*Development and Psychopathology,* **15** (2003), 811–832
Copyright © 2003 Cambridge University Press
Printed in the United States of America
DOI: 10.1017.S0954579403000385

# Conceptual, methodological, and statistical challenges in brain imaging studies of developmentally based psychopathologies

BRADLEY S. PETERSON
*Columbia College of Physicians & Surgeons and New York State Psychiatric Institute*

**Abstract**
Brain imaging studies in developmentally based psychopathologies most often use magnetic resonance imaging (MRI) to study regional volumes, task-related activity, neurometabolite concentrations, or the paths of fiber tracts within the brain. Methodological challenges for the use of MRI in studying these disorders include understanding the ultrastructural correlates of brain structure and function that are below the limits of resolution of this imaging modality and developing better methods for approximating the anatomical boundaries of the cytoarchitectonic units that are defined by those ultrastructural characteristics. Conceptual challenges include distinguishing findings that represent pathophysiologically central causes from compensatory and epiphenomenal effects, a difficulty that stems directly from the inherently correlational nature of imaging data. The promise of functional imaging studies must capitalize on the specificity of the cognitive and behavioral probes that are used to illuminate core features of the pathophysiology of developmental disorders, while recognizing the assumptions and limitations of the subtraction paradigms that are used to isolate the brain functions of interest. Statistical challenges include incorporating adequate statistical models for scaling effects within the brain, as well as modeling important demographic correlates that contribute to the substantial interindividual variability inherent in most imaging data. Statistical analyses need to consider the substantial intercorrelation of measures across the brain and the importance of correcting for multiple statistical comparisons, as well as the need for improved methods for brain warping and for assessing effective connectivity in functional imaging studies.

Brain imaging studies of developmentally based psychopathologies use magnetic resonance imaging (MRI) almost exclusively to study the brain in vivo, because other neuroimaging technologies, such as axial computerized tomography (CT scanning), positron emission tomography (PET), and single photon emission computed tomography (SPECT) all require exposure to radiation, which is considered ethically problematic in children, particularly healthy children. MRI provides exquisite images of brain anatomy and function. It can also provide measures of certain brain metabolite and neurotransmitter concentrations, and it most recently has been used to assess the integrity and orientation of nerve fiber tracts (Peterson, in press). Understanding the conceptual, methodological, and statistical challenges that these technologies pose for the study of developmental psychopathology first requires a familiarity with the basic elements of an image and with the brain processes or structures that these elements represent.

This work was supported in part by NIMH Grants MH01232, MH59139, and MH36197, grants from the Tourette Syndrome Association and the National Alliance for Research in Schizophrenia and Affective Disorders, and the Suzanne Crosby Murphy Endowment at Columbia University.

Address correspondence and reprint requests to: Bradley S. Peterson, MD, Columbia College of Physicians & Surgeons and New York State Psychiatric Institute, Unit 74, 1051 Riverside Drive, New York, NY 10032; E-mail: PetersoB@child.cpmc.columbia.edu.



https://doi.org/10.1017/S0954579403000385 Published online by Cambridge University Press

## Imaging Basics

### *Constructing an image*

The basic building block of a 2-dimensional (2-D) MR image is a tiny, 2-D square—a picture element, or pixel—that lies within the plane of the image. The pixel represents a corresponding 3-D cube of tissue—a volume element or voxel—in the brain. Many pixels assembled into a larger rectangular array constitute a 2-D image that represents a corresponding rectangular array within a 3-D "slice" of brain tissue. The complicated MRI scanning process turns the brain into a radio transmitter that encodes some physical feature of the tissue in each voxel as a numerical index within the corresponding pixel of the image. This physical feature may be, for example, the number of protons in the tissue (in an anatomical image), the relative concentration of deoxyhemoglobin (in a functional image), or the concentration of a chemical compound (in a metabolite image; Anderson & Gore, 1997; Buxton, 2002; Smith & Lange, 1998). The numerical index associated with a pixel, and therefore the physical feature in the brain voxel that it represents, is assigned a level of grayness ranging from black to white. The strength of the radio signal coming from any given voxel in the brain, and therefore the level of grayness of its corresponding pixel in the image, will differ from the neighboring pixels in the image if the average physical or chemical features of neighboring voxels in the brain differ. The differing level of grayness across the image is what we view as a picture of the physical property of the brain that we are imaging (Figure 1).

### *Measurement in differing imaging modalities*

Anatomical MRI in research studies is used primarily to measure the volumes and shapes of brain regions. These regions are delineated using manual or automated procedures which, despite their vastly varying degrees of complexity and sophistication, all generally classify pixels in the image as either belonging or not to the region of interest (Caviness, Meyer, Makris, & Kennedy, 1996; Clarke, Velthuizen, Camacho, Heine, Vaidyanathan, Hall, Thatcher, & Silbiger, 1995; Le Goualher, Procyk, Collins, Venugopal, Barillot, & Evans, 1999; MacDonald, Kabani, Avis, & Evans, 2000; Pham, Xu, & Prince, 2000; Toga & Thompson, 2002; Viergever, Maintz, Niessen, Noordmans, Pluim, Stokking, & Vincken, 2001). This classification is based on the gray scale values of the pixels and their spatial contiguity with other pixels having similar or differing gray scale values.

Functional MRI (fMRI) provides measures of change in the gray scale value of each pixel in an image across time. These changes in gray scale value reflect primarily time-dependent changes in the relative concentration of dexyhemoglobin in the voxel, which in turn are driven largely by changes in oxygen utilization. Changes in oxygen consumption are themselves caused by changes in neuronal activity within the voxel (Ogawa, Tank, Menon, Ellermann, Kim, Merkle, & Ugurbil, 1992). Thus, although changes in the fMRI radio signal provide an index of change in neuronal activity that is detected visually as changes in gray scale value within a pixel of the image, fMRI is not a direct, absolute, or specific measure of neuronal activity. Yoking these changes in gray scale to the performance of some kind of cognitive or behavioral task during acquisition of the images provides instead a good index of the *relative* change in neuronal activity accompanying the task. Maps of brain activity using fMRI are constructed by calculating the change in the level of grayness in each pixel when the brain is in two different states. These states are usually determined experimentally by instructing a subject to perform two different cognitive or behavioral tasks, one alternating with the other in an A–B–A–B–A–B fashion. More recently, so-called "event-related fMRI" has allowed the implementation of more naturalistic study designs in which the change in neural activity associated with discrete events can be discerned by averaging multiple fMRI time series segments that are time locked to the occurrence of discrete, single experimental trial or events, in a manner that is directly analogous to the signal detection and processing of event-related potentials using electroencephalography (EEG;

https://doi.org/10.1017/S0954579403000385 Published online by Cambridge University Press

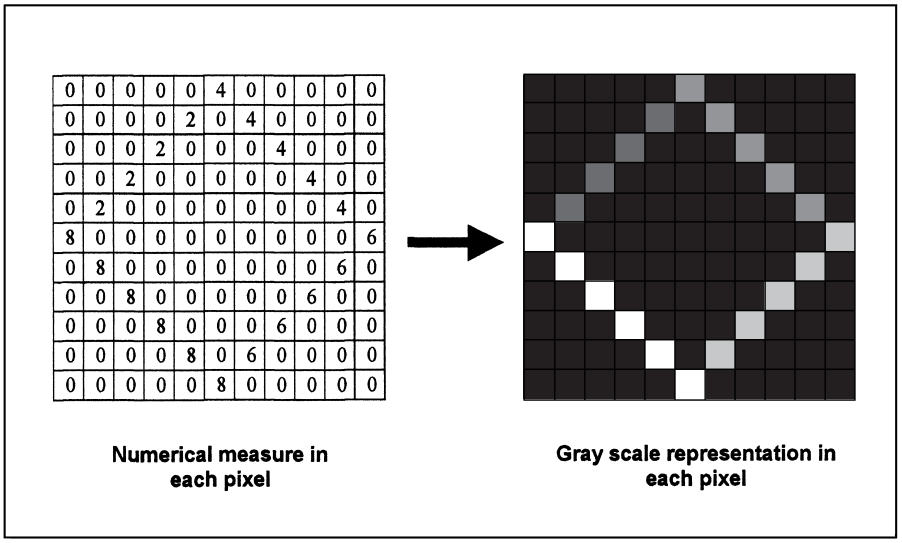



**Figure 1.** The construction of an image. Each 2-D picture element (pixel) has an associated numerical index that quantifies the physical property of the corresponding 3-D voxel in the brain (left panel). The numerical index in each of the pixels is assigned a gray scale value, ranging from white to black, to provide a visual representation of the spatial distribution of physical property in the brain that is being measured (right panel). This picture contains no information that is not already coded numerically in the pixel array on the left. It simply permits a more rapid visual assessment of the spatial distribution and quantification of the numerical measurements.

Leung, Skudlarski, Gatenby, Peterson, & Gore, 2000; Peterson, Kane, Alexander, Lacadie, Skudlarski, Leung, May, & Gore, 2002).

If the task-dependent variation in fMRI signal is to have a coherent physiological meaning, the two tasks for which the fMRI signals are being compared must be carefully and appropriately selected. The tasks must differ from one another by a single or, more realistically, a few physiological, affective, or cognitive-processing components, so that when brain activity is compared across tasks, the differential processing component across tasks is isolated and identified. One task therefore serves as the control or "baseline" task for the other. Brain activity associated with the control task is said in colloquial terms to "subtract out" from brain activity associated with the other, more primary, task of interest, thereby leaving only the brain activity associated with the cognitive processing components that are *unique* to the primary task. This general strategy of comparing brain activity associated with two or more tasks is typically termed a "subtraction paradigm."

Magnetic resonance spectroscopy (MRS) is used to quantify the concentrations of molecular compounds that are visible to the MRI scanner, which include molecules containing $^{1}H$, $^{13}C$, $^{23}Na$, $^{31}P$, $^{19}F$, or lithium nuclei (Novotny, Ashwal, & Shevell, 1998; Ross & Bluml, 2001). These compounds generate a characteristic "chemical shift" fingerprint in the spectrum of MRI radio signals emitted from the brain. The area under a specific peak of the spectrum estimates the concentrations of the corresponding compound within a voxel of brain tissue. Compounds that are typically measured include: *N*-acetyl aspartate (NAA), a putative index for the number and health of neurons (Arnold, de Stefano, Matthews, & Trapp, 2001) and a possible intercellular signaling compound between neurons and glia within the voxel (Lebon, Petersen, Cline, Shen, Mason, Dufour, Behar, Shulman, & Rothman, 2002); adenosine di- and triphosphate, molecules that are vitally important in brain metabolism; choline, a key metabolite of membrane phospholipids; lactate and pH, indices of anaerobic metabolism; and gluta-

https://doi.org/10.1017/S0954579403000385 Published online by Cambridge University Press

mate, glutamine, and gamma amino butyric acid, the brain's major excitatory and inhibitory neurotransmitters (Rothman, Sibson, Hyder, Shen, Behar, & Shulman, 1999; Sibson, Dhankhar, Mason, Behar, Rothman, & Shulman, 1997). Although MRS offers great promise for the study of the neurochemical underpinnings of developmental psychopathologies, it is still an evolving technology, the most prominent limitations of which are its poor spatial and temporal resolution when using the lower magnetic field strengths that are most often employed for clinical investigations in pediatric populations (Novotny et al., 1998; Ross & Bluml, 2001).

Diffusion tensor imaging (DTI) provides information about the orientation and integrity of the brain's fiber tracts, particularly in white matter (Pierpaoli, 2002). DTI therefore promises visualization of the anatomical processes of neurons that establish long-range neural connections in vivo, which can otherwise only be inferred from anatomical and fMRI data sets. DTI measures the degree of disruption that the diffusion of water molecules causes in the radio signal coming from within a brain voxel. If the diffusion of water molecules is constrained by membranes, organelles, proteins, and other cell structures in directionally oriented fibers within a voxel, then the degree and directional orientation of that constraint can be quantified and coded with gray scale values in the corresponding pixel of the MR image. Representation of these values across all pixels of the image provides a map of the magnitude and directionality of regional diffusivity across the brain (Conturo, Lori, Cull, Akbudak, Snyder, Shimony, McKinstry, Burton, & Raichle, 1999).

### *Prior technologies*

MRI affords considerable advantages over prior neuroimaging technologies. One of these is EEG, which has proved to be a time-tested, cost-efficient clinical tool for the diagnosis of, among other things, seizure disorders and encephalopathies. The attraction of EEG has always been its ability to measure the electrical activity of neurons directly, rather than using some more distant proxy for that activity, such as regional metabolism (with PET), blood flow (PET or SPECT), or deoxyhemoglobin (fMRI). EEG measures that activity, moreover, with exquisite temporal resolution, on the time scale of firing of individual neurons. Nevertheless, for a number of physical reasons, the spatial resolution of EEG is inherently poor, meaning that electrical sources close together in the brain are blurred together into a single signal when measured at the scalp. The use of many electrodes across the scalp and the development of new deblurring techniques have improved the resolution of EEG to several centimeters, but this is still considerably poorer than the resolution provided by MRI (Gevins, Le, Leong, McEvoy, & Smith, 1999). Localization of single point sources of electrical activity within the brain using EEG is also an inherently difficult and perhaps insoluble mathematical problem (Dale & Halgren, 2001; Gevins, Le, Brickett, Reutter, & Desmond, 1991), one that places obvious constraints on the confidence of pathophysiological interpretations of EEG data. Moreover, localization of the EEG signal currently assumes the existence of only a single source of brain activity associated with a cognitive task at any one point in time; this assumption seems implausible at face value, given that information processing in the brain, even for the simplest of tasks, is likely to be parallel and distributed rather than serial (Halpern, 2000; Mesulam, 1998; O'Brien & Opie, 1999; Zaborszky, 2002). Finally, EEG is able to measure electrical activity accurately only from the cortical surface, not from subcortical structures. This poses serious limitations on the study of developmentally based psychopathologies, many of which involve disturbances in cortical–subcortical circuitry.

First developed in the mid-1970s, CT scanning was the first technology to provide images of brain anatomy in vivo (Anderson & Gore, 1997). The images of brain anatomy that CT provides, however, are generally inferior to those of MRI, primarily in the attribute of the image that is referred to as *contrast,* the ability to discriminate among different types of tissue, such as gray matter, white matter, and

https://doi.org/10.1017/S0954579403000385 Published online by Cambridge University Press

cerebrospinal fluid (CSF). Also, whereas MRI can acquire images in any orientation, CT scans can be acquired only in the axial plane, limiting to some extent the research applicability of this imaging modality. Most importantly for the study of developmental psychopathologies, however, is that CT scans are acquired with X-radiation, and for research purposes, any radiation exposure, no matter how limited, is considered by many to be unethical. Thus, CT has had very limited application in the study of childhood disorders.

## Methodological Considerations

### *Macroscopic measurement*

Although the ability to measure brain structure, function, metabolites, and fiber tracts in the living brains of children is tremendously exciting, the finite, macroscopic size of the voxels that currently can be probed with MRI limits the biological conclusions that can be drawn about the ultrastructural features of brain tissue contained within the voxel. The sizes of voxels that can routinely be visualized is approximately 1 $mm^3$ in anatomical studies, 50–100 $mm^3$ in fMRI studies, and 1000–2000 $mm^3$ in MRS and DTI studies. At these sizes, the measures obtained with MRI from a voxel can tell us little about the ultrastructural features of tissue within the voxel that are of greatest interest in the study of developmental illnesses, such as the number, anatomical integrity, or connectivity of its cellular elements. Instead, all of the ultrastructural features of the voxel are blurred together into a single numerical index representing the property of interest for which the voxel is being probed. NAA levels, for example, are averaged together from all of the cellular sources of NAA within the voxel. FMRI signals, similarly, are averaged together from the activity of all of the synapses in the voxel, and the projections of neurons into and out of the voxel that contribute to those synapses may derive from very different regions within the brain. Similarly, in DTI studies, fiber trajectories that cross within a voxel will be averaged to yield a measure that disrupts the tracking of fibers across voxels.

### *Localization*

It is important to recognize the difficulty of identifying the location in the brain where measurement is actually being made. This problem with localization is particularly vexing for cortical regions, where gray scale provides little useful information for determining anatomically valid boundaries. Often, gray scale information is used to define cortical regions of interest (ROIs) by identifying and defining cortical sulci (Caviness et al., 1996) under the assumption that those sulci delimit the borders of underlying cytoarchitectonic units as defined by Brodmann and others (Braak, 1980; Brodmann, 1909; Flechsig, 1920). This assumption, however, is true only to a first approximation, as the spatial extent of these cytoarchitectonic units often do not follow sulcal boundaries (Dabringhaus, Schormann, Steinmetz, Schlaug, Zilles, & Roland, 1995; Zilles, Schleicher, Langemann, Amunts, Morosan, Palomero–Gallagher, Schormann, Mohlberg, Bürgel, Steinmetz, Schlaug, & Roland, 1997). In addition, the considerable variability in anatomical patterning of cortical sulci complicates the use of sulci for region definition (Ono, Kubik, & Abernathey, 1990), and sulci provide only a portion of the visible boundary for gyri lying at the cortical surface. Within the longitudinal extent of a gyrus, for example, the boundaries of regions lying within the gyrus must be defined, and the gray scale information provided by MRI is almost never helpful in visualizing them.

In practice, the longitudinal extents of these regions are most commonly defined using planes positioned through anatomical landmarks (common examples include the anterior and posterior commissures, the boundaries of the corpus callosum, and the positions of intersecting or neighboring sulci) that are assumed to be spatially invariant relative to the gyrus being transected. The use of anatomical landmarks in other brain regions to define an ROI, a procedure termed "planar stereotaxy," is a crude approximation of the underlying cytoarchitec-

https://doi.org/10.1017/S0954579403000385 Published online by Cambridge University Press

tural borders of the region, and the assumption that the spatial relationship of the landmarks to other structures within the brain is invariant is largely untested (Andreasen, Flashman, Flaum, Arndt, Swayze, O'Leary, Ehrhardt, & Yuh, 1994; Kates, Warsofsky, Patwardhan, Abrams, Liu, Naidu, Kaufmann, & Reiss, 1999; Peterson, Staib, Scahill, Zhang, Anderson, Leckman, Cohen, Gore, Albert, & Webster, 2001). In addition, these attempts at cortical subdivision all assume that cytoarchitecture provides the gold standard for cortical subdivision. Yet the standard itself is not ideal, in that the divisions of Brodmann's areas, far from being clear and distinct, tend to transition gradually from one area to another and do not have features that clearly and definitively distinguish one area from another (Barbas & Pandya, 1989; Fuster, 1997; Sanides, 1964). These problems with identifying valid anatomical subregions all conspire to limit the conclusions and inferences that can be drawn in imaging studies regarding the localization of anatomical and functional findings.

## Conceptual and Design Challenges

### *Distinguishing cause from effect*

Perhaps the major current challenge common to all brain imaging studies is distinguishing which findings represent core pathophysiological processes, which findings are simply epiphenomena, and which represent compensatory or adaptive changes of the nervous system. Adaptation to the presence of debilitating symptoms, with all of their physical, emotional, cognitive, and social ramifications, will likely alter broadly distributed adaptive systems throughout the brain, and these adaptive changes are difficult to distinguish from the events that set compensatory processes in motion to begin with. The source of the problem of distinguishing cause and effect is the inherently descriptive and correlational nature of in vivo imaging data and, perhaps surprisingly, this is true whether the data are acquired in cross-sectional or longitudinal studies.

*Cross-sectional correlates.* Cross-sectional imaging studies typically report the comparison of one particular brain imaging measure across two or more diagnostic groups, and indeed this is the most common kind of study reported in the imaging literature. These group differences are usually interpreted as representing an abnormality in the patient group that has contributed causally to the disease process. These diagnostic correlates, however, are almost always detected in individuals who are already affected with a particular illness. Cross-sectional findings from already-affected individuals inherently precludes us from discerning whether a given group difference is the cause or a consequence of the disorder with which it is associated. This difficulty with interpreting group differences is amply illustrated with findings from our own cross-sectional studies of regulatory systems in patients with Tourette syndrome (TS) and obsessive–compulsive disorder (OCD).

Subjects who voluntarily suppress motor tics activate broadly distributed cortical–subcortical circuits that include inferior prefrontal cortex, anterior cingulate cortex, and caudate nuclei (Peterson, Skudlarski, Anderson, Zhang, Gatenby, Lacadie, Leckman, & Gore, 1998). Similar findings of hypermetabolism and elevated blood flow to these same regions in resting subjects who have OCD have long been touted as strong evidence that these regions are the neural substrate of OCD symptoms; that is, most investigators claim that hypermetabolism in these regions produces obsessions and compulsions (Baxter, Schwartz, Mazziotta, Phelps, Pahl, Buze, & Fairbanks, 1988; Baxter, Phelps, Mazziotta, Guze, & Schwartz, 1987; Breiter, Rauch, Kwong, Baker, Weisskoff, Kennedy, Kendrick, Davis, Jiang, Cohen, Stern, Belliveau, Baer, O'Sullivan, Savage, Jenike, & Rosen, 1996; Rauch, Jenike, Alpert, Baer, Breiter, Savage, & Fischman, 1994; Swedo, Schapiro, Grady, Cheslow, Leonard, Kumar, Friedland, Rapoport, & Rapoport, 1989). This interpretation has been supported further by reports that successful antiobsessional therapies, either medication or cognitive behavioral treatments (Baxter, Schwartz, Bergman, Szuba, Guze, Mazziotta, Alazraki, Selin, Ferng, Munford, & Phelps, 1992; Benkelfat, Nordahl, Semple, King, Murphy, & Cohen, 1990; Rubin, Ananth, Villanueva–Meyer, Trajmar,

https://doi.org/10.1017/S0954579403000385 Published online by Cambridge University Press

& Mena, 1995; Schwartz, Stoessel, Baxter, Martin, & Phelps, 1996; Swedo, Pietrini, Leonard, Schapiro, Rettew, Goldberger, Rapoport, Rapoport, & Grady, 1992), normalized these regional metabolic "abnormalities."

Findings from our study of voluntary tic suppression, however, have called this interpretation of the OCD imaging data into question. The similarity in the pattern of elevated neural activity in subjects with TS who were suppressing their tics and in subjects with OCD who were presumably "at rest" suggested instead that the regional hypermetabolism in subjects with OCD was a consequence of the general requirement that those with OCD must suppress their compulsions either during the scan or during the period of radioligand uptake prior to the scan. This suppression of compulsory behaviors would thereby engage regulatory circuits similar to those that subjects with TS engage while suppressing their tics. It could therefore have been predicted that regional hypermetabolism in patients with OCD would normalize after successful anti-obsessional therapy, when the compulsions and the need to suppress them were reduced (Insel, 1992; Peterson & Thomas, 2000).

Our subsequent anatomical study of 155 patients with TS and 131 normal control subjects reported the presence of larger prefrontal cortices in children with TS and an inverse association of these regional volumes with the severity of tic symptoms (Peterson et al., 2001). Prefrontal volumes in adults, however, were smaller than control values, which we interpreted as likely resulting from the ascertainment bias of studying adults who were still symptomatic, rather than reflecting findings in a more representative sample of adults who have a lifetime history of tics and in whom symptoms usually substantially improve (Leckman, Zhang, Vitale, Lahnin, Lynch, Bondi, Kim, & Peterson, 1998). These findings, combined with our earlier findings that prefrontal regions activated significantly during the suppression of tics, suggested that compensatory responses involving activity-dependent plasticity and hypertrophy of prefrontal regions may be important in controlling tic symptoms. The smaller prefrontal volumes in symptomatic adults with TS suggested that compensatory control is probably impaired, thus accounting for their persistence of symptoms (Peterson et al., 2001). The possibility that the chronic suppression of tics induces a plastic response in prefrontal cortices that we can then visualize macroscopically with anatomical MRI is plausible in light of abundant clinical and preclinical studies demonstrating that the central nervous system (CNS) remodels itself in response to repeated activity and use. These plastic responses affect brain structure at all levels, from the remodeling of synaptic architecture to the organization of cortical maps (Buonomano & Merzenich, 1998; Cicchetti & Tucker, 1994a, 1994b; Elbert, Pantev, Wienbruch, Rockstroh, & Taub, 1995; Hann, Huffman, Lederhendler, & Meinecke, 1998).

Regulatory or compensatory circuits, similar to those engaged in the suppression of tics and compulsory behaviors, are active in other tasks in which either the explicit or implicit suppression of an automatic response is required. Activation of anterior cingulate, inferior frontal, and caudate nuclei have been reported, for example, in impulse control paradigms such as Stroop interference and go–no-go tasks (Bush, Whalen, Rosen, Jenike, McInerney, & Rauch, 1998; Carter, Mintun, & Cohen, 1995; Leung et al., 2000; Pardo, Pardo, Janer, & Raichle, 1990; Peterson, Kane, et al., 2002; Peterson, Skudlarski, Zhang, Gatenby, Anderson, & Gore, 1999), in studies of pain in which subjects must inhibit withdrawal of the painful extremity (Craig, Reiman, Evans, & Bushnell, 1996; Derbyshire, Jones, Gyulai, Clark, Townsend, & Firestone, 1997; Derbyshire, Vogt, & Jones, 1998; Talbot, Marrett, Evans, Meyer, Bushnell, & Duncan, 1991), in studies of itch induction in which subjects are not allowed to scratch (Hsieh, Hagermak, Stahle–Backdahl, Ericson, Eriksson, Stone–Elander, & Ingvar, 1994), in the regulation of sexual arousal (Beauregard, Lévesque, & Bourgouin, 2001), in studies of unsatisfied drug craving (Childress, Mozley, McElgin, Fitzgerald, Reivich, & O'Brien, 1999; Maas, Lukas, Kaufman, Weiss, Daniels, Rogers, Kukes, & Renshaw, 1998; Volkow, Wang, Fowler, Hitzemann, Angrist, Gatley, Logan, Ding, & Pappas, 1999), and in studies of anticipatory anxiety to electric shock (Chua, Krams, Toni, Passingham, & Dolan, 1999).

https://doi.org/10.1017/S0954579403000385 Published online by Cambridge University Press

Clearly, reduction in the noxious or motivating stimulus would also reduce the behavioral response that is intended to respond, remove, alleviate, or otherwise adapt to the stimulus. Therefore, normalization of these activations with successful treatment or intervention would not prove that the activations represented either the neural substrate of perceiving the stimulus or the behavioral response to them.

*Developmental and therapeutic correlates.* Cross-sectional studies often confuse and mislead us in our understanding of the true natural history and developmental correlates of an illness (Kraemer, Yesavage, Taylor, & Kupfer, 2000). A common assumption made in these cross-sectional studies is that members of differing age cohorts who have the same phenomenological diagnosis belong to a larger general population of subjects who have the same biological illness, and that therefore younger subjects will, with time, resemble their older counterparts. In most cases, this assumption will prove untrue, given that most childhood-onset illnesses differ from their adult-onset counterparts in phenomenology, degree of familial risk, associated comorbidities, and natural history.

The diametrically opposite morphological findings in the cortices of children and adults who have TS is an obvious example in which ascertainment bias most likely contributed to the detection of age correlates that may have very little resemblance to the developmental trajectories observed if these volumes were followed longitudinally in a cohort from childhood into adulthood. Rather than having provided genuine insight into the natural history of cortical development in subjects with TS, the cross-sectional age correlations more likely provided preliminary insight into the cortical regions and processes that contributed to the persistence of tic symptoms into adulthood, specifically in this instance, smaller prefrontal cortical volumes. These reduced prefrontal volumes presumably release tic symptoms from the inhibitory influences that reduce symptom severity for most individuals with TS in late adolescence and early adulthood (Bruun & Budman, 1992; Leckman et al., 1998). Thus, although cross-sectional studies of TS adults are important for the generation of pathophysiological hypotheses and the study of more severe disease outcome, they are nevertheless fraught with methodologic limitations that severely limit our ability to interpret and generalize the findings to the larger population of TS subjects. Similar difficulties are likely to plague cross-sectional studies of almost any process, capacity, or illness that has age-specific differences in prevalence.

Correlations of the magnitude of an abnormality in an imaging study with the severity of symptoms do not help to sort out cause from compensatory effects, because the need for compensation within the CNS is likely to be proportional to the severity of symptoms. Therefore, correlations of the magnitude of the abnormality with the severity of symptoms will be similar if the abnormality represents cause or if it represents compensation. In attempts to demonstrate causality in imaging findings, some investigators have elected to scan subjects before and after successful pharmaco- or psychotherapies (Baxter et al., 1992; Bench, Frackowiak, & Dolan, 1995; Benkelfat et al., 1990; Buchsbaum, Wu, Siegel, Hackett, Trenary, Abel, & Reynolds, 1997; Mayberg, Liotti, Brannan, McGinnis, Mahurin, Jerabek, Silva, Tekell, Martin, Lancaster, & Fox, 1999; Rubin et al., 1995; Schwartz et al., 1996; Swedo et al., 1992). Although this approach might initially seem appealing, it does not address the fundamental difficulty of dissociating cause from compensatory effect, because compensatory systems are no longer needed when the primary symptoms have been removed by the successful therapy.

Even longitudinal brain imaging studies of affected individuals do not fully address this seemingly intractable issue of discerning cause and effect. The longitudinal approach can allow us to track changes in brain structure, function, and neurochemistry through time. These measures of change in brain structure and function can then be correlated with measures of clinical change across time, and these measures of change can be helpful in allowing us to generate hypotheses concerning which of the baseline abnormalities are causing the illness and which are epiphenomena or compensatory responses. Unfortunately, however, im-

https://doi.org/10.1017/S0954579403000385 Published online by Cambridge University Press

aging changes that become evident in close temporal proximity to symptom onset or to times of exacerbation or remission of symptoms may still represent either the core process responsible for these events or a compensatory reaction to them. Assessing the correlation of changes in a brain measure with the changes in symptoms cannot prove whether the brain changes caused the changes in symptoms or whether the changes in symptoms caused the changes in the brain properties being measured (Gerard & Peterson, in press; Peterson & Thomas, 2000).

*Possible solutions.* One possible solution to this difficult problem of disentangling cause and effect in imaging studies is to extend our imaging techniques to progressively younger age groups or to high-risk cohorts in order to identify trait rather than state markers of CNS functioning that predispose individuals to particular illnesses. Studies of at-risk populations, such as the offspring or young siblings of affected individuals, would be uniquely able to identify trait vulnerability markers for an illness. These vulnerability markers would be relatively independent of compensatory changes because they would be observed in the absence of overt symptoms. Ideally, these high-risk samples would be followed through time to identify which pathophysiological processes are associated with the switch from trait vulnerabilities to overt illness. These prospective studies would then allow the study of state disturbances early in their course, before chronic compensatory mechanisms would exert their neurobiological effects, as well as during illness onset and the ensuing compensatory responses.

A second possible solution is to yoke imaging studies to clinical trials in which a putative, causally relevant variable is experimentally controlled and manipulated. This could be possible, for example, in appropriately designed imaging studies associated with prevention or intervention trials. As noted above, simply correlating imaging measures with measures of change in therapeutic trials is insufficient to sort out cause from compensation. Nevertheless, these effects can in principle be distinguished by using appropriate contrast groups, such as individuals who as placebo control subjects experience spontaneous remission of their illness. Although the active treatment group will contain individuals who remit either spontaneously or with placebo treatment, with a sufficiently large sample, the brain-based effects that are specific to treatment should be statistically discernable over the nonspecific effects of symptom reduction in the spontaneous remitters. This kind of design seems not to have been implemented yet in an imaging study.

A third possible solution is to image populations that are epidemiologically ascertained, rather than simply to study samples of convenience. This approach would, along with longitudinal studies, help to identify the true longitudinal trajectory of illness-related abnormalities in the brain, although it still cannot in itself distinguish abnormalities that are causal from those that are compensatory. In imaging studies of TS, this kind of study design would help us to know more definitively whether prefrontal volume reductions in adults are generally true of all adults who have a lifetime history of TS or whether they are present primarily, as has been postulated, in adults with TS who remain symptomatic. Imaging studies in epidemiological samples are unlikely to meet with resounding success, however, because recruitment into imaging studies, even for individuals who have a lifetime history of an illness, tends to be poor, thereby distorting the representativeness of the sample.

### *Conceptual and design challenges in fMRI subtraction paradigms*

Implicit in the subtraction methodology of functional imaging studies is the assumption that information processing in the brain is serial and additive. Subtraction paradigms have recently been popularized through the proliferation of functional imaging studies, and yet they have a long experimental tradition in cognitive neuroscience, particularly in studies of reaction times recorded during performance of a hierarchy of cognitive processing tasks. Subtraction paradigms make several strong assumptions: (a) the brain is a serial processor; (b) cognitive functions are linearly additive and can

https://doi.org/10.1017/S0954579403000385 Published online by Cambridge University Press

therefore be "subtracted," one from another; (c) cognitive processing is highly modular; and (d) serially ordered modules for information processing are spatially localized in the brain. Although these assumptions have been repeatedly and roundly criticized throughout the history of cognitive neuroscience (Dolan & Friston, 1997; Donders, 1969; Nemeroff, Kilts, & Berns, 1999; Poeppel, 1996; Sternberg, 1969; Wexler, 1986), the experimental results of these subtraction paradigms have proved to be reasonably robust and reproducible in imaging studies, at least in well-designed experiments (Leung et al., 2000; Peterson, Kane, et al., 2002). Furthermore, these paradigms have arguably produced, in some instances at least, results that support some of these underlying assumptions.

*Experimental control.* Probably the greatest challenge in constructing fMRI subtraction paradigms is to select the appropriate control task for the primary task of interest. Many fMRI tasks are poorly controlled, so that the differential brain activity being detected could be caused by any number of differences in task conditions, including differences in the physical properties of the stimulus presentation (e.g., the hue, saturation, brightness of color, shape, contour, or texture of visual stimuli); differences in comparative difficulties of the two tasks (producing systematic differences in alerting, orienting, vigilance, performance effort, uncertainty, confusion, search strategy, or timing of a motor response); differences in the cognitive strategy used to perform the tasks; differences in the thoughts and affects associated with each of the tasks; or differences in any number of other brain and mental functions associated with the tasks.

*Specificity of the probe.* The dependence on subtraction methodologies is both a strength and a weakness of fMRI studies of developmentally based psychopathologies. The strength of subtraction paradigms is that they permit probing of very specific mental processes in healthy and ill individuals. The weakness of subtraction paradigms is that they confine investigation to a very narrow domain of brain functioning. Thus, fMRI currently is not able to measure and compare brain activity "at rest" across individuals in the way that PET and SPECT can. The utility and power of fMRI in the study of developmental psychopathologies will therefore depend critically on the pathophysiological relevance of the subtraction paradigm and the mental process that it probes, as well as on the quality of the design and implementation of that subtraction paradigm in terms of how well it isolates the mental process of interest. Currently, our knowledge of the pathophysiology of any illness is too limited to provide a clear and adequate guide for the selection of a mental or behavioral probe that will illuminate central pathophysiological features of the illness, and not simply correlated or epiphenomenal ones.

*Performance differences in fMRI studies.* Another conceptual and methodological dilemma facing the implementation of fMRI subtraction paradigms involves the interpretation of results when subtraction paradigms employ tasks in which patient and control groups are known to perform differently. Most fMRI tasks, in fact, are selected precisely because previous behavioral studies have demonstrated that the patient group performs poorly on the task. Investigators then usually assume that an fMRI study that employs that task will demonstrate differences in brain activation across groups that will explain the origins of the performance differences.

Conducting and interpreting studies when performance in the task of interest differs across groups, however, is problematic, and the problem can be stated in lay terms: if we already know that two groups differ in the performance of some task, then we already know that the basis for that difference must be brain based. A fancy procedure and pretty pictures are not needed, critics will add, to know that brain activity differs between the two groups. The differences in brain activity may not represent a core abnormality in one group, but instead may depict an alternative pathway for information processing that may be adaptive, or they may deplete greater effort or frustration, or some other epiphenomenal feature associated with poorer performance on the task. Therefore, knowing where in the brain activity dif-

https://doi.org/10.1017/S0954579403000385 Published online by Cambridge University Press

fers across groups may not alone help us in understanding what those differences in brain activity represent, because the origin of group differences in brain activation when performance on the task differs across groups has numerous possible explanations. Of particular relevance for studies of developmental disorders is that performance often varies systematically with age even in healthy populations, and this can seriously complicate the design and interpretation of studies involving children.

When the differences in performance across groups are measured as differences in reaction time following presentation of a sensory stimulus, the differences in activation maps across groups can represent a kind of artifact of image processing, particularly in event-related fMRI studies. If the subtraction of images is time locked to the presentation of the sensory stimulus, images in the fMRI time series that follow presentation of the stimulus will obviously differ across groups simply because of the group differences in timing of the motor response. Differences in activation across groups may have little or no relationship to the neural processes that actually cause the group differences in reaction times. Similarly, if the subtraction of images is time-locked to performance of the motor response, images in the fMRI time series associated with processing of the sensory stimulus may differ across groups simply because they are no longer aligned in time. Neither method of time locking is necessarily correct, because the differences in the timing of information processing that underlies the differences in performance across groups will necessarily find their way into the fMRI subtraction paradigm and the images representing them.

*Possible solutions.* I do not pretend to have an easy answer for the difficulties posed by the specificity of fMRI behavioral probes and by the differences in performance measures that affect many studies of developmental psychopathologies. Nevertheless, one experimental approach that has not yet been implemented in any obvious way in existing functional imaging studies might help to mitigate some of these more intractable problems. Rather than trying to solve or sidestep the difficulties posed by the specificity of fMRI probes and by the group differences in performance measures, this approach not only accepts those limitations but also embraces and exploits them in novel study designs. These designs systematically probe neural systems according to an experimentally and theoretically well-founded understanding of the hierarchy of brain organization (Mesulam, 1998).

Instead of selecting a task because group differences in performance have been reliably demonstrated, I instead advocate selecting tasks based in part on the similarity of performance across groups. Group differences in brain activation on those tasks will then likely represent information processing pathways that are engaged in the service of providing normal overt behavioral performance, presumably as an alternate strategy to perform the task or as a compensatory neural system that is engaged because the normal pathway is somehow dysfunctional. Demonstrating the presence of alternate processing pathways in the context of normal behavioral performance, however, will not in itself, in any single subtraction paradigm, indicate which components of information processing are dysfunctional. Multiple fMRI studies using multiple subtraction paradigms must be conducted on the same individuals to probe information processing pathways from their earliest and most elementary components to the latest and most superordinate ones in order to determine where in the processing pathways dysfunction and group differences first appear.

In studies of developmental psychopathologies, therefore, initial studies of simple sensory processing, conducted separately across all sensory modalities, could be performed to demonstrate that these elementary sensory processing functions are in fact intact in the illness of interest. These tasks, which are usually very simple and which permit similar measures of performance across even the most disparate of patient and healthy comparison groups, would be expected to activate primary sensory regions. Next, studies of slightly higher sensory processing could be conducted to demonstrate the intactness of these processes within a single sensory modality. These tasks, which involve the progressive integration of features

https://doi.org/10.1017/S0954579403000385 Published online by Cambridge University Press

within a given sensory modality, would be expected to activate higher sensory association areas. Next, studies could be conducted to assess integration across sensory systems in the groups of interest. These tasks would be expected to activate heteromodal association areas in the parietal, temporal, or prefrontal cortices. Finally, superordinate information processing systems that are thought to help organize and integrate, in a top-down fashion, activity in lower order primary sensory, primary association, and heteromodal association areas could be probed in tests of executive and attentional functioning. These tasks would be expected to activate prefrontal, cingulate, and higher order association areas (Peterson et al., 1999).

This kind of study design will allow the determination of normal and pathological brain functioning in developmental psychopathologies, and it will help to localize more accurately where in the information processing stream abnormal neural functioning begins, even when performance on the tasks is normal. Both normal and abnormal patterns of activation, when taken together, will prove to be informative of brain function and pathophysiology in the patient group. Ironically, the most informative of these findings is likely to be the normal activation in the patient group, for it in particular helps to narrow the search for where in the brain abnormal functioning first originates. Normal activity, moreover, is likely to be more readily interpretable than is abnormal activation in a single task, which as we have seen can have multiple causes and wildly divergent explanations.

Implementing this proposed design will require an extended, systematic program of research that builds progressively on prior studies and their findings. This research paradigm would stand in stark contrast to current research strategies, which are mostly ad hoc and designed to identify a putative central locus of abnormality using a single fMRI experiment. Ideally, this proposed program of research would be conducted in a single experimental cohort, to minimize variability in findings across each of the stages of information processing. Multiple scanning sessions in the same patient and control groups would increase dramatically the practical difficulties of conducting this program of research. Nevertheless, this kind of programmatic approach is necessary if we are ever to make systematic progress in understanding the etiologies of developmentally based psychopathologies.

## Statistical Considerations

### *General*

Findings from imaging studies have been notoriously difficult to replicate. In many instances, poor reproducibility probably relates to the statistical instability associated with small sample sizes, particularly in light of the high interindividual variability of imaging data. PET and SPECT studies, for example, still report findings from samples of fewer than 10 subjects, which is most commonly necessitated by the expense of these particular imaging modalities. Other contributions to the poor reproducibility of imaging studies has been the general failure to incorporate into the statistical models covariates, such as sex, age, socioeconomic status, and comorbid illnesses, that contribute substantially to interindividual variability in the imaging measures being studied. Adequate modeling of these potentially confounding effects is possible only with appropriately large sample sizes.

### *Anatomical data*

The methods for statistical analyses of imaging data are relatively unique among biological measures, and this includes the analysis of anatomical imaging data. One of the most important effects for which statistical models of volumetric data must account is morphological scaling. Scaling refers to the fact that taller individuals will have bigger brains and individuals with bigger brains will have bigger subregions within their brains (Gould, 1981). Men, for example, on average have bigger brains than women, but this is largely attributable to larger body sizes and larger heads in men than in women. The same is true for older compared to younger children. Scaling typically accounts for the largest portion of interindividual variance in regional

https://doi.org/10.1017/S0954579403000385 Published online by Cambridge University Press

volumes, and failure to account for scaling as a potential confound can lead to erroneous results and inappropriate interpretations of imaging data.

Statistical analyses must therefore account for these scaling effects when comparing brain volumes across men and women or across age groups. Precisely how best to do this is a matter of some controversy, but most studies control the analyses statistically either for total body size (using a measure such as height) or for head size (using a measure such as head circumference, intracranial volume, or whole brain volume; Arndt, Cohen, Alliger, Swayze, & Andreasen, 1991; Mathalon, Sullivan, Rawles, & Pfefferbaum, 1993). Statistical control is generally preferred over the construction of "normalized" values in which the subregion's volume is divided by the value of the scalar. Older imaging studies of the ventricular system, for example, often compared the "ventricular to brain ratio" across diagnostic groups. The problem with use of ratios is that they employ a multiplication operator (ventricular volume $\times$ 1/total brain volume), which will multiply the measurement errors of the numerator and denominator. Consequently, linear correction using statistical covariation with the chosen scalar is generally preferred when correcting a subregion's volume for the effects of scaling (Andreasen, Arndt, Cohen, Alliger, & Swayze, 1993).

Another important consideration in the analysis of anatomical data is the intercorrelation among regional brain volumes, even after controlling for scaling effects within the brain. Imaging studies frequently report analyses for volumes of the same structure in the left and right hemispheres, for instance, or for multiple different regions within the same hemisphere. Because these measures are intercorrelated, the tests are not statistically independent, and yet their analysis using two separate statistical tests still requires correction for multiple comparisons. The standard for analysis of multiple regions is currently a repeated-measures analysis of variance, which takes into account the intercorrelation of the various dependent measures. Rather than considering intercorrelations of a repeated measure across time (which is the most common use of a repeated measures analysis), however, anatomical studies consider the intercorrelations of a repeated measure across the spatial domain of the brain. One disadvantage of this kind of analysis is that it is complicated for many readers to understand, in that hypothesis testing is performed simultaneously across these numerous brain regions. If the study hypothesizes, for example, that a group difference will be detected in one region but not another (i.e., that the abnormality is regionally specific, rather than present to the same degree in all regions measured), the hypothesis is tested using a statistical interaction (e.g., region by diagnosis). Another disadvantage is that determining the source of effect within an interaction (i.e., whether the significance of an interaction derives primarily from one region among several, or from one diagnosis, or from a combination of each) is not easy to assess. This is particularly true for the kinds of complicated, higher order interactions that are likely to be detected in data sets for developmentally based disorders, which might include, for example, interactions between some combination of brain region, cerebral hemisphere, diagnosis, age, and sex.

### *Data from fMRI*

Task-dependent variation in fMRI signal is tested statistically by acquiring many images in rapid succession while subjects alternate performance between two tasks. A typical fMRI experiment acquires an image in each slice every 1–2 s, totaling several hundred to several thousand images in each slice for each subject in an experiment. Whether the fMRI signal or, equivalently, the gray scale in a pixel, varies with the alternation between tasks at a level that is beyond chance is then tested using any number of more or less sophisticated statistical procedures. The simplest of these is to sort the signal values of the fMRI time series into two populations, those acquired during each of the two tasks, and then to compare the means of the two populations of signal values using a parametric or nonparametric test, most commonly a *t* test.

A common goal in studies of developmentally based psychopathologies is to compare

https://doi.org/10.1017/S0954579403000385 Published online by Cambridge University Press

differential fMRI signal change associated with a task across diagnostic groups or across differing ages. This comparison, however, is a nontrivial matter. Group comparisons of fMRI signal changes typically are performed either in a voxel-based or a ROI-based analysis. These approaches differ in many respects, the most important being, as their names imply, the unit of analysis—whether it is a single voxel or all voxels together within an ROI.

*Voxel-based approaches.* The voxel-based approach compares signal changes across groups on a voxel by voxel basis. This is done after "warping" each individual brain into a similar volume, shape, and spatial orientation, and then superimposing these brains one on top of another. Corresponding pixels in the images from each person are then assumed to represent the same voxels of tissue across the brains of individual subjects. The location of a pixel in the image of the motor cortex in one person, for example, is assumed to represent precisely the same anatomical and functional tissue as the corresponding pixel in the image of the motor cortex in another person. The warpings are usually performed with reference to a set of standard anatomical landmarks, the anterior and posterior commissures and an imaginary bounding box around the brain (Talairach & Tournoux, 1988). Mean signal changes associated with alternation of the tasks are calculated at each voxel for each individual in the study. Every voxel therefore has associated with it a population of signal changes for each diagnostic group in the study. These populations of signal change are then compared with one another using any number of standard parametric or nonparametric statistics, one test for each pixel in the image.

*ROI-based approaches.* In ROI-based comparisons of signal changes across groups, anatomical regions are first defined on each individual's unwarped images. These may be defined, for example, by manually circumscribing the boundary of the structure on the anatomical images, and upon these ROIs are superimposed the mean signal changes of the functional images at each voxel. The mean signal changes for each voxel within each ROI are then either summed or averaged to provide a descriptive statistic representing the overall task-dependent change in signal within the ROI. That measure of ROI-based signal change for each individual, a single continuous variable, is then used to compare activation in that ROI across diagnostic groups.

*Limitations of voxel- and ROI-based approaches.* Both voxel- and ROI-based approaches to comparing activations across groups have their limitations. Probably the greatest limitation of voxel-based approaches is the assumption that, except for differences in scaling (i.e., in overall volume) and head orientation, the anatomy and corresponding localization of function is the same across individuals, across ages, and across diagnostic groups, an assumption that anatomical imaging studies of normal and patient groups have unequivocally demonstrated to be false. Because of known differences from normal in the anatomy of brains of patients who have neuropsychiatric illnesses, the brains of patients would not be expected to superimpose well on each other or on healthy control brains when warped into the common coordinate system (Bookstein, 2001). Consider the kinds of effects that this misregistration of tissues across diagnostic groups is likely to have. The pixels representing gray matter in the frontal cortex of one group, for example, might superimpose on frontal white matter of the other group. Because signal changes should occur in the synapses located in cortical gray matter and not in the axons or myelin of the underlying white matter, a voxel-based comparison of signal changes across groups would demonstrate more activation in the first group than in the second, but this difference would simply be an artifact of misregistration in the brains of the patient and control groups.

The ROI-based approach, in contrast, is not particularly vulnerable to this kind of artifact because it defines boundaries of the ROI based on the relevant anatomy and anatomical landmarks of each individual in the study. A weakness of the ROI approach, on the other hand, is that it assumes that anatomically and functionally relevant regions can be reliably, validly, and precisely defined. As we have already seen for regions such as the cerebral

https://doi.org/10.1017/S0954579403000385 Published online by Cambridge University Press

cortex, the assumption that cytoarchitectonically valid regions can be defined using the gray scale values of MR images is a questionable one. This difficulty is only compounded when the ROIs are defined on functional MR images, which have much poorer resolutions (typically 70–100 $mm^3$) than do most current anatomical images. In these poor resolution data sets, cytoarchitectonic regions and their boundaries are smoothed across these large volumes.

*Multiple comparisons.* Another challenge in the statistical analysis of fMRI data is the correction for multiple statistical comparisons across voxels or ROIs. Because a typical fMRI experiment is performed across 10s of 1000s of pixels (50,000 pixels/subject is not uncommon), strict Bonferroni correction for multiple comparisons in voxel-based analyses would require $p$ values <.000001. This sets the statistical bar for most real-world experiments at a level that is, for several reasons, probably too high. First, brain regions do not activate in isolation in fMRI experiments, but rather in functional networks. Therefore, to set a statistical threshold with the expectation that only a single voxel will activate past such a stringent level of significance will predispose a study to missing real activations (Type II errors). Second, statistical tests in neighboring voxels are not independent of one another, because fMRI time series in neighboring voxels are intercorrelated. This intercorrelation derives in part from correlated noise across the voxels, as well as from correlation of real biological signal coming from multiple voxels lying within the same biological field of activation, such as the same Brodmann's area or the same functional unit that spans multiple adjacent voxels (Roland, 1993).

A number of statistical alternatives to strict Bonferroni correction have been proposed for fMRI experiments. One is simply to set a less stringent correction for voxel-based analyses. The less stringent statistical threshold usually is not adequately justified, and the actual threshold employed ranges widely across experiments and laboratories. A second approach has been to apply several logical operators simultaneously to determine whether a voxel is "activated." A commonly applied rule is that a voxel must meet some statistical threshold (a particular $t$ or $p$ value, for instance) and that the voxel must also be connected to some predesignated number of voxels that also pass the statistical threshold. Voxels captured in this "cluster filter" are presumed to represent a field of activation in the brain, and they are thought to be much less likely to represent a false positive statistical outlier (Forman, Cohen, Fitzgerald, Eddy, Mintun, & Noll, 1995; Goutte, Toft, Rostrup, Nielsen, & Hansen, 1999; Ledberg, Akerman, & Roland, 1998; Poline, Worsley, Evans, & Friston, 1997). Applying this cluster filter has an effect similar to spatial smoothing of the test statistic across voxels, in that both reduce the number of isolated, spurious false positive activations (Skudlarski, Constable, & Gore, 1999). A third approach is to use an ROI-based approach to statistical analysis, which effectively sums or averages across the activations within all of the individual voxels that the ROI comprises. Having several larger ROIs reduces significantly the number of comparisons that are performed in the experiment. Often this number is reduced even further by invoking a repeated measures statistical analysis of activation, in which the repeated measure is fMRI signal change measured across the spatial domain of multiple, intercorrelated ROIs (Peterson, Vohr, Kane, Whalen, Schneider, Katz, Zhang, Makuch, Gore, & Ment, 2002). Finally, the most common approach to the multiple comparisons problem in fMRI data sets is to use the theory of Gaussian random fields to model the distributions of both the signal changes across the data set and the interdependence of activations across voxels, allowing calculation of $p$ values that are corrected for the number of independent statistical tests across the data set (Ashburner & Friston, 2000; Worsley, Evans, Marrett, & Neelin, 1992). This calculation is nearly always performed using one of several common software packages that are available as both shareware and commercial software for the analysis of functional imaging data sets.

*Functional connectivity.* Despite the immense popularity of fMRI studies, they have not yet fulfilled their promise of identifying the func-

https://doi.org/10.1017/S0954579403000385 Published online by Cambridge University Press

tional anatomical basis of developmental psychopathologies. This shortcoming can be attributed to the naive assumption that psychopathologies arise in part from abnormalities in brain function that are spatially discrete and highly localized. Indeed, one of the tacit assumptions of fMRI subtraction paradigms is that neural activity associated with one task, relative to activity associated with its control task, will localize to a particular brain region. Countless published imaging studies purport to have localized the brain region that is the center or locus for a particular cognitive function, emotional process, or behavior. These claims constitute little more than a modern, albeit high-tech, phrenology (Cohen, 1996; Friston, 2002; Nemeroff et al., 1999).

Given contemporary understandings of brain organization and function, a more reasonable assumption is that developmental psychopathologies, rather than being regionally localized, are based within neural systems that are distributed across multiple, interactive brain regions (Cicchetti & Cannon, 1999; Halpern, 2000; Mesulam, 1998; O'Brien & Opie, 1999; Zaborszky, 2002). One of the great challenges for functional imaging studies is to devise ways to study this distributed, functional interaction. Improving our understanding of the interaction of multiple brain regions may require using functional imaging modalities that have a better temporal resolution than fMRI, which is on the order of 1–2 s. Currently, better temporal resolution is only possible using EEG or magnetocencephalography. Presumably, brain regions that increase their activity earlier in time in some way contribute to the activity of (and are "functionally connected" to) other brain regions that activate later during performance of a task (Dolan & Friston, 1997). Other approaches, however, have tried to characterize the intercorrelation of activity across brain regions using advanced statistical procedures, such as principal components (Alexander & Moeller, 1994; Peterson et al., 1999), independent components (McKeown, Makeig, Brown, Jung, Kindermann, Bell, & Sejnowski, 1998), variance components (Moeller & Strother, 1991), cross-temporal correlation (Biswal, Yetkin, Haughton, & Hyde, 1995; Hampson, Peterson, Skudlarski, Gatenby, & Gore, 2002), and mutual information analyses (Friston, 2002; McKeown, Varadarajan, Huettel, & McCarthy, 2002), or structural equation modeling (Horwitz & Sporns, 1994). These analytic techniques are novel, however, and their reproducibility and validity in imaging data sets are largely untested. Nevertheless, analyses like these are needed if functional imaging studies are ever to fulfill their promise of illuminating the etiologies of illnesses at a systems level of understanding.

### *MRS and DTI data*

Thus far, no good statistical methods have been developed to compare the orientation of fibers across groups of subjects in DTI data sets, and this will be necessary for the effective study of developmental psychopathologies using DTI. The difficulty with group comparisons arises from the complication of analyzing the directional component of the primary measure in DTI data sets, the diffusion tensor. The nondirectional component of the diffusion tensor has been quantified as a simple scalar measure in a number of different ways, and it can be used to compare mean diffusivity across groups (Pierpaoli, 2002). These nondirectional measures, however, tend to be better suited to detecting gross disruption of the integrity of brain tissue, as occurs in stroke, rather than the more subtle disruptions in organization of fiber tracts. Developing techniques to compare the full diffusion tensor, including its directional component, is one of the important statistical challenges for DTI.

In MRS, measurement of even the most abundant compounds, such as NAA, requires a volume of more than 1 cubic centimeter per voxel on most clinically available scanners. Measurement of less abundant compounds requires even larger voxels. Voxels of this size often include in their volume multiple tissue types, which seriously complicates interpretation of metabolite levels. The sizes of peaks in the MRS spectra will depend, for example, on the relative proportions of CSF, gray matter, and white matter in each voxel. These proportions can vary systematically across the diagnostic groups whose spectra are being compared. For example, a relatively thicker or thinner cortical rib-

https://doi.org/10.1017/S0954579403000385 Published online by Cambridge University Press

bon within the voxel could produce group differences in metabolite levels and be interpreted erroneously as representing differences in the concentrations of metabolites within the cortical ribbon itself. The improving spatial resolution of MRS with the increasing availability of higher field scanners will help to minimize these effects of including multiple tissue types within a voxel. In addition, mathematical techniques have been developed to help estimate and correct statistically for the proportion of various tissue types within a voxel. Improving these procedures for statistical correction is an important challenge for future MRS imaging studies.

## Conclusions

Neuroimaging technologies, by providing a window into the structure and function of the living brain, have revolutionized medicine and clinical neuroscience. We now have the ability to apply these technologies safely to the study of brain development in healthy and ill children. The technological advances in the two decades or so in which MRI has been available have been astounding, and we cannot yet even conceive of the new windows into brain structure and function that continuing technological advances in the coming decades will bring. These technological advances will undoubtedly help to address, to some extent, some of the challenges outlined here in the study of developmental psychopathologies, just as technological advances have helped to address the technological and methodological challenges historically faced in the fields of physics, astronomy, chemistry, molecular biology, and genetics. Technological advances have also helped to define and drive the scientific agendas of those fields, and they will no doubt do the same in studies of developmentally based psychopathologies.

Technology, however, is no substitute for sound theory, good experimental design, and rigorous methodology. Many of the conceptual, methodological, and statistical difficulties outlined here will still challenge future imaging studies that employ these emerging and still undeveloped technologies. Indeed, although new imaging technologies address some of the limitations of anteceding technologies, they always seem to bring their own new sets of limitations and always raise new methodological, conceptual, and statistical questions. Future imaging studies will continue to grapple with issues of adequate sample size, modeling of potentially confounding effects, disentangling cause from compensatory or epiphenomenal effects, valid anatomical localization, and the analysis and interpretation of findings at a neural systems level of conceptualization. It will be incumbent upon investigators to learn about the limitations of their technologies and the difficulties in interpretation that these limitations engender and then to discuss them unfailingly and candidly in their original, data-based publications.

The methodological, conceptual, and statistical challenges inherent in the application of imaging technologies are most pernicious and insidious when the technological aspects of the study are not well integrated with the neuroscientific and pathophysiological ones. The truth of this assertion may seem painfully obvious to many, but imaging studies are increasingly being conducted by psychologists, psychiatrists, and neuroscientists who have only a meager understanding not only of the strengths and limitations of the technologies and their appropriate implementation but also of the range of interpretations that are possible for the data that imaging technologies provide. These investigators are relying increasingly on experts in physics, engineering, or the radiological sciences to implement and safeguard the technical aspects of their studies. Experts in the physical and radiological sciences, on the other hand, usually have little or no knowledge of the neural or pathophysiolgical processes that neuroscientists and clinicians want to study or measure; thus, they can provide little or no informed advice on how to measure it.

Scientifically fruitful and valid applications of neuroimaging technologies to the study of developmental psychopathologies requires a genuinely integrated knowledge of physics, engineering, biostatistics, clinical neuroscience, human development, and psychopathology. Close communication among members of each of these disciplines can help to foster this integration during the course of a study, although more real-

https://doi.org/10.1017/S0954579403000385 Published online by Cambridge University Press

istically it will need to occur over the course of an entire academic career. Given the real-world demands of academic schedules and the difficulty of retaining personnel of multidisciplinary teams throughout their careers, however, the necessary level of integration across separate individuals or academic departments has been achieved only rarely, and the quality of the science has suffered accordingly.

The only way of ensuring the desired level of integration across disciplines is for one member of the study team to be trained in the theory and methods of each of the component disciplines. Despite the burgeoning number of imaging studies of developmental disorders that are published or underway, currently only a handful of investigators have received the kind of training and experience that permit this kind of multidisciplinary integration. The dearth of appropriately trained investigators is not surprising, given the magnitude of the task of training across these diverse disciplines, the duration of time involved, and the paucity of organized training programs that can appropriately educate and develop new investigators. Entering into training in this area of research is therefore a daunting prospect for most prospective young investigators. The added burden of remaining at a trainee's salary indefinitely can be the final consideration that dissuades even the most idealistic and intellectually passionate young clinical investigators from undertaking the necessary scientific training. Therefore, academic departments and government sponsored training programs need to provide better funding for the multidisciplinary training that is required to learn how to conduct imaging studies well, if prospective young investigators are ever to be encouraged to undertake it.

Despite these many challenges, pediatric neuroimaging will continue to make increasingly important contributions to understanding the etiologies and pathophysiologies of developmental disorders. Technological advances employed by well trained investigators in novel and rigorous experimental paradigms will help to drive these contributions to our understanding of both central etiological processes and compensatory responses, the balance of which helps to define neural functioning in health and illness. Empirical truths will never be revealed in a single experiment or a single investigative modality. Like the proverbial blind men and the elephant, we can only come to know an entity through multiple descriptions of its many aspects. Therefore, progress in understanding developmental psychopathologies will above all be made through careful, programmatic study across multiple neural systems, using multiple imaging modalities, and in conjunction with the discriminating use of investigative tools from molecular genetics, molecular biology, cognitive neuroscience, epidemiology, and phenomenology, all increasingly applied within studies of the same patient cohorts. Neuroimaging is likely to be the central investigative modality in these studies because imaging is the only one among these many modalities that samples and visualizes tissue directly in the living brain. It is therefore the one modality that can bridge and integrate molecular techniques on the one hand with behavioral methodologies on the other as they join to study developmental psychopathologies in living children.

## References

Alexander, G. E., & Moeller, J. R. (1994). Application of the scaled subprofile model to functional imaging in neuropsychiatric disorders: A principal component approach to modeling brain function in disease. *Human Brain Mapping, 2,* 79–94.

Anderson, A. W., & Gore, J. C. (1997). The physical basis of neuroimaging techniques. In B. S. Peterson (Ed.), *Child and adolescent psychiatric clinics of North America* (Vol. 6, pp 213–264). Philadelphia, PA: W. B. Saunders.

Andreasen, N. C., Arndt, S., Cohen, G., Alliger, R., & Swayze, V. W. (1993). Correction for head size. *Psychiatry Research, 50,* 283.

Andreasen, N. C., Flashman, L., Flaum, M., Arndt, S., Swayze, V. n., O'Leary, D. S., Ehrhardt, J. C., & Yuh, W. T. (1994). Regional brain abnormalities in schizophrenia measured with magnetic resonance imaging. *Journal of the American Medical Association, 272,* 1763–1769.

Arndt, S., Cohen, G., Alliger, R. J., Swayze, V. W., &

https://doi.org/10.1017/S0954579403000385 Published online by Cambridge University Press

Andreasen, N. C. (1991). Problems with ratio and proportion measures of imaged cerebral structures. *Psychiatry Research, 40,* 79–89.

Arnold, D. L., de Stefano, N., Matthews, P. M., & Trapp, B. D. (2001). *N*-Acetylaspartate: Usefulness as an indicator of viable neuronal tissue. *Annals of Neurology, 50,* 823–825.

Ashburner, J., & Friston, K. J. (2000). Voxel-based morphometry—The methods. *Neuroimage, 11,* 805–821.

Barbas, H., & Pandya, D. N. (1989). Architecture and intrinsic connections of the prefrontal cortex in the rhesus monkey. *Journal of Comparative Neurology, 286,* 353–375.

Baxter, L., Jr., Schwartz, J. M., Bergman, K. S., Szuba, M. P., Guze, B. H., Mazziotta, J. C., Alazraki, A., Selin, C. E., Ferng, H. K., Munford, P., & Phelps, M. E. (1992). Caudate glucose metabolic rate changes with both drug and behavior therapy for obsessive–compulsive disorder. *Archives of General Psychiatry, 49,* 681–689.

Baxter, L., Schwartz, J., Mazziotta, J., Phelps, M. E., Pahl, J. J., Buze, B. E., & Fairbanks, L. (1988). Cerebral glucose metabolic rates in nondepressed patients with obsessive–compulsive disorder. *American Journal of Psychiatry, 145,* 1560–1563.

Baxter, L. R. J., Phelps, J. M., Mazziotta, J. C., Guze, B. H., & Schwartz, J. M. (1987). Local cerebral glucose metabolic rates in obsessive-compulsive disorder: A comparison with rates in unipolar depression and normal controls. *Archives of General Psychiatry, 44,* 211–218.

Beauregard, M., Lévesque, J., & Bourgouin, P. (2001). Neural correlates of conscious self-regulation of emotion. *Journal of Neuroscience, 21,* 1–6.

Bench, C. J., Frackowiak, R. S. J., & Dolan, R. J. (1995). Changes in regional cerebral blood flow on recovery from depression. *Psychological Medicine, 25,* 247–251.

Benkelfat, C., Nordahl, T. E., Semple, W. E., King, A. C., Murphy, D. L., & Cohen, R. M. (1990). Local cerebral glucose metabolic rates in obsessive–compulsive disorder. Patients treated with clomipramine. *Archives of General Psychiatry, 47,* 840–848.

Biswal, B., Yetkin, F. Z., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic Resonance in Medicine, 34,* 537–541.

Bookstein, F. L. (2001). "Voxel-based morphometry" should not be used with imperfectly registered images. *Neuroimage, 14,* 1454–1462.

Braak, H. (1980). *Architectonics of the human telencephalic cortex*. Berlin: Springer–Verlag.

Breiter, H. C., Rauch, S. L., Kwong, K. K., Baker, J. R., Weisskoff, R. M., Kennedy, D. N., Kendrick, A. D., Davis, T. L., Jiang, A., Cohen, M. S., Stern, C. E., Belliveau, J. W., Baer, L., O'Sullivan, R. L., Savage, C. R., Jenike, M. A., & Rosen, B. R. (1996). Functional magnetic resonance imaging of symptom provocation in obsessive–compulsive disorder. *Archives of General Psychiatry, 53,* 595–606.

Brodmann, K. (1909). *Vergleichende lokalisationslehre der grosshirnrinde*. Leipzig: Barth.

Bruun, R. D., & Budman, C. L. (1992). The natural history of Tourette syndrome. *Advances in Neurology, 58,* 1–6.

Buchsbaum, M. S., Wu, J., Siegel, B. V., Hackett, E., Trenary, M., Abel, L., & Reynolds, C. (1997). Effect of sertraline on regional metabolic rate in patients with affective disorder. *Biological Psychiatry, 41,* 15–22.

Buonomano, D. V., & Merzenich, M. M. (1998). Cortical plasticity: From synapses to maps. *Annual Review of Neuroscience, 21,* 149–186.

Bush, G., Whalen, P. J., Rosen, B. R., Jenike, M. A., McInerney, S. C., & Rauch, S. L. (1998). The counting Stroop: An interference task specialized for functional neuroimaging—Validation study with functional MRI. *Human Brain Mapping, 6,* 270–282.

Buxton, R. B. (2002). *Introduction to functional magnetic resonance imaging*. Cambridge: Cambridge University Press.

Carter, C. S., Mintun, M., & Cohen, J. D. (1995). Interference and facilitation effects during selective attention: An $H_2{}^{15}O$ PET study of Stroop task performance. *Neuroimage, 2,* 264–272.

Caviness, V. S. J., Meyer, J., Makris, N., & Kennedy, D. N. (1996). MRI-based topographic parcellation of human neocortex: An anatomically specified method with estimate of reliability. *Journal of Cognitive Neuroscience, 8,* 566–587.

Childress, A. R., Mozley, P. D., McElgin, W., Fitzgerald, J., Reivich, M., & O'Brien, C. P. (1999). Limbic activation during cue-induced cocaine craving. *American Journal of Psychiatry, 156,* 11–18.

Chua, P., Krams, M., Toni, I., Passingham, R., & Dolan, R. (1999). A functional anatomy of anticipatory anxiety. *Neuroimage, 9,* 563–571.

Cicchetti, D., & Cannon, T. D. (1999). Neurodevelopmental processes in the ontogenesis and epigenesis of psychopathology. *Development and Psychopathology, 11,* 375–393.

Cicchetti, D., & Tucker, D. (1994a). Development and self-regulatory structures of the mind. *Development and Psychopathology, 6,* 533–549.

Cicchetti, D., & Tucker, D. (1994b). Neural plasticity, sensitive periods, and psychopathology [Special Issue]. *Development and Psychopathology, 6.*

Clarke, L. P., Velthuizen, R. P., Camacho, M. A., Heine, J. J., Vaidyanathan, M., Hall, L. O., Thatcher, R. W., & Silbiger, M. L. (1995). MRI segmentation: Methods and applications. *Magnetic Resonance Imaging, 13,* 343–368.

Cohen, M. (1996). Functional MRI: A phrenology for the 1990s? *Journal of Magnetic Resonance Imaging, 6,* 273–274.

Conturo, T. W., Lori, N. F., Cull, T. S., Akbudak, E., Snyder, A. A., Shimony, J. S., McKinstry, R. C., Burton, H., & Raichle, M. E. (1999). Tracking neuronal fiber pathways in the living brain. *Proceedings of the National Academy of Sciences USA, 96,* 10422–10427.

Craig, A. D., Reiman, E. M., Evans, A., & Bushnell, M. C. (1996). Functional imaging of an illusion of pain. *Nature, 384,* 258–260.

Dabringhaus, A., Schormann, T., Steinmetz, H., Schlaug, G., Zilles, K., & Roland, P. E. (1995). What is a standard brain? Intersubject variability of the shape of the human forebrain. *Human Brain Mapping, 1,* 64.

Dale, A. M., & Halgren, E. (2001). Spatiotemporal mapping of brain activity by integration of multiple imaging modalities. *Current Opinions in Neurobiology, 11,* 202–208.

Derbyshire, S. W. G., Jones, A. K. P., Gyulai, F., Clark, S., Townsend, D., & Firestone, L. L. (1997). Pain processing during three levels of noxious stimulation

https://doi.org/10.1017/S0954579403000385 Published online by Cambridge University Press

produces differential patterns of central activity. *Pain, 73,* 431–445.
Derbyshire, S. W. G., Vogt, G. A., & Jones, A. K. P. (1998). Pain and Stroop interference tasks separate processing modules in anterior cingulate cortex. *Experimental Brain Research, 118,* 52–60.
Dolan, R. J., & Friston, K. J. (1997). Functional imaging and neuropsychiatry. *Psychological Medicine, 27,* 1241–1246.
Donders, F. C. (1969). On the speed of mental processes. *Acta Psychologica, 30,* 412–431.
Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., & Taub, E. (1995). Increased cortical representation of the fingers of the left hand in string players. *Science, 270,* 305–307.
Flechsig, P. (1920). *Anatomie des menschlichen gehirns und rückenmarks auf myelogenetischer grundlage*. Leipzig: Thieme.
Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., & Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. *Magnetic Resonance in Medicine, 33,* 636–647.
Friston, K. (2002). Beyond phrenology: What can neuroimaging tell us about distributed circuitry? *Annual Review of Neuroscience, 25,* 221–250.
Fuster, J. M. (1997). *The prefrontal cortex: Anatomy, physiology, and neuropsychology of the frontal lobe* (3rd ed.). Philadelphia, PA: Lippincott–Raven.
Gerard, E., & Peterson, B. S. (in press). Developmental processes and brain imaging studies in Tourette syndrome. *Journal of Psychosomatic Research.*
Gevins, A., Le, J., Brickett, P., Reutter, B., & Desmond, J. (1991). Seeing through the skull: Advanced EEG's use MRI's to accurately measure cortical activity from the scalp. *Brain Topography, 4,* 125–131.
Gevins, A., Le, J., Leong, H., McEvoy, L. K., & Smith, M. E. (1999). Deblurring. *Journal of Clinical Neurophysiology, 16,* 204–213.
Gould, S. J. (1981). *The mismeasure of man*. New York: W. W. Norton.
Goutte, C., Toft, P., Rostrup, E., Nielsen, F. A., & Hansen, L. K. (1999). On clustering fMRI time series. *Neuroimage, 9,* 298–310.
Halpern, B. P. (2000). Sensory coding, decoding, and representations. Unnecessary and troublesome constructs? *Physiology and Behavior, 69,* 115–118.
Hampson, M., Peterson, B. S., Skudlarski, P., Gatenby, C., & Gore, J. C. (2002). Tracking connectivity in the language system revealed via temporal correlations in MR images. *Human Brain Mapping, 15,* 247–262.
Hann, D. M., Huffman, L. C., Lederhendler, I. I., & Meinecke, D. (Eds.). (1998). *Advancing research on developmental plasticity: Integrating the behavioral science and neuroscience of mental health*. Washington, DC: National Institute of Mental Health.
Horwitz, B., & Sporns, O. (1994). Neural modeling and functional neuroimaging. *Human Brain Mapping, 1,* 269–283.
Hsieh, J. C., Hagermak, O., Stahle–Backdahl, M., Ericson, K., Eriksson, L., Stone–Elander, S., & Ingvar, M. (1994). Urge to scratch represented in the human cerebral cortex during itch. *Journal of Neurophysiology, 72,* 3004–3008.
Insel, T. R. (1992). Toward a neuroanatomy of obsessive–compulsive disorder. *Archives of General Psychiatry, 49,* 739–744.
Kates, W. R., Warsofsky, I. S., Patwardhan, A., Abrams, M. T., Liu, A. M., Naidu, S., Kaufmann, W. E., & Reiss, A. L. (1999). Automated Talairach atlas-based parcellation and measurement of cerebral lobes in children. *Psychiatry Research, 91,* 11–30.
Kraemer, H. C., Yesavage, J. A., Taylor, J. L., & Kupfer, D. (2000). How can we learn about developmental processes from cross-sectional studies, or can we? *American Journal of Psychiatry, 157,* 163–171.
Lebon, V., Petersen, K. F., Cline, G. W., Shen, J., Mason, G. F., Dufour, S., Behar, K. L., Shulman, G. I., & Rothman, D. L. (2002). Astroglial contribution to brain energy metabolism in humans revealed by 13C nuclear magnetic resonance spectroscopy: Elucidation of the dominant pathway for neurotransmitter glutamate repletion and measurement of astrocytic oxidative metabolism. *Journal of Neuroscience, 22,* 1523–1531.
Leckman, J. F., Zhang, H., Vitale, A., Lahnin, F., Lynch, K., Bondi, C., Kim, Y.-S., & Peterson, B. S. (1998). Course of tic severity in Tourette's syndrome: The first two decades. *Pediatrics, 102,* 14–19.
Ledberg, A., Akerman, S., & Roland, P. E. (1998). Estimation of the probabilities of 3D clusters in functional brain images. *Neuroimage, 8,* 113–128.
Le Goualher, G., Procyk, E., Collins, D. L., Venugopal, R., Barillot, C., & Evans, A. C. (1999). Automated extraction and variability analysis of sulcal neuroanatomy. *IEEE Transactions on Medical Imaging, 18,* 206–217.
Leung, H.-C., Skudlarski, P., Gatenby, J. C., Peterson, B. S., & Gore, J. C. (2000). An event-related fMRI study of color and word interference. *Cerebral Cortex, 10,* 552–560.
Maas, L. C., Lukas, S. E., Kaufman, M. J., Weiss, R. D., Daniels, S. L., Rogers, V. W., Kukes, T. J., & Renshaw, P. F. (1998). Functional magnetic resonance imaging of human brain activation during cue-induced cocaine craving. *American Journal of Psychiatry, 155,* 124–126.
MacDonald, D., Kabani, N., Avis, D., & Evans, A. C. (2000). Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI. *Neuroimage, 12,* 340–356.
Mathalon, D. H., Sullivan, E. V., Rawles, J. M., & Pfefferbaum, A. (1993). Correction for head size in brain imaging measurements. *Psychiatry Research: Neuroimaging, 50,* 121–139.
Mayberg, H. S., Liotti, M., Brannan, S. K., McGinnis, S., Mahurin, R. K., Jerabek, P. A., Silva, J. A., Tekell, J. L., Martin, C. C., Lancaster, J. L., & Fox, P. T. (1999). Reciprocal limbic–cortical function and negative mood: Converging PET findings in depression and normal sadness. *American Journal of Psychiatry, 156,* 675–682.
McKeown, M. J., Makeig, S., Brown, G. G., Jung, T.-P., Kindermann, S. S., Bell, A. J., & Sejnowski, T. J. (1998). Analysis of fMRI data by blind separation into independent spatial components. *Human Brain Mapping, 6,* 160–188.
McKeown, M. J., Varadarajan, V., Huettel, S., & McCarthy, G. (2002). Deterministic and stochastic features of fMRI data: Implications for analysis of event-related experiments. *Journal of Neuroscience Methods, 118,* 103–113.
Mesulam, M.-M. (1998). From sensation to cognition. *Brain, 121,* 1013–1052.
Moeller, J. R., & Strother, S. C. (1991). A regional covar-

https://doi.org/10.1017/S0954579403000385 Published online by Cambridge University Press

iance approach to the analysis of functional patterns in positron emission tomographic data. *Journal of Cerebral Blood Flow and Metabolism, 11,* A121-A135.

Nemeroff, C. B., Kilts, C. D., & Berns, G. S. (1999). Functional brain imaging: Twenty-first century phrenology of psychological advance for the millenium? *American Journal of Psychiatry, 156,* 671–673.

Novotny, E., Ashwal, S., & Shevell, M. (1998). Proton magnetic resonance spectroscopy: An emerging technology in pediatric neurology research. *Pediatric Research, 44,* 1–10.

O'Brien, G., & Opie, J. (1999). A connectionist theory of phenomenal experience. *Behavior and Brain Science, 22,* 127–148; discussion, 148–196.

Ogawa, S., Tank, D. W., Menon, R., Ellermann, J. M., Kim, S.-G., Merkle, H., & Ugurbil, K. (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*. New York: Thieme Medical.

Pardo, J. V., Pardo, P. J., Janer, K. W., & Raichle, M. E. (1990). The anterior cingulate cortex mediates processing selection in the Stroop attentional conflict paradigm. *Proceedings of the National Academy of Sciences, 87,* 256–259.

Peterson, B. S. (in press). Brain imaging modalities and methodologies in the study of childhood neuropsychiatric disorders. *Trends in Evdence-Based Neuropsychiatry.*

Peterson, B. S., Kane, M. J., Alexander, G. M., Lacadie, C., Skudlarski, P., Leung, H.-C., May, J., & Gore, J. C. (2002). An event-related functional MRI study comparing interference effects in the Simon and Stroop tasks. *Cognitive Brain Research, 13,* 427–440.

Peterson, B. S., Skudlarski, P., Anderson, A. W., Zhang, H., Gatenby, J. C., Lacadie, C. M., Leckman, J. F., & Gore, J. C. (1998). A functional magnetic resonance imaging study of tic suppression in Tourette syndrome. *Archives of General Psychiatry, 55,* 326–333.

Peterson, B. S., Skudlarski, P., Zhang, H., Gatenby, J. C., Anderson, A. W., & Gore, J. C. (1999). An fMRI study of Stroop word–color interference: Evidence for cingulate subregions subserving multiple distributed attentional systems. *Biological Psychiatry, 45,* 1237–1258.

Peterson, B. S., Staib, L., Scahill, L., Zhang, H., Anderson, C., Leckman, J. F., Cohen, D. J., Gore, J. C., Albert, J., & Webster, R. (2001). Regional brain and ventricular volumes in Tourette syndrome. *Archives of General Psychiatry, 58,* 427–440.

Peterson, B. S., & Thomas, P. (2000). Functional brain imaging in Tourette's syndrome: What are we really imaging? In M. Ernst, & J. Rumsey (Eds.), *Functional neuroimaging in child psychiatry* (pp 242–265). Cambridge: Cambridge University Press.

Peterson, B. S., Vohr, B., Kane, M., Whalen, D., Schneider, K., Katz, K., Zhang, H., Makuch, R., Gore, J., & Ment, L. (2002). A functional magnetic resonance imaging study of language processing and its cognitive correlates in prematurely born children. *Pediatrics, 110,* 1153–1162.

Pham, D. L., Xu, C., & Prince, J. L. (2000). Current methods in medical image segmentation. *Annual Review of Biomedical Engineering, 2,* 315–337.

Pierpaoli, C. (2002). Inferring structural and architectural features of brain tissue from DT-MRI measurements. *CNS Spectrums, 7,* 510–515.

Poeppel, D. (1996). A critical review of PET studies of phonological processing. *Brain and Language, 55,* 317–351.

Poline, J.-B., Worsley, K. J., Evans, A. C., & Friston, K. J. (1997). Combining spatial extent and peak intensity to test for activations in functional imaging. *Neuroimage, 5,* 83–96.

Rauch, S. L., Jenike, M. A., Alpert, N. M., Baer, L., Breiter, H. C., Savage, C. R., & Fischman, A. J. (1994). Regional cerebral blood flow measured during symptom provocation in obsessive–compulsive disorder using oxygen 15-labeled carbon dioxide and positron emission tomography. *Archives of General Psychiatry, 51,* 62–70.

Roland, P. E. (1993). *Brain activation* (pp. 105–140). New York: Wiley–Liss.

Ross, B., & Bluml, S. (2001). Magnetic resonance spectroscopy of the human brain. *Anatomical Record, 265,* 54–84.

Rothman, D. L., Sibson, N. R., Hyder, F., Shen, J., Behar, K. L., & Shulman, R. G. (1999). In vivo nuclear magnetic resonance spectroscopy studies of the relationship between the glutamate–glutamine neurotransmitter cycle and functional neuroenergetics. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 354,* 1165–1177.

Rubin, R. T., Ananth, J., Villanueva–Meyer, J., Trajmar, P. G., & Mena, I. (1995). Regionial [133] xenon cerebral blood flow and cerebral [99m]Tc-HMPAO uptake in patients with obsessive–compulsive disorder before and during treatment. *Biological Psychiatry, 38,* 429–437.

Sanides, R. (1964). Structure and function of the human frontal lobe. *Neuropsychologia, 2,* 209–219.

Schwartz, J. M., Stoessel, P. W., Baxter, L. R., Martin, K. M., & Phelps, M. E. (1996). Systematic changes in cerebral glucose metabolic rate after successful behavior modification treatment of obsessive–compulsive disorder. *Archives of General Psychiatry, 53,* 109–113.

Sibson, N. R., Dhankhar, A., Mason, G. F., Behar, K. L., Rothman, D. L., & Shulman, R. G. (1997). In vivo 13C NMR measurements of cerebral glutamine synthesis as evidence for glutamate–glutamine cycling. *Proceedings of the National Academy of Sciences USA, 94,* 2699–2704.

Skudlarski, P., Constable, R. T., & Gore, J. C. (1999). ROC analysis of statistical methods used in functional MRI: Individual subjects. *Neuroimage, 9,* 311–329.

Smith, R. C., & Lange, R. C. (1998). *Understanding magnetic resonance imaging*. Boca Raton, FL: CRC Press.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donder's method. In W.G. Koster (Ed.), *Attention and performance II.* Amsterdam: North–Holland.

Swedo, S. E., Pietrini, P., Leonard, H. L., Schapiro, M. B., Rettew, D. C., Goldberger, E. L., Rapoport, S. I., Rapoport, J. L., & Grady, C. L. (1992). Cerebral glucose metabolism in childhood-onset obsessive–compulsive disorder. Revisualization during pharmacotherapy. *Archives of General Psychiatry, 49,* 690–694.

Swedo, S. E., Schapiro, M. B., Grady, C. L., Cheslow, D. L., Leonard, H. L., Kumar, A., Friedland, R., Rapoport, S. I., & Rapoport, J. L. (1989). Cerebral glucose metabolism in childhood-onset obsessive–compulsive disorder. *Archives of General Psychiatry, 46,* 518–523.

https://doi.org/10.1017/S0954579403000385 Published online by Cambridge University Press

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

Talbot, J. D., Marrett, S., Evans, A. C., Meyer, E., Bushnell, M. C., & Duncan, G. H. (1991). Multiple representations of pain in human cerebral cortex. *Science, 251,* 1355–1358.

Toga, A. W., & Thompson, P. M. (2002). New approaches in brain morphometry. *American Journal of Geriatric Psychiatry, 10,* 13–23.

Viergever, M. A., Maintz, J. B., Niessen, W. J., Noordmans, H. J., Pluim, J. P., Stokking, R., & Vincken, K. L. (2001). Registration, segmentation, and visualization of multimodal brain images. *Computerized Medical Imaging & Graphics, 25,* 147–151.

Volkow, N. D., Wang, G.-J., Fowler, J. S., Hitzemann, R., Angrist, B., Gatley, S. J., Logan, J., Ding, Y.-S., & Pappas, N. (1999). Association of methylphenidate-induced craving with changes in right striato-orbitofrontal metabolism in cocaine abusers: Implications in addiction. *American Journal of Psychiatry, 156,* 19–26.

Wexler, B. E. (1986). A model of brain function: its implications for psychiatric research. *British Journal of Psychiatry, 148,* 357–362.

Worsley, K. J., Evans, A. C., Marrett, S., & Neelin, P. (1992). A three-dimensional statistical analysis for CBF activation studies in human brain. *Journal of Cerebral Blood Flow and Metabolism, 12,* 900–918.

Zaborszky, L. (2002). The modular organization of brain systems. Basal forebrain: The last frontier. *Progress in Brain Research, 136,* 359–372.

Zilles, K., Schleicher, A., Langemann, C., Amunts, K., Morosan, P., Palomero–Gallagher, N., Schormann, T., Mohlberg, H., Bürgel, U., Steinmetz, H., Schlaug, G., & Roland, P. E. (1997). Quantitative analysis of sulci in the human cerebral cortex: Development, regional heterogeneity, gender difference, asymmetry, intersubject variability, and cortical architecture. *Human Brain Mapping, 5,* 218–221.

https://doi.org/10.1017/S0954579403000385 Published online by Cambridge University Press