistical Methods in Medical Research, 12*(5), 419-446.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*(3), 231-259.

Northoff, G., Qin, P., & Feinberg, T. E. (2011). Brain imaging of the self --Conceptual, anatomical and methodological issues. *Consciousness and Cognition, 20*(1), 52-63.

Ochsner, K. N., Bunge, S. A., Gross, J. J., & Gabrieli, J. D. E. (2002). Rethinking feelings: An FMRI study of the cognitive regulation of emotion. *Journal of Cognitive Neuroscience, 14*(8), 1215-1229.

Olson, I. R., Plotzker, A., & Ezzyat, Y. (2007). The enigmatic temporal pole: A review of findings on social and emotional processing. *Brain, 130*(7), 1718-1731.

Pascual-Marqui, R.D., Michel, C.M., & Lehmann, D. (1994). Low resolution electromagnetic tomography: A new method for localizing electrical activity in the brain. *International Journal of Psychophysiology, 18*, 49-65.

Pfeifer, J. H., Masten, C. L., Borofsky, L. A., Dapretto, M., Fuligni, A. J., & Lieberman, M. D.

(2009). Neural correlates of direct and reflected self-appraisals in adolescents and adults: when social perspective-taking informs self-perception. *Child Development, 80*(4), 1016-1038.

Pierce, K. (2011). Early functional brain development in autism and the promise of sleep fMRI. *Brain Research, 1380*(C), 162-174.

Plichta, M. M., Schwarz, A. J., Grimm, O., Morgen, K., Mier, D., Sauer, C., et al. (in press). Test–retest reliability of evoked BOLD signals from a cognitive–emotive fMRI test battery. *NeuroImage*, 1-13.

Plonsey, R. (1963). Reciprocity applied to volume conductors and the EEG. *IEEE Transactions on Biomedical Engineering, 19*, 9-12.

Poldrack, R. (2011, Oct 04). The iPhone and the Brain. *http://www.nytimes.com/2011/10/05/opinion/the-iphone-and-the-brain.html*

Poldrack, R., Fletcher, P., Henson, R., Worsley, K., Brett, M., & Nichols, T. (2008). Guidelines for reporting an fMRI study. *NeuroImage, 40*(2), 409-414.

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences, 10*(2), 59-63.

Poldrack, R. A. (2010). Mapping mental function to brain structure: How can cognitive neuroimaging succeed? *Perspectives on Psychological Science, 5*(6), 753-761.

Poldrack, R. A., Kittur, A., Kalar, D., Miller, E., Seppa, C., Gil, Y., et al. (2011). The cognitive atlas: Toward a knowledge foundation for cognitive neuroscience. *Frontiers in Neuroinformatics, 5*, 1-11.

Poldrack, R. A., Mumford, J. A., & Nichols, T. E. (2011). *Handbook of Functional MRI Data Analysis*. New York, NY: Cambridge University Press.

Raemaekers, M., Vink, M., Zandbelt, B., van Wezel, R. J. A., Kahn, R. S., & Ramsey, N. F. (2007). Test–retest reliability of fMRI activation during prosaccades and antisaccades. *NeuroImage, 36*(3), 532-542.

Redcay, E., Dodell-Feder, D., Pearrow, M.J., Mavros, P.L., Kleiner, M., Gabrieli, J.D.E., Saxe, R. (2010). Live face-to-face interaction during fMRI: A new tool for social cognitive neuroscience. *Neuroimage,* 50, 1639-1647.

Redcay, E., Kennedy, D. P., & Courchesne, E. (2007). fMRI during natural sleep as a method to study brain function during early childhood. *NeuroImage, 38*(4), 696-707.

Rugg, M. D., Schloerscheidt, A. M., Doyle, M. C., Cox, C. J., & Patching, G. R. (1996). Event-related potentials and the recollection of associative information. *Cognitive Brain Research, 4*, 297-304.

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.

Sareen, J., Campbell, D. W., Leslie, W. D., Malisza, K. L., Stein, M. B., Paulus, M. P., et al. (2007). Striatal function in generalized social phobia: A functional magnetic resonance imaging study. *Biological Psychiatry, 61*(3), 396-404.

Saxe, R., & Powell, L. (2006). It's the thought that counts: Specific brain regions for one component of theory of mind. *Psychological Science, 17*(8), 692-699.

Scherg, M. & Ebersole, J.S. (1993). Models of brain sources. *Brain Topography, 5*, 419-423.

Shafir, E. B., Smith, E. E., & Osherson, D. N. (1990). Typicality and reasoning fallacies. *Memory & Cognition, 18*(3), 229-239.

Shattuck, D. W., Mirza, M., Adisetiyo, V., Hojatkashani, C., Salamon, G., Narr, K. L., et al.

(2008). Construction of a 3D probabilistic atlas of human cortical structures. *NeuroImage, 39*(3), 1064-1080.

Skolnick Weisberg, D., Keil, F. C., Goodstein, J., Rawson, E., & Gray, J. R. (2008). The seductive allure of neuroscience explanations. *Journal of Cognitive Neuroscience, 20*(3), 470-477.

Small, D. M., Gerber, J. C., Mak, Y. E., & Hummel, T. (2005). Differential neural responses evoked by orthonasal versus retronasal odorant perception in humans. *Neuron*, *47*(4), 593–605.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage, 23*, S208-S219.

Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage, 44*(1), 83-98.

Somerville, L. H., Wagner, D. D., Wig, G. S., Moran, J. M., Whalen, P. J., & Kelley, W. M. (in press). Interactions between transient and sustained neural signals support the generation and regulation of anxious emotion. *Cerebral Cortex*.

Spunt, R.P., & Lieberman, M.D. (2012). Dissociating modality-specific and supramodal neural systems for action understanding. *Journal of Neuroscience, 32*(10), 3575-3583.

Stice, E., Yokum, S., Burger, K. S., Epstein, L. H., & Small, D. M. (2011). Youth at risk for obesity show greater activation of striatal and somatosensory regions to food. *Journal of Neuroscience, 31*(12), 4360-4366.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*: Thieme.

Taylor, S.E., Eisenberger, N.I., Saxbe, D., Lehman, B.J., & Lieberman, M.D. (2006). Neural responses to emotional stimuli are associated with childhood family stress. *Biological Psychiatry, 60*(3), 296-301.

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science 315*(5811), 515-518.

Tomarken, A.J., Davidson, R.J., Wheeler, R.E., & Doss, R.C. (1992). Individual differences in anterior brain asymmetry and fundamental dimensions of emotion. *Journal of Personality and Social Psychology, 62,* 676-687.

Tulving, E., Schacter, D. L., & Stark, H. A. (2003). Priming effects in word-fragment completion are independent of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*(3), 336-342.

Uddin, L. Q., Clare Kelly, A. M., Biswal, B. B., Xavier Castellanos, F., & Milham, M. P. (2009). Functional connectivity of default mode network components: Correlation, anticorrelation, and causality. *Human Brain Mapping, 30*(2), 625-637.

Urry, H. L., Nitschke, J. B., Dolski, I., Jackson, D. C., Dalton, K. M., Mueller, C. J., Rosenkranz, M. A., Ryff, C. D., Singer, B. H., & Davidson, R. J., (2004). Making a life worth living: Neural correlates of well-being. *Psychological Science, 15*, 367-372.

van Veen, V., & Carter, C.S. (2002). The timing of action-monitoring processes in the anterior cingulate cortex. *Journal of Cognitive Neuroscience, 14*, 593–602.

Viswanathan, A., & Freeman, R. D. (2007). Neurometabolic coupling in cerebral cortex reflects synaptic more than spiking activity. *Nature Neuroscience, 10*(10), 1308-1312.

Vul, E., Harris, C., Winkielman, P., & Pashler, H. (2009). Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspectives on Psychological*

Science, 4(3), 274-290.

Wager, T. D., & Nichols, T. E. (2003). Optimization of experimental design in fMRI: a general framework using a genetic algorithm. *NeuroImage, 18*(2), 293-309.

Weiskopf, N., Hutton, C., Josephs, O., Turner, R., & Deichmann, R. (2007). Optimized EPI for fMRI studies of the orbitofrontal cortex: compensation of susceptibility-induced gradients in the readout direction. *Magnetic Resonance Materials in Physics, Biology and Medicine, 20*(1), 39-49.

Williams, K. D., Cheung, C. K., & Choi, W. (2000). Cyberostracism: Effects of being ignored over the internet. *Journal of Personality and Social Psychology, 79*(5), 748-762.

Wilms, M., Schilbach, L., Pfeiffer, U., Bente, G., Fink, G. R., & Vogeley, K. (2010). It's in your eyes--using gaze-contingent stimuli to create truly interactive paradigms for social cognitive and affective neuroscience. *Social Cognitive and Affective Neuroscience*, *5*(1), 98–107.

Xu, J., Monterosso, J., Kober, H., Balodis, I. M., & Potenza, M. N. (2011). Perceptual load-dependent neural correlates of distractor interference inhibition. *PLoS ONE, 6*(1), e14552.

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods, 8*(8), 665-670.

Figure captions

Figure 1. The predicted BOLD response for three stimuli of varying duration. The response for a very brief stimulus of 1 second or shorter, called a "stick function" or "punctate event" (solid line). The response for a stimulus in the 6-8 second range, called an "epoch" (dotted line). And the response for a stimulus of 16 seconds or longer, called a "block" (dashed line).

Figure 2. The main types of designs used in social neuroscience studies. (Top) Blocked designs feature the same kind of trial for prolonged periods (e.g., 30-50 seconds) followed by alternating periods of baseline. Blocked designs maximize power to detect differences among conditions, but provide no information about the shape of the hemodynamic response function. (Middle) Slow event-related designs feature brief epochs of a single even followed by extended rest to allow the BOLD response to return to baseline. These designs are not often used because of their low power to detect differences between conditions (though they provide excellent estimation of the hemodynamic response function). (Bottom) Rapid event-related designs feature multiple trial types interleaved in a random or optimized sequence along with a baseline condition. These kinds of designs are popular because they offer a balance between contrast detection power and estimation power.

Figure 3. Whole-brain images are constructed from a series of two-dimensional "slices". The figure shows 22 axial (horizontal) slices through a template brain in Montreal Neurological Institute (MNI) coordinates. The slices are arranged from inferior (top left) to superior (bottom), with the sagittal (side view) image on the bottom right indicating the position of each slice.

Figure 4. (a) An activation "blob" overlaid on a group mean anatomical image (an MP-RAGE) provides good spatial localization and is preferred over using (b) a template overlay which can be misleading about the resolution of the acquired data.

Figure 5. Data flow of a typical social neuroscience fMRI study, including cleaning, preprocessing, analysis and reporting steps. Note that the order of motion correction and slice timing correction are often swapped, and that SPM computes normalization before smoothing and first-level statistical analysis whereas FSL and AFNI normalize after those two steps. QC = (Additional) quality checks.

Figure 6. Three planes corresponding to the *x*, *y*, and *z* axes. (Top left) The coronal plane at $y = 0$. (Top right) The sagittal plan at $x = 0$. The anterior and posterior commissure (AC and PC, respectively) are visible along the horizontal crosshair. (Bottom left). The axial plane at $z = 0$. The crosshairs on all three axes are centered on the AC.

Figure 7. Statistical models for the three designs shown in Figure 2 (known as "design matrices"). (Left) Blocked design with two blocks of condition "a" and one block of condition "b". (Middle) Slow event-related design with alternating epochs of conditions "a" and "b". (Right) Rapid event-related design with a random ordering of "a", "b", and null events. By convention, time is shown from top-to-bottom. The rightmost column in each design matrix is the constant, and corresponds to β_0 in the regression model.

Figure 8. (A) Continuous EEG data sampled at 250Hz. (B) ERP components for new and old memory judgments. The waveforms shown are averages across many trials of each type, time-locked to the onset of the trials (depicted as vertical bars in A). ERP components are labeled for positive vs. negative deflections. Unpublished data for figure courtesy of Per Sederberg's Computational Memory Lab at http://memory.osu.edu.

Figure 9. (A) Simulated ERP data with a P100 and N170 component associated with face processing. (B) Simulated data moving the onset. (C) Grand average data combining the ERPs with different onsets compared against the modal ERP.

Figure 10. The relationship between *p*-value threshold and statistic threshold (e.g., correlation *r*) is moderated by sample size. (A) The *t*-distribution dictates the relationship between an alpha level and the *t* (and thus *r*) cutoff. It becomes less conservative at larger sample sizes. (B) Simulation results plotting the number of false-positive clusters as a function of sample size (N) at two different thresholds assuming no true effect. Note the steep decline in Type I error rate until N reaches ~20, and the effect of the threshold in the difference between the two lines.

Figure 11. The brain-as-predictor approach. Traditionally, social and personality psychologists have been interested in, among other things, mapping the relationship between psychological processes (e.g., cognitions, emotions) and real-world outcomes (e.g., health behavior, discrimination). In contrast, others use neuroimaging tools to map the relationship between psychological process and brain mechanisms. The *brain-as-predictor* approach integrates these by using brain systems (that have been identified with a specific psychological process) to

predict meaningful outcomes beyond the confines of the laboratory.

Figure 1.

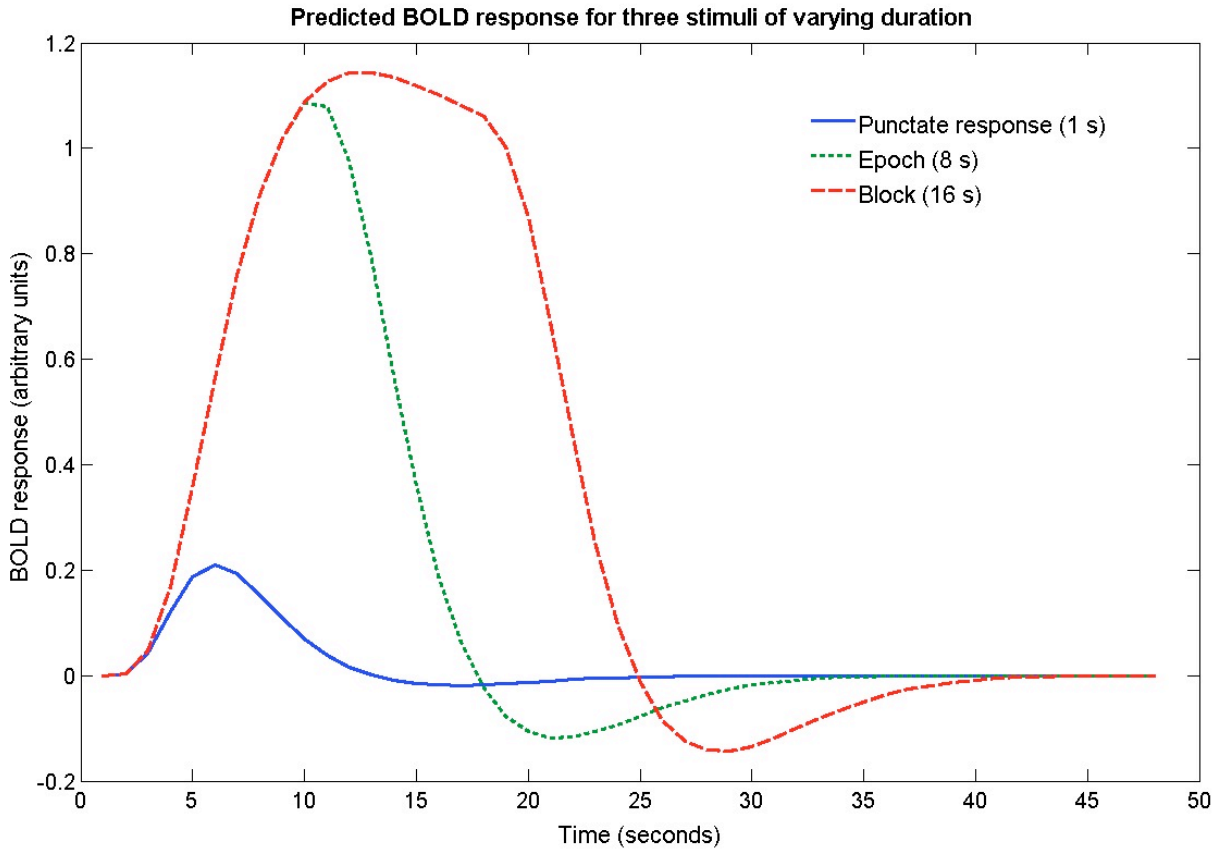

Predicted BOLD response for three stimuli of varying duration
Punctate response (1 s)
Epoch (8 s)
Block (16 s)
BOLD response (arbitrary units)
Time (seconds)
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
0
5
10
15
20
25
30
35
40
45
50

Figure 2.

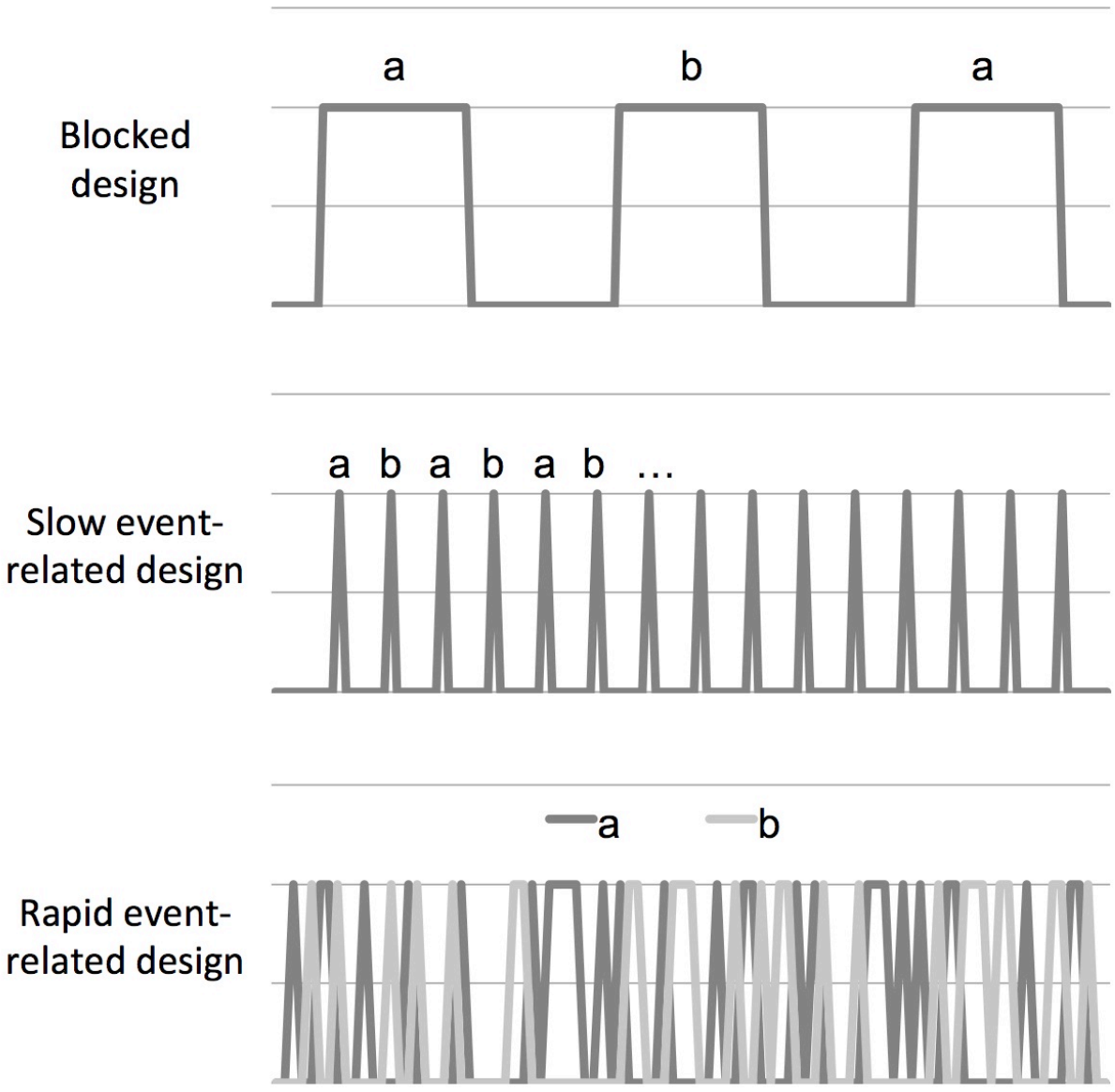

a
b
a
Blocked design
a b a b a b …
Slow event-related design
a
b
Rapid event-related design

Figure 3.

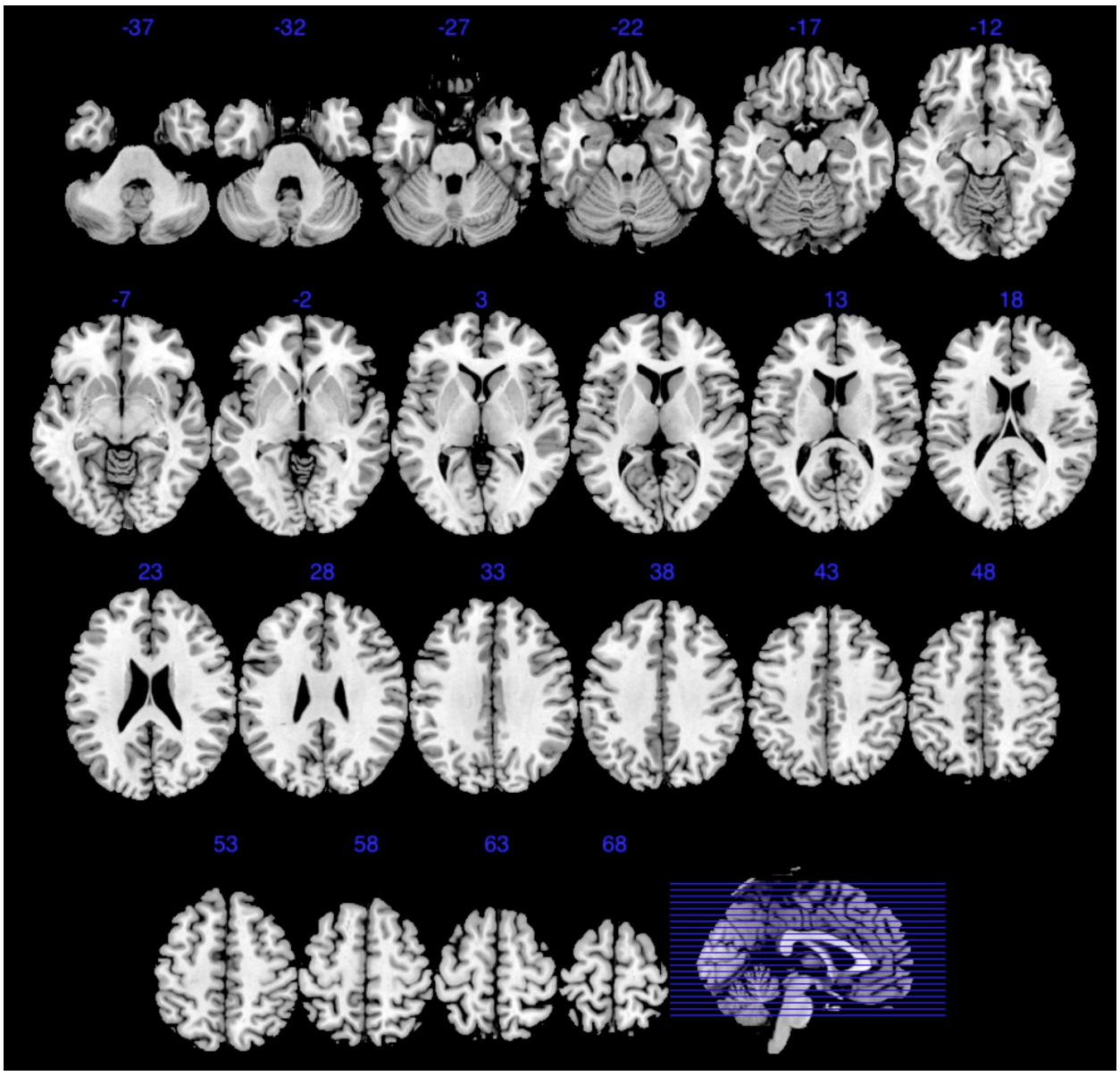


Figure 4.

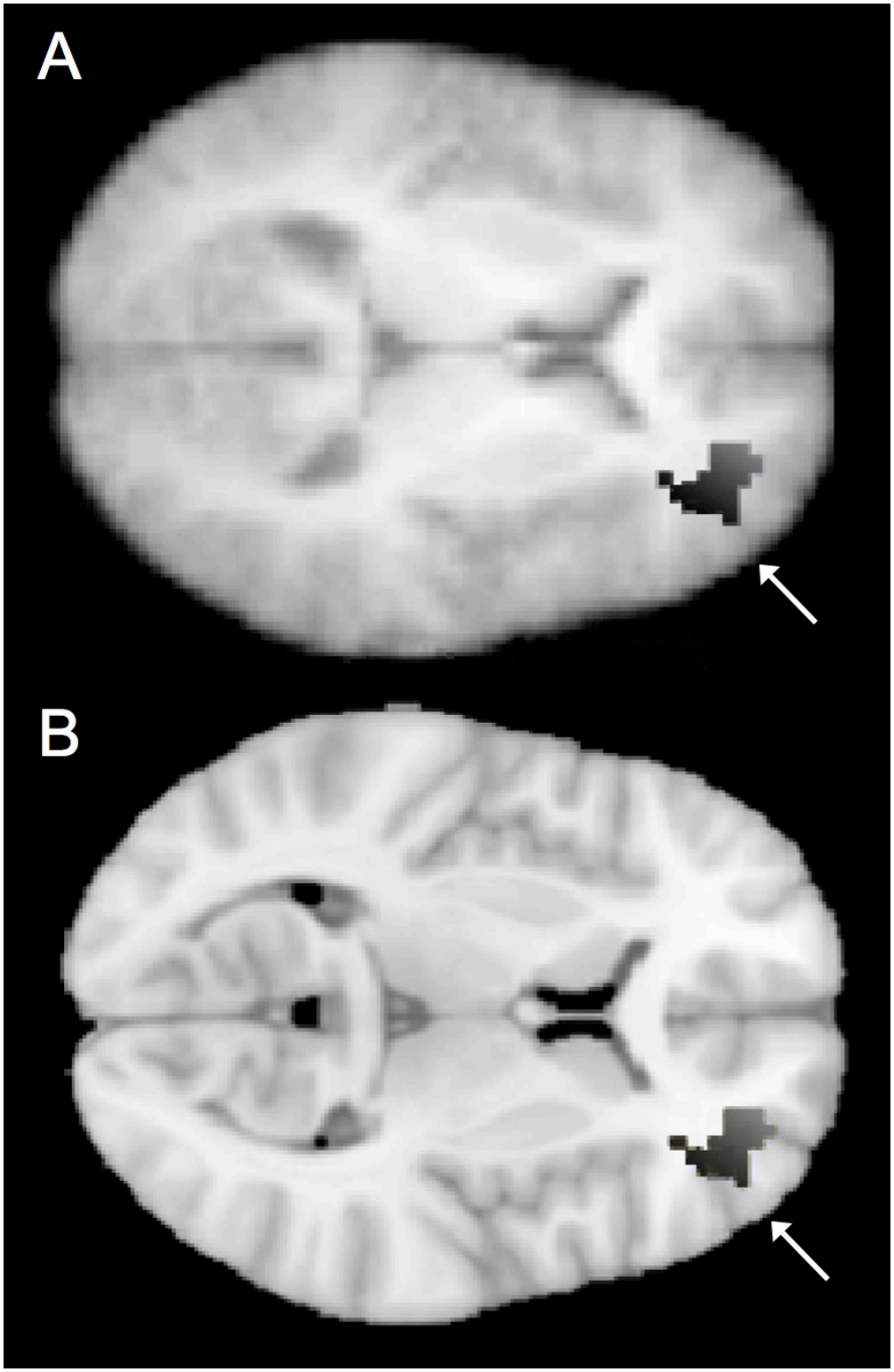


Figure 5.

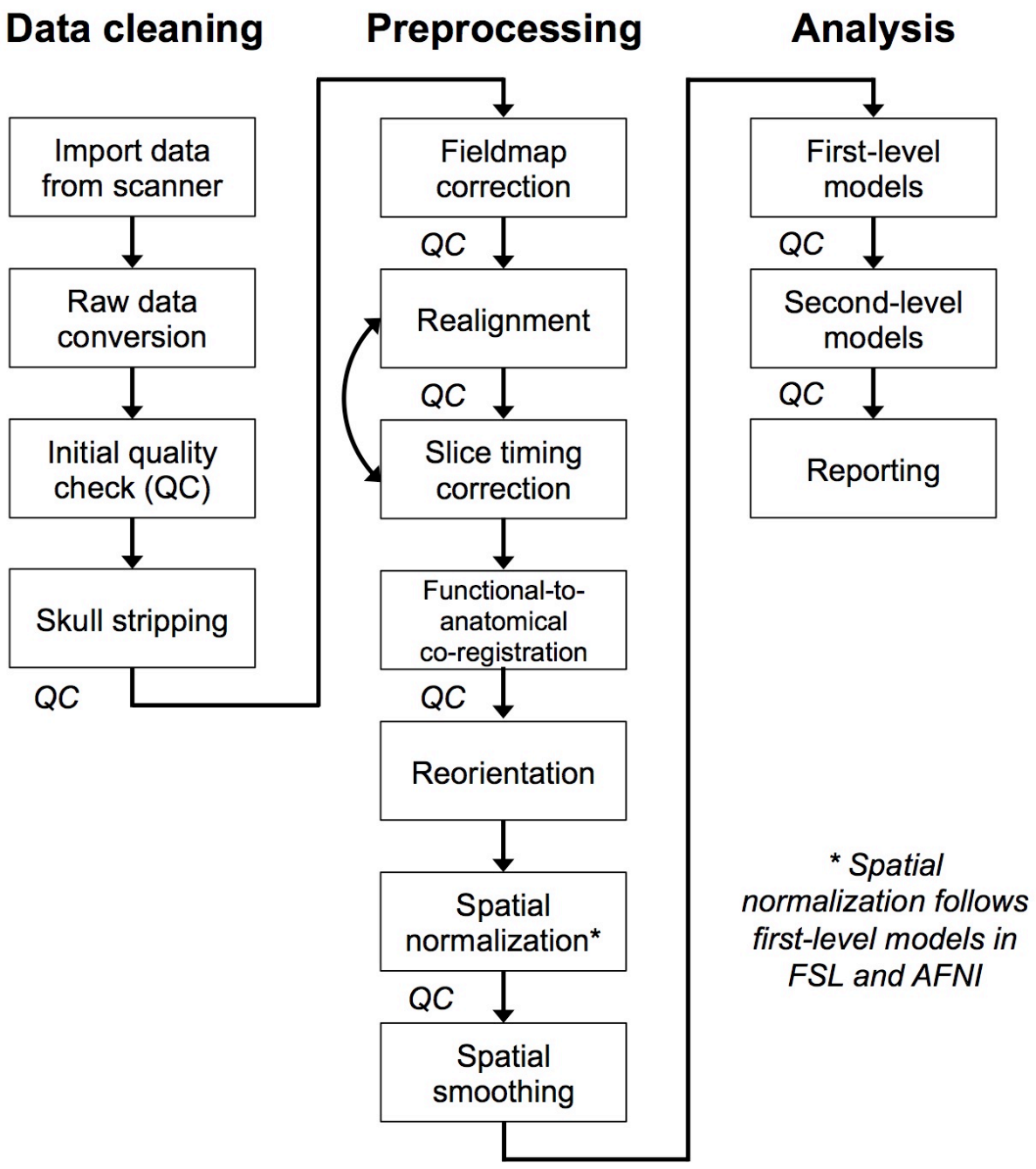

Data cleaning
Preprocessing
Analysis
Import data from scanner
Raw data conversion
Initial quality check (QC)
Skull stripping
QC
Fieldmap correction
QC
Realignment
QC
Slice timing correction
Functional-to-anatomical co-registration
QC
Reorientation
Spatial normalization*
QC
Spatial smoothing
First-level models
QC
Second-level models
QC
Reporting
* Spatial normalization follows first-level models in FSL and AFNI

Figure 6.

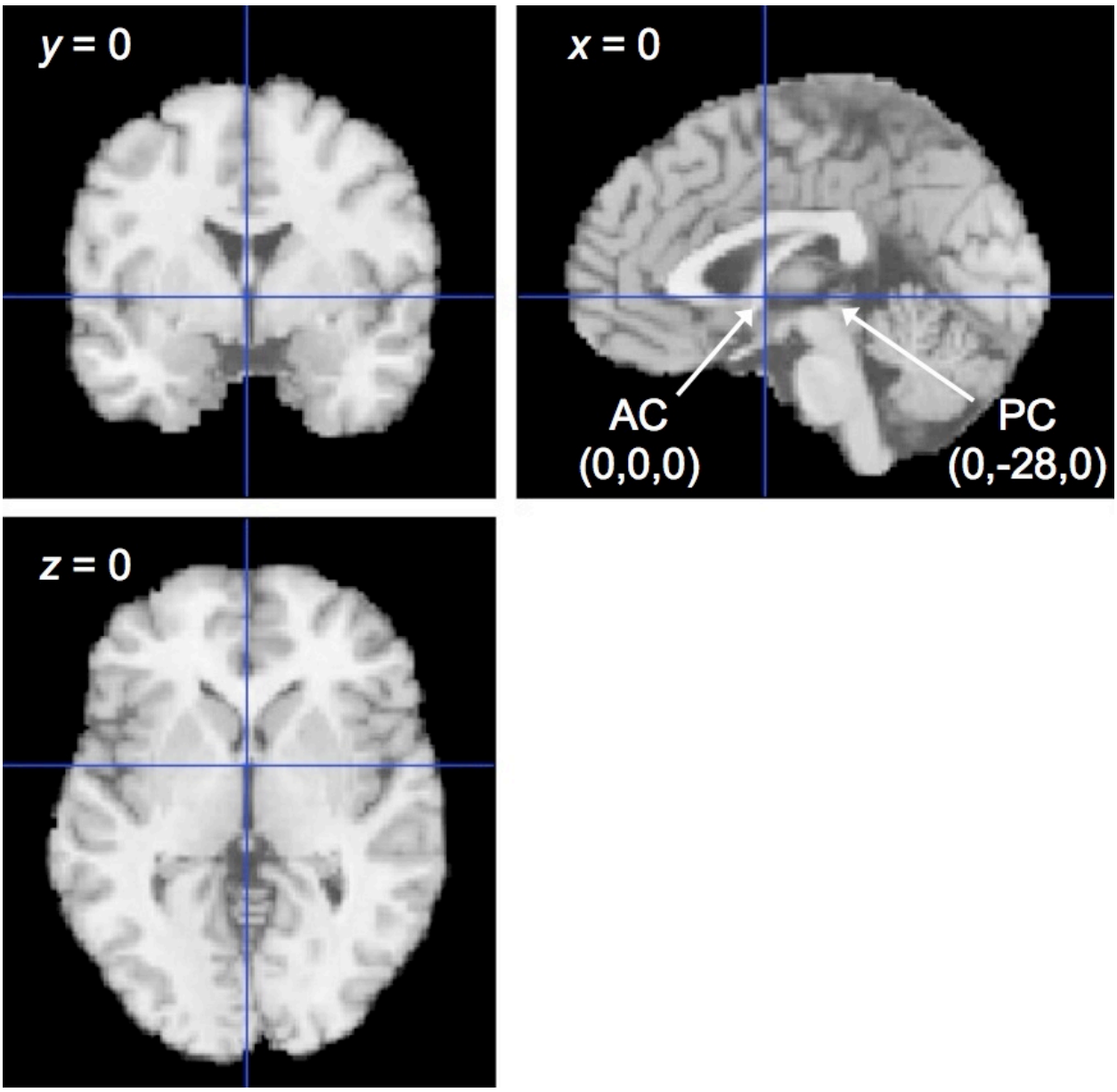


Figure 7.

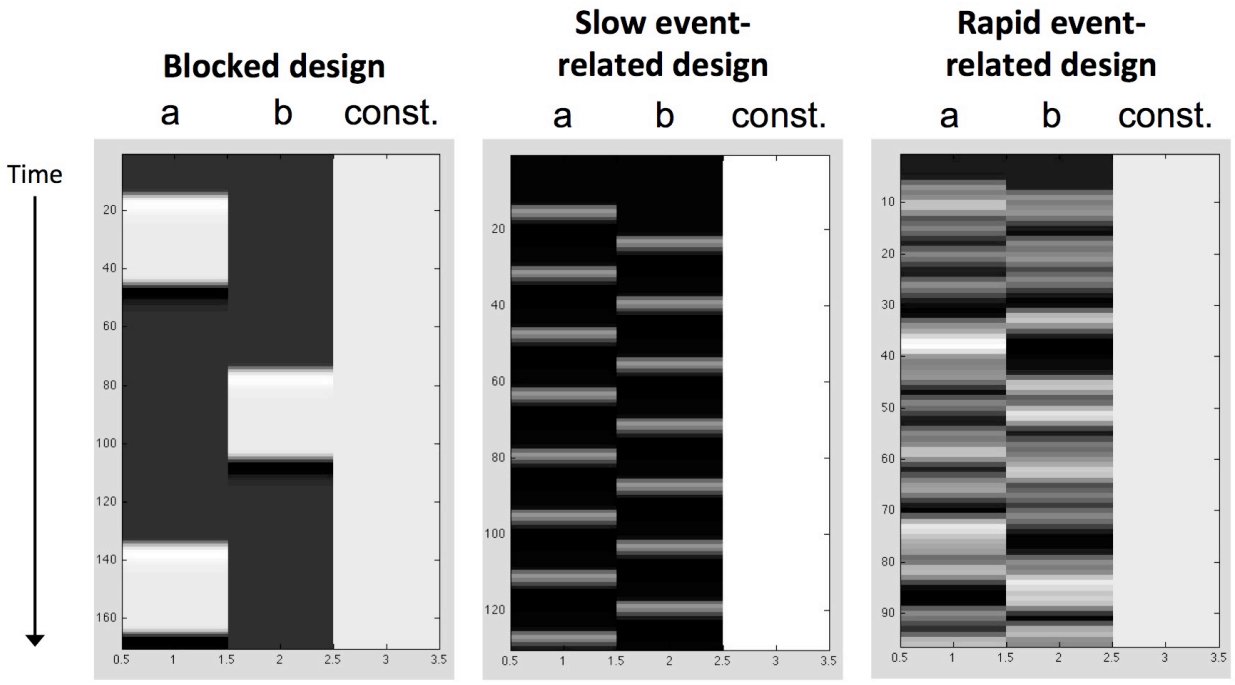

Blocked design
a b const.
Slow event-related design
a b const.
Rapid event-related design
a b const.
Time

Figure 8.

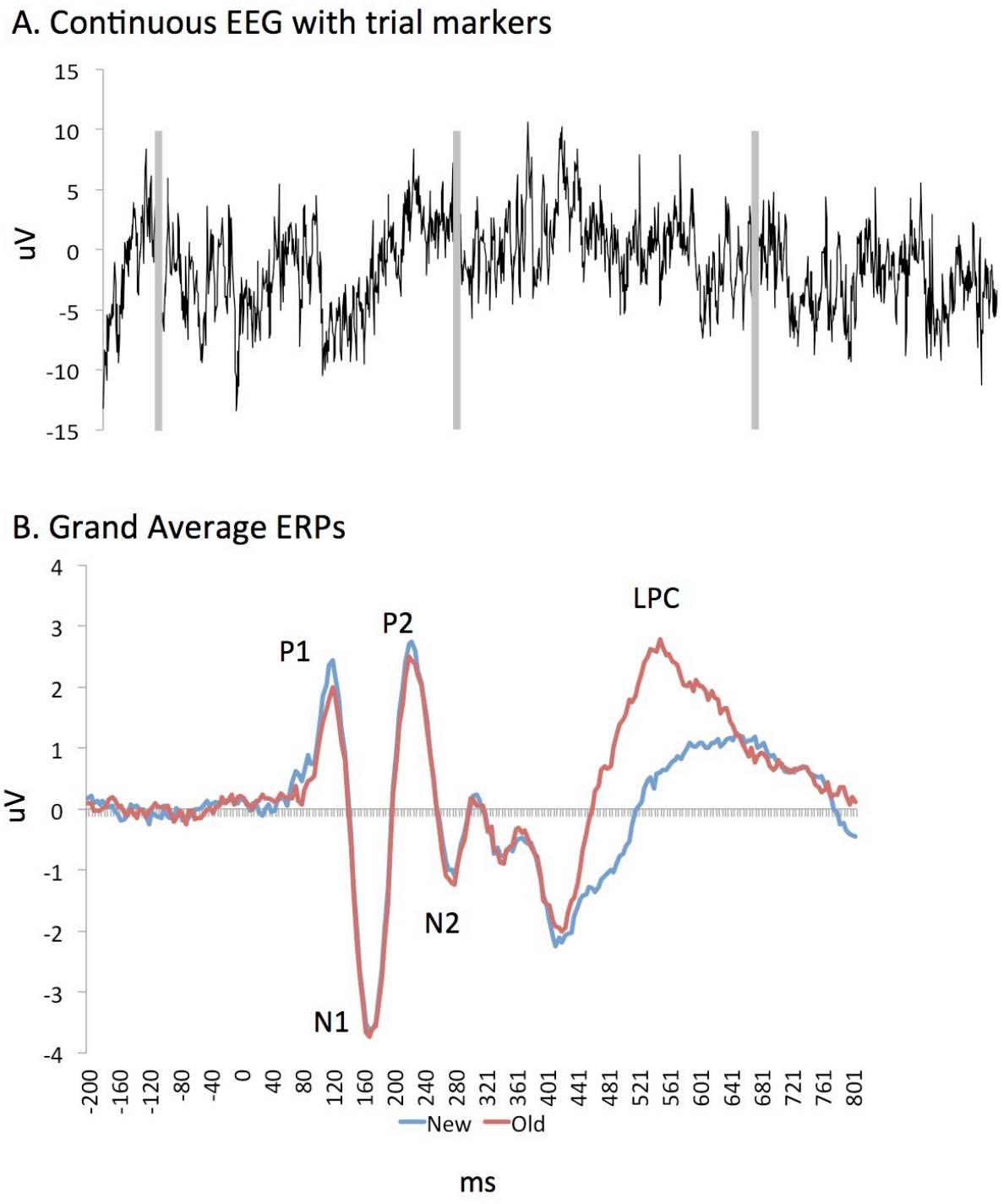

A. Continuous EEG with trial markers
15
10
5
0
-5
-10
-15
uV
B. Grand Average ERPs
4
3
2
1
0
-1
-2
-3
-4
uV
P1
P2
LPC
N1
N2
-200
-160
-120
-80
-40
0
40
80
120
160
200
240
280
321
361
401
441
481
521
561
601
641
681
721
761
801
New
Old
ms

Figure 9.

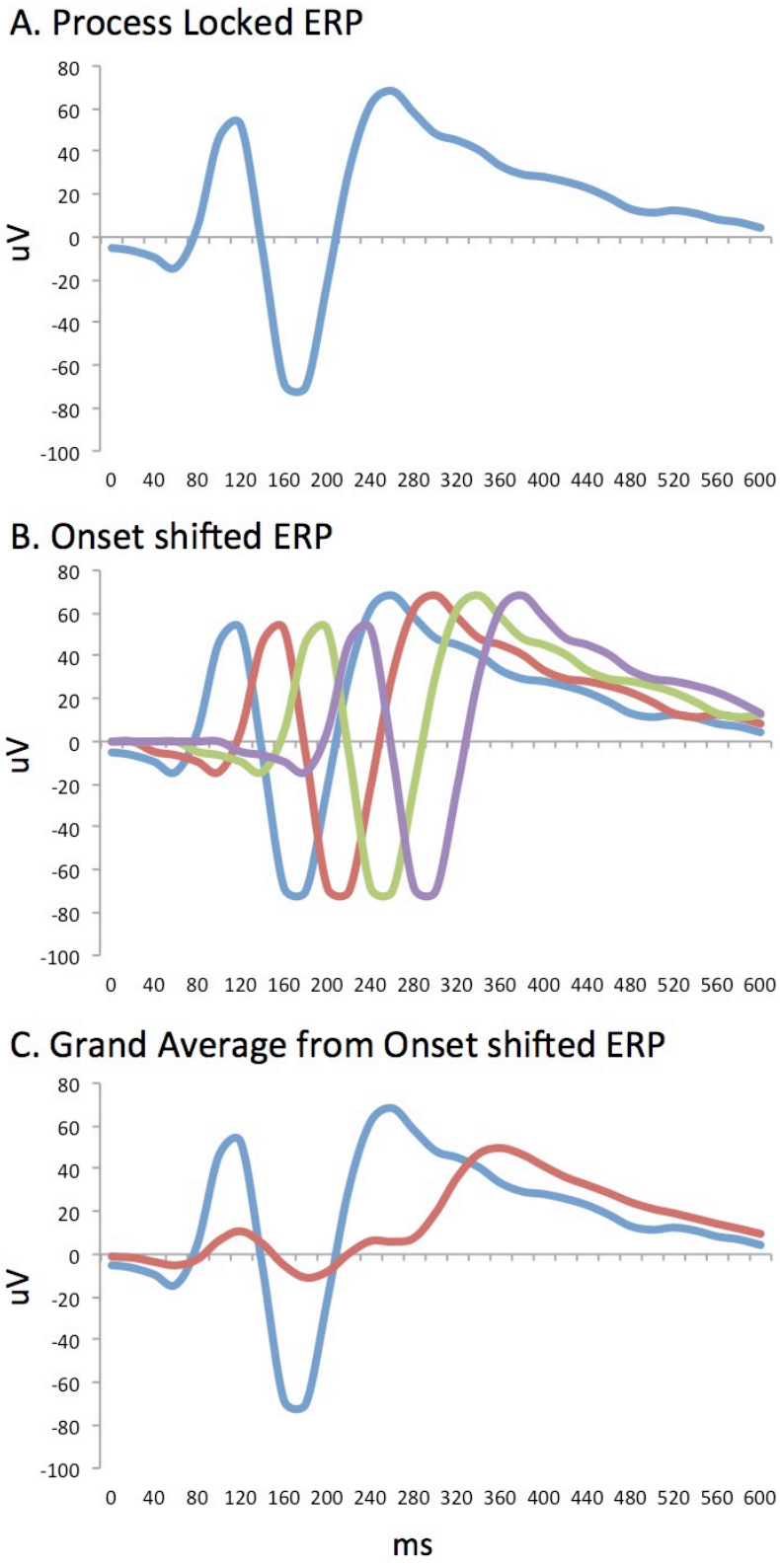

A. Process Locked ERP
80
60
40
20
0
-20
-40
-60
-80
-100
uV
0 40 80 120 160 200 240 280 320 360 400 440 480 520 560 600
B. Onset shifted ERP
80
60
40
20
0
-20
-40
-60
-80
-100
uV
0 40 80 120 160 200 240 280 320 360 400 440 480 520 560 600
C. Grand Average from Onset shifted ERP
80
60
40
20
0
-20
-40
-60
-80
-100
uV
0 40 80 120 160 200 240 280 320 360 400 440 480 520 560 600
ms

Figure 10.

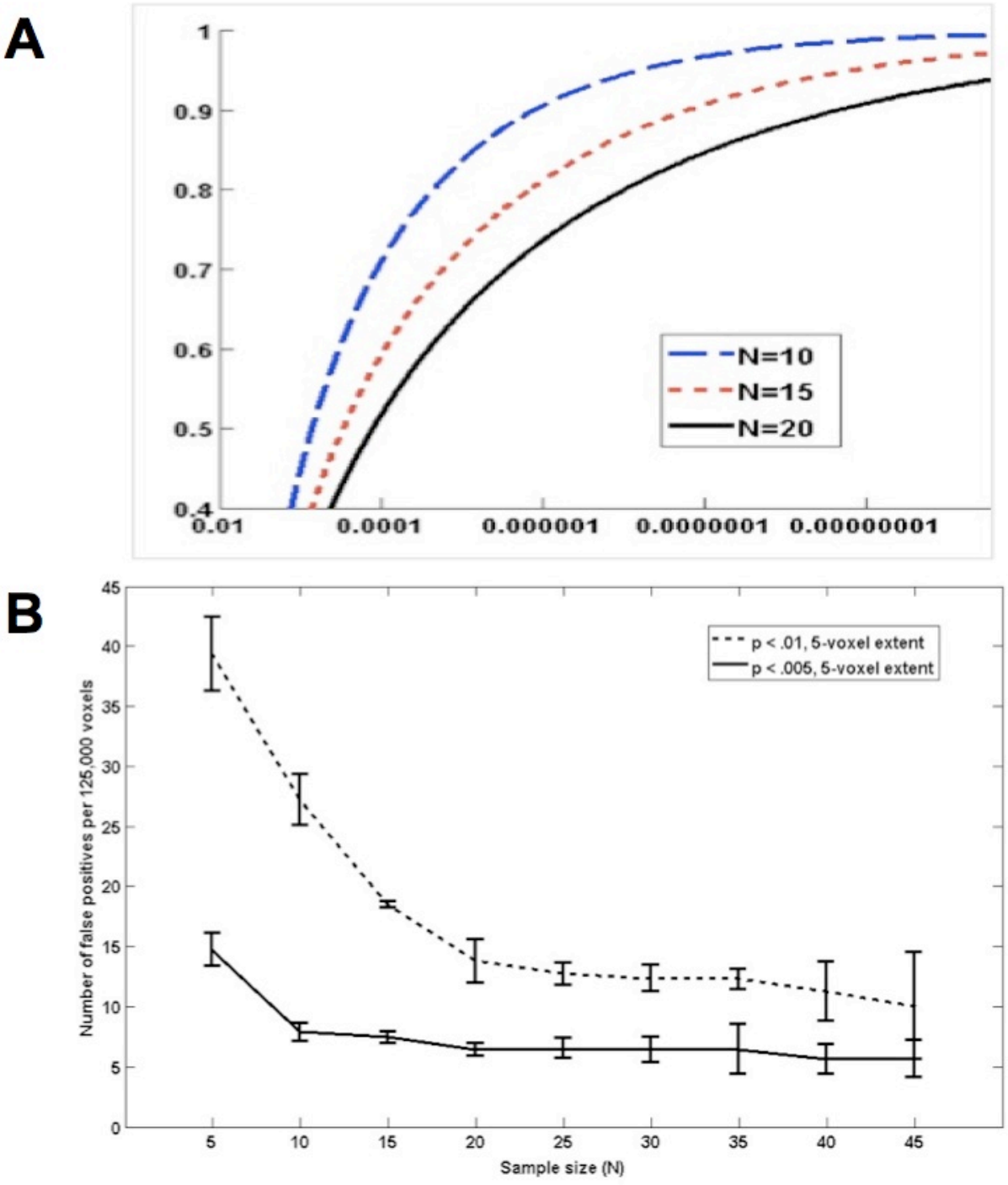
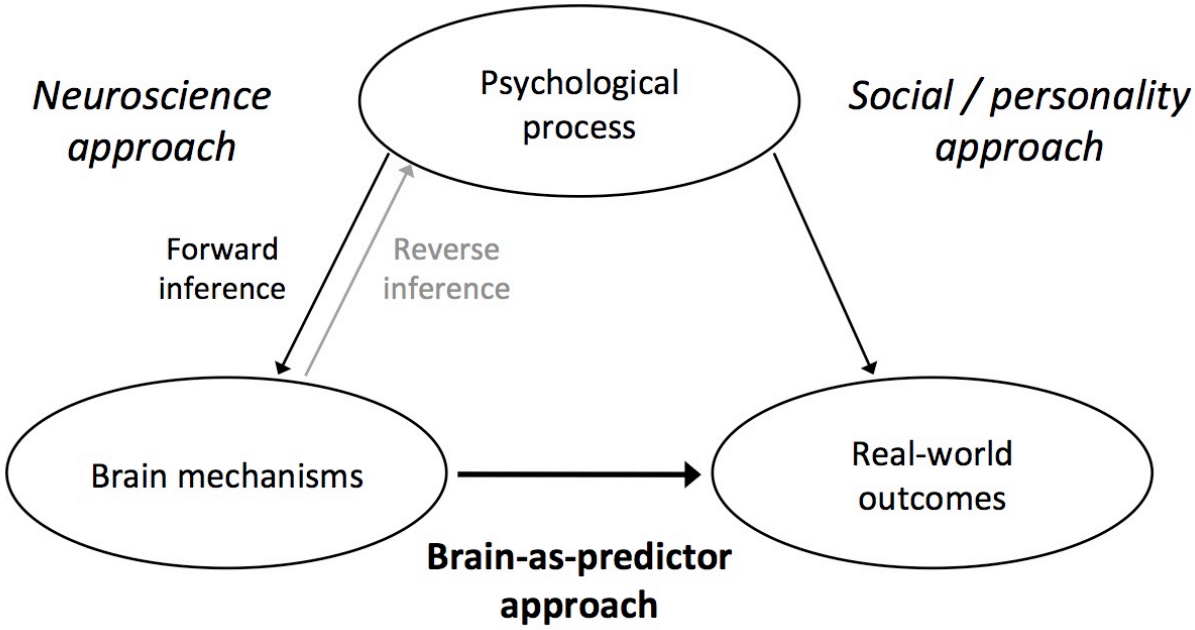

A
1
0.9
0.8
0.7
0.6
0.5
0.4
0.01
0.0001
0.000001
0.0000001
0.00000001
N=10
N=15
N=20
B
Number of false positives per 125,000 voxels
Sample size (N)
p < .01, 5-voxel extent
p < .005, 5-voxel extent
0
5
10
15
20
25
30
35
40
45

Figure 11.


Neuroscience approach
Psychological process
Social / personality approach
Forward inference
Reverse inference
Brain mechanisms
Real-world outcomes
Brain-as-predictor approach