NEUROIMAGE **8,** 249–261 (1998)
ARTICLE NO. NI980360

Reproducibility of fMRI Results across Four Institutions Using a Spatial Working Memory Task

B. J. Casey,*,[1] Jonathan D. Cohen,*,† Kathy O'Craven,‡,§ Richard J. Davidson,¶ William Irwin,¶ Charles A. Nelson,‖ Douglas C. Noll,* Xiaoping Hu,‖ Mark J. Lowe,** Bruce R. Rosen,‡ Charles L. Truwitt,‖ and Patrick A. Turski¶,[2]

**University of Pittsburgh Medical Center, 3811 O'Hara Street, Pittsburgh, Pennsylvania 15213; †Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, Pennsylvania 15213; ‡Massachusettes General Hospital, NMR Center, 13th Street, Charlestown, Massachusettes 02129; §Rowland Institute for Science, 13th Street, Charlestown, Massachusettes 02129; ¶University of Wisconsin at Madison, 1202 West Johnson Street, Madison, Wisconsin 53706; ‖University of Minnesota, 51 East River Road, Minneapolis, Minnesota 55455; and **Indiana University, CL 120, 540 Clinical Drive, Indianapolis, Indiana 46202*

Received January 5, 1997

Four U.S. sites formed a consortium to conduct a multisite study of fMRI methods. The primary purpose of this consortium was to examine the reliability and reproducibility of fMRI results. FMRI data were collected on healthy adults during performance of a spatial working memory task at four different institutions. Two sets of data from each institution were made available. First, data from two subjects were made available from each site and were processed and analyzed as a pooled data set. Second, statistical maps from five to eight subjects per site were made available. These images were aligned in stereotactic space and common regions of activation were examined to address the reproducibility of fMRI results when both image acquisition and analysis vary as a function of site. Our grouped and individual data analyses showed reliable patterns of activation in dorsolateral prefrontal cortex and posterior parietal cortex during performance of the working memory task across all four sites. This multisite study, the first of its kind using fMRI data, demonstrates highly consistent findings across sites.

INTRODUCTION

Functional magnetic resonance imaging (fMRI) has revolutionized our understanding of human brain function. Never before have we had the spatial resolution in a relatively noninvasive methodology that allows neuroscientists to examine the function of the human brain, especially the developing human brain. Yet, with the emergence of each new technological advance, considerable time and effort are required to validate and deem reliable these new techniques. Such a step is critical if we are to move this exciting methodology into the forefront of cognitive and clinical neuroscience.

A recent example of a number of institutions pooling their resources to examine the reproducibility of brain imaging results was reported by Poline *et al.* (1996). This study included positron emission tomography (PET) data from 12 European PET centers (Cologne, Copenhagen, Dusseldorf, Essen, London, Groningen, Leuven, Liege, Lyon, Milano, Orsay, and Stockholm) during performance of a verbal fluency test. Each center's data set and pooled data sets were analyzed using statistical parametric mapping and consistency across sites was reported in addition to methodological issues believed to influence sensitivity in detecting brain activation. Other similar efforts have been reported by Senda *et al.* (1993) and Watson *et al.* (1995). All of these studies examined PET data.

Accordingly, across similar lines and with similar goals, four U.S. sites formed a consortium to conduct a multi-site study of fMRI methods. The primary purpose of this consortium was not only to examine the reliability of results, but also to establish the potential and importance of this methodology with pediatric populations. The data presented here focus exclusively on the reproducibility and comparability of results from fMRI data collected on healthy adults during performance of the same behavioral paradigm at four different institutions.

The behavioral paradigm selected was a spatial/motor working memory task. This task was developed by the consortium and based on paradigms used previously by members of the Consortium to activate sensorimotor systems (e.g., motor cortex and cerebellum) as well as higher cognitive systems (e.g., prefrontal cor-

[1] To whom correspondence and reprint requests should be addressed. Fax: (412) 624-3521. E-mail: bjcasey@pitt.edu.
[2] Authors are listed in alphabetical order.

1053-8119/98 $25.00

tex) using fMRI (Cohen *et al.,* 1994; Casey *et al.,* 1995; Kim *et al.,* 1994).

This article addresses three specific issues: First, how robust are activations of sensorimotor systems and association cortices during performance of the behavioral paradigm? Second, how reproducible are results from different sites despite differences in imaging hardware, postprocessing software, and analytic strategy? Third, how do the results obtained when pooling data across sites compare with results from individual sites? To our knowledge, this is the first effort to investigate the robustness and reproducibility of fMRI results across sites.

METHODS

Data Acquisition and Collection

The fMRI data were acquired at four different sites (Carnegie Mellon University/University of Pittsburgh Medical Center, Massachusetts General Hospital/ Harvard University, University of Minnesota, and the University of Wisconsin at Madison) using 1.5-T scanners and volume head coils. However, specific hardware, software, and pulse sequences varied slightly across sites. All scanners were made by General Electric (GE) except for the scanner at Minnesota which was a Seimens Magnetom 1.5-T Vision system with echo planar imaging capabilities. The Pittsburgh and Boston sites both used Advanced NMR systems for acquisition of echo planar images, while the Madison site's 1.5-T GE Echo Speed scanner was outfitted with fast gradients to allow for whole body echo planar imaging. To restrict head motion, the Boston and Madison sites used bite bars (Madison; Mock and Irwin, 1997) and Minnesota and Pittsburgh used foam padding around the head. Although all sites acquired 7-mm slices, pulse sequences varied across sites. Boston used an asymmetric spin echo (ASE) pulse sequence (TR/TE = 5/70; Matrix = 128 × 64; FOV = 40 × 20) to acquire echo planar images, while gradient echo pulse sequences were used at Minnesota (TR/TE = 3000/60, Matrix = 128 × 128; FOV = 22), Pittsburgh (TR/TE = 6000/40, Matrix = 128 × 64; FOV = 40 × 20; NEX = 1; Flip = 90), and Madison (TR/TE = 2000/50; Matrix = 64 × 64; FOV = 24; NEX = 1; Flip = 90). All sites collected 3-D volume scans for the purposes of coregistration and localization of activity.

Two sets of data from each institution were made available to the Pittsburgh site for final processing. Data from two subjects[3] were made available from each site and were processed and analyzed as a pooled data set. This allowed us to examine patterns of brain activation when image acquisition varied across subjects, but image processing and analysis were identical. Second, statistical maps from five to eight right-handed healthy adults[4] per site were made available. These images were aligned in stereotactic space and common regions of activation were examined to address the reproducibility of fMRI results when both image acquisition and analysis vary as a function of site. We have a very strong argument for the reproducibility and comparability of our fMRI results if we observe the same general findings when the data are acquired and processed differently across sites.

Behavioral Paradigm

Subjects monitored a visual display for a dot[5] that was randomly presented in one of four adjacent box locations every 2 s (refer to Fig. 1). The stimulus duration was 500 ms and the interstimulus interval was 1500 ms. The three task conditions (visual, motor, and memory) were identical except for the instructions to the subject. In the visual condition, the subject simply maintained fixation and did not respond. In the motor condition, the subject was asked to press one of four buttons corresponding to the location of the dot for the current trial. In the memory condition, the subject was asked to press the button corresponding to the location at which the colored dot appeared two trials back. All subjects were instructed to maintain fixation on a central cross bar during performance of the task.

All subjects were right-handed and responded using their right hand and a hand-held response box with fiber-optic connections to a Macintosh computer in the control room running Psyscope software (Cohen *et al.*, 1993). At the Boston, Minnesota, and Pittsburgh sites, trials were blocked by condition, with each block lasting 60 s. The condition order was varied pseudorandomly, with the constraint that each of the three conditions occurred once in every set of three blocks. Blocks were separated by 20-s baseline intervals during which only the four adjacent boxes were presented between each condition and the subject was instructed to rest. A brief instruction screen preceded each condition. Thirty blocks were presented with 10 blocks per condition and a total of 400 trials per condition. Each site made slight modifications to the paradigm to optimize the task to site specific scanning parameters. Due to a constraint in image acquisition, the Minnesota site excluded the rest condition for the last five subjects to increase the number of acquisitions during the three conditions of memory, motor, and visual. The Madison site implemented an off-on block design, rather than the randomized block design, with three separate runs

[3] These subjects were typically the first two subjects scanned at each site.

[4] The subjects consisted of four females and four males ages 19–27 from Pittsburgh, three females and three males ages 22 to 36 from Minnesota, two females and three males ages 25 to 36 from Boston, and five males from Madison ages 20 to 43.

[5] Subjects monitored a visual display for an X at the Madison site.

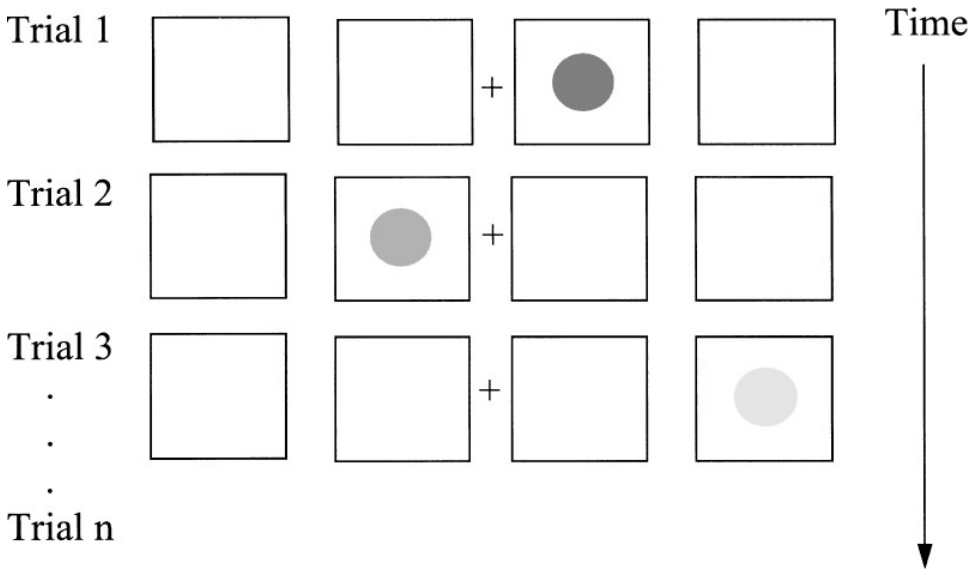

FIG. 1. Illustration of behavioral paradigm in which a dot was randomly presented in one of four adjacent box locations.

of rest–vision, vision–motor, and then motor–memory (refer to Fig. 2). Each run consisted of 11 blocks of 16 trials alternating between task conditions and beginning and ending with "off" blocks.

Group Analysis

The data from two subjects per site were first registered in 3-D space using an automated image registration algorithm (Woods *et al.*, 1992) to correct for head motion and then analyzed using a two-factor multifactorial design treating runs as one factor, thereby controlling for transients over time and condition as the second factor. An analysis of variance was performed on each slice of each individual subject's data and then thresholded for significance using a cluster-size algorithm that takes account of the spatial extent of activation in correcting for multiple comparisons (Forman *et al.*, 1995). A contiguity threshold of three contiguous pixels and $P < 0.005$ were used as criteria for significant activity corresponding to a corrected pixel-wise α protection of $P < 0.00001$ (Forman *et al.*, 1995). The resulting *F*-maps from each subject were then standardized into Talairach stereotactic coordinates (Talairach and Tournoux, 1988) by rescaling images relative to the positions of the anterior and posterior commissure and the edges of the brain (determined manually on a registered anatomical volume image) using AFNI 2.10 software (Cox, 1996). To accommodate individual differences in gyral patterns, the *F*-maps were passed through a 5-mm FWHM Gaussian

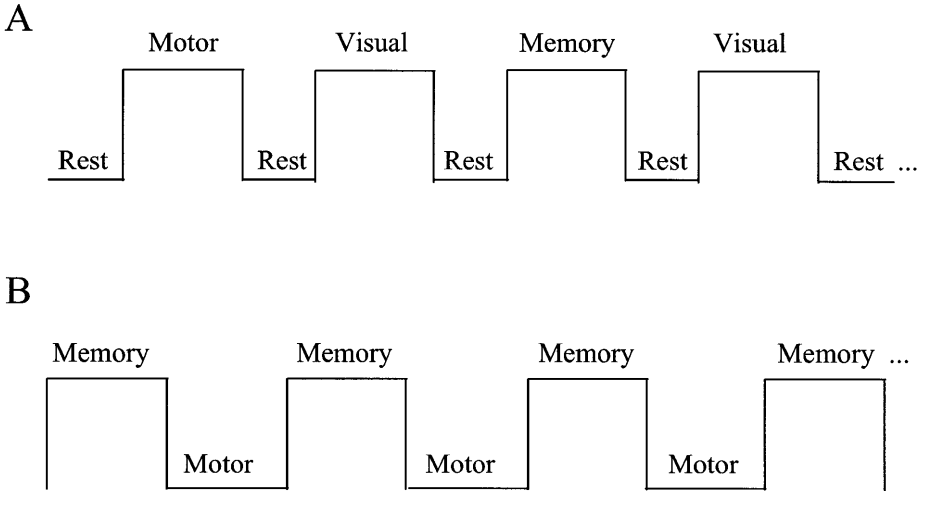

FIG. 2. Depiction of (A) pseudorandom block design and (B) off–on block design.

filter. Next, a count analysis was performed using AFNI 2.10 software to determine how many of the subjects had common areas of activation (i.e., how many subjects had significant activity in a given voxel). The results reported are those regions where at least six of the eight subjects had overlapping activity for the comparisons of Memory vs Motor and for Motor vs Visual. For the comparison of Visual vs Rest, only seven of the eight subjects had data for the rest condition; therefore results are reported for regions where at least five of the seven subjects had overlapping activity. The significant regions of activation are reported by gyri, Brodmann's areas, and Talairach coordinates.

Individual Data Sets

The independent samples from each of the four sites addresses the reproducibility of the results when both acquisition and analysis vary as a function of site. All sites' data were corrected for head motion in either 2-D (Madison and Minnesota) or 3-D (Boston and Pittsburgh) space and analyzed prior to making the data available to the Pittsburgh site. The Boston site generated Kolmogorov–Smirnov (KS) statistical maps of signal change in individual functional image voxels correlated to task performance for five adult subjects. The Minnesota and Madison sites both generated statistical maps based on voxel-wise Student *t* tests for the functional images of each of five subjects per site. The statistical maps from each of these three sites were first standardized into Talairach stereotactic coordinates (Talairach and Tournoux, 1988) by rescaling images relative to the positions of the anterior and posterior commissures and the edges of the brain (determined manually on a registered anatomical volume image) using AFNI software (Cox, 1996). The number of subjects showing significant activation at each voxel of the aligned images was computed for the data set after blurring the images with a 5-mm FWHM Gassian filter to accommodate between-subject anatomical variability. The Pittsburgh site generated statistical maps based on images that were coregistered and pooled across subjects. This was done by coregistering the structural images for each subject to a common reference brain using a 12-parameter automated algorithm (Woods *et al.*, 1992) and then smoothing the images using a 3-D Gassian filter (8 mm FWHM) to accomodate between subject differences in anatomy. *F*-maps were then generated by performing multifactorial analyses of variance on the pooled data set of eight subjects. The resulting *F*-maps were then standardized into Talairach stereotactic coordinates as described above. Statistical maps from all sites were thresholded ($P < 10^{-5}$) to adjust for multiple voxel-wise comparisons. The observed patterns of activity across sites were localized by gyri, Brodmann's area, and Talairach coordinates using the Talairach atlas (Talairach and Tournoux, 1988).

RESULTS

Group Analysis

Three analyses of variance were performed to identify regions of activation for the comparisons of Motor vs Visual conditions, Memory vs Motor conditions, and Visual vs Rest conditions. Significant regions of brain activity are presented in Table 1 for each of these three contrasts. The locations of significant activity observed in overlapping regions across subjects are reported. The number of subjects reported in Table 1 is the number of subjects with overlapping activity in these regions. The volume of activity reported is the observed extent of overlap in cubic millimeters.

Memory vs Motor Contrast

Significant activation for the contrast of the memory and motor conditions was observed in the middle, medial, and superior frontal gyri (BA 46, 10, and 9), left insular, both inferior and superior parietal lobules (BA 40 and 7), right precuneus (BA 7), and posterior cingulate gyrus (BA 31) (refer to Table 1). The most reliable regions of activity (i.e., the greatest number of subjects with overlapping activity in these regions) were observed in the right dorsolateral prefrontal cortex (BA 46) and right superior parietal lobule with 7 of 8 subjects showing overlapping activity in these regions.

Motor vs Visual Contrast

Seven of eight subjects showed overlapping activity for the contrast of motor and visual conditions in supplementary motor, premotor, and primary motor cortices as well as somatosensory cortex. This activity was predominantly lateralized to the left hemisphere and was specifically observed in middle frontal (BA 8), superior frontal (BA 6), precentral (BA 4), and postcentral (BA 3/5) gyri. Significant activity was also observed in the right cerebellum, fusiform gyrus (BA 18), and anterior cingulate gyrus (BA 32) for six of eight subjects.

Visual and Rest Contrast

For the contrast of the Visual and Rest conditions, five of seven subjects showed bilateral activity in the middle occipital gyri and cuneus (BA 18/19) and left lateralized activity in the lingual gyrus (BA 18).

Individual Data Sets

Significant regions of brain activity are presented in Tables 2–3 for each of the contrasts of Memory vs Motor conditions and Motor vs Visual conditions for each site.

TABLE 1

Location of Activation for the Three Comparisons of Memory and Motor Conditions, Motor and Visual Conditions, and Visual and Rest Conditions for the Pooled Data Set

Region of interest	Hemisphere	Brodmann's area	Volume in mm^3	No. of subjects	Talairach coordinates		
					X	*Y*	*Z*
Memory vs Motor (n = 8)							
Medial prefrontal cortex	L	10	57	6	12	57	2
	R	10	30	6	−2	58	4
Superior frontal gyrus	L	9	364	6	4	57	17
Middle frontal gyrus	L	46	51	6	38	25	35
	R	46	103	7	−44	16	41
Insular	L		19	6	35	15	7
Inferior parietal lobule	L	40	10	6	53	−27	36
Cingulate gyrus	L/R	31	61	6	0	−43	28
Superior parietal lobule	L	7	462	6	36	−46	62
	R	7	213	7	−24	−62	38
Precuneus	R	7	70	6	−15	−66	41
Motor vs Rest (n = 8)							
Superior frontal gyrus	L	8	25	6	2	31	38
	R	8	78	6	−3	30	40
Cingulate gyrus	R	32	79	6	−1	4	41
Precentral gyrus	L	4	261	7	50	−7	30
	R	4	26	6	−44	−10	46
Medial/Sup. frontal gyrus	L	6	202	7	5	−12	60
	R	6	13	6	−3	−4	53
Postcentral gyrus	L	3/5	45	6	17	−43	69
	R	3/5	117	7	−25	−37	70
Cerebellum	R		26	6	−4	−57	0
Fusiform gyrus	R	18	68	6	−28	−83	−4
Visual vs Rest (n = 7)							
Middle occipital gyrus	L	19	80	5	39	−78	8
	R	19	147	5	−25	−81	15
Cuneus	L	18/19	29	5	12	−83	26
	R	18	46	5	−6	−92	14
Lingual gyrus	L	18	71	5	4	−67	2

TABLE 2a

Location and Volume of Activation for the Comparison of Memory Minus Motor for the Boston Site (n = 5)

Region of interest	Hemisphere	Brodmann's area	Volume in mm^3	Maximum −ln (p)	Talairach coordinates		
					X	*Y*	*Z*
Middle frontal gyrus	L	46	844	10.89	32	43	13
	R	46	4316	12.78	−39	38	25
	R	8	1139	15.48	−42	18	48
	R	6	146	12.71	−29	11	55
Inferior frontal gyrus	L	45	493	11.16	42	29	17
Insular	L		343	12.34	36	18	1
Superior frontal gyrus	L/R	8	904	12.42	1	15	51
Precentral gyrus	R	4	395	14.94	−23	−27	57
Inferior parietal lobule	L	40	411	13.32	48	−29	48
	R	40	645	12.60	−46	−32	43
Superior parietal lobule	L	7	322	12.78	19	−56	54
	R	7	1412	18.16	−39	−40	60
Precuneus	L	7	3530	15.76	5	−41	56

TABLE 2b

Location and Volume of Activation for the Comparison of Memory Minus Motor for the Madison Site ($n = 5$)

Region of interest	Hemisphere	Brodmann's area	Volume in mm^3	Maximum t value	Talairach coordinates		
					X	Y	Z
Middle frontal gyrus	R	46	4672	6.26	−37	38	21
	R	6	3611	5.39	−31	3	46
	L	6	458	4.92	26	−2	46
Cingulate gyrus	R	32	1176	6.59	−2	13	42
Posterior cingulate gyrus	L	23	1258	6.66	2	−43	24
Superior parietal lobule	R	7	17376	5.92	−11	−51	45

Location and volume of activity are reported in addition to maximum statistical values. For the comparison of Visual vs Rest conditions, only three sites collected data during the rest condition and of these, only two sites had significant activity after thresholding images to account for multiple comparisons. These results are presented in Tables 4a and 4b. The behavioral data for the memory and motor conditions from each site are reported in Table 5.

Memory vs Motor Contrast

The most reliable regions of activation were observed in right dorsolateral prefrontal cortex and in the right superior parietal lobule (BA 7) similar to the findings from our group analysis. Specifically, all four sites found significant activity in the right middle frontal gyrus in Brodmann's area 46 (refer to Fig. 3a) and Brodmann's Area 6. Three of the four sites observed activation in the left insular, left middle frontal gyrus (BA 6), left or right superior frontal gyri (BA 8), and left or right posterior cingulate gyri (BA 23 and 30). Although less reliable across sites, activation was also observed in the putamen, pulvinar, precuneus, and portions of the occipital (BA 18 and 19) and temporal cortices (BA 20, 21, 37, and 39).

Motor vs Visual Contrast

The most reliable regions of significant activation for the contrast of motor and visual conditions were observed in the left supplementary motor and premotor cortex (BA 6), left primary motor (refer to Fig. 3b), left somatosensory cortex (BA 1, 2, and 5), and right

TABLE 2c

Location and Volume of Activation for the Comparison of Memory Minus Motor for the Minnesota Site ($n = 6$)

Region of interest	Hemisphere	Brodmann's area	Volume in mm^3	Maximum t value	Talairach coordinates		
					X	Y	Z
Medial prefrontal cortex	L	10	126	9.36	7	63	15
Middle frontal gyrus	R	46	266	9.88	−24	36	33
Superior frontal gyrus	L	8	284	16.64	23	31	47
Orbital frontal gyrus	R	11	185	9.10	−9	28	−23
Middle frontal gyrus	L	6/8	300	9.88	23	19	34
	R	6/8	475	9.88	−25	19	41
Putamen	R		148	7.54	−17	−10	−5
Insular	L		194	13.26	39	−13	14
Fusiform gyrus	L	20	212	7.54	47	−20	−34
Precentral gyrus	L	4	127	7.80	7	−26	69
Middle temporal gyrus	L	21	549	17.16	63	−26	−16
Cingulate gyrus	R	23	130	9.62	−1	−33	29
Inferior temporal gyrus	L	20	390	12.22	47	−34	−18
Pulvinar	R		116	9.62	−17	−34	6
Postcentral gyrus	L	5	101	6.24	23	−39	76
Hippocampal gyrus	L	30	102	9.36	15	−40	3
Superior parietal lobule	R	7	200	7.24	−10	−47	62
Superior temporal gyrus	R	39	124	7.54	−49	−60	29
Cerebellum	L		991	14.56	30	−60	−35
	R		190	10.40	−17	−52	−15
Cuneus	L	19	203	8.84	7	−74	32
Lingual gyrus	R	18	371	12.22	−17	−84	15

TABLE 2d

Location and Volume of Activation for the Comparison of Memory Minus Motor Conditions for the Pittsburgh Site ($n = 8$)

Region of interest	Hemisphere	Brodmann's area	Volume in mm^3	Maximum F ratio	Talairach coordinates		
					X	Y	Z
Medial prefrontal cortex	L	10	477	52.20	3	50	0
Superior frontal gyrus	R	8	536	70.47	−10	40	50
	L	6	186	21.75	9	24	61
	R	6	273	39.15	−29	−5	49
Middle frontal gyrus	L	6	341	40.02	32	15	56
	R	46	382	67.87	−36	38	27
Insular	R		173	33.93	−29	10	8
	L		192	21.75	42	6	5
Precentral gyrus	R	4	192	43.50	−30	−6	43
Precuneus	L	7	1229	53.94	7	−52	58
	R	7	212	20.88	−12	−48	66
Posterior cingulate gyrus	R	30	114	33.06	−4	−50	10
Inferior parietal lobule	R	40	392	51.33	−44	−53	35
	L	40	117	26.10	57	−51	34
Superior parietal lobule	L	7	555	30.45	36	−58	59
	R	7	1404	25.23	−25	−60	54
Middle temporal gyrus	R	21/37	145	28.71	−55	−52	−2
Superior occipital gyrus	L	19	187	33.06	29	−76	32
Middle occipital gyrus	L	19	434	31.32	36	−78	11
Cuneus	L	18	413	23.49	2	−88	19
Lingual gyrus	R	18	196	24.36	−25	−86	−15

cerebellum. At least three of the four sites found activation in the right anterior cingulate gyrus (BA 24 and 32) and fusiform gyri (BA 37). Less reliable activity was observed in the striatum, thalamus, ventral, and medial prefrontal cortices (BA 11 and 10) and temporal cortices (BA 20, 21, and 41).

Visual vs Rest Contrast

As noted above, only two sites observed significant activity which was localized in occipital cortex, specifically in the cuneus and lingual gyri (BA 18 and 19).

DISCUSSION

Overall Findings

Our results revealed reliable patterns of activation across all three task conditions, with the most reliable patterns of activity observed for the memory and motor task conditions. During performance of the spatial working memory task, activity was observed in the right dorsolateral prefrontal cortex and the right superior parietal lobule individually across the four sites as well as in the group analysis. The pattern of activity in

TABLE 3a

Location and Volume of Activation for the Comparison of Motor Minus Visual for the Boston Site ($n = 5$)

Region of interest	Hemisphere	Brodmann's area	Volume in mm^3	Maximum $-\ln(p)$	Talairach coordinates		
					X	Y	Z
Superior frontal gyrus	L	6	8435	16.34	12	3	58
	R	8	230	13.3	−29	24	63
	R	6	1907	15.39	−21	16	52
Anterior cingulate gyrus	R	32	44	10.07	−19	13	40
Precentral gyrus	L	4	1452	16.72	22	−12	58
Postcentral gyrus	L	2	173	10.45	33	−34	60
Cerebellum	L		330	10.83	13	−54	−7
	R		1032	12.16	−34	−68	−23
Fusiform gyrus	R	37	6496	19.34	−38	−53	1
Lingual gyrus	L	18	2364	14.06	2	−69	6
	R	18	252	10.83	−22	−74	−12

TABLE 3b

Location and Volume of Activation for the Comparison of Motor Minus Visual for the Madison Site ($n = 5$)

Region of interest	Hemisphere	Brodmann's area	Volume in mm^3	Maximum *t* value	Talairach coordinates		
					X	*Y*	*Z*
Superior frontal gyrus	L	6	301	7.59	18	−2	62
Medial sup. frontal gyrus	L	6	700	5.74	2	−7	57
Precentral gyrus	L	4	24020	10.26	37	−13	52
Postcentral gyrus	R	1	3144	7.59	−43	−23	37
Superior parietal lobule	R	7	643	5.95	−37	−42	53
Cerebellum	R		340	7.18	−20	−48	−10
Fusiform gyrus	L	37	271	7.18	26	−49	−11

prefrontal cortex was similar to that observed previously by Cohen *et al.* (1994, 1997) and Casey *et al.* (1995) using a nonspatial working memory task with fMRI. The observed activity in dorsolateral prefrontal cortex and posterior parietal cortex is also consistent with nonhuman primate neurophysiological studies (Fuster, 1989; Goldman-Rakic, 1987) and both PET and fMRI studies (Courtney *et al.*, 1996; Feiz *et al.*, 1996; Jonïdes *et al.*, 1993; McCarthy *et al.*, 1994; Petrides *et al.*, 1993; Swartz *et al.*, 1995). Furthermore, the observed right lateralization of activity is consistent with previous studies of working memory for spatial information (Smith *et al.*, 1996). Other regions of activity observed in the group analysis as well as across several of the sites individually included left motor areas (SMA and premotor cortex) and the left insula. Activity in both of these regions has been observed in previous imaging studies using very similar nonspatial working memory tasks (Awh *et al.*, 1996; Braver *et al.*, 1997). An interpretation of this result put forth by Braver *et al.* is that the observed activity in motor regions may reflect motor readiness. Since the memory task condition is likely to demand more processing time, this leaves less time for response preparation. The system may adapt by maintaining a higher level of motor readiness, producing a tonic increase in activity within motor cortex. While the observed left insular activity has been observed in working memory tasks it has also been observed in tasks requiring retrieval of information from long-term memory (Buckner and Petersen, 1996). Both situations require a comparison of internally stored representations with externally presented stimuli. Thus, the insula may be involved in matching these two types of information.

The most reliable regions of activation for the motor task were observed in the left motor areas (supplementary motor, premotor, primary motor cortex), left somatosensory cortex, and right cerebellum. The activity observed in motor areas reflect the subjects' preparation, execution, and control of purposeful movements as they made a right-handed button press that corresponded to the dot location. The fact that this activity was lateralized predominantly to the left hemisphere in motor cortex and the right hemisphere in the cerebellum is consistent with cortical organization as contralateral and cerebellar organization as ipsilateral to the limb engaged in the movement. The observed cerebellar activity replicates previous studies implicating this

TABLE 3c

Location and Volume of Activation for the Comparison of Motor Minus Visual for the Minnesota Site ($n = 6$)

Region of interest	Hemisphere	Brodmann's area	Volume in mm^3	Maximum *t* value	Talairach coordinates		
					X	*Y*	*Z*
Superior frontal gyrus	L	9	333	14.26	35	51	37
Orbital frontal gyrus	R	11	262	15.19	−25	25	−17
Cingulate gyrus	R/L	24	345	26.35	−1	25	17
Inferior frontal gyrus	L	44	112	15.50	62	13	4
Middle frontal gyrus	L	8	152	21.70	55	11	42
	L	6	198	22.60	39	9	50
Precentral gyrus	L	4	593	24.80	23	−14	63
Middle temporal gyrus	L	21	106	13.64	58	−17	−8
Postcentral gyrus	L	5/7	124	15.50	39	−35	61
Fusiform gyrus	L	37	103	20.77	31	−55	−7
Cerebellum	L		119	14.26	63	−70	−35
	R		188	19.53	−17	−55	−25

TABLE 3d

Location and Volume of Activation for the Comparison of Motor Minus Visual for the Pittsburgh Site ($n = 8$)

Region of interest	Hemisphere	Brodmann's area	Volume in mm^3	Maximum F value	Talairach coordinates		
					X	Y	Z
Medial prefrontal cortex	R	10	396	22.04	−1	47	1
Superior frontal gyrus	L	8	366	29.64	24	37	17
Orbital frontal gyrus	R	11	244	36.48	−16	26	−8
Caudate nucleus	R		579	19.76	−8	22	1
Superior temporal gyrus	L	38	108	31.16	23	14	−37
Inferior frontal	L	45	982	27.20	43	14	10
	R	45	988	41.04	−38	15	3
Cingulate gyrus	L	24	1135	34.20	9	−1	42
	R	24	339	21.28	−8	7	24
Precentral gyrus	L	4	2302	76.80	33	−13	52
	R	4	678	27.36	−49	−13	33
Putamen	L		2179	46.36	27	−4	9
	R		845	21.28	−22	−5	1
Thalamus	R		400	37.24	−2	−4	11
Superior temporal gyrus	R	41	309	27.34	−46	−23	7
Inferior parietal lobule	R	40	653	28.88	−49	−29	32
Postcentral gyrus	L	5	4440	55.48	42	−29	57
	R	5	222	35.72	−22	−35	73
Inferior temporal gyrus	R	20	312	22.80	−61	−36	−5
Middle temporal gyrus	L	21	298	24.32	38	−38	4
Superior parietal lobule	L	7	302	51.85	39	−44	50
Inferior parietal lobule	L	40	351	23.56	57	−52	31
	R	40	223	34.96	−32	−58	41
Cerebellum	L		209	25.08	20	−38	−41
	R		203	21.28	−42	−54	−35
	L/R		745	25.08	−1	−53	−25
	L		226	23.56	40	−62	−33
Precuneus	L	7	1084	34.96	12	−71	36
Cuneus	R	18/19	249	20.52	−10	−78	31

region in the sequencing and timing of movements (Kim *et al.*, 1994). The activity observed in the left somatosensory cortex most likely reflects the sensation of pressure as the subjects repetitively pressed the buttons to respond during the motor task relative to resting their fingers continuously on the buttons during the visual task.

During performance of the visual task, activity was observed in Brodmann's areas 18 and 19 of the occipital cortex, although this pattern of activity was not observed across all four sites. The lack of robustness of visual cortical activity may be due to the combination of the slow rate of stimulus presentation (once every 2 s) and the wide visual angle of the stimulus display (8°). The stimulus presentation rate was intentionally held constant across task conditions and the selection of the slower stimulus presentation rate was necessary for the subjects to be able to perform the memory task condition. Alternatively, the lack of activity in visual cortex may be due to the similarity in the visual stimulus during the Rest and Visual conditions. The boxes and fixation appeared throughout each condition and only the dot was added in the Visual condition.

In sum, our results suggest that the paradigm used in this study is a fairly robust behavioral probe for activating both sensorimotor and association cortices.

TABLE 4a

Location and Volume of Activation for the Comparison of Visual Minus Rest for the Boston Site ($n = 5$)

Region of interest	Hemisphere	Brodmann's area	Volume in mm^3	Maximum $-\ln(p)$	Talairach coordinates		
					X	Y	Z
Fusiform gyrus	R	19	429	9.19	−36	−55	5
Lingual gyrus	R	18	157	9.16	−5	−70	9
Cuneus	R	19	113	9.36	−11	−74	33

TABLE 4b

Location and Volume of Activation for the Comparison of Visual Minus Rest for the Pittsburgh Site ($n = 8$)

Region of interest	Hemisphere	Brodmann's area	Volume in mm^3	Maximum F value	Talairach coordinates		
					X	Y	Z
Cingulate gyrus	L	24	148	30.21	4	−11	43
Postcentral gyrus	L	7	272	48.76	16	−45	73
Cingulate gyrus	L	31	258	43.46	2	−51	42
Middle occipital gyrus	L	19	152	45.58	39	−80	10
Cuneus	L	18/19	102	36.57	7	−88	31
	R	18/19	249	39.22	−10	−90	34
Lingual gyrus	L	18	96	21.20	13	−96	−2

The design of the behavioral task is hierarchical with each condition adding well specified cognitive demands to the simpler one, while holding other factors constant. Thus, each condition acts as a control for the next more complex one, permitting within-subject comparisons of all conditions.

Reproducibility of the Results

While the observed patterns of activation observed during the three task conditions are interesting, the primary objective of this study was to examine the robustness and reproducibility of results across sites. The similarity in location and distribution of activity is shown in Fig. 3 revealing almost exact matches in stereotactic coordinates for activation of right dorsolateral prefrontal cortex and left primary motor cortex across the four sites. In sum, even when different image acquisitions and analytic tools were used we observed the same general findings across sites, and in our group analysis. This provides strong evidence for the reproducibility, reliability, and comparability of our fMRI results.

Differences across Sites

While the similarities observed across sites are impressive, differences were also observed. These differences are most clearly shown in Tables 2–4. For example, one obvious difference is fewer, although typically larger, regions of activity for some sites relative to others (refer also to Fig. 3). This difference may be due to differences in contrast to noise across the various MR centers, number of subjects, number of acquisitions, data processing, differences in motion artifact across sites given the use of bite bars at some and foam padding at others, or even individual differences in subject's behavioral performance. It did appear to be the case that those sites with a large number of active regions did not use a bite bar, but rather relied on foam padding around the head. Thus, activity associated with motion may account for some of the differences observed in number of activated regions (Hajnal *et al.*, 1994). Another case of differences across sites was the observation of lateralized activity at some sites but not others. For example, two of the four sites reported only right dorsolateral prefrontal activity (BA 46) during the memory task condition, while the other two sites reported bilateral dorsolateral prefrontal activity with less overall left dorsolateral activity. This pattern of activity could be explained by differences across sites in image acquisition, processing, or analysis. One possible factor may be the procedures used by different investigators to determine the threshold for considering brain activity to be significant. In the current study, statistical maps were corrected for multiple comparisons in the same way across sites. By increasing or lowering

TABLE 5

Mean Reaction Times and Mean Accuracy for the Memory and Motor Conditions for Each Site[a]

Site	No. of subjects	Memory task		Motor task	
		Reaction time	Accuracy	Reaction time	Accuracy
Pittsburgh	8	399.75 (63.36)	0.95 (0.33)	562.50 (113.42)	0.98 (0.15)
Boston	5	435.89 (213.17)	0.88 (0.37)	485.63 (128.92)	0.90 (0.21)
Minnesota	6	429.27 (190.71)	0.86 (0.24)	542.98 (76.89)	0.97 (0.04)

[a] Due to technical problems, the behavioral data from the Madison site is not available. However, four of their five subjects were given practice on the task in a simulated MRI environment until they reached a 70–90% accuracy rate prior to scanning.

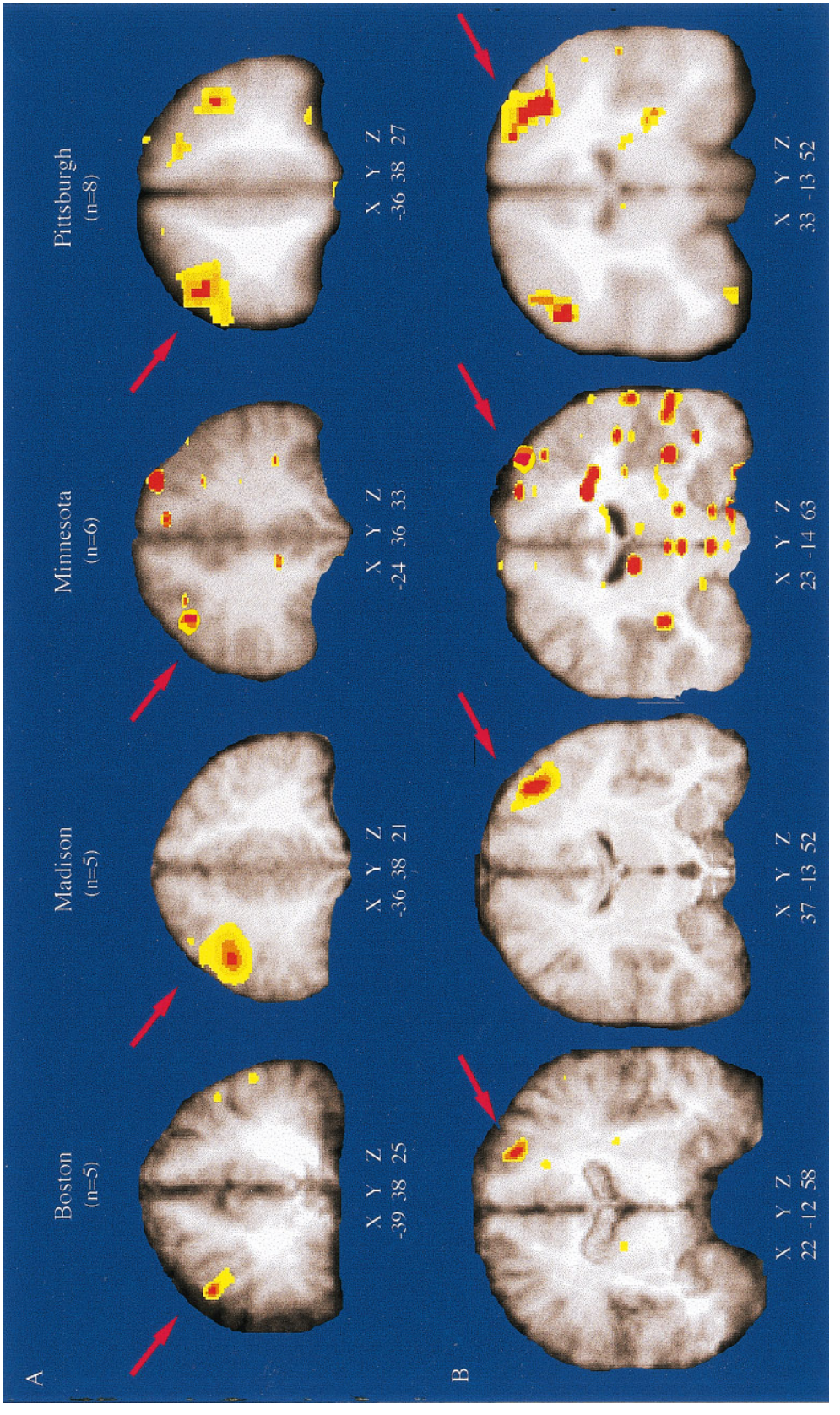

FIG. 3. (A) Observed activation within dorsolateral prefrontal cortex across all four sites for the comparison of Memory vs Motor conditions. (B) Observed activation within primary motor cortex across all four sites for the comparison of Motor vs Visual conditions.

that threshold, the laterality effects may change, especially given the different statistical tests applied across sites. Similar conflicting results have been reported in the imaging literature using very similar working memory paradigms, with some studies reporting highly lateralized prefrontal activity (Smith *et al.*, 1996; D'Esposito, *et al.*, 1997) and others reporting primarily bilateral activity (Braver *et al.*, 1997).

Limitations

There are limitations in the current study. First, although a strong effort was made to standardize the design of the experiment and the data processing from site to site, there were a number of variations across sites. While the many descrepancies in collecting and analyzing images across the four sites make the reliability all the more impressive, they make any interpretation of the differences impossible. Two main sources of variability were mixed in the present study. These include the image processing and statistical analysis as well as the data acquisition and differences between groups of subjects. Whether the variability between the different sites is due to acquisitions or due in part to the analysis is unclear. Even for the most robust difference of Motor versus Visual conditions, several anatomic regions of interest were observed across the centers (refer to Tables 3a–3d) in addition to the nine regions found in at least three of the four sites. These results suggest that there is either a sensitivity problem in the ability to detect true activity across sites or a specificity problem in the inclusion of false positives from some sites. Further, the type of statistical analyses performed (e.g., mixed and random effects model versus fixed effect models) may call into question the generalizability of the current findings. However, the reproducibility of patterns of activation across sites is nonetheless impressive.

CONCLUSIONS

This multisite study, the first of its kind using fMRI data, demonstrates highly consistent findings across sites. In conjunction with other, even larger, multicenter studies (e.g., Poline *et al.*, 1996), it demonstrates the beginnings of a new era in collaborative scientific research among institutions due to the ease and accessibility of information and data over the internet. Finally, an important goal for our consortium was to establish the hallmark for collaborative studies using this relatively noninvasive methodology with pediatric populations. Although the data that are presented in this paper focus exclusively on the reproducibility and comparability of results from fMRI data collected on healthy adults, the preliminary findings from a parallel study in children are yielding very similar results (Casey *et al.*, 1997). Clearly, fMRI promises a whole new era in cognitive neuroscience.

ACKNOWLEDGMENTS

The authors thank and acknowledge other members of the Consortium: Aaron Berkowitz, Peter Franzen, Steven King, Leigh Nystrom, Jennifer Orendi, Rolf Trainor, and Anne Schubert at the University of Pittsburgh Medical Center; Aiping Jiang, Rodney Look, and Robert Savoy at Rowland Institute for Science and Massachusetts General Hospital; Leslie Carver, Joseph Linn, Chris Monk, Kathleen Thomas, Kamil Ugurbil at the University of Minnesota; and Wil Irwin, Bryan Mock, Jennifer Orendi, and James Sorenson at the University of Wisconsin at Madison. This work was supported by the John D. and Catherine T. MacArthur Foundation.

REFERENCES

Awh, E., Jonides, J., Smith, E. E., Schumacher, E. H., Koeppe, R., and Katz, S. 1996. Dissociation of storage and rehearsal in verbal working memory: Evidence from PET. *Psychol. Sci.* **7:**25–31.

Braver, T. S., Cohen, J. D., Nystrom, L. E., Jonides, J., Smith, E. E., and Noll, D. C. 1997. A parametric study of prefrontal cortex involvement in human working memory. *Neuroimage* **5:**49–62.

Buckner, R. L., and Petersen, S. E. 1996. What does neuroimaging tell us about the role of prefrontal cortex in memory retrieval? *Semin. Neurosci.* **8:**47–55.

Casey, B. J., Cohen, J. D., Jezzard, P., Turner, R., Noll, D. C., Trainor, R. J., Giedd, J., Kaysen, D., Hertz-Pannier, L., and Rapoport, J. L. 1995. Activation of prefrontal cortex in Children during a nonspatial working memory task with functional MRI. *Neuroimage* **2:**221–229.

Casey, B. J., Cohen, J. D., King, S. W., Franzen, P. L., Nystrom, L. E., Badgaiyan, R. D., Schubert, A. B., and Noll, D. C. 1997. A developmental fMRI study of cortical activation during a spatial working memory task. *Proceedings of the International Conference on Functional Mapping of the Human Brain.* (abstract)

Cohen, J. D., Forman, S. D., Braver, T. S., Casey, B. J., Servan-Schreiber, D., and Noll, D. C. 1994. Activation of the prefrontal cortex in a nonspatial working memory task with functional MRI. *Hum. Brain Mapp.* **1:**293–304.

Cohen, J. D., MacWhinney, B., Flatt, M. R., and Provost, J. 1993. Psyscope: A new graphic interactive environment for designing psychology experiments. *Behav. Res. Methods Instrum. Comput.* **25:**257–271.

Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., and Smith, E. E. 1997. Temporal dynamics of brain activation during a working memory task. *Nature* **386:**604–608.

Courtney, S. M., Ungerleider, L. G., Keil, K., and Haxby, J. V. 1996. Object and spatial visual working memory activate separate neural systems in human cortex. *Cereb. Cortex* **6:**39–49.

Cox, R. W. 1996. AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* **29:**162–173.

D'Esposito, M., Aguirre, G. K., Zarahn, E., and Thompson, C. 1997. Activation of identical prefrontal regions by delay and non-delay components of a working memory task: A fMRI study. *Abst. Soc. Neurosci.* **22:**968.

Fiez, J. A., Raife, E. A., Balota, D. A., Schwarz, J. P., Raichle, M. E., and Peterson, S. E. 1996. A positron emission tomography study of the short-term maintenance of verbal information. *J. Neurosci.* **16:**808–822.

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun,

M. A., and Noll, D. C. 1995. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. *Magn. Reson. Med.* **33:**636–647.

Fuster, J. M. 1989. *The Prefrontal Cortex, Anatomy, Physiology, and Neuropsychology.* Raven Press, New York.

Goldman-Rakic, P. S. 1987. Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. *Handb. Physiol. Nerv. Syst.* **5:**373–417.

Hajnal, J. V., Myers, R., Oatridge, A., Schwieso, J. E., Young, I. R., and Bydder, G. M. 1994. Artifacts due to stimulus correlated motion in functional imaging of the brain. *Magn. Reson. Med.* **31:**(3)283–291.

Jonides, J., Smith, E. E., Koeppe, R. A., Awh, E., Minoshima, S., and Mintun, M. A. 1993. Spatial working memory in humans as revealed by PET. *Nature* **363:**623–625.

Kim, S. G., Ugurbil, K., and Strick, P. L. 1994. Activation of a cerebellar output nucleus during cognitive processing. *Science* **265:**949–951.

McCarthy, G., Blamine, A. M., Puce, A., Nobre, A. C., Bloch, G., Hyder, F., Goldman-Rakic, P., and Shulman, R. G. 1994. Functional magnetic resonance imaging of human prefrontal cortex during a spatial working memory task. *Proc. Natl. Acad. Sci. USA* **91:**8690–8694.

Mock, B. J., and Irwin, W. 1997. *Head Restarint Method and Apparatus for Use in MRI.* Wisconsin Alumni Research Foundation. [Patent pending]

Petrides, M. E., Alivisatos, B., Meyer, E., and Evans, A. C. 1993. Functional activation of the human frontal cortex during the performance of verbal working memory tasks. *Proc. Natl. Acad. Sci. USA* **90:**878–882.

Poline, J. B., Vandenberghe, R., Holmes, A. P., Friston, K. J., and Frackowiak, R. S. J. 1996. Reproducibility of PET activation studies: Lessons from a multi-center European experiment. *Neuroimage* **4:**34–54.

Senda, M., Kanno, I., Yonekura, Y., Fujita, H., Ishii, K., Lyshkow, H., Miura, S., Oda, K., Sadato, N., and Toyama, H. 1993. Comparison of three anatomical standardization methods regarding foci localization and its between subject variation in the sensorimotor activation. In *Quantification of Brain Functions: Tracer Kinetics and Image Analysis in Brain PET* (K. Uemura, N. A. Lassen, T. Jones, and I. Kanno, Eds.), pp. 439–445. Excerpta Medica, Amsterdam.

Smith, E. E., Jonides, J., and Koeppe, R. A. 1996. Dissociating verbal and spatial working memory using PET. *Cereb. Cortex* **6:**11–20.

Swartz, B. E., Halgren, E., Fuster, J. M., Simpkins, E., Gee, M., and Mandelkern, M. 1995. Cortical metabolic activation in humans during a visual memory task. *Cereb. Cortex* **5:**205–214.

Talairach, J., and Tournoux, P. 1988. *Co-planar stereotaxic atlas of the human brain,* Theime, New York.

Watson, J. D. G., Coltheart, M., O'Keefe, G. J., O'Sullivan, B. T., Ehan, G. F., Tochon-Danguy, H. J., Barrett, N. A., Large, M., Bierlangieri, S. U., and Miekle, S. R. 1995. *17th International Symposium on Cerebral Blood Flow and Metabolism,* Cologne, Germany, 2–6 July, p. S51.

Woods, R. P., Cherry, S. R., and Mazziotta, J. C. 1992. Rapid automated algorithm for aligning and reslicing PET images. *J. Comput. Assisted Tomogr.* **16:**620–633.