ACADEMIC PRESS

Available online at www.sciencedirect.com
SCIENCE DIRECT®

NeuroImage

NeuroImage 20 (2003) 615–624

www.elsevier.com/locate/ynimg

Commentary

Predicting performance from functional imaging data: methods matter

John J. Sidtis,[a,b,*] Stephen C. Strother,[c] and David A. Rottenberg[c,d]

[a] *Geriatrics Division, Nathan Kline Institute, Orangeburg, NY 10962, USA*
[b] *Department of Psychiatry, New York University Medical School, New York, NY, USA*
[c] *Department of Radiology, University of Minnesota Medical School and Minneapolis VA Medical Center, Minneapolis, MN, USA*
[d] *Department of Neurology, University of Minnesota Medical School and Minneapolis VA Medical Center, Minneapolis, MN, USA*

Received 18 March 2003; revised 3 June 2003; accepted 3 June 2003

Abstract

In the standard approach to functional imaging studies, brain-behavior relationships are studied by contrasting data obtained during different behavioral states. It is generally assumed that relative change yields meaningful data about relevant brain processes, and that the magnitude of the change reflects the extent of a region's involvement in the behavior being studied. The present study takes a different approach by asking the question, Can functional imaging data predict performance? Regional cerebral blood flow was measured using positron emission tomography in a group of 13 right-handed, normal volunteers during speech production and quiet baseline. A number of methodological assumptions were addressed by examining the relationships between different imaging measures derived from the same raw data and performance on the speech task. The results demonstrate that several common assumptions are not necessarily true. First, although measures based on "activated" scans alone had predictive value with respect to speech rate, measures based on contrasts between "baseline" and "activated" states did not. This was true regardless of whether the contrast was based on subtraction or covariance analyses. Second, while many regions demonstrated large signal increases during speech, speech rate could be predicted by a linear combination of data from two regions, neither of which had the highest "activation" peak, and one of which had a negative relationship with performance. The results demonstrate that contrasting experimental conditions do not necessarily isolate or enhance brain activity related to performance, and that the current assumptions about activation in functional imaging need to be reconsidered.

Introduction

Functional imaging has assumed a major role in the study of human brain-behavior relationships. Unlike lesion studies, which establish relationships based on selective deficits following damage to specific brain regions (Teuber, 1955), functional imaging studies seek to establish such relationships based on indirect observations of brain activity while subjects are engaged in specific experimental tasks. Whereas the view of brain organization resulting from each of these approaches has been shaped in part by pragmatic limitations and in part by their respective methodological and theoretical assumptions, an understanding of the effects of many of the assumptions underlying functional imaging is relatively rudimentary.

Despite their limitations, lesion studies can be considered the best approximation to a "gold standard" for brain-behavior relationships. If damage to a specific brain region routinely results in a specific deficit, then the damaged region can be considered to play a significant role in the behavior in question (Benson, 1994). One of the strongest demonstrations of functional localization in lesion studies is the "double dissociation" where localization in each of two different brain areas is established by demonstrating complementary patterns of impairment and sparing on two different functions (Teuber, 1955). For example, following left temporal-parietal area lesions, speech sound discrimination is impaired but complex-pitch discrimination is spared whereas the reverse is true following right temporal-parietal area lesions (Sidtis and Volpe, 1988). More generally, the ability to diagnose the site and type of neurological lesion

* Corresponding author. Geriatrics Division, Nathan Kline Institute, 140 Old Orangeburg Road, Orangeburg, NY 10962. Fax: +1-845-398-5575.

E-mail address: john.sidtis@nyu.edu (J.J. Sidtis).

1053-8119/$ – see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/S1053-8119(03)00349-5

from clinical signs and symptoms demonstrates that this approach to establishing brain-behavior relationships has predictive value. The limits of the lesion approach in the study of highly circumscribed behaviors include the possible confounding of specific and nonspecific behavioral changes after brain damage, the difficulty in separating the degree to which the damaged region contributes directly to a behavior versus the degree to which it contributes to a network of other functionally significant regions, and the possibility that the study of performance after neurological damage will yield little insight into how behavior is normally organized in the brain.

Functional imaging techniques like positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) have the potential to overcome the limitations of the lesion approach to understanding human brain-behavior relationships as the physiological significance of what is being measured is better understood. The standard method of studying both normal and brain-injured subjects using functional imaging is based on a series of assumptions: task performance results in increases in neuronal activity; this activity is seen as an increase or "activation" in the signal detected by the imaging method (Raichle, 1987, 1991); and the brain areas that are most "activated" are the areas most involved in task execution (Friston et al., 1991). An alternative, but not mutually exclusive, view characterizes activation in terms of a relationships among brain areas (e.g., Moeller et al., 1987; Moeller and Strother, 1991; Horwitz, 1991; Strother et al., 1995; McIntosh et al., 1996; Taylor et al., 2000). In this approach, activation is represented by a significant relationship among areas rather than by an increased signal in individual areas.

Whether the search for brain-behavior relationships follows the path of regional activation or the path of regional patterns, functional imaging studies also typically assume that during any given task, brain activity is a composite, usually an additive combination, of constituent conditions. In simple experimental situations, the additivity assumption allows the subtraction of "baseline" images from activated images to identify brain areas where changes associated with task performance occur (Fox et al., 1988a; Fox, 1991). In a more complex form, the additivity assumption supports experimental methods that assume that a behavior and its underlying brain processes can be decomposed into constituent components by a series of contrasts involving conditions that differ along one or more dimensions. Although the assumption that complex tasks can be accurately decomposed by subtraction has recently been questioned (Friston et al., 1996; Jennings et al., 1997; Sidtis et al., 1999; Newman et al., 2001), in some form this assumption remains central to much of the work in functional imaging. For most functional imaging studies, then, evidence for brain-behavior correlation is commonly sought in changes in regional activity or in multi-regional covariance measures. It is further believed that the largest increases are observed in areas most involved in the task under study, and that sensitivity to such effects can be improved by using methodological approaches that incorporate some contrast of experimental conditions.

The standard functional imaging method, then, correlates brain and behavior by identifying activation that is reliably observed in association with the contrast of two behavioral conditions. However, a converse approach can be taken in which functional imaging data are used to predict a specific behavioral measure. With this second method, behavioral specificity is directly established with functional imaging data rather than indirectly inferred from the experimental design. Whereas this approach can utilize image data derived from contrasting behavioral conditions, it can also look for correlates in single conditions. This study used the second approach to compare the brain-behavior relationships derived using different forms of image data commonly used in standard functional studies. Each form of image data incorporates different assumptions commonly made in studies using the standard approach. The question and analyses were simple: Which functional imaging measures predict a measure of task performance? Using the converse approach to standard methods, the results of this study suggest that subtraction or covariance-based task contrasts between task and baseline conditions do not necessarily isolate or enhance "activation" effects related to performance. Further, these results suggest that the current assumptions about activation may need to be reconsidered.

Methods

Subjects

A group of 13 right-handed normal volunteers (eight females and five males aged 43 $\pm$ 11 years) participated in this study. All participants were native speakers of English. Subjects were screened to exclude confounding neurologic, psychiatric, and medical disorders, and to exclude current psychotropic medication or recreational drug use. After study procedures and their possible consequences were explained, informed consent was obtained for each study.

Behavioral task

Subjects were studied with eyes covered, room lights dimmed, and insert earphones placed in each auditory canal. During each baseline scan, subjects were required to remain awake and quiet. The task consisted of the repetition of the syllables *pa, ta,* and *ka,* produced as quickly as possible. Subjects were instructed to take a deep breath and then produce as many syllables as possible during expiration. Subjects repeated this process for 60 s. Syllable repetitions were recorded during the scans for subsequent acoustic analyses.

PET scanning

Bolus injection of [^{15}O]water was used as a marker of regional cerebral blood flow (rCBF) (Silbersweig et al., 1993). Each study consisted of eight 90-s scans (four baseline scans alternating with four activated scans), separated by an interscan interval of approximately 9 min, acquired using a Siemens-ECAT 953B tomograph in 3D mode. Each task was performed for 60 s beginning at tracer injection. Subjects were engaged in the task for 10 to 15 s prior to the initiation of each scan, which was triggered when radioactivity reached the brain. Preprocessing procedures, including intra- and intersubject alignment and transformation into Talairach space (Talairach and Tournoux, 1988) can be found in Strother et al. (1995).

Functional imaging metrics

A set of 22 regions of interest (ROIs) was generated based on evidence of a significant regional change during this task or either of two other speech tasks (sustained phonation, repetitive lip closure) performed by this group of subjects (Sidtis et al., 1999). Areas of change were determined by subtracting group mean volumes for control scans from the group mean volumes for each of the tasks. A threshold was applied to each ROI such that ROI values represented the mean of the upper 25% of voxel values (Rottenberg et al., 1991). The following regions (left and right) were examined: inferior, middle, and superior portions of cerebellum in horizontal planes, superior temporal gyrus, transverse temporal gyrus, putamen, caudate, thalamus, inferior frontal lobe, pre- and postcentral gyri, and supplementary motor area.

Five different measures were derived from these data. The data obtained from each of these regions provided a simple surrogate measure of regional cerebral blood flow (rCBF). A volume-mean normalized (VN) measure of regional cerebral blood flow (VNrCBF) was created by multiplying each subject's ROI value by the ratio of the highest volume mean (all voxels) in the dataset divided by that individuals volume mean. The VNrCBF measure is similar to rCBF with the additional assumptions that intersubject differences in volume means are not functionally significant and that volume-mean normalization will reduce irrelevant intersubject variability (Arndt et al., 1996). The VNrCBF values were also used to create simple subtraction values (SSrCBF) representing the differences between consecutive baseline and activated scans. This measure incorporates the assumption that subtraction will remove the nonspecific activation effects due to the scanning procedure itself, resulting in activation specific to the experimental task.

To examine regional covariance, the rCBF values were also subjected to a scaled subprofile model (SSM) analysis (Moeller et al., 1987; Moeller and Strother, 1991) using the volume mean as the estimate of the global scaling factor. This analysis produces a series of principal components representing regional covariance patterns. The loading of each subject's scan on each pattern is represented by a scalar value referred to as a subject scaling factor (SSF). This measure assumes that relevant information in the functional image data can be found in a principal components analysis of the residuals after the removal of mean effects due to factors like subject and region. The SSM analysis was performed twice, once including all activation (a) and baseline (b) scans (SSMab), and a second time including only the activation scans (SSMa). The results of the SSMab analysis were further subjected to a linear discriminant analysis that produces a weighted combination of principal components (canonical discriminant function) to yield the best discrimination between activation and baseline scans. This analysis is referred to as a scaled subprofile model with canonical variables analysis (SSMab/CVA), and it assumes that a linear combination of factors will result in a functionally significant measure of brain activity that distinguishes scanning states (Rottenberg et al., 1996; Strother et al., 1995).

Regression analyses

Multiple linear regression analysis was used to determine the ability of each functional imaging measure to predict the syllable production rate (Norusis, 1988). Separate analyses with stepwise replacement were conducted for each measure using the following criteria: probability of F to enter (0.05), probability of F to remove (0.10), and tolerance (0.01).

Results

Speech rate

Analysis of the recorded speech samples revealed a production rate of 4.0 ± 0.2 syllables/second (group mean $\pm$ 1 standard error of the mean (SEM)).

rCBF

The raw regional data from the four activation scans per subject produced no significant linear relationship with speech rate across subjects. The group mean rCBF values $\pm$ 1 SEM are presented in Fig. 1a. There are no clear peaks, and the highest regions represent the putamen and thalamus while the lowest represent the sensory-motor strip and supplementary motor areas, bilaterally.

VNrCBF

A significant linear model [$F(2,49) = 10.26$; $P = 0.0002$] was found to relate speech rate to the VNrCBF measures derived from the activated scans. The solution

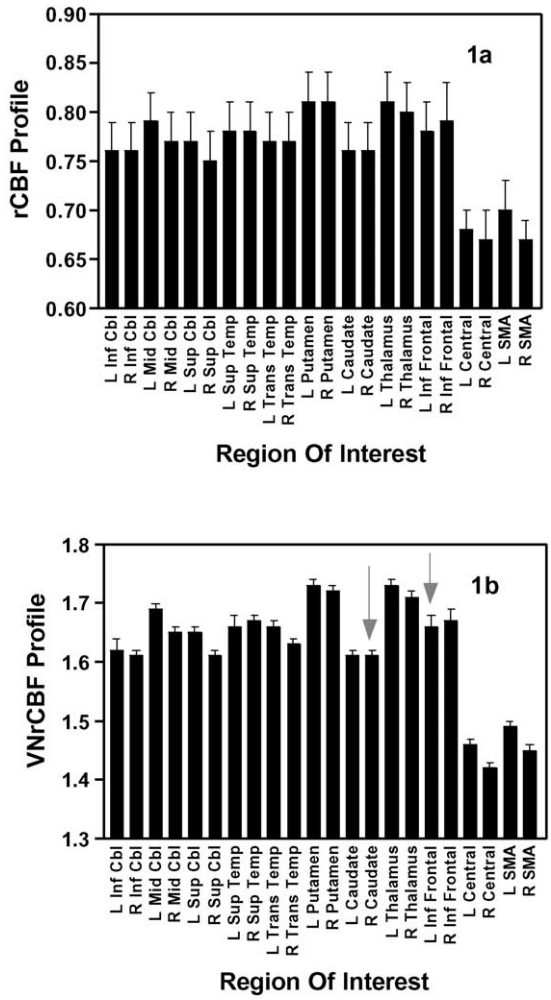

Fig. 1. Group mean regional values for the regional cerebral blood flow (rCBF) (a) and volume-normalized regional cerebral blood flow measures (VNrCBF) (b). Error bars represent one standard error of the mean (SEM). Solid arrows in b indicate the regions that have a linear relationship with speech rate. Regions are presented in left (L) and right (R) pairs, and refer to the following areas: Inf CbI (inferior portion of the cerebellum); Mid CbI (middle portion of the cerebellum); Sup CbI (superior portion of the cerebellum); Sup Temp (superior temporal); Trans Temp (transverse temporal); Putamen; Caudate; Thalamus; Inf (inferior) Frontal; Central (pre- and postcentral gyri); SMA (supplementary motor area). A scatter plot demonstrating the relationship between the syllable production rate predicted by the VNrCBF measure and the actual syllable production rate is presented in Fig. 2a.

contained two regions, the left inferior frontal area and the right caudate, represented in the following equation:

$$\text{Syllable rate} = (-3.55 * \text{right caudate}) + (2.51 * \text{left inferior frontal}) + 5.60.$$

Fig. 1b depicts the group mean VNrCBF values ± 1 SEM. As expected, the normalization did not appreciably alter the relative positions of the regions, but did substantially reduce the intersubject variability. With respect to the current notion of activation, it is important to note that neither the left inferior frontal cortex nor the right caudate regions represented a maximum value or peak. Fig. 2a presents the syllable rates predicted by the VNrCBF equation above plotted against the actual syllable rates. These values were significantly correlated ($r = 0.54$; $P < 0.000$).

SSrCBF

The simple subtraction values derived from the four pairs of sequentially adjacent baseline and activated scans did not yield a significant relationship with speech rate. This was the case both when SSrCBF was analyzed as a difference score and when it was analyzed as a percentage change from baseline. The group data are presented as mean percentage change scores ± 1 SEM in Fig. 3a. Subtraction introduces substantial changes in the relative positions of the ROIs compared to the unsubtracted data. Significantly, neither the left inferior frontal region nor the right caudate, predictors of speech rate before subtraction, represent peak increases after subtraction. The activated left inferior frontal cortex shows only a small percentage increase from baseline, while

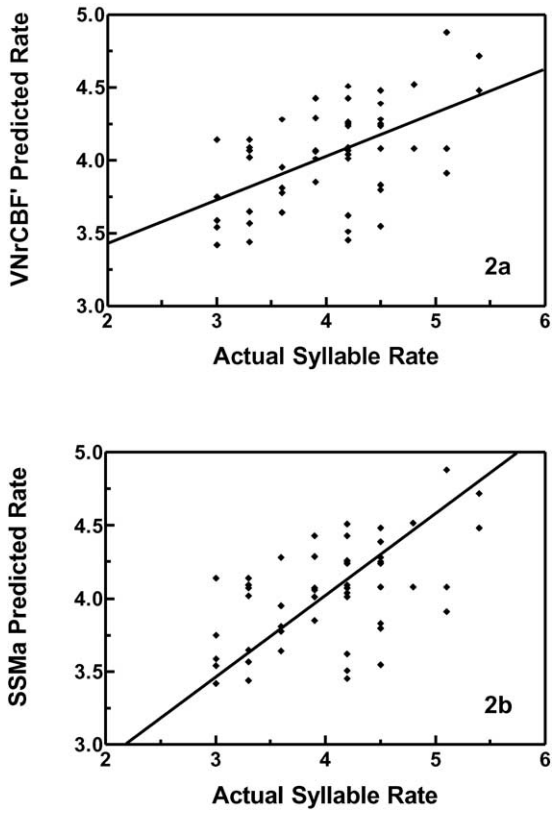

Fig. 2. Scatter plots presenting the relationships between the predicted and the observed syllable rates for the VNrCBF (a) and SSMa (b) measures. The correlation between the predicted and observed values using the VNrCBFN measure is 0.54 ($P < 0.000$). The correlation between the predicted and observed values using the SSMa measure (Fig. 4) is 0.75 ($P < 0.000$).

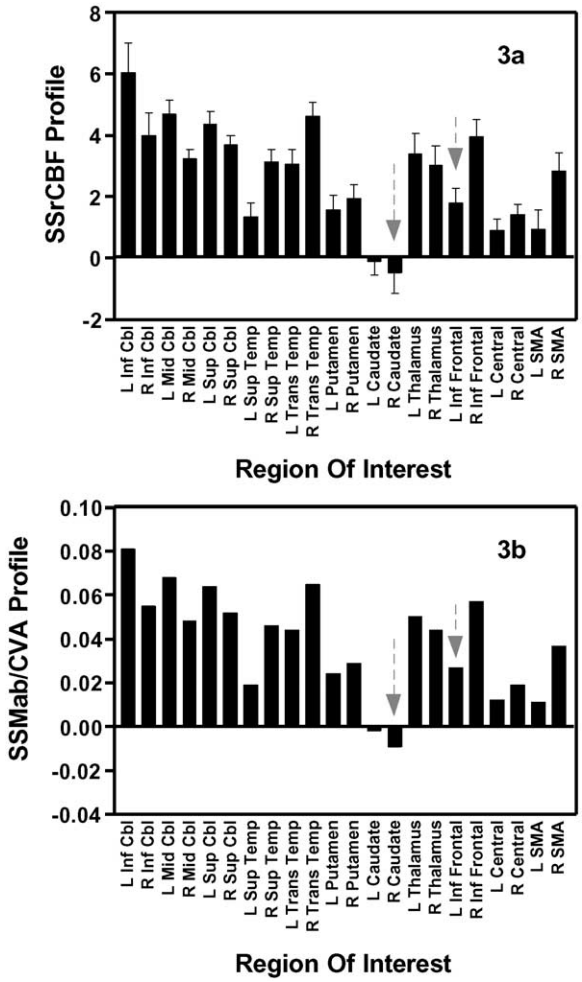

Fig. 3. Group mean regional values ± 1 SEM for the subtraction values (SSrCBF) based on the percentage difference between paired "activated" and baseline VNrCBF values (a). Error bars represent one standard error of the mean (SEM). (b) The covariance profile that represents the best separation of "activated" and "baseline" scans is presented. The regional profiles extracted using these two techniques are essentially identical. The dashed arrows point to the regions that predict the syllable production rate using the VNrCBF measure.

the activated right caudate shows the largest decrease compared to "baseline."

SSMab/CVA (activated and baseline scans)

Consistent with the negative result for the SSrCBF measure, the scores of the canonical variate that optimized the separation of baseline and activated scans were not correlated with speech rate. The results of the canonical variables analysis across ROIs is presented in Fig. 3b. Comparison of the profiles in Fig. 3a and b reveals that the simple subtraction and SSMab/CVA analysis, both of which contrast the activated and baseline conditions, yielded similar profiles. Derived using two different techniques, these profiles agree and likely represent the difference between activated and control states, yet they fail to predict task performance.

SSMa (activated scans only)

An additional SSM analysis was run including only the activated scans. This analysis yielded a significant linear relationship [$F(7,44) = 8.031$; $P < 0.000$] with speech rate that includes seven subject scaling factors:

$$\begin{aligned}\text{Syllable rate} &= (-3.90 * \text{SSFa6}) + (-2.78 * \text{SSFa3}) \\ &+ (-6.66 * SSFa13) + (5.07 * SSFa11) \\ &+ (-10.74 * SSFa18) + (-6.47 * SSFa14) \\ &+ (-8.75 * SSFa17) + 4.05.\end{aligned}$$

The results of applying the regression weights to the appropriate SSFa's are depicted in Fig. 4. As with the plot in Fig. 3b, this profile represents a set of covariance relationships. At present, the neurophysiological interpretation of covariance patterns is not as obvious as the interpretation that can be applied to measures of regional activity. As in the regional measures, however, neither the left inferior frontal nor the right caudate occupy extreme positions in the covariance profile. Fig. 2b presents the syllable rates predicted by the SSMa equation above plotted against the actual syllable rates. These values were significantly correlated ($r = 0.75$; $P < 0.000$).

Discussion

Using an alternative approach to studying brain-behavior relationships with functional imaging, the present results

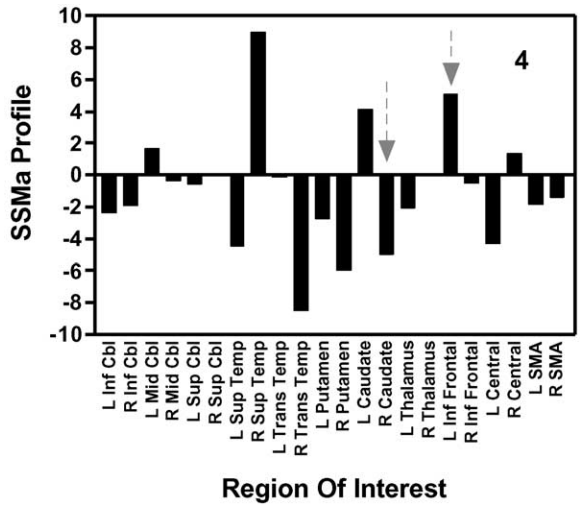

Fig. 4. The composite covariance profile significantly related to speech rate based on an SSM analysis of "activated" scans alone. Each regional value is the weighted sum of its values across seven principle components. Weights were determined by multiple linear regression. The dashed arrows point to the regions that predict the syllable production rate using the VNrCBF measure. Fig. 2b represents the relationship between the syllable production rate predicted by this measure and the actual production rate.

demonstrate that performance on a speech production task can be predicted from some forms of functional imaging data but not others. On a regional basis, normalized data from speech scans yielded a significant association with speech rate, but when the same normalized data from paired speech and baseline scans were subtracted, this significant relationship was lost. Using a covariance analysis that included only speech scans, a combination of scaling factors also yielded a significant relationship with speech rate. As with the regional analysis, however, a covariance analysis that contrasted speech and baseline scans produced a regional pattern similar to that found with subtraction, but like the subtraction result, this covariance measure also failed to predict speech rate. Finally, the two regions significantly associated with speech rate were not the most activated brain regions under current use of the term (i.e., they did not contain the largest values) using any form of the data. These results raise several questions about assumptions commonly made in functional imaging studies.

Do task contrasts identify task-specific activations?

The failure of both subtraction and covariance measures derived from a contrast between speech and baseline conditions in a data set where the noncontrasted speech data alone produced a significant predictive model raises questions about the utility of the baseline contrast. It is generally assumed that the removal of signal due to nonspecific effects of the experimental situation from the image should enhance the relationship between the residual data and the specific experimental task. However, this is only true if a number of other conditions are met. Chief among these are that brain processes are additive and that baseline and activation states do not interact. In addition to the present results, there is growing evidence that these assumptions may not be valid (Sidtis et al., 1999; Jennings et al., 1997; Friston et al., 1996). For example, Newman et al. (2001) reported that the use of different baseline conditions produced different levels of activation, but, more importantly, that these differences were not uniform across regions. Although these authors suggest that covariance approaches provide one solution to the baseline problem, the present results suggest that a covariance approach that contrasts baseline and activated conditions provides results similar to subtraction (Fig. 3a and b). In the present study, the absence of a predictive relationship was not due to the use of subtraction, but, rather, to the use of measures based on contrasts between conditions.

The failure of the contrast measures to predict performance suggests that contrasting conditions do not necessarily isolate task-specific activation. This is an important consideration not only for PET, but for other functional imaging methods that rely on behavioral contrasts. The present results demonstrated that performance could be predicted from data taken while the performance occurred when those data were normalized to reduce global effects and not contrasted with any other condition. Whereas PET has a long tradition of dealing with the issue of normalization, both at the level of whole brain image reconstruction and in the process of data analysis, this has received far less attention with fMRI methods. Similarly, while there is a significant experience with examining single conditions in PET because of the long half-lives of some of the isotopes employed, fMRI methods have relied on within subject comparisons across conditions. Relative change measures are well suited to MRI data acquisition, and under some conditions, may reduce the need to reduce global effects. However reliable such changes are, though, the present results suggest that inferring that the largest relative changes are the most functionally significant with respect to specific behaviors may not always be accurate.

Comparison with the lesion experience

Although the regions where normalized data predicted speech rate were neither identified as activation peaks nor isolated by contrasts with the baseline condition, they do fare well when compared to the gold standard of the lesion experience. The importance of the left inferior frontal region in speech production in right-handed individuals has long been documented (see Kertesz, 1979; Duffy, 1995; Kent and Tjaden, 1997, for reviews). Damage in this area can produce speech production problems such as aphemia (Schiff et al., 1983), anarthria (Mohr et al., 1978), and oral apraxia (Tognola and Vignolo, 1980). Damage to the right caudate has also been associated with dysarthria and repetition. In a series of caudate infarcts, Caplan et al. (1990) reported that dysarthria was associated with damage to either the left or the right caudate, but was more common after damage to the right caudate. Metter et al. (1988) correlated the performance of 11 aphasic patients on a range of language tests with regional cerebral glucose metabolism obtained with fluorodeoxyglucose PET. Performance on the repetition subtest of the Porch Index of Communicative Ability was significantly correlated with metabolism in the caudate, bilaterally, and in Broca's areas and its homologue in the right hemisphere. However, repetition was the only area of performance for which a significant correlation was found for the right caudate. Further, a factor analysis of subtest scores produced a "speaking" factor. As in the present study, there was an inverse relationship between right caudate activity and weights on the "speaking" factor, but this relationship did not reach significance in the Metter et al. study. If the lesion experience is used as the standard, the results from speech scans alone produce better agreement with known functional anatomy than the results from contrasting speaking with not-speaking scans.

Activation: Does a larger increase mean greater functional significance?

To the extent that one accepts a role for Broca's area and the caudate in speech production, the present results also address the assumption that the degree of involvement of a brain structure in a particular task is directly reflected by the magnitude of its numerical value. In the present study, volume-normalized activity in two brain regions predicted speech rate, but neither area demonstrated the largest activation during speech compared to the baseline condition. Further, the changes in the caudate represented what would be considered "negative activation" or "deactivation" in subtraction terminology. Had the analysis strategy in this study simply focused on regions that represented an absolute or relative maximum value, or peak, the caudate would have been ignored, and it is likely that the left inferior frontal region (Broca's area) would have been ignored as well. These results raise several issues about our current conceptualization of activation (Sidtis, 2000).

In the case of cerebral blood flow measures, the belief that greater activation in a region of interest reflects greater involvement of that region in the behavior under investigation can be viewed, in part, as a consequence of two assumptions: first, that areas most involved in task performance will exhibit the highest levels of neuronal activity, and, second, that cerebral blood flow as measured by PET or fMRI is strongly coupled to metabolism during activation. The first assumption ignores possible efficiencies resulting from cerebral specialization. It may well be that when specialized areas are involved in a task for which they are specialized (e.g., speech production), the functional imaging signal will not be as high as that produced by a nonspecialized area. Regarding the second assumption, "strong" may be a relative term in some circumstances (see Fox and Raichle, 1986; Fox et al., 1988b).

The assumption that areas most involved in task performance will exhibit the most activity is also contrary to functional imaging data obtained during the course of learning. In a range of "activation" studies that involve both motor (Mazziotta et al., 1991; Friston et al., 1992) and cognitive tasks (Raichle et al., 1994), improved performance over the course of practice has been associated with decreases in functional imaging signals. A similar reduction occurs as stimuli become more familiar (Vandenberghe et al., 1995). Speech production is not only a highly practiced skill, it is one for which cerebral specialization is well established. Whether skilled behaviors are associated with relatively smaller activations than unskilled behaviors is an open question, and one that is difficult to address if activation is defined only in terms of peak measures.

Brain-behavior correlates vs functional image analysis

Given the present results, the interface between the search for behavioral correlates and the analysis of the data structure in functional images should be addressed. There has been real progress in approaches that extract reliable structure from image data sets without imposing hypotheses about where activation should occur and what it should look like (e.g., Kjems et al., 2002; Levin and Uftring, 2001; Hansen et al., 1999). One of the valid concerns of this approach is that finding an "expected" result (i.e., activation in an area thought to be involved in a specific process) is not a suitably rigorous means of validating an analysis technique (e.g., Strother et al., 2002). Ideally, one would prefer to conduct an unbiased analysis of a functional imaging dataset and extract reliable structure that bore a close neurobiological correspondence to the process being studied. However difficult the problem of extracting reliable structure from imaging data is, neurobiological correspondence is likely far more difficult. Whether the analyses focus on regions of interest or networks, most studies assume a combinatorial process in the brain that is veridical or at least a close approximation to the contrasts imposed by the experimental design. If this were true, the contrast of "speaking" and "not speaking" conditions in the present study should have enhanced the speaking-related information, but it did not. The problem exists not only for baseline contrasts, but for task contrasts as well (Newman et al., 2001; Sidtis et al., 1999; Jennings et al, 1997; Friston et al., 1996). The sophistication of functional image acquisition has grown substantially, but it cannot be safely assumed that the capacity for finding the function in functional images has increased in tandem.

Other approaches to functional imaging

Although the specifics of the problem differ in electrophysiological imaging, the conceptual question remains the same: Whether the measurement is in terms of evoked potentials, evoked magnetic fields, or number of dipoles, which transformations of the signal are functionally meaningful? As an event-related rather than a "peak" picking approach (e.g., Simos et al., 2000), magnetoencephalography (MEG) measures are derived by association with temporal features of a task over a shorter time frame than methods relying on changes in blood flow. Thus, there is some direct association between behavior and physiologic change independent of experimental design. However, behavioral contrasts have also been employed with this method (Sekihara et al., 1998) and effects that change with repetition are less likely to be observed. MEG also has a strong tradition of application in clinical settings, comparing its results to those of invasive measures like Wada testing (e.g., Papanicolaou et al., 1999), intracerebral recording in epilepsy (e.g., Schwartz et al., 2003), and presurgical mapping (Schiffbauer et al., 2002), and the results thus far have been promising. However, the signals of interest are small, the reconstruction problem is subject to multiple solutions, and there are significant effects of signal processing assumptions (e.g., Jerbi et al., 2002; Vrba and

Robinson, 2001; Darvas et al., 2001; Kajihara et al., 2000; Huang et al., 1999). As in the case of functional techniques like PET and fMRI, caution should be used in assuming that the magnitude of an electrophysiological response is a reflection of its functional significance since the measures are heavily model dependent and their properties as physiological scales are not well understood.

Conclusion: Finding the appropriate method

Although obvious, the appropriateness of the imaging method depends on the question being addressed. The predominant approach is aimed at identifying reliable maps of significant signal change when two or more conditions are contrasted. Whereas it is clear that such maps can be successfully generated, the present work indicates that the neurophysiological relevance of such maps to specific components of behavior should not be assumed.

If contrasting conditions is not the answer to the problem of identifying functionally relevant portions of the image signal, what is? A number of studies have identified brain areas where activity is correlated with parametric experimental manipulation (e.g., Grafton et al., 1992; Price et al., 1992; Dettmers et al., 1995; Blinkenberg et al., 1996; Sadato et al., 1996; Price et al., 1996). This is probably the most direct way to establish functionally significant measures from imaging data. It should be noted that as in the present study, it is not necessary to parametrically vary the task to relate intersubject variability in performance to intersubject variability in the functional signal. However, establishing a functional relationship between performance and imaging data depends on finding the relevant signals. In the present study, region selection was based on evidence of signal change across several speech-related tasks (Sidtis et al., 1999): the caudate was included not because of a large signal change in this task, but rather because of its response in a related articulatory task. The image volume can be searched for regions or pixels that correlate with performance, but this may not adequately identify functional relationships. Relationships may involve multiple regions, and decreases as well as increases in activity may be important. Further, even simple behaviors can be characterized by numerous performance measures, and other performance measures (e.g., syllable durations) are related to other regions in the functional imaging profile (Sidtis et al., 2001). At present, there is no single methodological solution to the problem. What seems apparent, though, is that current assumptions about the meaning of functional imaging signals in the standard approach are not sensitive to a wide range of possible brain-behavior relationships.

Although the present results represent only a single study, they do demonstrate one of the oldest associations in neurological localization with functional imaging data, the relationship between Broca's area and speech production. In doing so, questions are raised about the use of task contrasts to isolate behavior-specific activations, as well as about the current definition of activation. The interpretation of functional imaging data may well benefit from exploring a range of possible brain-behavior relationships broader than those embodied in the current assumptions about activation, and investigating methodological approaches that do not depend on task contrasts.

Acknowledgments

Part of this work was originally presented at the 4th International Conference on Functional Mapping of the Human Brain, June 1998. Support was provided by the Bob Allison Ataxia Research Center, by the Program in Hereditary Ataxia (PO1NS33718), and by the Human Brain Project Program in Visualization of Functional Connectivity in the Brain (P20 MH57180). The helpful comments of the reviewers, of Drs. Diana Sidtis, and Costantino Iadecola on earlier versions of this manuscript, and of Dr. Babak Ardekani on normalization in fMRI are gratefully acknowledged. The assistance of J. Anderson, D. Daly, M. Kneer, C. Farmer, D. Hamm, C. Erickson, and J. Mahowald in scanning, data collection, and processing are also gratefully appreciated.

References

Arnt, S., Cizadlo, T., O'Leary, D., Gold, S., Andreasen, N.C., 1996. Normalizing counts and cerebral blood flow intensity in functional imaging studies of the human brain. NeuroImage 3, 175–184.

Benson, D.F., 1994. The Neurology of Thinking. Oxford Univ. Press, New York.

Blinkenberg, M., Bonde, C., Holm, S., Svarer, C., Andersen, J., Paulson, O.B., Law, I., 1996. Rate dependence of regional cerebral activation during performance of a repetitive motor task: a PET study. J. Cereb. Blood Flow Metab. 16, 794–803.

Caplan, L.R., Schmahmann, J.D., Kase, C.S., Feldmann, E., Baquis, G., Greenberg, J.P., Gorelick, P.B., Helgason, C., Hier, D.B., 1990. Caudate infarcts. Arch. Neurol. 47, 133–143.

Darvas, F., Schmitt, U., Louis, A.K., Fuchs, M., Knoll, G., Buchner, H., 2001. Spatio-temporal current density reconstruction (stCDR) from EEG/MEG-data. Brain Topogr. 13, 195–207.

Dettmers, C., Fink, G.R., Lemon, R.N., Stephan, K.M., Passingham, R.E., Silbersweig, D., Holmes, A., Ridding, M.C., Brooks, D.J., Frackowiak, S.J., 1995. Relation between cerebral activity and force in the motor areas of the human brain. J. Neurophysiol. 74, 802–815.

Duffy, J.R., 1995. Motor Speech Disorders. Mosby, St. Louis.

Fox, P.T., 1991. Physiological ROI definition by image subtraction. J. Cereb. Blood Flow Metab. 11, A79–A82.

Fox, P.T., Mintun, M.A., Reiman, E.M., Raichle, M.E., 1988a. Enhanced detection of focal brain responses using intersubject averaging and change-distribution analysis of subtracted PET images. J. Cereb. Blood Flow Metab. 8, 642–653.

Fox, P.T., Raichle, M.E., 1986. Focal physiological uncoupling of cerebral blood flow and oxidative metabolism during somatosensory stimulation in human subjects. Proc. Natl. Acad. Sci. USA 83, 1140–1144.

Fox, P.T., Raichle, M.E., Mintun, M.A., Dence, C., 1988b. Nonoxidative glucose consumption during focal physiological neural activity. Science 241, 462–464.

Friston, K.J., Frith, C.D., Liddle, P.F., Frackowiak, R.S.J., 1991. Comparing functional (PET) images: the assessment of signal change. J. Cereb. Blood Flow Metab. 11, 690–699.

Friston, K.J., Frith, C.D., Passingham, R.E., Liddle, P.F., Frackowiak, R.S.J., 1992. Motor practice and neurophysiological adaptation in the cerebellum: a positron emission tomographic study. Proc. R. Soc. London. [Biol.] 248, 223–228.

Friston, K.J., Price, C.J., Fletcher, P., Moore, C., Frackowiak, R.S.J., Dolan, R.J., 1996. The trouble with cognitive subtraction. NeuroImage 4, 97–104.

Grafton, S., Mazziotta, J., Presty, S., Friston, K.J., Frakowiak, R.S.J., Phelps, M., 1992. Functional anatomy of human procedural learning determined with regional cerebral blood flow and PET. J. Neurosci. 12, 2542–2548.

Hansen, LK., Larsen, J., Nielsen, F.A., Strother, S.C., Rostrup, E., Savoy, R., Lange, N., Sidtis, J., Svarer, C., Paulson, O.B., 1999. Generalizable patterns in neuroimaging: how many principal components? NeuroImage 9, 534–544.

Horwitz, B., 1991. Functional interactions in the brain: use of correlations between metabolic rates. J. Cereb. Blood Flow Metab 11, A114–A120.

Huang, M.X., Mosher, J.C., Leahy, R.M., 1999. A sensor-weighted overlapping-sphere head model and exhaustive head model comparison for MEG. Phys. Med. Biol. 44, 423–440.

Jennings, J.M., McIntosh, A.R., Kapur, S., Tulving, E., Houle, S., 1997. Cognitive subtractions may not add up: the interaction between semantic processing and response mode. NeuroImage 5, 229–239.

Jerbi, K., Mosher, J.C., Baillet, S., Leahy, R.M., 2002. On MEG forward modeling using multipolar expansions. Phys. Med. Biol. 47, 523–555.

Kajihara, S., Tornita, S., Kondo, Y., Arakawa, A., Okamura, S., Tomita, T., Takanashi, Y., 2000. Moving mesh method for reconstructing some spread sources in the brain. Brain Topogr. 12, 283–292.

Kent, R.D., Tjaden, K., 1997. Brain functions underlying speech. in: Hardcastle, W.J., Laver, J. (Eds.), Handbook of Phonetic Sciences. Blackwell, Oxford, pp. 220–255.

Kertesz, A., 1979. Aphasia and Associated Disorders. Grune & Stratton, New York.

Kjems, U., Hansen, L.K., Anderson, J., Frutiger, S., Muley, S., Sidtis, J., Rottenberg, D., Strother, S.C., 2002. The quantitative evaluation of functional neuroimaging experiments: mutual information learning curves. NeuroImage 15, 772–786.

Levin, D.N., Uftring, S.J., 2001. Detecting brain activation in fmri data without prior knowledge of mental event timing. NeuroImage 13, 153–160.

Mazziotta, J.C., Grafton, S.T., Woods, R.C., 1991. The human motor system studied with PET measurements of cerebral blood flow: topography and motor learning, in: Lassen, N.A., Ingvar, D.H., Raichle, M.E., Friberg, L. (Eds.), Alfred Benzon Symposium 31, Brain Work and Mental Activity, Munksgaard, Copenhagen, pp. 289–290.

McIntosh, A.R., Bookstein, F.L., Haxby, J.V., Grady, C.L., 1996. Spatial pattern analysis of functional brain images using partial least squares. NeuroImage 3, 143–157.

Metter, J.E., Riege, W.H., Hanson, W.R., Phelps, M.E., Kuhl, D.E., 1988. Evidence for a caudate role in aphasia from FDG positron computed tomography. Aphasiology 2, 33–43.

Moeller, J.R., Strother, S.C., 1991. A regional Covariance Approach to the analysis of functional patterns in positron emission tomographic data. J. Cereb. Blood Flow Metab. 11, A121–A135.

Moeller, J.R., Strother, S.C., Sidtis, J.J., Rottenberg, D.A., 1987. Scaled subprofile model: a statistical approach to the analysis of functional patterns in positron emission tomographic data. J. Cereb. Blood Flow Metab. 7, 649–658.

Mohr, J.P., Pessin, M.S., Finkelstein, S., Funkenstein, H.H., Duncan, G.W., Davis, K.R., 1978. Broca aphasia: Pathological and clinical . Neurology 28, 311–324.

Newman, S.D., Tweig, D.B., Carpenter, P.A., 2001. Baseline conditions and subtractive logic in neuroimaging. Hum. Brain Mapp. 14, 228–235.

Norusis, M.J., 1988. SPSS/PC+ V2.0. SPSS Inc., Chicago.

Papanicolaou, A.C., Simos, P.G., Breier, J.I., Zouridakis, G., Wilmore, L.J., Wheeles, J.W., Constantinou, J.C., Gormley, W., Maggio, W.W., 1999. Magnetoencephalographic mapping of the language-specific cortex. J. Neurosurg. 90, 85–93.

Price, C., Moore, C.J., Frackowiak, R.S.J., 1996. The effect of varying stimulus rate and duration on brain activity during reading. NeuroImage 3, 40–52.

Price, C., Wise, R.J.S., Ramsay, S., Friston, K.J., Howard, D., Patterson, K., Frackowiak, R.S.J., 1992. Regional response differences within the human auditory cortex when listening to words. Neurosci. Lett. 146, 179–182.

Raichle, M.E., 1987. Circulatory and metabolic correlates of brain function in normal humans. in: Mountcastle, V.B., Plum, F. (Eds.), Handbook of Physiology: The Nervous System. Am. Physiol. Soc, Bethesda, MD.

Raichle, M.E., 1991. The metabolic requirements of functional activity in the human brain: a positron emission tomography study. Adv. Exp. Med. Biol. 291, 1–4.

Raichle, M.E., Fiez, J.A., Videen, T.O., MacLeod, A.M., Pardo, J.V., 1994. Practice- related changes in human brain functional anatomy during non-motor learning. Cereb. Cortex 4, 8–26.

Rottenberg, D.A., Moeller, J.R., Strother, S.C., Dhawan, V., Sergi, M.L., 1991. Effects of percent thresholding on the extraction of [18f]fluorodeoxyglucose positron emission tomographic region-of-interest data. J. Cereb. Blood Flow Metab. 11, A83–A88.

Sadato, N., Ibanez, V., Deiber, M-P., Campbell, G., Leonardo, M., Hallett, M., 1996. Frequency-dependent changes of regional cerebral blood flow during finger movements. J. Cereb. Blood Flow Metab. 16, 23–33.

Schiff, H.B., Alexander, M.P., Naesser, M.A., Galaburda, A.M., 1983. Aphemia: clinical-anatomic correlations. Arch. Neurol. 40, 720–727.

Schiffbauer, H., Berger, M.S., Ferrari, P., Freudenstein, D., Rowley, H.A., Roberts, T.P., 2002. Preoperative magnetic source imaging for brain tumor surgery: a quantitative comparison with intraoperative sensory and motor mapping. J. Neurosurg. 97, 1333–1342.

Schwartz, D.P., Badier, J.M., Vignal, J.P., Toulouse, P., Scarabin, J.M., Chauvel, P., 2003. Non-supervised spatio-temporal analysis of interictal magnetic spikes: comparison with intracerebral recordings. Clin. Neurophysiol. 114, 438–449.

Sekihara, K., Poeppel, D., Marantz, A., Koizumi, H., Miyashita, Y., 1998. Comparison of covariance-based and waveform-based subtraction methods in removing the interference from button-pressing finger movements. Brain Topogr. 11, 95–102.

Sidtis, J.J., 2000. From chronograph to functional image: What's next? Brain Cogn. 42, 75–77.

Sidtis, J.J., Anderson, J.R., Strother, S.C., Rottenberg, D.A., 2001. Establishing behavioral correlates of functional imaging signals, in: Gjedde, A., Hansen, S.B., Knudsen, G.M., Paulson, O.B. (Eds.), Physiological Imaging of the Brain with PET, Academic Press, San Diego.

Sidtis, J.J., Strother, S.C., Anderson, J.R., Rottenberg, D.A., 1999. Are brain functions really additive? NeuroImage 5, 490–496.

Sidtis, JJ., Volpe, BT., 1988. Selective loss of complex-pitch or speech discrimination after unilateral cerebral lesion. Language 34, 235–245.

Silbersweig, D.A., Stern, E., Frith, C.D., Cahill, C., Schnorr, L., Grootoonk, S., Spinks, T., Clark, J., Frackowiak, R., Jones, T., 1993. Detection of thirty-second cognitive activations in single subjects with positron emission tomography: a new low-dose $H_2{}^{15}O$ regional cerebral blood flow three-dimensional imaging technique. J. Cereb. Blood Flow Metab. 13, 617–629.

Simos, P.G., Papanicolaou, A.C., Breier, J.I., Fletcher, J.M., Wheless, J.W., Maggio, W.W., Gormley, W., Constantinou, J.E.C., Kramer, L.,

2000. Insights into brain function and neural plasticity using magnetic source imaging. J. Clin. Neurophysiol. 17, 143–162.

Strother, S.C., Anderson, J.R., Hansen, L.K., Kjems, U., Kustra, R., Sidtis, J., Frutiger, S., Muley, S., LaConte, S., Rottenberg, D.A., 2002. The quantitative evaluation of functional neuroimaging experiments: the NPAIRS data analysis framework. NeuroImage 15, 747–771.

Strother, S.C., Anderson, J.R., Schaper, K.A., Sidtis, J.J., Liow, J-S., Woods, R.P., Rottenberg, D.A., 1995. Principal component analysis and the scaled subprofile model compared to intersubject averaging and statistical parametric mapping. I. "Functional connectivity" of the human motor system studied with [^{15}O]water PET. J. Cereb. Blood Flow Metab. 15, 738–753.

Talairach, J., Tournoux, P., 1988. Co-planar Stereotaxic Atlas of the Human Brain. Thieme, New York.

Taylor, J.G., Krause, B., Shah, N.J., Horwitz, B., Mueller-Gaertner, H.-W., 2000. On the relation between brain images and brain neural networks. Hum. Brain Mapp. 9, 165–182.

Teuber, H.L., 1955. Physiological psychology. Annu. Rev. Psychol. 6, 267–296.

Tognola, G., Vignolo, L.A., 1980. Brain lesions associated with oral apraxia in stroke patients: a clinico-neuroradiological investigation with the CT scan. Neuropsychologia 18, 257–272.

Vandenberghe, R., Dupont, P., Bormans, G., Mortelmans, L., Orban, G., 1995. Blood flow in human anterior temporal cortex decreases with stimulus familiarity. NeuroImage 2, 306–313.

Vrba, J., Robinson, S.E., 2001. Signal processing in magnetoencephalography. Methods 25, 249–271.