Functional Neuroimaging: Experimental Design and Analysis

Jody C. Culham

Final Submission to Roberto Cabeza and Alan Kingstone (Eds.),
Handbook of Functional Neuroimaging of Cognition (2nd ed.)

December 31, 2004

Correspondence should be directed to:

Jody Culham
Department of Psychology
Social Science Centre
University of Western Ontario
London Ontario
Canada N6A 5C2

Phone: 519-661-2111 x85643
Fax: 519-661-3961
E-mail: culham@imaging.robarts.ca
Web Page: http://defiant.ssc.uwo.ca/Jody_web

Number of words:	13,685
Body of text	9,679
References	3,358
Figure Captions	648
Number of Figures	3

"...the single most critical piece of equipment is still the researcher's own brain. All the equipment in the world will not help us if we do not know how to use it properly, which requires more than just knowing how to operate it. Aristotle would not necessarily have been more profound had he owned a laptop and known how to program. What is badly needed now, with all these scanners whirring away, is an understanding of exactly what we are observing, and seeing, and measuring, and wondering about."

Endel Tulving, interview in Cognitive Neuroscience
(2002, Gazzaniga , Ivry & Mangun, Eds., NY: Norton, p. 323)

INTRODUCTION

With the advent of positron emission tomography (PET) in the 1980s (Fox et al., 1986) and functional magnetic resonance imaging (fMRI) in the 1990s (Kwong et al., 1992; Ogawa et al., 1992), neuroimaging became a keystone for the growing field of cognitive neuroscience. Historically, our knowledge of the human brain has been far more limited than for other species (Crick & Jones, 1993), primarily due to the restriction against using invasive techniques, such as single neuron recording, in humans. Although human neuropsychological studies have been very enlightening, newer non-invasive neuroimaging methods, particularly fMRI, enable exploration of the normal rather than disordered human brain and allow resolution at a fine spatial scale which lesions rarely provide (Savoy, 2001). These new imaging methods have identified dozens of functionally-specific areas in the human brain, some of which seem comparable to areas in other species, particularly the macaque monkey, and some of which may be uniquely human (Culham & Kanwisher, 2001; Duncan & Owen, 2000; Grill-Spector & Malach, 2004; Tootell et al., 1998). Within numerous human regions that have been identified, detailed explorations have revealed the underlying computational processes (e.g., Wandell, 1999).

The growth of neuroimaging has been phenomenal. Neuroimaging publications continue to increase exponentially (Fox, 1997), with recent estimates of four published papers each day (Tootell et al., 2003). The caliber of imaging papers has improved considerably with the growth of the field, primarily due to the theory-driven approaches and high standards of experimental design expected in other disciplines. However, for many newcomers, fMRI methodology can seem overwhelmingly complex, leading to an increasing demand for resources to learn neuroimaging techniques. Numerous resources have suggested *how* neuroimaging experiments should be designed but with limited consideration of the philosophy behind experimental design. One exception is an article by Steve Kosslyn whose clever title posed the question, "*If neuroimaging is the answer, what is the question?*" (Kosslyn, 1999) Another exception is William Uttal's book, *The New Phrenology (Uttal, 2001)*, which posed numerous pessimistic criticisms of the entire brain imaging enterprise (for a rebuttal, see Donaldson, 2004).

Here I intend to present a brief overview of neuroimaging design principles followed by a somewhat opinionated review of my thoughts regarding the types of questions for which neuroimaging is (and is not) suited. As well, I will outline some of the principles which can lead to better questions and better experimental designs. The emphasis will be on

fMRI but many of the same principles apply to studies performed with PET. Space does not permit detailed recommendations for design issues, but the reader will find many useful resources in print (Aguirre & D'Esposito, 1999; Buckner & Logan, 2001; Chein & Schneider, In press; Huettel et al., 2004; Jezzard et al., 2001) and online (such as my website, *fMRI for Dummies*, http://defiant.ssc.uwo.ca/Jody_web/fmri4dummies.htm). By no means do I intend to present a cynical view of the brain imaging enterprise. Rather, I hope to prompt newcomers to the field to think carefully about their approaches and the caveats, such that their contribution to the enterprise will become more fruitful.

FUNCTIONAL NEUROIMAGING: EXPERIMENTAL DESIGN AND ANALYSIS

In neuroimaging, brain activation levels must always be considered relative to another condition. The signal strength in a particular area depends on many factors such as inherent metabolic rate and location with respect to the coil (particularly true for surface coils which sample some regions of the brain with a higher signal-to-noise ratio, or SNR, than other areas). Thus the absolute level of signal is relatively meaningless on its own. As a consequence, all neuroimaging experiments rely on *subtraction logic* to make sense of the data.

Subtraction logic was originally developed in the 19th century for the study of reaction times to infer differences in cognitive processing (Donders, 1868 reprinted in Donders, 1969). In subtraction logic, one compares two events that differ putatively by only one factor. For example, consider two reaction times[1]. In condition A, the subject must press a button whenever a light of any color turns on. In condition B, the subject must press a button whenever a red light turns on. Assuming that condition B differs from A only in the need to discriminate the color of the light, a subtraction of the reaction time for A from that for B should reveal the cognitive processing time to make the discrimination (with components common to both tasks, say visual processing time, being "subtracted out"). Subtraction logic relies on the *assumption of pure insertion*, namely the belief that two conditions differ in *one and only one* critical component. If this assumption is false and multiple differences exist, then it is impossible to distinguish between them.

Subtraction logic forms the basis for all neuroimaging experiments. The neuroimager must choose the best baseline condition that will subtract out all activation other than the process of interest. At first, this can be harder than it seems. The final section of this chapter will consider some of the issues involved in choosing appropriate baselines and some alternative designs (such as parametric designs) that are less vulnerable to the problems of subtraction logic.

The vast majority of neuroimaging experiments are considered *hypothesis-driven*. That is, the researcher plans the experiment with certain expectations regarding what types of differences will be observed between critical conditions. Statistical tests are then performed to determine whether such differences exist and are unlikely to be due to

[1] Although some authors differentiate between reaction times and response times, for simplicity, I will use the single term, reaction time, to indicate the time between initial stimulus presentation and the pressing of a response key.

chance. In recent years, an alternative has been provided in the form of *data-driven* approaches which make no *a priori* assumptions about the expected patterns in the data. Newer approaches use sophisticated statistical techniques, such as independent component analysis (ICA), to extract the patterns that account for the most variability within the data (e.g., Biswal & Ulmer, 1999; McKeown et al., 1998). Such techniques have advantages in that they require fewer assumptions about the data and they may discover sources of variability unexpected by the experimenter. These techniques have the potential to be useful for generating new hypotheses about processing; however, to date, data-driven approaches have been quite limited, and in the vast majority of studies, the experimenters have definitive hypotheses to test. Thus the remainder of this chapter will focus on standard hypothesis-driven approaches.

Among hypothesis-driven analyses, there are two main approaches, the *voxelwise approach* and the *region of interest (ROI) approach*. In the voxelwise approach, the experimenter performs a statistical comparison between two (or more) conditions of interest on a voxel-by-voxel basis. No prior assumptions are needed regarding which specific brain areas will be differentially activated. Based on statistical evaluation, a list of areas in which significant activation differences were observed can be generated. This approach requires that every subject's data be transformed into a standard space that enables averaging across different brains, The two most common stereotaxic systems are the Talairach atlas (Talairach & Tournoux, 1988) and the Montreal Neurological Institute (MNI) space (Evans et al., 1993). This is a powerful approach in that it enables the researcher to investigate the whole brain (or as much of the volume as can be scanned given other constraints), without the need for detailed hypotheses regarding expected loci of activation. Because the approach is highly contingent on solid statistical methods, it is essential for the user to be quite savvy about the statistics which are employed and to have access to an analysis package which offers the appropriate statistics and corrections.

The alternative region-of-interest approach focuses on the role of previously-described regions in a novel experiment. ROIs may be defined based on anatomical criteria. For example, in a memory study, regions of the hippocampus may be identified and selected based on the anatomical slice data. More commonly, ROIs are well-established areas defined by their functional responses. For example, in vision science, common ROIs include the middle temporal complex (MT+) which responds more to motion than still images (Culham, He et al., 2001), the fusiform face area (FFA) which responds more to faces than other categories of objects (Kanwisher et al., 1997), and striate and extrastriate visual areas defined by boundaries in retinotopic maps (Wandell, 1999). In the ROI approach, regions are first identified using a *localizer* based on prior studies which have reported reliable activation for a particular comparison. Activation in these ROIs is then evaluated in an independent line of experiments designed to test a new hypothesis. For example, a localizer for MT+ would compare activation during visual motion to a stationary state, possibly at low contrast, to isolate the MT+ complex from other motion-related areas (Tootell, Reppas, Kwong et al., 1995). After the ROI has been identified in each subject, the activation time courses are extracted and analyzed statistically to determine whether there are significant differences in the comparison for the new experimental task. For example, one study investigated whether area MT+, as defined by

a standard localizer, was more activated in static images that implied motion than those that did not (Kourtzi & Kanwisher, 2000a).

The standards for appropriate neuroimaging statistics (used in both voxelwise and ROI approaches) have been evolving over the years. Many statistical developments have become more commonplace with their introduction into standard statistical packages for neuroimaging data. In the early days of brain imaging, simple statistics such as t-tests and correlations were applied to data, sometimes without preprocessing, and statistical thresholds were determined either arbitrarily or with conservative corrections for the problem that one was simultaneously examining tens of thousands of individual voxels (namely the Bonferroni correction for multiple comparisons). Currently, standard procedures often include preprocessing (such as correcting for motion of the subject's head, spatial filtering and temporal filtering), corrections for serial correlations (to correct the false assumption that activation at one point in time is unrelated to activation at the preceding and following points in time), less conservative corrections for multiple comparisons (such as false discovery rate corrections, Genovese et al., 2002), and random effects analyses which allow generalization from the group of tested subjects to the population as a whole. In addition, many statistical packages now include general linear model analyses. General linear models can decompose the data into various factors, both *predictors of interest* to the experimenter (such as the differences between two conditions) as well as those *predictors of no interest* which add noise to the data (such as head movements which may be spuriously correlated with the signal).

One of the most useful advances in neuroimaging has been the development of event-related designs (Buckner & Braver, 1999; Donaldson & Buckner, 2001). In PET, and in the early years of fMRI, researchers used only block designs with long continuous periods (typically 10-30 sec) for each condition. Even when individual trials were brief, many trials of the same type were presented sequentially, because researchers believed that the slow rise and fall of the hemodynamic response would prevent the resolution of individual events. Several key breakthroughs in the mid-1990s challenged this assumption and led to the development of event-related designs (Blamire et al., 1992; Buckner et al., 1996; Dale & Buckner, 1997). Unlike block designs, event-related designs allow the presentation of trials in an unpredictable order. In slow event-related designs, trials are widely separated in time (ideally 12-14 sec or longer) to ensure that the signal returns to baseline between events (Bandettini & Cox, 2000). The trials can be subjected to time-locked averaging to extract the common signal change related to the event, while smoothing out the noise. In rapid event-related designs (Dale & Buckner, 1997), trials of various types can be intermixed and presented rapidly (e.g., every 2 sec). The analysis of rapid event-related designs is more complex and requires either careful counterbalancing of trial sequences or deconvolution of the fMRI signal using "jittered" designs with irregular intervals. Although event-related designs have a somewhat reduced statistical power to detect activation relative to block designs, they offer many valuable advantages. They make it possible to see the unfolding of sequential processes, for example in paradigms with a delay interval interposed between the stimulus and response. The randomized and unpredictable sequences of events prevents subjects from locking into a mental set. Trials can also be categorized by subjects' performances, for

example, into correct and incorrect trials. Particularly in some domains, such as memory research, event-related approaches have opened many possibilities for elegant designs. Mixed designs, which incorporate blocks of variable trials, also hold promise for distinguishing between transient and sustained processing (Donaldson, 2004).

Although standard fMRI designs allow the experimenter to evaluate only the responses of an entire population of neurons within an imaging voxel, recently-developed fMR adaptation techniques hold excellent potential for evaluating the nature of the mental representations within an area (Grill-Spector et al., 1999; Grill-Spector & Malach, 2001). While behavioral adaptation has been called the "psychophysicist's microelectrode" (Frisby, 1979), fMR adaptation may become the "neuroimager's microelectrode." Both behavioral and fMR adaptation work on the same principle: with continued stimulation, neurons show reduced responses. These reduced responses can be used to determine the dimensions to which a neuron (or neuronal subpopulation or voxel or ROI) is sensitive. For example, in object-selective cortex, including the lateral occipital cortex and the FFA, the response is lower for prolonged exposure to repeated images of the same face than for varying images of different faces. Grill-Spector and colleagues (1999) investigated whether neurons within the FFA are sensitive to viewpoint by comparing (1) the response to varying images of *different faces* seen from the *same viewpoint*; (2) the responses to repeated images of the *same face* seen from the *same viewpoint* ; and (3) the response to repeated images of the *same face* seen from *different viewpoints*. A basic adaptation effect was observed based on a reduced response to condition 2 (same face, same viewpoint) relative to condition 1 (different faces, same viewpoint). The experiments also tested whether object-selective regions were viewpoint dependent or viewpoint invariant. If the regions were selective for viewpoint, the same faces viewed from different viewpoints would be processed by different subpopulations of neurons and thus adaptation would be weak. If the regions were invariant to viewpoint, the same faces viewed from different viewpoints would be processed by the same subpopulation of neurons. Adaptation would then be strong, and activation levels would be expected to drop by a comparable amount for conditions 2 and 3 relative to condition 1. Object-selective regions showed little adaptation to condition 3, suggesting viewpoint dependence.

Regardless of the exact design selected (data-driven vs. hypothesis-driven, voxelwise vs. ROI, block vs. event-related, adaptation), it is essential to optimize the imaging parameters for the question being asked (for extended discussion, see Huettel et al., 2004). The experimenter has many decisions to make regarding the type of coil to employ (surface coil for high regional SNR vs. head coil for moderate global SNR vs. phased array coils for high global SNR) and tradeoffs between high spatial resolution (up to 1 mm), high temporal resolution (up to 100s of ms) and the number of slices covered (Menon & Goodyear, 2001). The experimenter must also decide how many subjects to test, as well as the number of runs and their duration, based on assumptions about how large the difference between conditions is likely to be. The order of conditions within a run must be decided and is often constrained by counterbalancing issues. Some experiments are best-suited to the statistical power of block designs while others can benefit from the elegance of event-related paradigms. Even for event-related designs,

there is the choice between slow and rapid trial presentation. Once the design has been established, many more analysis decisions follow – preprocessing (motion correction, spatial smoothing, temporal filtering), statistical analysis (correlations, t-tests, general linear models, Fourier analysis; parametric vs. non-parametric statistics), group analysis vs. single subject analysis, brain presentation (2D slice data, 3D volumetric data, or cortical surface rendering) and the biggest bugaboo, statistical thresholds and the application of corrections for inappropriate assumptions (multiple comparisons, serial correlations, fixed vs. random effects). It is not possible to recommend a single optimal design because the decisions will strongly depend on the priorities of the given experiment and the overall approach (voxelwise vs. ROI).

ISSUES

In ten years of doing neuroimaging and teaching others how to design successful experiments, I have noticed many common mistakes which I made myself and now see other newcomers frequently making. In some ways, it may be easier to explain how **NOT** to do an imaging experiment, based on some of the common pitfalls, than it is to explain how to do an imaging experiment well. Perhaps the most common mistake of newcomers to brain imaging is to take a well-established paradigm that has been used with other techniques (e.g., cognitive psychology or neurophysiology) and to simply run *exactly* the same paradigm in the scanner, without optimizing the design or considering whether the possible outcomes can enlighten theories about the process in question. In the worst examples, the experimenter does not consider changing any of the parameters which may be critical in neuroimaging designs, such as event timing. In the simplest experiments, two conditions are compared and lists (sometimes gigantic ones) of activation foci are published, followed by a post hoc "just so story" about why the task recruits the "subcortico-occipito-parieto-temporo-frontal network" (known to laypeople as "the brain"). Often, the areas that are activated are likely to include numerous parietal and frontal cortical areas that seem to be activated across a whopping variety of heterogeneous tasks from visuomotor control (Culham & Kanwisher, 2001) to meditation (e.g., Lazar et al., 2000) to gum-chewing (e.g., Takada & Miyamoto, 2004). Alternatively, poorly considered experiments may fail to find any meaningful activation. For instance, newcomers are often tempted to examine very subtle behavioral effects in brain areas where more robust effects have not yet been established. In such cases, there is little hope of an enlightening result.

Here, I make the case that careful consideration of the reasons and methods for doing neuroimaging can lead to more successful imaging experiments, making more valuable contributions to the field of cognitive neuroscience. I suggest consideration of the following points prior to starting new neuroimaging projects.

1. Do the thought experiment first

The cost of neuroimaging is exorbitant with scan time typically costing between US$200 and $1000 per hour, not including additional expenses such as subject fees, equipment and staff salaries. In an environment with limited resources, particularly grant funds, it

becomes critical to plan effective studies with a high probability of generating meaningful conclusions. Thought experiments are essentially free and can save considerable wasted effort in real experiments.

The first question that should be asked in a thought experiment is whether the information about the anatomical foci that are activated during a particular comparison would constrain theories about the mechanisms involved. With the exception of rare dualists, most cognitive neuroscientists today believe that all behaviors have some neural correlate in the brain. The question then is whether knowing something about the *particular* neural correlate constrains possible theories of cognitive function. Certainly, in clinical research, such as presurgical planning, localization is very valuable; however, in cognitive neuroscience, localization is valuable only insofar as it informs theories of cognitive function.

There is no single good reason for doing an imaging study; nevertheless, I would like to suggest certain types of studies where neuroimaging can make a unique contribution above and beyond what can be learned from other techniques. Where possible, I have provided specific examples, often taken from my own domain of vision and visual cognition.

a. **Comparisons of activation across multiple tasks**. It can be quite informative to investigate whether two tasks involve similar or different networks of areas. As one example, there has been much debate about whether subjects can pay attention to locations other than where they are directly looking (Klein, 1980; Rizzolatti et al., 1994). An elegant study by Maurizio Corbetta and colleagues (1998) demonstrated that attention and eye movements evoke activation in nearly equivalent brain networks. This fMRI experiment did not definitively resolve whether the two functions can be dissociated (as behavioral evidence suggests they can) but it certainly supported a close linkage between them. Comparisons can also provide interesting dissociations, or even double dissociations based on the same logic used in neuropsychological studies. For example, theories about dissociations between vision for perception and vision for action (Milner & Goodale, 1995) have received some support from neuroimaging (Culham et al., 2003; James et al., 2003).
b. **Characterization of a single ROI's responses**. Neuroimaging can be used to identify one particular ROI and then to systematically investigate the stimuli and tasks that drive that region. Excellent examples come from seminal studies that identified key areas in the recognition of faces (Kanwisher et al., 1997) and other categories of objects (Malach et al., 1995). Subsequent experiments have categorized the activation of these regions across a wide range of stimuli and tasks (for a review see Kanwisher et al., 2001), and have led to much more detailed hypotheses about face and object processing in the normal and disordered human brain (e.g., Hadjikhani et al., 2004; Hasson, Avidan et al., 2003).
c. **Correlation between brain and behavior**. Studies that acquire concurrent behavioral data with the imaging acquisitions have the ability to examine the neural correlates of human behavior. Such neural correlates have been investigated in other species (e.g., Logothetis & Schall, 1989), but it is more

challenging to infer subjective states in other species and the results may not generalize to humans. Numerous interesting experiments investigated the relationship between a particular neural response in an area and conscious awareness (Rees, 2001; Ress & Heeger, 2003; Tong et al., 1998), performance, or subsequent memory for items (e.g., Brewer et al., 1998; Wagner et al., 1998).

d. **Evaluation of the role of experience**. Studies of subjects with different levels of training can help distinguish the degree of innate hard-wiring versus experience-driven plasticity in an area. Due to the non-invasiveness of fMRI, subjects can be scanned multiple times with an intervening training period (e.g., Gauthier et al., 1999). Alternatively, activation patterns can be compared between experts and novices (e.g., Gauthier et al., 2000; Maguire et al., 1997).
e. **Comparisons between species**. Cognitive neuroscience has a wealth of information about brain processing in other species, particularly macaque monkeys, and tends to assume that similar mechanisms exist in the human brain; neuroimaging allows us to test this assumption. In some cases, there are strong cases for homologies between species, such as in early visual areas (Tootell et al., 1998). In other cases, discrepancies have been reported (Tootell et al., 1997). In higher order cognitive areas, there are some speculative suggestions for homologies (e.g., Culham, 2003; Rizzolatti & Arbib, 1998).
f. **Exploration of uniquely human functions**. Many functions may be considerably more developed in humans than other species and thus cannot be explored using comparative techniques. Neuroimaging has been used to investigate such topics including mathematical cognition (e.g., Naccache & Dehaene, 2001), theory of mind (e.g., Saxe et al., 2004), and tool usage (e.g., Chao & Martin, 2000). The rising new field of social cognitive neuroscience uses neuroimaging to investigate the links between human social behavior and brain mechanisms (Blakemore et al., 2004; Ochsner, 2004).
g. **Derivation of general organizational principles**. The ability of neuroimaging studies to investigate activation across many stimuli and tasks may provide some insight into the organization of the human brain. For example, there has been debate about whether object-selective regions of the temporal lobe have discrete or distributed representations of object categories (Ishai et al., 1999). One group has proposed an interesting theory that accounts for why category-selective subregions would exist in the arrangement that is observed across numerous fMRI studies (Hasson, Harel et al., 2003).
h. **Examination of irregular brain function**. Neuroimaging offers the opportunity to study the disordered human brain in action. In addition to using anatomical scans to determine lesion foci, functional neuroimaging can reveal which brain areas are impaired and which areas remain intact (e.g., Hasson, Avidan et al., 2003; James et al., 2003; Steeves et al., 2004). This can elucidate which brain areas may be necessary or sufficient for a particular cognitive function. In addition, neuroimaging may be able to determine how connectivity has been affected by lesions (Maguire et al., 2001). Neuroimaging studies may also help explain the unusual functioning of patient populations. For example, one fascinating study found activation of the auditory cortex in schizophrenic patients when they heard voices (Dierks et al., 1999). Neuroimaging may also become

> increasingly useful in the diagnosis of brain disorders and in the evaluation of treatment outcomes (Matthews & Jezzard, 2004).

Assuming that an interesting question does exist, one might consider whether neuroimaging is the best approach and whether other techniques might be able to address the same issue in a better way. Other techniques may offer benefits such as better temporal resolution (as with event-related potentials or transcranial magnetic stimulation) or may be more cost-effective (as with behavioral studies). In my opinion, neuroimaging experiments are most successful when they are based on a strong foundation of research from other domains, particularly behavioral studies. For instance, neuroimaging of human visual function has been quite successful (as reviewed in Wandell, 1999), likely because it is based on over a century of psychophysics, as well as considerable neurophysiology, neuropsychology, and modelling. Neuroimaging of poorly understood phenomena may be largely futile.

Assuming that the neural substrates subserving some aspect of cognition form an interesting question, the next step should be to generate plausible hypotheses (including the null hypothesis) and the conclusions that would be derived if those hypotheses were true. A worthwhile experiment demands more than one plausible hypothesis. If all theories predict the same outcome, there's no point in doing the experiment. For each hypothesis, the experimenter should consider the likelihood of that hypothesis being supported. Even with a potentially groundbreaking hypothesis, the experiment may not be worth doing if that outcome is highly improbable. The best experiments are the ones in which several possible hypotheses (including the null hypothesis) would each lead to interesting and publishable results. The results of a study should be worthy of at least one "publicon" (one unknown scientist's name for "the smallest unit of publishable matter") that advances the existing knowledge in the field. It's worth thinking in advance about what the headline of that publicon might be and whether it would be worth the time, effort and funding required.

Considering the expense of brain imaging, it is not always recommended to go straight from the thought experiment to the full data set. Often, it is worth running several iterations of pilot experiments and fine-tuning the design as necessary. Many newcomers plan to test hypotheses based on very subtle effects from other paradigms that are based on a complex series of assumptions. Sometimes, it may be worthwhile to establish basic effects and validate assumptions before moving on to subtleties. The temptation is always to include as many conditions as possible to provide the most stringent controls. Often, however, for the optimal statistical power during pilot testing, it is best to run a couple of subjects with the maximum number of trials in only the critical comparison conditions. The pilot results can be used to evaluate the minimum amount of data necessary to extract the basic effect. If the basic effect requires intensive data, it may be better to add control conditions in separate sessions.

Another issue worth considering is the acquisition of behavioral data. It may be beneficial to obtain behavioral data during pilot testing and/or during the actual scan session itself. Behavioral data such as response times, accuracy measures, and eye

movement monitoring can be useful in interpreting neuroimaging data. Often, it is desirable to have such measures equated across conditions, in which case, behavioral piloting can be very important. In other cases, the protocol used in basic behavioral testing may need to be modified to optimize neuroimaging designs. Here it is critical to ensure that the basic effect still holds, despite any changes to the paradigm. In event-related designs, behavioral data can provide additional information that can be correlated with brain activation on a trial-by-trial basis.

Consider the pros and cons of voxelwise vs. ROI approaches when deciding how to design your experiment

Surprisingly in brain imaging, scientists tend to exclusively follow either the voxelwise approach or the ROI approach, sometimes fervently. Often the choice is made not by careful thought, but rather by the experimenter's field (with the voxelwise approach more common from groups with a background in PET and the ROI approach more common from groups who have only utilized fMRI), geographical area (with the voxelwise approach more common in Europe and the ROI approach more common in North America), and statistical package (some of which facilitate one approach over the other).

Each approach has its pros and cons. Voxelwise approaches are very useful during an initial foray into the neural substrates of a behavior; whereas, ROI approaches are useful for characterizing a broad range of responses of a given area. Voxelwise approaches are useful when one has few hypotheses or very general hypotheses (e.g., Task X will activate Lobe Y); whereas, ROI approaches force one to generate specific hypothesis about known areas. Voxelwise approaches can lead to short experiments that focus on the contrasts of interest; whereas, experiments that use functionally-defined ROIs require that time be spent on localizer runs. Users of voxelwise approaches must be very knowledgeable about the current and ever-changing norms for statistics and may need to employ overly conservative corrections; whereas users of ROI approaches can use very basic statistics and more liberal thresholds because the hypotheses and regions are very limited. Voxelwise approaches are useful for large areas of activation that are relatively consistent between individual subjects. When anatomical foci are large and overlap well between subjects, voxelwise approaches make it easier to observe overall patterns of activity across the group without the confusion of intersubject variability. However, ROI approaches are preferable for areas that are small, are adjacent to other distinct areas, or require precise functional localization (e.g., retinotopic cortex). ROI approaches allow regions to be tailored to individual neuroanatomy. This can be quite important given that activation tends to be more consistent with respect to sulcal landmarks than stereotaxic coordinates (Watson et al., 1993) and sulci can vary considerably between subjects (Ono et al., 1990). A compromise between the two approaches may become more feasible with the development of algorithms that use sulcal landmarks and/or functionally-defined foci to constrain intersubject averaging (Fischl et al., 1999).

The two approaches are not mutually exclusive and there may be benefits to combining them. An ROI approach forces the experimenter to consider full hypotheses for at least one area. ROI approaches can be invaluable for establishing the role of critical regions

across many variants of tasks and stimuli. However, overdependence on the ROI approach can lead scientists to overlook other regions which may also play a critical role. As an example, the vast majority of studies on motion perception have focused on the most robust "motion area," the middle temporal complex (MT+), despite the fact that more than one dozen other areas have been implicated in motion processing (Sunaert et al., 1999). Neuroimaging has an advantage over single neuron recording by being able to easily investigate multiple regions simultaneously; yet many researchers ignore all but the clearest activations. In fact, it's not unheard of for experimenters to deliberately exclude regions whose activations they can't understand or that may detract from the areas they are interested in (with the cerebellum being a very common target for exclusion!) ROI approaches have an advantage in being able to facilitate cross-talk between laboratories. With stereotaxic coordinates, it's never certain as to whether two labs that report coordinates in the same vicinity are actually studying the same area. Monkey neurophysiology suggests that the human cortex may consist of a mosaic of small, closely packed regions finer than the resolution of group-averaged data. With ROI approaches, groups that use the same localizers and contrasts within individual subjects can be more confident that they are evaluating the same functional area. One approach that can be very useful is the development of a functional brain bank for a pool of commonly scanned subjects. Using modern analysis packages, each individual subject's data from each session can be aligned to a standard anatomical image from that subject. That way, data from that subject can be compared across many subjects, each with numerous localizers. This approach has been quite successful in many labs that utilize retinotopic mapping where the localizers are very time-intensive to collect (e.g., Tootell et al., 1998).

Think carefully about subtraction logic

As described earlier, neuroimaging studies must rely on subtraction logic to measure *differences* in activation rather than absolute levels of activation per se. The assumption of pure insertion requires that the two critical conditions being compared differ in only one important aspect. Ideally, the two conditions to be compared should differ either in the stimulus or in the task, but not both. Stimulus manipulations are best for cognitive processes that are largely automatic (e.g., visual processing, face perception) while task manipulations are better for controlled processes (e.g., attention).

Even "simple" baselines may not be simple (Gusnard & Raichle, 2001). Early PET studies often used a passive rest condition as a baseline, but even with rest, the brain is still active. Indeed, this has posed problems in domains such as memory research where the ongoing mental processes during rest may actually lead to higher activation than active memory tasks. Some suggest that other active baseline tasks that distract the subject from memory encoding and retrieval may be preferable (Stark & Squire, 2001). As yet, deactivations observed across many diverse tasks remain difficult to explain (Shulman et al., 1997).

The assumption of pure insertion is likely often false because components may not summate with simple addition. Consider an example of measuring the neural activation

for saccadic (i.e., rapid and sudden) eye movements to targets. The experimental condition may consist of eye movements to briefly flashed targets. Two control conditions are possible. In the most common design for a saccade system localizer, the control condition would be fixation on a central point. When subtracting fixation from saccades, the assumption is that the purely inserted component is the eye movement. Of course, in reality, other components are also being inserted. The flash of the target provokes a visual transient and draws attention. An alternative control condition would be to have fixation of a central point with target flashes that are ignored. Superficially, in the subtraction of this control from saccades, target transients and stimulus-driven attention would subtract out, leaving "purely inserted" saccades. The situation may not be so simple, however. Some neurons in saccade areas may respond to both stimulus-driven attention and to eye movements (e.g., Colby et al., 1996) such that the activation during the experimental condition depends on whether these components summate and whether they do so using straightforward (linear) addition. There may be other differences between the components of the tasks. Fixation during flashing targets may involve preparation and suppression of saccades and could in fact yield more signal than pure saccades.

More sophisticated designs can sometimes reduce the problems of selecting the correct baseline. One particularly attractive design is the parametric study, developed in cognitive psychology to avoid some of the similar problems with subtraction logic that occur with reaction time experiments (Sternberg, 1969). Parametric imaging studies search for brain activation that scales with the *amount* of a particular component that is added rather than searching for brain activation that increases with the first addition of that component. In our example of saccadic eye movements, an example would be to compare conditions with saccades at different rates (e.g., 2 vs. 1 vs. 0.5 saccades per second) on the assumption that areas involved in saccades would respond in proportion to saccade frequency (Beauchamp et al., 2000). Even with parametric designs, there may be reasons to expect nonlinear relationships between the stimulus manipulation and brain activation that can be used to address important theoretical issues (e.g., Culham, Cavanagh et al., 2001; Price et al., 1994). Another approach is the factorial design where at least two factors are manipulated in the same experiment to evaluate main effects and interactions (Price & Friston, 1997). In our example of a saccade task, we could have a 2x2 design with one manipulation being saccades vs. fixation, and the other being the use of static or flashed targets. That would lead to four conditions: (1) Fixation on a point with only static targets present; (2) saccades between static targets; (3) fixation on a point with flashed targets; and (4) saccades between flashed targets. A region with a role in saccadic eye movements would be expected to show a main effect of eye movements (saccades vs. fixation, collapsed across flashing and static events). A region with a role in stimulus-driven attention would be expected to show a main effect of stimulus transience (flashed vs. static, collapsed across saccade and. fixation events). More complex regions might show interactions (e.g., a much bigger response for saccades to flashed targets than would be predicted from the two main effects alone). Because factorial designs include comparisons across multiple conditions, they are less subject to the problems of a single subtraction between just two conditions.

Variations in task difficulty, or the degree of attention required to do the task, frequently violate the assumption of pure insertion. Many studies have demonstrated that brain activation is strongly modulated by the degree of attention devoted to the task (e.g., O'Craven et al., 1997). Attention is a common confound in many neuroimaging studies and may account for the fact that a surprising number of very diverse studies seem to find activation in the network of areas that subserves attention – including a large extent of the intraparietal sulcus, as well as the frontal eye fields (Corbetta et al., 1998; Culham & Kanwisher, 2001; Wojciulik & Kanwisher, 1999).

The common wisdom to avoid confounding differences in attention is to equate stimuli for attentional salience and to equate tasks for attentional performance. This is often easier in theory than in practice. Many stimuli are inherently more interesting than their control condition counterparts. A common solution is to introduce a task that demands attention for all stimuli such as a *1-back task* where the subject must press a button whenever the same stimulus appears twice in a row. This technique works well when the 1-back task has comparable difficulty in all conditions. Unfortunately, that isn't always the case. For example, 1-back tasks are common in localizers for object-selective visual areas in which objects and scrambled objects are compared. The 1-back task is easier for objects and subjects may use different strategies for objects (e.g., remembering semantic labels) than scramble objects (e.g., attending to spatial relationships). The 1-back approach works well in temporal cortex, but not so well in parietal cortex where object-selective areas may also be involved in attention and spatial processing.

In other cases, attention may be a fundamental aspect of neuronal function. For example, much debate has been sparked regarding the relationship between attention and awareness. In studies of awareness, then, would it be appropriate to even try to control for attention? Such issues have been raised in the context of motion illusions which have been suggested to relate to the perception of motion (Culham et al., 1999; He et al., 1998; Tootell, Reppas, Dale et al., 1995) vs. attention to motion (Huk et al., 2001). Given that the strength of the perceived motion is strongly related to the degree of attention directed to the motion (e.g., Beauchamp et al., 1997; Rees et al., 1997), the dichotomy may not be so straightforward.

Be aware of the caveats of neuroimaging

fMRI has been an invaluable tool in cognitive neuroscience, but the technique is still so new that the underlying mechanisms are to date, quite poorly understood. It is well-established that the blood-oxygenation-level dependent (BOLD) signal, measured by fMRI, indirectly reflects neuronal activity levels. What remains poorly understood is the relationship between various aspects of neuronal processing and the changes in BOLD activity. Much recent research has addressed this relationship and is well summarized elsewhere (Song, Huettel and McCarthy, this volume) (Logothetis & Wandell, 2004).

At present, the predominant hypothesis, proposed by Logothetis and colleagues (2001), is that BOLD activation may reflect synaptic inputs and local processing within an area, more so than action potentials. This may help to account for some discrepancies between

the results from human fMRI studies and monkey neurophysiology studies. For instance, human fMRI results indicate that attention modulates activity in primary visual cortex (e.g., Gandhi et al., 1999), in contrast to monkey neurophysiology where negligible attentional effects have been reported (e.g., Luck et al., 1997). This difference could arise from the different mechanisms of BOLD vs. action potentials or from the reduced spatial resolution of fMRI which cannot disentangle feedforward vs. feedback signals. Furthermore, both neurophysiology and BOLD imaging have inherent biases. Neurophysiological recordings are biased toward selecting larger neurons (such as pyramidal cells in cortex) while bypassing smaller neurons (Logothetis & Wandell, 2004). BOLD imaging depends on vascular density which is correlated with the number of synapses rather than the number of neurons in an area and varies considerably between brain areas (Duvernoy et al., 1981). For instance, primary visual cortex (V1) has more dense vascularization than adjacent extrastriate cortex (namely V2). Some have suggested that it may be easier to observe BOLD activation in some regions compared to others because of differences in vasculature (Harrison et al., 2002). This may account for why some regions, for example in frontal cortex (Duncan & Owen, 2000), are notoriously hard to activate, while others, such as the intraparietal sulcus, seem to be activated by practically anything (Culham & Kanwisher, 2001).

Cognitive neuroscientists tend to assume that neuronal action potentials measured by single neuron recording are the gold standard in interpreting brain processes; however, the BOLD signal may provide information about additional mechanisms beyond the action potential that are a fundamental component of neural processing. Given the biases of each technique, multiple approaches may provide a more complete picture.

In addition to the uncertainty regarding the relationship between neuronal spikes and BOLD activity, there are problems in using single neuron activity to predict the response of the population that will be measured by neuroimaging. Scannell and Young (1999) provide an excellent exposition of these issues. They used computational models of single unit activity in the MT+ motion complex to predict the population response that would be recorded with BOLD. They found cases in which it was possible to observe a subpopulation of neurons that responded vigorously to a stimulus without a corresponding difference in BOLD signal. The population response was a function of several factors including baseline firing rate, the modulation of the firing rate, the tuning function for each attribute and the number of attributes encoded. Their models suggested that effects which induce small changes in the firing rates of large numbers of neurons, specifically attention, have a much greater population result than effects which induce large changes in a small number of neurons. They also suggested that as experience leads to narrowing of tuning functions, BOLD population response differences become smaller.

Other problems arise in the actual modelling of the BOLD signal. Although the hemodynamic response profile has been well-characterized, data suggest that it can vary considerably from subject-to-subject (Aguirre et al., 1998) and from area-to-area. Thus many of the models employed may be accurate on average but subject to error for an individual case.

The paradigm of fMR adaptation holds great promise for expanding the types of questions that can be addressed with fMRI; however, it remains poorly understood and may be far more complex than it appears on the surface. The many names for the phenomenon, ranging from repetition suppression (Henson & Rugg, 2003) to adaptation (Grill-Spector & Malach, 2001) to priming (Henson, 2003), illustrate the confusion regarding possible mechanisms. In some brain regions, particularly the object-selective lateral occipital complex, the technique has been very successful (e.g., Grill-Spector et al., 1999; James et al., 2002; Kourtzi & Kanwisher, 2000b). In other regions however, even those where adaptation is known to occur from psychophysical techniques, some experimenters have failed to find it with fMRI (Boynton & Finney, 2003; Murray & Wojciulik, 2004). At this point, it is not clear why some paradigms produce better fMR adaptation than others. Differences may depend on the time scale of adaptation (Henson et al., 2000; Henson, 2003). Although some have suggested the adaptation approach avoids the problem of attentional confounds that occur with regular paradigms (Huk et al., 2001), priming does indeed appear to be highly dependent on attention (Eger et al., 2004; Murray & Wojciulik, 2004). Just as with regular fMRI, the specificity of adaptation is subject to differences in the mechanisms of blood supply across brain regions, and the technique may be particularly sensitive to synaptic inputs and local processing rather than outgoing action potentials (Tolias et al., 2001). Finally, the correspondence of such repetition effects with behavioral measures of priming is often incomplete (Henson, 2003; Henson & Rugg, 2003; Thiel et al., 2001). Given these concerns, while fMR adaptation remains an exciting tool for neuroimagers, a better understanding is necessary to make sense of the results it produces.

Look at your data and understand what your preprocessing and statistics are doing

Since the advent of neuroimaging, many sophisticated analysis packages for neuroimaging data have been developed. Many of these packages offer a bewildering array of options to the novice neuroimager. The temptation can be to simply follow the standard recipe to "see what blobs light up," without ever looking at the raw data. This can be risky, particularly with temperamental fMRI magnets where days of "dysfunctional" MRI are not uncommon. Even simple exploration, such as *voxel-surfing* – viewing the time courses of voxels selected at random -- can be enlightening. For example, some data sets may have spikes in the data that can lead to spurious activation or can prevent real activation from passing the statistical threshold. In my opinion (and I realize large camps of neuroimagers would disagree), the ability to easily view time course data, either for random voxels or for regions of interest, is an underrated tool in statistical analysis packages. Voxel-surfing (or region-surfing) can be a very effective way to flag problems in the data and violations of statistical assumptions. Viewing of movies of the brain over subsequent points in time can also reveal possible problems with the data (e.g., spikes may appear as a brightening of the entire image) or with the subject (e.g., head motion will appear as shifts in the movie). Movie viewing *following* motion correction can sometimes reveal residual problems that remain. Analysis of single subjects and/or single runs can sometimes indicate problematic data that should be removed from the final analysis (of course, this should only be done when there is a valid

objective reason for doing so, not simply because a subject's data is inconsistent with your hypothesis).

It is often worth analyzing data in several different ways to evaluate the effect of various manipulations. In the case of robust data, the conclusions will be the same, regardless of the specific analysis. However, in some cases, different analyses may produce different results, and the reasons for the differences may themselves be enlightening. For example, one might want to evaluate an effect with both an ROI approach and a voxelwise approach. If the two approaches lead to similar conclusions, there is no problem. If they lead to different conclusions, the underlying reasons may be interesting. If a result is significant in an ROI approach but not a voxelwise approach, it may indicate that the effect is weak and the voxelwise approach is too stringent. If a voxelwise analysis shows activation within a subregion of the ROI, it may be that functional subdivisions exist within the region. I am *not* advocating an approach of trying every possible combination of analyses to find *something* that reaches significance. It's still important to have an *a priori* plan about how best to proceed, and to maintain standards of acceptable data analysis.

Some neuroimaging statistics packages offer formulaic steps for data analysis that, although seemingly simple, warrant careful consideration. One excellent example is the application of motion correction algorithms. There is a tendency to assume that these algorithms are a panacea for possible head motion artifacts and that, as long as the algorithms are run as a default, concerns about motion have been sufficiently addressed. This is not always true and in some cases, motion correction can even induce artifacts that were not present in the original data (Freire & Mangin, 2001). In my experience, the success of these algorithms is quite limited and often does not suffice to fix all motion-related problems, particularly ones with large abrupt movements. The algorithms assume that the head can move with six degrees of freedom (three translations and three rotations) and that as long as the shifted position of the head is reset to the starting point, the problem will have been fixed. There are several problems with this approach. First, the data must be resampled, thus information is lost because it is unavailable (particularly in the top and bottom slices) or because it has been interpolated. Prospective motion correction, in which slices are replanned on-the-fly with each volume adjusted based on detected motion, is now being provided by some scanner manufacturers and appears to provide superior output that traditional post hoc corrections (Lee et al., 1996). Second, although adjustment of six parameters works well with PET data, the situation is more complex in fMRI. Unfortunately in fMRI, when the head moves, it not only changes its position relative to the frame of the image, but it also changes relative to the magnetic field. In an inhomogeneous field, this means that the overall signal level can remain distorted when the head is replaced back to its original position in the image (Jezzard & Clare, 1999). A procedure known as *shimming* attempts to make the magnetic field as homogeneous as possible, but it is never completely successful. The pattern of inhomogeneities is dependent on the distribution of mass in the bore. With movement of the head, this distribution changes, such that, not only is the head in a different part of the field, but the field itself may be different and thus signal levels throughout the volume may be affected. Simple coregistration of volumes over time ignores this problem.

Furthermore, field distortions can arise due to the changing distribution of mass from body parts other than the head. In cases where movements are correlated with the paradigm (e.g., studies of speech, swallowing or grasping), slow, event-related paradigms can be used to dissociate the artifacts (which occur with no lag) from brain activation (which occurs with the appropriate hemodynamic lag, Birn et al., 1999). Figure 1 illustrates a case where movement of mass within the field leads to profound artifacts that lead to faulty motion detection and correction. Such artifacts occur whenever any mass moves in the field and is worst for movements closest to the head (as in head movements themselves but also movements of the tongue, throat and jaw, as in swallowing artifacts, and with movements of the chest during respiration). New subjects are always instructed not to move the head; however, we have found much better results with more detailed instructions that discourage swallowing and changes to mouth and body posture during functional scans (Breathing, however, is encouraged). Although motion correction algorithms may sometimes improve data quality, nothing works as well as *preventing* motion in the first place. Although the development of high field strength and impressive new technology (e.g., phased array coils) has greatly improved signal-to-noise, so-called *physiological noise* (of which motion is a large component) remains a major limiting factor in data quality.

The researcher should be aware of common fallacies in statistical logic and the representation of statistical data. Here are three examples:

a. One very common mistake is the indirect comparison. The researcher assumes that if a particular region responds more in one condition than the control state, but not more in a second condition than the control state, then the activation must be higher in the first state than the second. To understand why this logic is false, recall that statistical significance depends not only on the size of the effect, but also on the threshold selected and the variability in the data. As shown in Figure 2, one contrast may be more significant than another because of a larger effect size, or alternatively, because of small differences in statistical significance or because of reduced variability. In sum, "Differences in significance do not necessarily indicate significant differences," (as succinctly stated by Nancy Kanwisher, personal communication). The straightforward solution is to perform *direct* comparisons between critical conditions.

b. Another common mistake is to misinterpret interactions in factorial designs. An interaction occurs whenever a manipulation has different effect sizes depending on the status of another manipulation. When graphing factorial effects, an interaction can be seen by any pair (or more) of lines which do not run parallel to one another. Consider a hypothesis that attention to a stimulus will amplify activity in brain areas known to process that stimulus, but not other areas. ROIs selective to faces (FFA) and scenes (parahippocampal place area, or PPA) could be identified using a localizer scan. Subjects could then be shown photographs of either faces or places (main effect of object category) and ask them to attend to the stimulus or to an irrelevant attribute (e.g., the color of the fixation point). The hypothesis leads to the prediction that in the FFA, the difference between

attending vs. not attending to the stimulus will be higher for faces than for places (see Figure 3). Conversely, the hypothesis predicts that in the PPA, the difference between attending vs. not attending to the stimulus will be higher for places than for faces. For simplicity, let's consider only the FFA. The obvious way to test whether the hypothesis holds true would be to perform a voxelwise 2x2 factorial analysis (stimulus category x attentional state) and to look for a significant interaction. If an interaction is found, is the hypothesis correct? Not necessarily. Just as in conventional statistics, interactions can occur for many reasons. To phrase it another way, in the graphs for a factorial design, the lines can be non-parallel for several reasons. Post hoc t-tests would help confirm the pattern (e.g., attend face > attend elsewhere but not attend place > attend elsewhere).

c. A third common problem is that of "proving the obvious" in ROI approaches. When you use statistics to select a region, by necessity, that region must demonstrate the effect indicated in the statistics. For example, if you identify the motion complex, MT+, by contrasting moving vs. stationary images and then do a post hoc test to compare moving minus stationary conditions, it should come as little surprise that they differ significantly. This seems obvious, but it is not uncommon to see manuscripts that comment on significant differences within an area defined by a statistical test of that difference. In cases where post hoc analyses are performed on the same comparison used to localize a region, it should be clarified whether they would by necessity be significant or whether the analysis adds new information.

Future directions

Neuroimaging is a very young science experiencing a combination of enthusiasm, cynicism, and growing pains. Progress in the past decade or two has been phenomenal. Basic scanning techniques have improved considerably and many recent capabilities of fMRI, such as event-related designs and fMR adaptation, were unanticipated in the early years. Recent developments hold much promise for solutions to many of the problems described here, and for adding powerful new techniques to the cognitive neuroscientist's toolbox. Some of the most promising recent developments include (1) the use of diffusion tensor imaging (e.g., Basser et al., 2000) and functional connectivity studies (Friston et al., 1997; McIntosh, 1999) to determine the wiring of functional brain areas; (2) the use of mental chronometry to push the temporal resolution of fMRI (Formisano & Goebel, 2003); (3) the development of imaging genomics to correlate genetic markers, brain activation, and behavior (Hariri & Weinberger, 2003); (4) the combination of neuroimaging with other techniques such as ERPs, magnetoencephalography (MEG) and or TMS to benefit from the combined strengths of each technique (Paus et al., 1997); and (5) the combination of monkey physiology, monkey neuroimaging, and human neuroimaging to better understand differences between techniques and between species (Logothetis et al., 1999; Orban et al., 2004).

Acknowledgements

I would like to thank many colleagues over the years who have provided stimulating discussions and debates about experimental design. In particular, I would like to thank Nancy Kanwisher who first gave me the opportunity to become involved in brain imaging and shared many valuable insights into experimental design, some of which I have included here. I thank my assistant, Ken Valyear, for comments, for assistance with the preparation of Figure 1 and for valuable discussion of fMR adaptation caveats. I also thank Matt Brown, Christine Carter, Cristiana Cavina Pratesi, Stefan Kohler, Mary Ellen Large, Anthony Singhal and an anonymous reviewer for comments on the manuscript. Any interpretations, omissions or errors are my own. This work was supported by the Canadian Institutes for Health Research and the National Sciences and Engineering Research Council of Canada.

References

Aguirre, G. K., & D'Esposito, M. (1999). Experimental design for fMRI. In C. T. W. Moonen & P. A. Bandettini (Eds.), *Functional MRI* (pp. 369-380). Berlin: Springer-Verlag.

Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1998). The variability of human, BOLD hemodynamic responses. *Neuroimage, 8*, 360-369.

Bandettini, P. A., & Cox, R. W. (2000). Event-related fMRI contrast when using constant interstimulus interval: theory and experiment. *Magnetic Resonance in Medicine, 43*, 540-548.

Basser, P. J., Pajevic, S., Pierpaoli, C., Duda, J., & Aldroubi, A. (2000). In vivo fiber tractography using DT-MRI data. *Magnetic Resonance in Medicine, 44*, 625-632.

Beauchamp, M. S., Cox, R. W., & DeYoe, E. A. (1997). Graded effects of spatial and featural attention on human area MT and associated motion processing areas. *Journal of Neurophysiology, 78*, 516-520.

Beauchamp, M. S., Petit, L., Ellmore, T. M., Ingeholm, J., & Haxby, J. V. (2000). A parametric fMRI study of overt and covert shifts of visuospatial attention. *Society for Neuroscience Abstracts, 26*, 1586.

Birn, R. M., Bandettini, P. A., Cox, R. W., & Shaker, R. (1999). Event-related fMRI of tasks involving brief motion. *Human Brain Mapping, 7*, 106-114.

Biswal, B. B., & Ulmer, J. L. (1999). Blind source separation of multiple signal sources of fMRI data sets using independent component analysis. *Journal of Computational Assisted Tomography, 23*, 265-271.

Blakemore, S. J., Winston, J., & Frith, U. (2004). Social cognitive neuroscience: where are we heading? *Trends in Cognitive Sciences, 8*, 216-222.

Blamire, A. M., Ogawa, S., Ugurbil, K., Rothman, D., McCarthy, G., Ellermann, J. M., et al. (1992). Dynamic mapping of the human visual cortex by high-speed magnetic resonance imaging. *Proceedings of the National Academy of Sciences (USA), 89*, 11069-11073.

Boynton, G. M., & Finney, E. M. (2003). Orientation-specific adaptation in human visual cortex. *J Neurosci, 23*, 8781-8787.

Brewer, J. B., Zhao, Z., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1998). Making memories: brain activity that predicts how well visual experience will be remembered. *Science, 281*, 1185-1187.

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 (USA), 93*, 14878-14883.

Buckner, R. L., & Braver, T. S. (1999). Event-related functional MRI. In C. T. W. Moonen & P. A. Bandettini (Eds.), *Functional MRI* (pp. 441-452). Berlin: Springer-Verlag.

Buckner, R. L., & Logan, J. M. (2001). Functional neuroimaging methods: PET and fMRI. In C. R. & A. Kingstone (Eds.), *Handbook of Functional Neuroimaging of Cognition* (Vol. 1, pp. 27-48). Cambridge, MA: MIT Press.

Chao, L. L., & Martin, A. (2000). Representation of manipulable man-made objects in the dorsal stream. *Neuroimage, 12*, 478-484.

Chein, J. M., & Schneider, W. (In press). Designing effective fMRI experiments. In J. Grafman & I. Robertson (Eds.), *Handbook of Neuropsychology*. Amsterdam: Elsevier Science.

Colby, C. L., Duhamel, J.-R., & Goldberg, M. E. (1996). Visual, presaccadic, and cognitive activation of single neurons in monkey lateral intraparietal area. *Journal of Neurophysiology, 76*, 2841-2851.

Corbetta, M., Akbudak, E., Conturo, T. E., Snyder, A. Z., Ollinger, J. M., Drury, H. A., et al. (1998). A common network of functional areas for attention and eye movements. *Neuron, 21*, 761-773.

Crick, F., & Jones, E. (1993). Backwardness of human neuroanatomy. *Nature, 361*, 109-110.

Culham, J., He, S., Dukelow, S., & Verstraten, F. A. (2001). Visual motion and the human brain: what has neuroimaging told us? *Acta Psychol (Amst), 107*, 69-94.

Culham, J. C. (2003). Human brain imaging reveals a parietal area specialized for grasping. In N. Kanwisher & J. Duncan (Eds.), *Attention and Performance XX: Functional Brain Imaging of Human Cognition*. Oxford, U.K.: Oxford University Press.

Culham, J. C., Cavanagh, P., & Kanwisher, N. G. (2001). Attention response functions. Characterizing brain areas using fMRI activation during parametric variations of attentional load. *Neuron, 32*, 737-745.

Culham, J. C., Danckert, S. L., DeSouza, J. F., Gati, J. S., Menon, R. S., & Goodale, M. A. (2003). Visually guided grasping produces fMRI activation in dorsal but not ventral stream brain areas. *Experimental Brain Research, 153*, 180-189.

Culham, J. C., Dukelow, S. P., Vilis, T., Hassard, F. A., Gati, J. S., Menon, R. S., et al. (1999). Recovery of fMRI activation in motion area MT following storage of the motion aftereffect. *Journal of Neurophysiology, 81*, 388-393.

Culham, J. C., & Kanwisher, N. G. (2001). Neuroimaging of cognitive functions in human parietal cortex. *Current Opinion in Neurobiology, 11*, 157-163.

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

Dierks, T., Linden, D. E., Jandl, M., Formisano, E., Goebel, R., Lanfermann, H., et al. (1999). Activation of Heschl's gyrus during auditory hallucinations. *Neuron, 22*, 615-621.

Donaldson, D. I. (2004). Parsing brain activity with fMRI and mixed designs: what kind of a state is neuroimaging in? *Trends in Neurosciences, 27*, 442-444.

Donaldson, D. I., & Buckner, R. L. (2001). Effective paradigm design. In P. Jezzard, P. M. Matthews & S. M. Smith (Eds.), *Functional MRI: An Introduction to Methods* (pp. 177-195). Oxford UK: Oxford University Press.

Donders, F. C. (1969). On the speed of mental processes. *Acta Psychologica, 30*, 412-431.

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences, 23*, 475-483.

Duvernoy, H. M., Delon, S., & Vannson, J. L. (1981). Cortical blood vessels of the human brain. *Brain Research Bulletin, 7*, 519-579.

Eger, E., Henson, R. N., Driver, J., & Dolan, R. J. (2004). BOLD repetition decreases in object-responsive ventral visual areas depend on spatial attention. *Journal of Neurophysiology, 92*, 1241-1247.

Evans, A. C., Collins, D. L., Mills, S. R., Brown, E. D., Kelly, R. L., & Peters, T. M. (1993). 3D statistical neuroanatomical models from 305 MRI volumes. *Proceedings of the IEEE-Nuclear Science Symposium and Medical Imaging Conference*, 1813-1817.

Fischl, B., Sereno, M. I., Tootell, R. B., & Dale, A. M. (1999). High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Mapping, 8*, 272-284.

Formisano, E., & Goebel, R. (2003). Tracking cognitive processes with functional MRI mental chronometry. *Current Opinion in Neurobiology, 13*, 174-181.

Fox, P. T. (1997). The growth of human brain mapping. *Human Brain Mapping, 5*, 1-2.

Fox, P. T., Mintun, M. A., Raichle, M. E., Miezin, F. M., Allman, J. M., & Van Essen, D. C. (1986). Mapping human visual cortex with positron emission tomography. *Nature, 323*, 806-809.

Freire, L., & Mangin, J. F. (2001). Motion correction algorithms may create spurious brain activations in the absence of subject motion. *Neuroimage, 14*, 709-722.

Frisby, J. P. (1979). *Seeing: Illusion, Brain, and Mind*. Oxford: Oxford University Press.

Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage, 6*, 218-229.

Gandhi, S. P., Heeger, D. J., & Boynton, G. M. (1999). Spatial attention affects brain activity in human primary visual cortex. *Proceedings of the National Academy of Sciences (USA), 96*, 3314-3319.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience, 3*, 191-197.

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nature Neuroscience, 2*, 568-573.

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

Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzchak, Y., & Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron, 24*, 187-203.

Grill-Spector, K., & Malach, R. (2001). fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychol (Amst), 107*, 293-321.

Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience, 27*, 649-677.

Gusnard, D. A., & Raichle, M. E. (2001). Searching for a baseline: functional imaging and the resting human brain. *Nature Reviews Neuroscience, 2*, 685-694.

Hadjikhani, N., Joseph, R. M., Snyder, J., Chabris, C. F., Clark, J., Steele, S., et al. (2004). Activation of the fusiform gyrus when individuals with autism spectrum disorder view faces. *Neuroimage, 22*, 1141-1150.

Hariri, A. R., & Weinberger, D. R. (2003). Imaging genomics. *British Medical Bulletin, 65*, 259-270.

Harrison, R. V., Harel, N., Panesar, J., & Mount, R. J. (2002). Blood capillary distribution correlates with hemodynamic-based functional imaging in cerebral cortex. *Cerebral Cortex, 12*, 225-233.

Hasson, U., Avidan, G., Deouell, L. Y., Bentin, S., & Malach, R. (2003). Face-selective activation in a congenital prosopagnosic subject. *J Cogn Neurosci, 15*, 419-431.

Hasson, U., Harel, M., Levy, I., & Malach, R. (2003). Large-scale mirror-symmetry organization of human occipito-temporal object areas. *Neuron, 37*, 1027-1041.

He, S., Cohen, E. R., & Hu, X. (1998). Close correlation between activity in brain area MT/V5 and the perception of a visual motion aftereffect. *Current Biology, 8*, 1215-1218.

Henson, R., Shallice, T., & Dolan, R. (2000). Neuroimaging evidence for dissociable forms of repetition priming. *Science, 287*, 1269-1272.

Henson, R. N. (2003). Neuroimaging studies of priming. *Progress in Neurobiology, 70*, 53-81.

Henson, R. N., & Rugg, M. D. (2003). Neural response suppression, haemodynamic repetition effects, and behavioural priming. *Neuropsychologia, 41*, 263-270.

Huettel, S. A., Song, A. W., & McCarthy, G. (2004). *Functional magnetic resonance imaging*

Huk, A. C., Ress, D., & Heeger, D. J. (2001). Neuronal basis of the motion aftereffect reconsidered. *Neuron, 32*, 161-172.

Ishai, A., Ungerleider, L. G., Martin, A., Schouten, J. L., & Haxby, J. V. (1999). Distributed representation of objects in the human ventral visual pathway. *Proceedings of the National Academy of Sciences (USA), 96*, 9379-9384.

James, T. W., Culham, J., Humphrey, G. K., Milner, A. D., & Goodale, M. A. (2003). Ventral occipital lesions impair object recognition but not object-directed grasping: an fMRI study. *Brain, 126*, 2463-2475.

James, T. W., Humphrey, G. K., Gati, J. S., Servos, P., Menon, R. S., & Goodale, M. A. (2002). Haptic study of three-dimensional objects activates extrastriate visual areas. *Neuropsychologia, 40*, 1706-1714.

Jezzard, P., & Clare, S. (1999). Sources of distortion in functional MRI data. *Human Brain Mapping, 8*, 80-85.

Jezzard, P., Matthews, P. M., & Smith, S. M. (2001). *Functional MRI: An Introduction to Methods*. Oxford UK: Oxford University Press.

Kanwisher, N., Downing, P., Epstein, R., & Kourtzi, Z. (2001). Functional neuroimaging of visual recognition. In R. Cabeza & A. Kingstone (Eds.), *Handbook of Functional Neuroimaging of Cognition* (pp. 109-152). Cambridge, MA: MIT Press.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*, 4302-4311.

Klein, R. M. (1980). Does oculomotor readiness mediate cognitive control of visual attention? In R. S. Nickerson (Ed.), *Attention and Performance VII* (pp. 259-276). Hillsdale, NJ: Erlbaum.

Kosslyn, S. M. (1999). If neuroimaging is the answer, what is the question? *Philos Trans R Soc Lond B Biol Sci, 354*, 1283-1294.

Kourtzi, Z., & Kanwisher, N. (2000a). Activation in human MT/MST by static images with implied motion. *Journal of Cognitive Neuroscience, 12*, 48-55.

Kourtzi, Z., & Kanwisher, N. (2000b). Cortical regions involved in perceiving object shape. *Journal of Neuroscience, 20*, 3310-3318.

Kwong, K. K., Belliveau, J. W., Chesler, D. A., Goldberg, I. E., Weisskoff, R. M., Poncelet, B. P., et al. (1992). Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proceedings of the National Academy of Sciences (USA), 89*, 5675-5679.

Lazar, S. W., Bush, G., Gollub, R. L., Fricchione, G. L., Khalsa, G., & Benson, H. (2000). Functional brain mapping of the relaxation response and meditation. *Neuroreport, 11*, 1581-1585.

Lee, C. C., Jack, C. R., Jr., Grimm, R. C., Rossman, P. J., Felmlee, J. P., Ehman, R. L., et al. (1996). Real-time adaptive motion correction in functional MRI. *Magnetic Resonance in Medicine, 36*, 436-444.

Logothetis, N. K., Guggenberger, H., Peled, S., & Pauls, J. (1999). Functional imaging of the monkey brain. *Nature Neuroscience, 2*, 555-562.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature, 412*, 150-157.

Logothetis, N. K., & Schall, J. D. (1989). Neuronal correlates of subjective visual perception. *Science, 245*, 761-763.

Logothetis, N. K., & Wandell, B. A. (2004). Interpreting the BOLD signal. *Annu Rev Physiol, 66*, 735-769.

Luck, S. J., Chelazzi, L., Hillyard, S. A., & Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *Journal of Neurophysiology, 77*, 24-42.

Maguire, E. A., Frackowiak, R. S. J., & Frith, C. D. (1997). Recalling routes around London: activation of the right hippocampus in taxi drivers. *Journal of Neuroscience, 17*, 7103-7110.

Maguire, E. A., Vargha-Khadem, F., & Mishkin, M. (2001). The effects of bilateral hippocampal damage on fMRI regional activations and interactions during memory retrieval. *Brain, 124*, 1156-1170.

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., et al. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences (USA), 92*, 8135-8139.

Matthews, P. M., & Jezzard, P. (2004). Functional magnetic resonance imaging. *Journal of Neurology, Neurosurgery and Psychiatry, 75*, 6-12.

McIntosh, A. R. (1999). Mapping cognition to the brain through neural interactions. *Memory, 7*, 523-548.

McKeown, M. J., Jung, T.-P., Makeig, S., Brown, G., Kinderman, S. S., Lee, T.-W., et al. (1998). Spatially independent activity patterns in functional magnetic resonance imaging data during the Stroop color-naming task. *Proceedings of the National Academy of Sciences (USA), 95*, 803-810.

Menon, R. S., & Goodyear, B. G. (2001). Spatial and temporal resolution in fMRI. In P. Jezzard, P. M. Matthews & S. M. Smith (Eds.), *Functional MRI: An Introduction to Methods* (pp. 145-158). Oxford UK: Oxford University Press.

Milner, A. D., & Goodale, M. A. (1995). *The Visual Brain in Action*. Oxford, England: Oxford University Press.

Murray, S. O., & Wojciulik, E. (2004). Attention increases neural selectivity in the human lateral occipital complex. *Nature Neuroscience, 7*, 70-74.

Naccache, L., & Dehaene, S. (2001). The priming method: imaging unconscious repetition priming reveals an abstract representation of number in the parietal lobes. *Cerebral Cortex, 11*, 966-974.

Ochsner, K. N. (2004). Current directions in social cognitive neuroscience. *Current Opinion in Neurobiology, 14*, 254-258.

O'Craven, K. M., Rosen, B. R., Kwong, K. K., Treisman, A., & Savoy, R. L. (1997). Voluntary attention modulates fMRI activity in human MT-MST. *Neuron, 18*, 591-598.

Ogawa, S., Tank, D. W., Menon, R., Ellermann, J. M., Kim, S.-G., Merkle, H., et al. (1992). Intrinsic signal changes accompanying sensory stimulation: Functional brain mapping with magnetic resonance imaging. *Proceedings of the National Academy of Sciences (USA), 89*, 5951-5955.

Ono, M., Kubik, S., & Abernathey, C. D. (1990). *Atlas of the Cerebral Sulci*. Stuttgart: Thieme Medical Publishers.

Orban, G. A., Van Essen, D., & Vanduffel, W. (2004). Comparative mapping of higher visual areas in monkeys and humans. *Trends in Cognitive Sciences, 8*, 315-324.

Paus, T., Jech, R., Thompson, C. J., Comeau, R., Peters, T., & Evans, A. C. (1997). Transcranial magnetic stimulation during positron emission tomography: a new method for studying connectivity of the human cerebral cortex. *Journal of Neuroscience, 17*, 3178-3184.

Price, C. J., & Friston, K. J. (1997). Cognitive conjunction: a new approach to brain activation experiments. *Neuroimage, 5*, 261-270.

Price, C. J., Wise, R. J., Watson, J. D., Patterson, K., Howard, D., & Frackowiak, R. S. (1994). Brain activity during reading. The effects of exposure duration and task. *Brain, 117 (Pt 6)*, 1255-1269.

Rees, G. (2001). Neuroimaging of visual awareness in patients and normal subjects. *Current Opinion in Neurobiology, 11*, 150-156.

Rees, G., Frith, C. D., & Lavie, N. (1997). Modulating irrelevant motion perception by varying attentional load in an unrelated task. *Science, 278*, 1616-1619.

Ress, D., & Heeger, D. J. (2003). Neuronal correlates of perception in early visual cortex. *Nature Neuroscience, 6*, 414-420.

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences, 21*, 188-194.

Rizzolatti, G., Riggio, L., & Sheliga, B. M. (1994). Space and selective attention. In C. Umilta & M. Moscovitch (Eds.), *Attention and Performance XV* (pp. 231-265). Cambridge: MIT Press.

Savoy, R. (2001). History and future directions of human brain mapping. *Acta Psychologica*.

Saxe, R., Carey, S., & Kanwisher, N. (2004). Understanding other minds: linking developmental psychology and functional neuroimaging. *Annual Review of Psychology, 55*, 87-124.

Scannell, J. W., & Young, M. P. (1999). Neuronal population activity and functional imaging. *Proceedings of the Royal Society of London, B., 266*, 875-881.

Shulman, G. L., Fiez, J. A., Corbetta, M., Buckner, R. L., Miezin, F. M., Raichle, M. E., et al. (1997). Common blood flow changes across visual tasks: II. decreases in cerebral cortex. *Journal of Cognitive Neuroscience, 9*, 648-663.

Stark, C. E., & Squire, L. R. (2001). When zero is not zero: the problem of ambiguous baseline conditions in fMRI. *Proceedings of the National Academy of Sciences (USA), 98*, 12760-12766.

Steeves, J. K. E., Humphrey, G. K., Culham, J. C., Menon, R. S., Milner, A. D., & Goodale, M. A. (2004). Behavioral and neuroimaging evidence for a contribution of color and texture information to scene classification in a patient with visual form agnosia. *Journal of Cognitive Neuroscience, 16*, 955-965.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30*, 276-315.

Sunaert, S., Van Hecke, P., Marchal, G., & Orban, G. A. (1999). Motion-responsive regions of the human brain. *Experimental Brain Research, 127*, 355-370.

Takada, T., & Miyamoto, T. (2004). A fronto-parietal network for chewing of gum: a study on human subjects with functional magnetic resonance imaging. *Neurosci Lett, 360*, 137-140.

Talairach, J., & Tournoux, P. (1988). *Co-Planar Stereotaxic Atlas of the Human Brain*. New York: Thieme Medical Publishers.

Thiel, C. M., Henson, R. N., Morris, J. S., Friston, K. J., & Dolan, R. J. (2001). Pharmacological modulation of behavioral and neuronal correlates of repetition priming. *Journal of Neuroscience, 21*, 6846-6852.

Tolias, A. S., Smirnakis, S. M., Augath, M. A., Trinath, T., & Logothetis, N. K. (2001). Motion processing in the macaque: revisited with functional magnetic resonance imaging. *Journal of Neuroscience, 21*, 8594-8601.

Tong, F., Nakayama, K., Vaughan, J. T., & Kanwisher, N. (1998). Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron, 21*, 753-759.

Tootell, R. B., Tsao, D., & Vanduffel, W. (2003). Neuroimaging weighs in: humans meet macaques in "primate" visual cortex. *Journal of Neuroscience, 23*, 3981-3989.

Tootell, R. B. H., Hadjikhani, N. K., Mendola, J. D., Marrett, S., & Dale, A. M. (1998). From retinotopy to recognition: fMRI in visual cortex. *Trends in Cognitive Sciences, 2*, 174-183.

Tootell, R. B. H., Mendola, J. D., Hadjikhani, N. K., Ledden, P. J., Lui, A. K., Reppas, J. B., et al. (1997). Functional analysis of V3A and related areas in human visual cortex. *Journal of Neuroscience, 17*, 7060-7078.

Tootell, R. B. H., Reppas, J. B., Dale, A. M., Look, R. B., Sereno, M. I., Malach, R., et al. (1995). Visual motion aftereffect in human cortical area MT revealed by functional magnetic resonance imaging. *Nature, 375*, 139-141.

Tootell, R. B. H., Reppas, J. B., Kwong, K. K., Malach, R., Born, R. T., Brady, T. J., et al. (1995). Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *Journal of Neuroscience, 15*, 3215-3230.

Uttal, W. R. (2001). *The new phrenology : the limits of localizing cognitive processes in the brain*. Cambridge, Mass. ; London: MIT Press.

Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A. M., et al. (1998). Building memories: remembering and forgetting of verbal experiences as predicted by brain activity. *Science, 281*, 1188-1191.

Wandell, B. A. (1999). Computational neuroimaging of human visual cortex. *Annual Review of Neuroscience, 22*, 145-173.

Watson, J. D. G., Myers, R., Frackowiak, R. S. J., Hajnal, J. V., Woods, R. P., Mazziotta, J. C., et al. (1993). Area V5 of the human brain: Evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cerebral Cortex, 3*, 79-94.

Wojciulik, E., & Kanwisher, N. (1999). The generality of parietal involvement in visual attention. *Neuron, 23*, 747-764.

Figure Captions

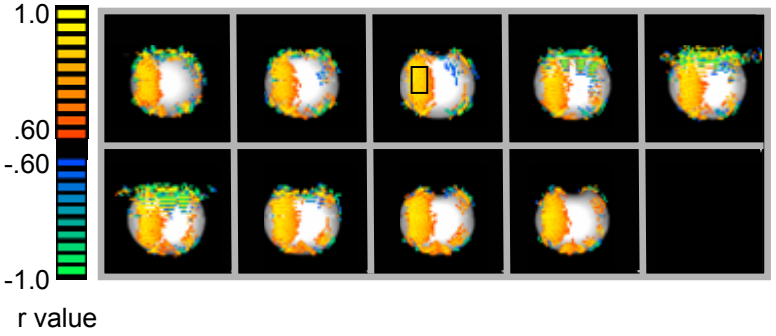
Figure 1. Experiment illustrating the contamination of data by the movement of mass within the magnetic field. In this experiment, a spherical phantom was placed in a head coil and remained completely stationary throughout the acquisition of T2*-weighted "functional" scans. Four 1-L (1 kg) bags of saline were placed end-to-end on a wooden pole to approximate the mass of a human arm. Every 30 s, the mass moved between two positions to the left and right of the phantom and then remained stationary for the remainder of the interval. Results are shown for a 24 cm movement at a distance of 10 cm from the phantom. Similar, though reduced effects were seen with smaller masses (as low as 2 kg) or the mass of a real arm, with smaller movements (as small as 5 cm) and with larger distances from the phantom (up to 30 cm). A. A comparison between the signal when the mass was to the right vs. left of the phantom illustrates widespread false "activation". The color table on the left shows the statistical correlation (r) value. B. The time course within a region denoted by the black box in the third slice of A is shown by the black line. Note the changes in overall signal level and the brief spikes whenever the external mass changed position. The reference time course used to generate the map in A is indicated by the blue line. C. A statistical map showing regions where spiking occurred at the time of movement. D. The time course within the region indicated by the black box in the third slice of C is shown by the black line. The reference time course used to generate the map in C is indicated by the blue line. E. The comparison between the signal with the mass at each position is shown after motion correction was applied. Note there is no reduction of false activation. F. Output of the motion detection and correction algorithm provided in Brain Voyager (Brain Innovation, Maastricht, The Netherlands). Similar results were obtained with various forms of motion correction in the Statistical Parametric Mapping (SPM) package (Functional Imaging Labs, University College London, London UK). The resulting three translation parameters are shown by blue lines and the three rotation parameters are shown by red lines. Note that although there was no movement whatsoever in the phantom, the algorithms detect and falsely attempt to correct for spurious motion created by the moving external mass.

Figure 2. Schematic illustration of faulty statistical logic regarding differences of significance. In all three hypothetical data sets, the statistics indicate that Condition I is *significantly* different from Condition III while Condition II is *not significantly* different from Condition III. Is it therefore true that Condition I is different from Condition II in all cases? (A). The difference in significance between the two comparisons (I vs. III and II vs. III) occurs primarily because of differences in the magnitude of activation. (B). The difference in significance also occurs because of differences in the magnitude of activation, but may or may not be present depending on the statistical threshold chosen. (C). The difference in significance occurs primarily because of differences in the variability of activation. Error bars indicate 95% confidence intervals.

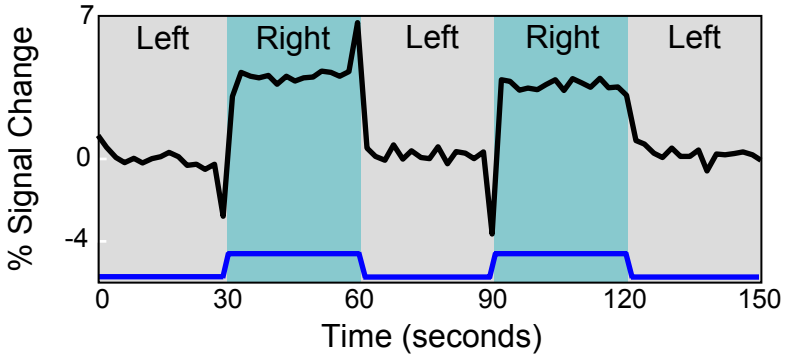
Figure 3. Schematic illustration of faulty statistical logic regarding interactions. Graphs show hypothetical activation in the fusiform face area when the subject attends to a particular stimulus (a face or a place) or attends to a central fixation point. A reasonable hypothesis may be that attention will enhance activation for the preferred stimulus

category (faces) but not for other categories (e.g., places). In (A), the interaction appears consistent with the hypothesis. In (B) and (C), similar interactions occur but they are not consistent with the hypothesis. This example illustrates that the interaction term alone is not sufficient to understand the outcome of a factorial design.

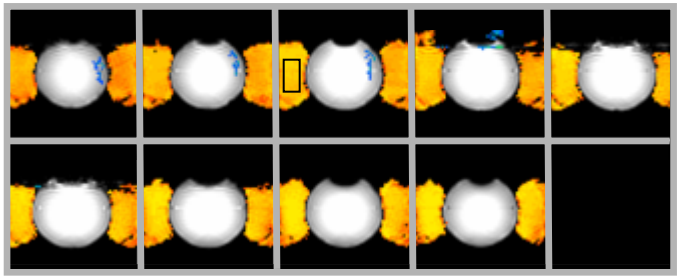
A. Pre-corrected Statistical Map 1:



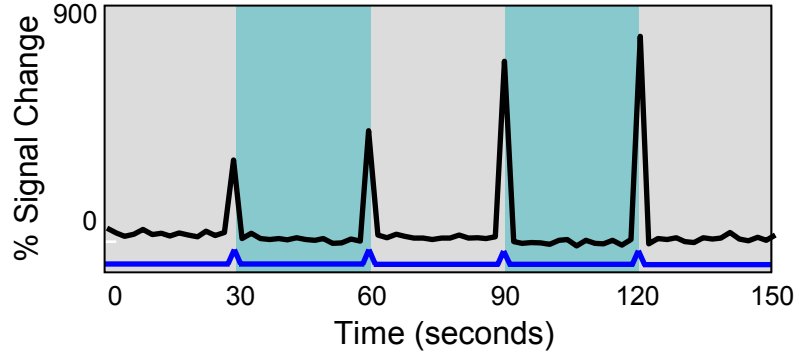
B. Time Course 1:



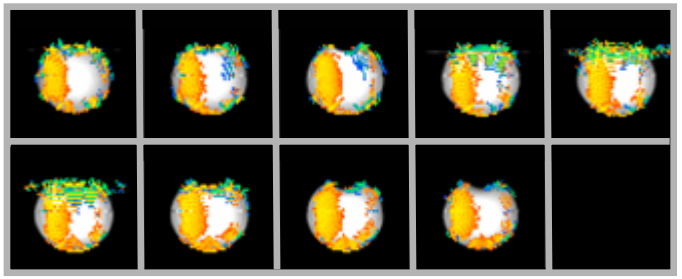
C. Pre-corrected Statistical Map 2:



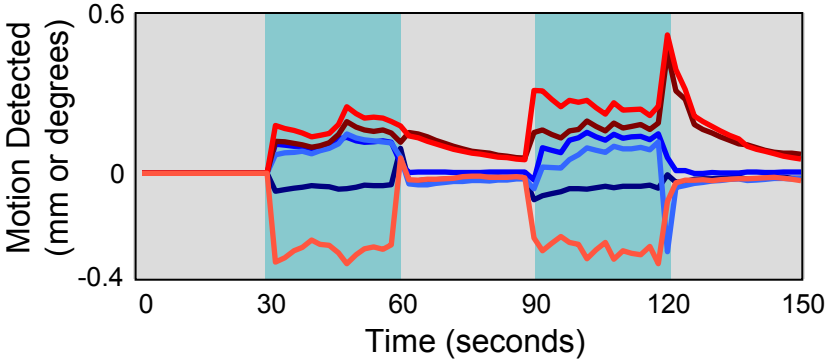
D. Time Course 2:



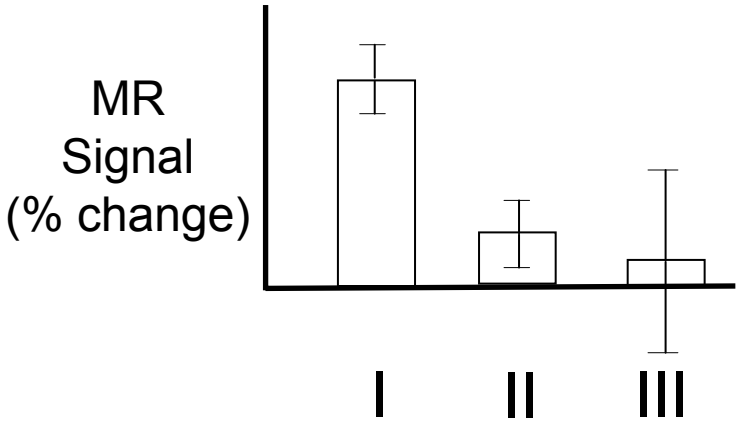
E. Post-corrected Statistical Map 1:



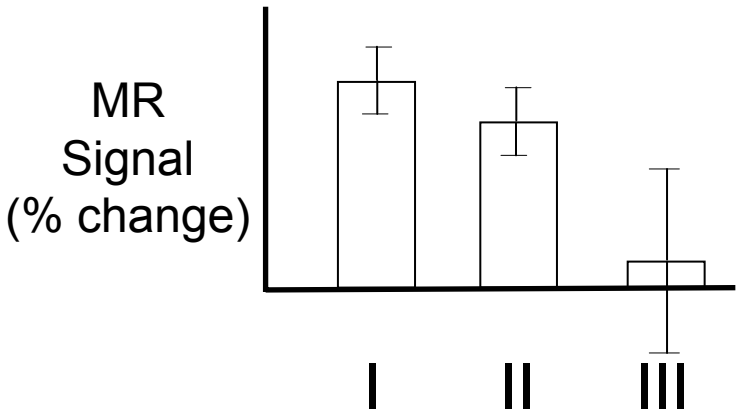
F. Motion Correction Parameters



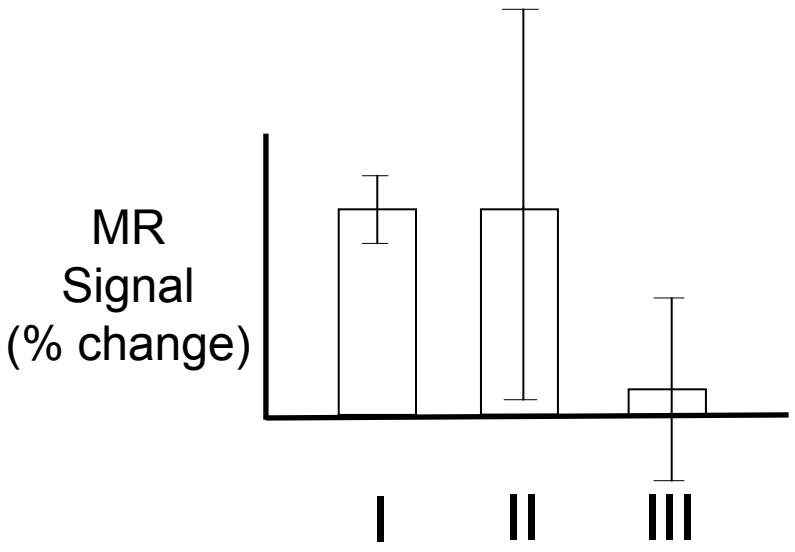
A.


MR
Signal
(% change)
I
II
III

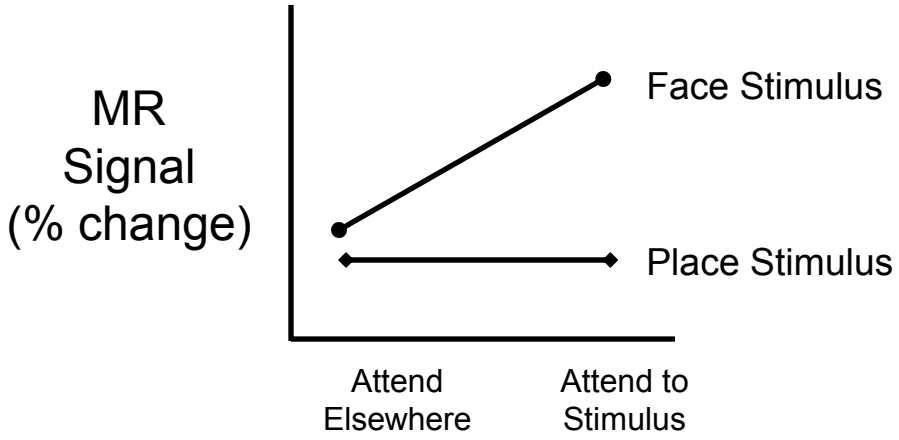
B.


MR
Signal
(% change)
I
II
III

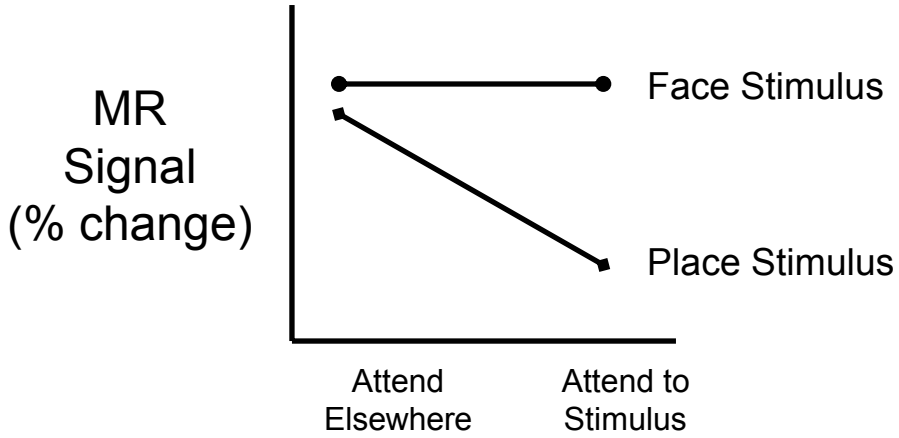
C
.


MR
Signal
(% change)
I
II
III

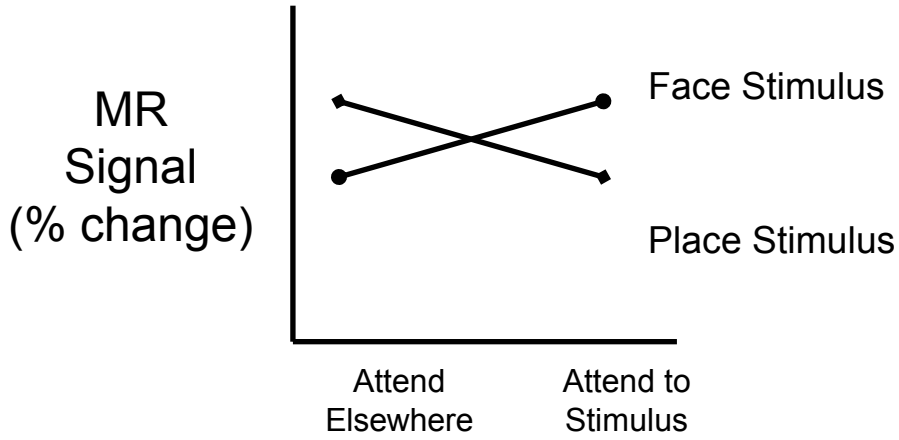
A.


MR
Signal
(% change)
Face Stimulus
Place Stimulus
Attend
Elsewhere
Attend to
Stimulus

B.


MR
Signal
(% change)
Face Stimulus
Place Stimulus
Attend
Elsewhere
Attend to
Stimulus

C.


MR
Signal
(% change)
Face Stimulus
Place Stimulus
Attend
Elsewhere
Attend to
Stimulus