Brain Metabolism Patterns Are Sensitive to Attentional Effort Associated with a Tone Recognition Task

Henry H. Holcomb, Barry Gordon, Harry L. Loats, Edward Gastineau, Zuo Zhao, Deborah Medoff, Robert F. Dannals, Roger Woods, and Carol A. Tamminga

Using positron emission tomography with the tracer 18-fluoro-D-deoxyglucose, we assessed regional cerebral glucose utilization patterns (rCMRglu) associated with three performance levels in a forced choice, tone recognition task. Four normal subjects responded with one hand when they heard a high-frequency tone (1500 Hz), and with the other hand when they recognized a low-frequency tone (750 Hz). The EASY (EAS) condition accuracy average was 96%, the INTERMEDIATE level accuracy averaged 89%, and the DIFFICULT (DIF) recognition task accuracy average was 77%. Statistical parametric mapping (SPM94) analysis revealed that the DIF minus EAS contrast is associated with a marked metabolic elevation in the right middle and inferior temporal gyri and the gyrus fusiformis. The EAS minus DIF contrast revealed greater rCMRglu in the right medial geniculate body. Enhanced activity in right temporal lobe structures may reflect a role in auditory memory and "image" visualization. The medial geniculate enhancement may reflect tone frequency assessment.

Key Words: Tone recognition, cortex, thalamus, FDG, attention, positron emission tomography

BIOL PSYCHIATRY 1996;39:1013–1022

Introduction

The neural mechanisms associated with different aspects of task performance, such as accuracy, skill, practice, attention, and fatigue, are currently the subject of intense investigation. Among the problems complicating these investigations is their interaction in a complex fashion (Merzenich et al 1988). For example, attention and fatigue can be expected to vary in the course of a repetitive task lasting for 20–30 min, yet skill and practice can also be expected to vary. Diminished novelty will modify the neural response to various stimuli (Raichle et al 1994). Neural response, whether measured by single-unit electrode methods or positron emission tomography (PET), is sensitive to these interactions. Spitzer (1988, 1991), Moran (1985), and Richmond (1987) have demonstrated that

From the Maryland Psychiatric Research Center, University of Maryland School of Medicine, Department of Psychiatry, Baltimore, Maryland (HHH, EG, ZZ, DM, CAT); Johns Hopkins Medical Institutes, Departments of Radiology, Nuclear Medicine Division (HHH, RFD), and Neurology (BG), Baltimore, Maryland; Loats Associates, Westminster, Maryland (HLL); and Department of Neurology, University of California Los Angeles, Los Angeles, California (RW).

Address reprint requests to Henry H. Holcomb, MD, Maryland Psychiatric Research Center, P.O. Box 21247, Baltimore, MD 21228.

Received July 19, 1993; revised April 27, 1995.

0006-3223/96/$15.00
SSDI 0006-3223(95)00303-X

neuronal responsiveness in the visual system varies with the amount of attention or cognitive effort required of the subject. Using visual and somatosensory stimuli, Pardo (1991) identified a multimodal "attention" system in human subjects, using local blood flow as a measure of neural activity. Similarly, Meyer (1991) used blood flow measures to demonstrate that mental distraction reduces physiological activity in the somatosensory cortex of volunteers receiving vibrotactile stimulation. He also found that focused attention enhanced blood flow to that area when the stimulation was yoked to a perceptual task. Blood flow patterns have been shown to reflect changes in visual attention by Corbetta (1990, 1991, 1993). He and his colleagues have demonstrated that attention to a particular feature in a field significantly changes the subject's cerebral blood flow pattern. Stimulus relevance is defined by the subject's attention to the stimulus; the extent a subject focuses on a particular stimulus will have a significant impact on his neurophysiological activity patterns. Investigators have used several methods to determine whether a subject finds a given stimulus relevant or irrelevant. By making a stimulus ambiguous and by promoting uncertainty in the subject's decision, the investigator increases the relevance of the stimulus.

Stimulus degradation can be used to decrease perceptual discriminability, increase performance difficulty, and enhance decision uncertainty. Nuechterlein (1983), Nestor (1991), and Buchsbaum (1990) have used degraded visual stimuli to assess visual discrimination accuracy in schizophrenic and normal subjects. Spitzer (1988, 1991) used graded task difficulty to study primate visual discriminability and neuronal activity patterns. She and her colleagues demonstrated a marked enhancement of attention on behavioral performance and neuronal firing activity. In these studies increased attentional effort appears to be associated with enhanced behavioral responsiveness and sharpened selectivity of the neurons subserving the task. In some instances attention to a distracter is also associated with selective inhibition of particular neurons (Moran and Desimone 1985). Most studies support the generalization that as a task becomes simple and the stimuli become "less relevant" neuronal populations that once exhibited high activity in response to a particular task tend to become less active. Stimulus "irrelevance" is similar to stimulus "familiarity" in some respects. To the extent that a stimulus requires little or no "effort" from the observer and to the extent that the observer has had numerous previous experiences with a particular stimulus, then the stimulus becomes familiar and less relevant than one that is novel and commanding high effort (Raichle et al 1994; Tulving et al 1994). Tulving (1994) and Raichle (1994) have recently provided important studies that illustrate how some brain regions become less, and others become more, "active" when a subject is exposed to familiar versus novel stimuli. These are especially relevant to the study presented here. They emphasize that attention to familiar, irrelevant, or low-demand stimuli will be associated with specific blood flow patterns, different from those observed with novel, relevant, and high-demand stimuli.

Electrophysiological and blood flow studies of attention have used tasks lasting 2–120 sec. PET studies using the tracer 18-fluoro-D-deoxyglucose (FDG) are constrained by the requirement for a prolonged (20–30 min) physiological steady state. Subjects able to perform a sensorimotor task tend to become inattentive, bored, or fatigued when required to participate for 20–30 min, especially if the task is easy. In this study we have compared three auditory recognition conditions. One required a high level of effort (DIF); a second recognition task was easy, but the subjects had not practiced it (INT). A third condition was extremely easy (EAS), owing both to the absence of background noise and the extensive practice provided each subject. These contrasts, made in the context of PET/FDG, permitted us to assess the impact of task difficulty and performance on activity patterns driven by auditory recognition.

Methods

Experimental Design

The basic design was done within subjects, using three task conditions INTERMEDIATE (INT), DIFFICULT (DIF), and EASY (EAS). The DIF and INT conditions were presented first, 1–2 weeks apart; the EAS condition was studied 3–6 months after the first two studies. Three subjects received INT first, and one received DIF first. The hiatus between the first two and the third study reflected our need to wait for a new calendar year in order to administer additional radioactivity to our subjects. Uptake of FDG during the performance of the tasks was measured using PET.

Subjects

The subjects were four normal volunteers, two men and two women. Their ages ranged from 25 to 35. All were right-handed. Three had completed four years of college; one had completed two years of college. They did not have any personal or family history of psychiatric illness or drug abuse as determined by a structured interview (Structured Clinical Interview for DSM-III-R). They were paid for their participation. Informed consent was obtained. The protocol was reviewed and approved by the Johns Hopkins Joint Commission on Clinical Investigation and the University of Maryland Institutional Review Board.

Behavioral Task during FDG Uptake

The basic task was tone recognition of two different frequencies, low (750 Hz) and high (1500 Hz). The tones were presented sequentially, binaurally, through headphones (Realistic Nova 34). Subjects' eyes were covered and they sat upright in a cushioned chair in a quiet room. The task was to press a microswitch held in one hand in response to a low-pitched tone, and in the other hand in response to a high-pitched tone (pitch to hand association was randomized). Subjects were asked to respond as quickly and as accurately as possible. Tone presentation and response recording (both the response and the response time) were registered by a computer (IBM XT), which also generated the tones. Each tone lasted 100 msec. Tones were not followed by a subsequent tone until the subject either made a choice or 2.5 sec elapsed. The intertrial interval was 500 msec. In general, subjects were able to perform 700–1000 trials during a 20-min block of time. Tone recognition began 1 min prior to FDG infusion. Subjects rested quietly in the sitting position until the completion of a 30-min uptake phase was completed. Immediately thereafter subjects urinated and walked to the PET scanner. The three conditions of the experiment were defined by the amount of pseudo-white noise added to the task and the target accuracy level predetermined by the investigator. For the DIF condition the tones were presented at a 65-dB level but mixed with broad band noise of 80–85 dB. In the INT condition tones presented at 65 dB in the presence of broad band noise of 60–70 dB. For the EAS condition, the tones were presented at 65 dB, and mixed with background noise of 50 dB. Presentation levels were measured at the earphone–ear interface. The noise was modified for maintenance of a constant performance level, as described below.

To maintain a constant accuracy over the course of the experiment, the background noise was altered depending upon the subject's performance. For the EAS condition, the target accuracy was 98% ± 2; for the INT condition, 90% ± 5%; for the DIF condition, 80% ± 5. The loudness of the background noise was modified in 5-dB steps, up or down, whenever the average accuracy exceeded the predetermined ranges of 75–85% for DIF, 85–95% for INT, or 96–100% for EAS, based on a 20-trial running average. Successive adjustments were made at five-trial intervals until the subject reached an accuracy of 80%, 90%, or 98%, respectively.

Training

For the DIF and INT conditions subjects received two 10-min training sessions in the hour prior to the uptake period. In the EAS condition, subjects received a minimum of 1500 trials in 500-trial blocks during the 5 days preceding the scan. A single block of 500 trials was presented on any given day.

Magnetic Resonance Imaging (MRI) Scanning

Magnetic resonance scanning was done 1–2 weeks prior to the initial PET procedure. Each subject was fitted with his own molded thermoplastic face mask in order to optimize MRI–PET registration. In the MRI scanner, the subject's head was tilted and cushioned to create a parallel alignment between the subject's glabella–inion axis and the machine's landmark laser. A line was drawn on the mask along the landmark laser, which was made to coincide with the glabella–inion axis. This became the superior–inferior "zero" reference for the scanner. Directly over this marked line a 20-cm flexible plastic tube containing vegetable oil was taped. The oil-filled reference tube appeared as a bright object in the MRI images and was used as an internal marker to help assess accuracy in estimating the glabella–inion axis.

MRI scanning was done on a General Electric Signa Magnetom with a 1.5-T magnet. T1 weighted images were obtained, using relaxation time (TR) = 600, echo time (TE) = 20, sagittal orientation, 3-mm thickness, 24-cm field of view (FOV), 256 × 256 pixel matrix, and 1 nexus (NEX). A set of 50 images was acquired, extending ear-to-ear. One subject was scanned a second time using the three-dimensional Fourier transform, spoiled gradient recalled acquisition (SPGR). That protocol used the following specifications: TR = 35 msec, TE = 9 msec, 45° flip angle, FOV = 24 cm, 1.0 NEX, thickness = 1.5 mm, matrix = 256 × 256 × 124 (planes). Voxel dimensions were 0.975 × 0.975 × 1.5 mm. This set of volumetric images was acquired in the transaxial plane.

PET Scan Technique

A 5-mCi dose of FDG, synthesized from ^{18}F-fluoride and produced in a Sanditronix MC-16F biomedical cyclotron (Uppsala, Sweden), was injected intravenously. Image acquisition began 40–45 min later. Twelve images, which were on 8-mm centers and parallel to the anterior commissure–posterior commissure axis (AC-PC) line were acquired using the NEUROECAT scanner (Computer Technology Incorporated, Knoxville, TN), which has a spatial resolution of 8.0 mm within-plane and 15 mm axially at full width half maximum (FWHM). Scans were corrected for attenuation (using an ellipse) and radioactive decay. After injection of the FDG, 1.5 mL venous blood samples were obtained at intervals throughout the study for plasma glucose levels and radioactivity.

Image Analysis and Reconstruction

Absolute regional cerebral metabolic rates for glucose utilization (rCMRglu) were calculated using published methods and rate constants (Phelps et al 1979). Global metabolic activity of each subject was rescaled to a mean rCMRglu whole brain gray (WBG) value of 8.0 mg/100 g/min. PET and MRI data were transformed into volumetric sets using a trilinear interpolation routine.

Image Registration

The strategy for volumetric image registration used two steps: 1) operator-controlled, landmark-based alignment; and 2) machine-controlled, iterative algorithm-driven alignment, which maximizes the correlation between corresponding planes from separate image sets (Woods et al 1992, 1993).

Operator-controlled registration of PET, MRI, and the Statistical Parametric Mapping (SPM) software program's internal PET template data (Friston et al 1991a) was accomplished using the MIRA System (Multimodal Image Registration and Analysis; Loats Associates, Westminster, MD). PET and MRI data were manually translated and rotated to effectively superimpose (midvolume orthogonal planes) corresponding projections. Grossly aligned with one another, the PET and MRI data sets were then registered algorithmically with one another, within subject, using the Automatic Image Registration (AIR) program developed by Roger Woods (Woods et al 1992, 1993). Within that program the three aligned PET studies of each subject were averaged. The averaged set of *each subject* was then manually superimposed against the internal SPM PET template (three orthogonal projections). Translations and rotations were performed on the averaged individual sets to further optimize their fit to the template. These spatial displacements were then applied to each of the three PET scans of that subject and these were subjected to a second set of algorithmic registrations. The fit of each scan to its companion study was measured using the root mean square error method (Pelizzari et al 1989).

Statistical Analysis

Data were analyzed using the SPM software program (MRC Cyclotron Unit, Hammersmith Hospital; Friston et al 1989, 1990, 1991a and b). The statistical significance of a focal activation was assessed using the spatial extent of the cluster volume above a "high" threshold (Friston et al 1994; Worsley et al 1992).

The SPM analysis proceeds by three steps: 1) stereotaxic normalization (Friston et al 1989, 1991b); 2) analysis of covariance (Friston et al 1990); and 3) t statistics computed for each voxel for the specified comparisons (Friston et al 1991a). Image interpolation and stereotaxic normalization utilizes a nonlinear transformation to "rewrite" or format each volume into the stereotaxic space specified by Talairach and Tournoux (1988). Following stereotaxic normalization the image data were smoothed by a three-dimensional Gaussian filter, 6 mm × 6 mm × 6 mm. Global and intersubject variance in activity were removed by applying a voxel by voxel analysis of covariance. Global metabolic rates were not significantly different between any contrasting set of conditions. In the final SPM step, at each voxel t statistics were computed for three comparisons: 1) DIF minus INT ($p < .01$, two-tailed); 2) DIF minus EAS ($p < .01$, two-tailed); and 3) INT minus EAS ($p < .01$, two-tailed).

Spatial Extent Assessment, Theory

In an SPM{Z} map the likelihood that a cluster of voxels (spatially contiguous) above a high Z score threshold (u) indicates significant change can be determined from several parameters associated with the map (Friston et al 1991a, 1994; Worsley et al 1992). These include the data set's dimensionality (D), smoothness (W), (Worsley et al 1992; Friston et al 1991a, 1994), and search volume (S) (Friston et al 1994; Worsley et al 1992). As the smoothness (full width half maximum, FWHM) increases, the number of connected voxels in a cluster above a specified high statistical threshold must also increase in order for that cluster to be statistically significant. The greater the smoothness the less likely it is that two proximal voxels represent independent measurements. The larger the search volume, the more likely it is that a small cluster is a chance occurrence (Poline and Mazoyer 1993). Furthermore, the higher the Z score threshold u the more significant a cluster of a given size is likely to be. For any given set of values for W, D, u, and S, a large cluster is less likely to occur by chance alone than a small cluster.

Spatial Extent Assessment, Application

At a relatively high statistical threshold, $p < .01$, $Z = 2.33$, and alpha = .05, the required cluster volume was determined by the method described (Friston et al 1994). Dimensionality, D, equals 3; the search volume, S, was determined from the SPM{Z} map. Clusters were identified as spatially contiguous voxels above a Z score threshold of 2.33. Only voxels with adjoining faces were considered contiguous. Edge and corner connections were not sufficient. The smoothness parameter, W, was calculated from the DIF versus INT SPM{Z} map because this comparison contained no significant change clusters and therefore best approximated a Gaussian random field. The

Table 1. Data Summarizing the Performance Characteristics of Each Normal Subject for Each of the Three Conditions

Subject	Condition	Trials completed	% accurate	MED ACC RT
11166	3	1127	96.7169	251.582
62552	3	1437	96.1726	194.203
70963	3	1054	97.2486	428.13
92654	3	1215	92.2634	180.962
Mean	EAS	1208	95.6004	263.721
11166	1	993	94.6626	313.374
62552	1	750	86.2667	550
70963	1	1300	87.3077	611.299
92654	1	694	88.0403	366.338
Mean	INT	934.25	89.0693	460.253
11166	2	1043	91.6587	441.371
62552	2	425	70.3529	556.128
70963	2	847	74.026	816.537
92654	2	943	71.2619	626.747
Mean	DIF	814.5	76.8249	610.196

The number of trials completed and the accuracy fall for each subject when the subject goes from EASY (EAS) to INTERMEDIATE (INT) to DIFFICULT (DIF). The median accurate response time (MED ACC RT) rises when the subject encounters progressively more difficult levels of signal recognition.

smoothness parameter was determined along each axis x, y, z (Friston et al 1994; Worsley et al 1992). Cluster centroids were determined in coordinates consistent with the Talairach and Tournoux (1988) atlas. The probability of finding a cluster of a particular volume above the specified Z score threshold of 2.33, having an alpha level equal to .05, was then determined (Friston et al 1994). Worsley's method of assessing the likelihood of finding even a single voxel above a high threshold in a specified search volume with a particular smoothness was also applied to small, high Z score voxels (Worsley et al 1992).

Results

Auditory Recognition Performance

Table 1 and Figure 1A–D illustrate the behavioral data obtained from the four subjects in each of the three behavioral conditions. The EAS condition was associated with the fastest median accurate response time (mean = 263.7 msec), the largest number of trials (mean = 1208), and the highest average accuracy (95.6%). INT and DIF conditions were characterized by progressively slower median accurate response times (460.3 and 610 msec), fewer trials (934 and 815), and lower average accuracy scores (89.1% and 76.8%). The graphs illustrate the performances of the subjects during the FDG uptake phase of the experiment.

Subject #11166 (Figure 1A) was able to perform with high accuracy in the DIF condition. Nonetheless a marked response-time slowing indicates that this condition was more difficult than the other two in spite of the high score.

Spatial Extent Parameters

Smoothness (FWHM) was calculated using the DIF–INT contrast SPM{Z} (Friston et al 1991a; Worsley et al

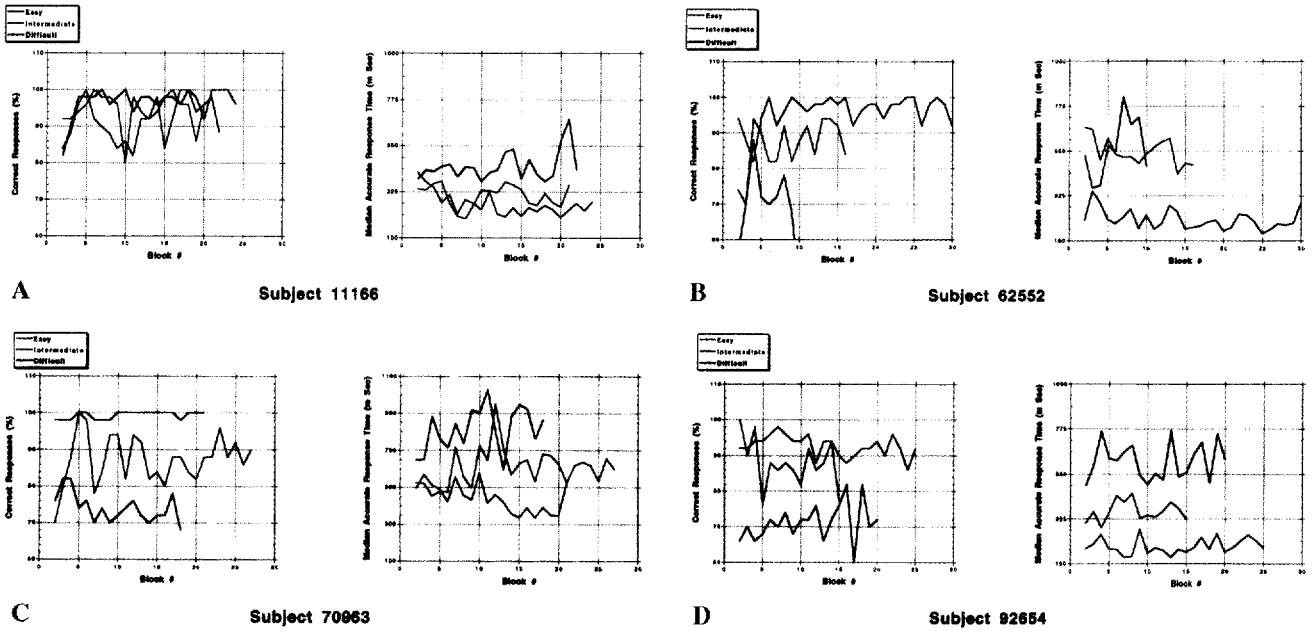

Figure 1A–D. These four pairs of graphs illustrate the accuracy and median accurate response times for each of the four normal subjects performing the auditory recognition task during EASY, INTERMEDIATE, and DIFFICULT signal:noise conditions. Each data point represents the mean of a block containing 20 trials. Only correct trials were used to calculate response time.

Table 2. Spatial Extent of Right Temporal Lobe Cluster

z	N	Medial	Lateral	Post.	Ant.	xm	ym	Zmax
−28	12	(34,−62)	(42,−54)	(38,−66)	(40,−52)	36	−62	3.09
−24	99	(34,−22)	(52,−12)	(36,−64)	(50,−12)	46	−34	4.11
−20	129	(36,−50)	(56,−36)	(36,−58)	(42,−12)	40	−50	3.72
−16	98	(40,−12)	(56,−36)	(48,−50)	(40,−12)	52	−38	3.50
−12	31	(40,−18)	(56,−40)	(54,−42)	(42,−16)	48	−18	2.79

Voxel volume: 369 (5.904 cm^3). *Z* score mean: 2.82. Geometric centroid coordinates: 46, −30, −20.

z: plane level (Talairach and Tournoux 1988). *N*: number of cluster pixels within the specified plane. Medial, Lateral, Post., Ant.: the x,y coordinates of voxels forming the boundaries of the cluster within the plane. xm, ym: x,y coordinates of maximum values within the cluster, at the specified plane level. Zmax: maximum *Z* value within the plane's portion of the cluster.

Table 3. Spatial Extent of Right Medial Geniculate Cluster

z	N	Medial	Lateral	Post.	Ant.	xm	ym	Zmax
−8	4	(16,−34)	(18,−34)	(16,−36)	(16,−34)	18	−34	3.77
−4	33	(6,−34)	(20,−30)	(16,−40)	(12,−30)	18	−34	4.82
0	36	(6,−34)	(20,−30)	(10,−40)	(10,−30)	18	−34	8.01
4	24	(10,−28)	(20,−28)	(10,−36)	(10,−28)	18	−34	3.93

Voxel volume: 97 (1.552 cm^3). *Z* score mean: 3.26. Geometric centroid coordinates: 14, −32, −4.

z: plane level (Talairach and Tournoux 1988). *N*: number of cluster pixels within the specified plane. Medial, Lateral, Post., Ant.: indicate the x,y coordinates of voxels forming the boundaries of the cluster within the plane. xm, ym: x,y coordinates of maximum values within the cluster, at the specified plane level. Zmax: maximum *Z* value within the plane's portion of the cluster.

1992). This map contained no significant clusters by the spatial extent model. Smoothness estimates (units are voxels 2 mm × 2 mm × 4 mm) were as follows: FWHMx = 5.20, FWHMy = 5.46, FWHMz = 2.95, FWHM = 4.37 (global). A search volume (S) was determined to include 51,162 voxels (600 Resels) and dimensionality, D = 3. An alpha level of .05 was applied to an SPM{Z} map thresholded at a $Z = 2.33$, which corresponds to $p < .01$ (by χ^2 criteria applied to the overall change pattern). Using these parameters, a cluster volume of 229 connected voxels was found to be the critical spatial extent necessary for statistical significance. A search routine that identifies clusters of contiguous voxels was subsequently applied to each contrast: DIF–EAS, EAS–DIF, DIF–INT, INT–DIF, INT–EAS, and EAS–INT. None of the four contrasts containing the INT condition revealed clusters large enough to reach significance; nor were there small clusters with especially high *Z* scores, over 4.0. Only the DIF–EAS contrast contained a cluster that met the criteria for spatial extent. Examination of the clusters and their coordinates also revealed two activations (EAS–DIF) located in an area of the thalamus consistent with the medial geniculate body, a structure too small to meet the underlying premises of the spatial extent model (Friston et al 1994; Worsley et al 1992).

Spatial Characteristics of the Activation Foci

Tables 2 and 3 identify the spatial/anatomical characteristics of the right temporal lobe cluster and the right medial geniculate focus; all coordinates are mapped in Talairach space (Talairach and Tournoux 1988). The medial geniculate cluster is significant inasmuch as it contains a small group of high-change score voxels ($Z = 8.0$) that have an extremely low probability of occurring by chance alone, $p = 5.59 \times 10^{-11}$. The medial geniculate change cluster generated by the EAS–DIF contrast extends from 8 mm below to 4 mm above the AC-PC plane. These patterns are illustrated in Figure 2.

The right temporal lobe cluster, generated by the DIF–EAS contrast, extends 16 mm in the z axis and contains contiguous voxels in the middle, inferior, and fusiform gyri. Figure 3 illustrates the extent of the cluster 20 mm below the AC-PC plane.

Discussion

Behavioral Performance

Table 1 and Figure 1A–D illustrate our success in creating three different levels of difficulty based on three different levels of performance, which is defined in terms of accuracy and response time for accurate trials. As the conditions became harder the accuracy fell, the response times increased, and the number of trials completed diminished. The extensive training provided to subjects prior to the EAS condition presumably helped them to be faster and more accurate. This training was specifically provided to eliminate novelty and consolidate familiarity.

Image Analysis

These data were first analyzed and presented (Holcomb et al 1992) without the benefit of automatic registration algorithms (Woods et al 1992, 1993), automated statistical parametric mapping routines (Friston et al 1991a), or the spatial extent model of statistical significance (Friston et al 1994; Poline and Mazoyer 1993, 1994; Worsley et al 1992). It is, however, noteworthy that the largest cluster, with the highest *Z* score, found in the DIF–EAS contrast in the initial analysis, was also found in the analysis presented here, right middle and inferior temporal gyri plus the right fusiform gyrus. No significant differences had been found in the first EAS−DIF subtraction; whereas in this analysis we have clearly demonstrated a robust focus of change in an area of the posterior thalamus consistent with the medial geniculate body. The substantial improvements in registration and the systematic stereotaxic transformations used in the Hammersmith Group's SPM routines (Friston et al 1989, 1991b) have improved the

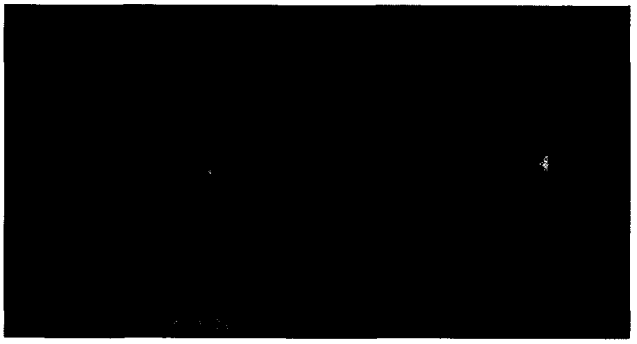

Figure 2. The medial geniculate change cluster generated by the EAS–DIF contrast, extending from 8 mm below to 4 mm above the anterior commissure–posterior commissure (ACPC) plane. This image is taken from the plane passing through the ACPC plane, 00 mm. This plane contains the 2 × 2 × 4 mm voxel with the highest change score, 8 standard deviations.

sensitivity of this analysis. Both of these procedures, registration and stereotaxic transformation, improve the signal-to-noise ratio in group average images (Fox et al 1988).

The application of the spatial extent assessment for significance improves the likelihood that one is not "finding" an activation focus by chance alone (Friston et al 1994). It does not, however, resolve the problem when one encounters a large change value in a small structure like the medial geniculate. Worsley and colleagues (1992) provide an approach that calculates the likelihood of encountering high t values within a specified search volume having a specified number of Resels. Because the right medial geniculate contains at least one voxel with a Z value of 8.0, it is calculated (equation #5, in Worsley et al 1992) that this occurred at an extremely low level of chance, $p = 5.59 \times 10^{-11}$.

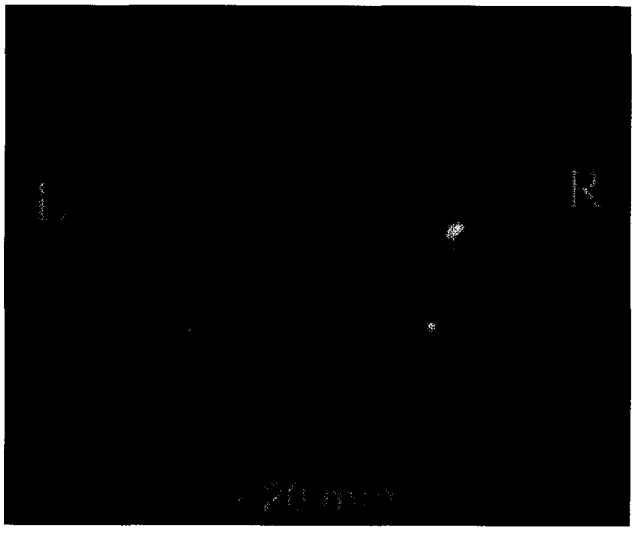

Figure 3. The right temporal lobe cluster, generated by the DIF–EAS contrast, extends 16 mm in the z axis and contains contiguous voxels in the middle and inferior temporal gyri and the fusiform gyrus. This portion of the cluster is found 20 mm below the anterior commissure–posterior commissure (ACPC) plane.

Task Characteristics and Performance Change Brain Activity Patterns Background

Previous reports provide information regarding change patterns with auditory performance; each used a within-subject design in which two or more conditions were administered to each participant. Using nonverbal auditory stimuli to assess different aspects of acoustic processing, Zatorre (1992) compared a silent baseline against bursts of noise. His results suggest the localization of noise perception alone. Mazziotta (1983) compared a silent condition against one requiring the subject to indicate whether two sequentially presented chords were identical. These data suggest the localization of noise perception, attention, and judgment. Knopman (1982) imaged resting subjects on one occasion and the same subjects detecting a low loudness tone on another occasion. These data suggest the localization of threshold noise perception. Roland (1981) compared subjects matching tones against their quiet, resting condition. These data suggest the functional localization of auditory task performance: noise perception, attention, judgment, discrimination, and memory.

Although imaging technologies differed substantially in these four reports (Xenon-133 in two, FDG in one, dynamic oxygen-15 in one), some generalizations are apparent. In contrast to Zatorre (1992), who imposed no decision task on his subjects listening to noise bursts, the other three investigators required some decision regarding the auditory characteristics of the signals. Only Zatorre (1992) found right–left symmetry in the functional response of the primary auditory cortices, suggesting that perception alone is bilaterally symmetric. Mazziotta's study used tone pattern discrimination and observed marked right-sided asymmetry in glucose utilization. This was most pronounced in the superior temporal gyrus and right inferior frontal cortex (Mazziotta et al 1983). Knopman (1982) used a simple task involving loudness discrimination and found increased regional cerebral blood flow (rCBF) in the left posterior Sylvian area when compared against a quiet baseline. Roland (1981) found substantial elevations in the right hemispheric rCBF of subjects making tone pattern judgments. These right-sided elevations were especially prominent in primary auditory cortex, middle temporal gyrus, inferior frontal gyrus, and temporo-occipital regions.

Other studies emphasize the physiological relationships between temporal lobe and frontal cortex in auditory

discrimination behavior. Cohen (Cohen et al 1992) used FDG-PET to compare normal subjects actively detecting the lowest presentation tone against subjects hearing the same tones but making no decisions regarding their characteristics. A marked elevation in the right anterior frontal region at midlevel was confirmed against previous studies (Cohen et al 1988). The role of the right midfrontal cortex in pitch discrimination was extensively supported by a recent report by Zatorre (Zatorre et al 1994). Furthermore, the decisive impact of attention on auditory cortical activity has been vividly demonstrated by neuromagnetic measurements (Woldorf et al 1993). This group has extended the work of evoked response research; they clarify the impact of attention on auditory stimulus processing. When subjects attend to tones presented to one ear and ignore those presented to the other, significantly greater magnetic fields are recorded from the supratemporal plane of the attended side. These three reports clearly emphasize the role of attentive decision making on auditory stimulus processing.

Temporal Lobe and Medial Geniculate Body Observations

Observations reported in this paper are consistent with those presented in the literature reviewed. In both contrasts, DIF–EAS and EAS–DIF, brain activity patterns exhibited a strong right-sided asymmetry. The auditory recognition task performed in association with a low signal-to-noise ratio enhanced the activity of the right inferior temporal lobe and fusiform gyrus. This task may also have caused reduced activity in the medial geniculate body, a structure intimately associated with frequency detection (Heffner and Heffner 1984; Lennartz and Weinberger 1992; Motomura et al 1986). When activity associated with this difficult recognition task is subtracted from that generated by an easy condition, only the right medial geniculate body (MGB) exhibits a highly significant difference. This may reflect the enhanced role of the MGB during an undemanding prolonged frequency recognition task, or it may be consistent with marked reduction in activity during an especially difficult prolonged frequency assessment task. The present data do not distinguish between these two possibilities.

The heightened metabolic activity in the right inferior temporal lobe may reflect greater neuronal activity occasioned by performing the DIF auditory discrimination task over a prolonged time period. The change pattern is consistent with enhanced "neuronal work" in areas implicated in auditory and speech discrimination (Brugge and Reale 1985; Colombo et al 1990; Pandya and Yeterian 1985; Phillips 1989; Seldon 1985). It is, however, noteworthy that the right temporal lobe cluster overlaps extensively with a large cluster found by Haxby and colleagues (Haxby et al 1994) in a study that compared face recognition with location matching. This overlap is consistent with the temporal lobe's role in pattern recognition. The same neurons may play an important role in auditory and visual stimulus identification.

The prolonged tracer uptake period, 25 min, may complicate physiological interpretations of these data, especially when viewed in association with other measurements made in milliseconds (electrophysiology and functional MRI) or seconds (rCBF-PET). Future efforts may take advantage of oxygen-15's short half-life of 2 min. One might, for example, obtain multiple short scans of 60 sec at 10-minute intervals during a 30-min behavioral condition. In the context of the study reported here it is useful to note that all four subjects exhibited shifts in accuracy and response time consistent with a sustained, effortful state.

Behavioral Normalization Using Variable Degraded Stimuli

Functional neuroimaging has historically used the same behavioral task with diverse patient groups (Buchsbaum et al 1990; Cohen et al 1988); however, in many studies patients and normal volunteers exhibit "ceiling" and "floor" performance effects. We have demonstrated here that it is possible to control the difficulty of a single task by reducing the discriminability of the target stimulus. This manipulation may substantially reduce individual differences in task performance. Thus, graded error rate task performance could be used to standardize behavioral performance state between different groups.

To the extent that behavior changes during the tracer uptake phase of a glucose metabolism scan, the integrated image will reflect an integration of different behaviors. The behavioral pattern will be neither constant nor linear over time. One response to this situation is to modify the level of target stimulus degradation from one block of trials to another in accordance with the subject's accuracy, as was carried out here. The investigator can increase or decrease the discriminability of the stimuli and maintain a relatively constant, mean level of accuracy throughout the tracer uptake phase. Thus, a graded error-rate-task protocol can be used to maintain a critical level of performance during a testing phase.

In summary, these data suggest that the right inferior temporal lobe and the right MGB play important roles in tone frequency recognition, especially when the task occurs over a prolonged period of time. High task difficulty intensifies labeled glucose utilization in an area

recently associated with visual recognition (Haxby et al 1994). The enhanced activity in the MGB may indicate an important role for this thalamic structure during routine, nondemanding frequency detection occurring over an extended time period. These data also suggest the utility of graded error rate tasks for use in functional imaging procedures to standardize task performance between and within groups.

This research was supported by grants from the National Institute of Mental Health (Neuroscience Center for Research in Schizophrenia Grant #MH44211; and Clinical Research Center Grant #MH40279).

We thank Jeff Sieracki for the auditory discrimination computer software. Hayden Rayvert, David Clough, and Bill Boitnott provided invaluable technical support for FDG tracer synthesis and PET image acquisition. Joe Maisogue, Nick Lange, and Craig Formby also provided important statistical and audiological assistance.

References

Brugge JF, Reale RA (1985): Auditory cortex. In: Peters A, Jones EG (eds), *Cerebral Cortex, Association and Auditory Cortices.* New York: Plenum Press, pp 229–271.

Buchsbaum MS, Nuechterlein KH, Haier RJ, Wu J, Sicotte N, Hazlett E, Asarnow R, Potkin S, Guich S (1990): Glucose metabolic rate in normals and schizophrenics during the continuous performance test assessed by positron emission tomography. *Br J Psychiatry* 156:216–227.

Cohen R, Semple WE, Gross M, Holcomb HH, Dowling SM, Nordahl TE (1988): Functional localization of sustained attention: Comparison to sensory stimulation in the absence of instruction. *Neuropsychiatr Neuropsychol Behav Neurol* 1:3–20.

Cohen R, Semple WE, Gross M, King AC, Nordahl TE (1992): Metabolic brain pattern of sustained auditory discrimination. *Exp Brain Res* 92:165–172.

Colombo M, D'Amato MR, Rodman HR, Gross CG (1990): Auditory association cortex lesions impair auditory short-term memory in monkeys. *Science* 247:336–338.

Corbetta M, Miezin FM, Dobmeyer S, Shulman GL, Petersen SE (1990): Attentional modulation of neural processing of shape, color, and velocity in humans. *Science* 248:1556–1559.

Corbetta M, Miezin FM, Shulman GL, Petersen SE (1991): Selective and divided attention during visual discrimination of shape, color, and speed: Functional anatomy by positron emission tomography. *J Neurosci* 11:2383–2402.

Corbetta M, Miezin FM, Shulman GI, Petersen SE (1993): A PET study of visuospatial attention. *J Neurosci* 13:1202–1226.

Fox PT, Mintun MA, Reiman EM, Raichle ME (1988): 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.

Friston KJ, Passingham RE, Nutt JG, Heather JD, Sawle GV, Frackowiak RSJ (1989): Localisation in PET images: Direct fitting of the intercommissural (AC-PC) line. *J Cereb Blood Flow Metab* 9:690–695.

Friston KJ, Frith CD, Liddle PF, Dolan RJ, Lammertsma AA, Frackowiak RSJ (1990): The relationship between global and local changes in PET scans. *J Cereb Blood Flow Metab* 10:458–466.

Friston KJ, Frith CD, Liddle PF, Frackowiak RSJ (1991a): Comparing functional (PET) images: The assessment of significant change. *J Cereb Blood Flow Metab* 11:690–699.

Friston KJ, Frith CD, Liddle PF, Frackowiak RSJ (1991b): Plastic transformation of PET images. *J Comput Assist Tomogr* 15:634–639.

Friston KJ, Worsley KJ, Frackowiak RSJ, Mazziotta JC, Evans AC (1994): Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping* 1:210–220.

Haxby JV, Horwitz B, Ungerleider LG, Maisog JM, Pietrini P, Grady C (1994): The functional organization of human extrastriate cortex: A PET-rCBF study of selective attention to faces and location. *J Neurosci* 14:6336–6353.

Heffner RS, Heffner HE (1984): Hearing loss in dogs after lesions of the brachium of the inferior colliculus and medial geniculate. *J Comp Neurol* 230:207–217.

Holcomb HH, Gordon B, Loats HL, Gastineau EA, Medoff D, Dannals RF, Tamminga CA (1992): Increased attentional effort during a pitch discrimination task elevates regional brain glucose utilization. *Soc Neurosci Abstr* 18(part 1):934.

Knopman DS, Rubens AB, Klassen AC, Meyer MW (1982): Regional cerebral blood flow correlates of auditory processing. *Arch Neurol* 39:487–493.

Lennartz RC, Weinberger NM (1992): Frequency selectivity is related to temporal processing in parallel thalamocortical auditory pathways. *Brain Res* 583:81–92.

Mazziotta JC, Phelps ME, Halgren E (1983): Local cerebral glucose metabolic response to audiovisual stimulation and deprivation: Studies in human subjects with positron CT. *Human Neurobiol* 2:11–23.

Merzenich MM, Recanzone G, Jenkins WM, Allard TT, Nudo RJ (1988): Cortical representational plasticity. In Rakic P, Singer W (eds), *Neurobiology of Neocortex.* Chichester, U.K.: John Wiley & Sons Limited, pp 41–67.

Meyer E, Ferguson SSG, Zatorre RJ, Alivisatos B, Marrett S, Evans AC, Hakim AM (1991): Attention modulates somatosensory cerebral blood flow response to vibrotactile stimulation as measured by positron emission tomography. *Ann Neurol* 29:440–443.

Moran J, Desimone R (1985): Selective attention gates visual processing in the extrastriate cortex. *Science* 229:782–784.

Motomura N, Yamadori A, Mori E, Tamaru F (1986): Auditory agnosia. Analysis of a case with bilateral subcortical lesions. *Brain* 109:379–391.

Nestor PG, Faux SF, McCarley RW, Sands SF, Horvath TB, Peterson A (1991): Neuroleptics improve sustained attention in schizophrenia: A study using signal detection theory. *Neuropsychopharmacology* 4:145–149.

Nuechterlien KH, Parasuraman R, Jiang Q (1983): Visual sustained attention: Image degradation produces rapid sensitivity decrement over time. *Science* 220:327–329.

Pandya DN, Yeterian EH (1985): Architecture and connections

of cortical association areas. In Peters A, Jones EG (eds), *Cerebral Cortex, Association and Auditory Cortices.* New York: Plenum Press, pp 13–20.

Pardo JV, Fox PT, Raichle ME (1991): Localization of a human system for sustained attention by positron emission tomography. *Nature* 349:61–64.

Pelizzari CA, Chen GTY, Spelbring DR, Weichselbaum RR, Chen C-T (1989): Accurate three-dimensional registration of CT, PET and MR images of the brain. *J Comput Assist Tomogr* 13:20–26.

Phelps ME, Huang SC, Hoffman EJ, et al (1979): Tomographic measurement of local cerebral glucose metabolic rate in humans with [F-18]1-fluoro-2-deoxy-D-glucose: Validation of method. *Ann Neurol* 6:371–388.

Phillips DP (1989): The neural coding of simple and complex sounds in the auditory cortex. In Lund JS (ed), *Sensory Processing in the Mammalian Brain: Neural Substrates and Experimental Strategies.* New York: Oxford University Press, pp 173–203.

Poline JB, Mazoyer BM (1993): Analysis of individual positron emission tomography activation maps by detection of high signal-to-noise-ratio pixel clusters. *J Cereb Blood Flow Metab* 13:425–437.

Poline JB, Mazoyer BM (1994): Analysis of individual brain activation maps using hierarchical description and multiscale detection. *IEEE Trans Med Imag* 13:702–710.

Raichle ME, Fiez JA, Videen TO, MacLeod AMK, Pardo JV, Fox PT, Petersen SE (1994): Practice-related changes in human brain functional anatomy during nonmotor learning. *Cerebral Cortex* 4:8–26.

Richmond BJ, Sato T (1987): Enhancement of inferior temporal neurons during visual discrimination. *J Neurophysiol* 58: 1292–1306.

Roland PE, Skinhoj E, Lassen NA (1981): Focal activation of human cerebral cortex during auditory discrimination. *J Neurophysiol* 45:1139–1151.

Seldon L (1985): The anatomy of speech perception, human auditory cortex. In Peters A, Jones EG (eds), *Cerebral Cortex, Association and Auditory Cortices.* New York: Plenum Press, pp 273–327.

Spitzer H, Richmond BJ (1991): Task difficulty: Ignoring, attending to, and discriminating a visual stimulus yield progressively more activity in inferior temporal neurons. *Exp Brain Res* 83:340–348.

Spitzer H, Desimone R, Moran J (1988): Increased attention enhances both behavioral and neuronal performance. *Science* 240:338–340.

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

Tulving E, Markowitsch, Kapur S, Habib R, Houle S (1994): Novelty encoding networks in the human brain: Positron emission tomography data. *NeuroReport* 5:2525–2528.

Woldorf MG, Gallen CC, Hampson SA, Hillyard SA, Pantev C, Sobel D, Bloom FE (1993): Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proc Natl Acad Sci USA* 90:8722–8726.

Woods RP, Cherry SR, Mazziotta JC (1992): Rapid automated algorithm for aligning and reslicing PET images. *J Comput Assist Tomogr* 16:620–633.

Woods RP, Mazziotta JC, Cherry SR (1993): MRI-PET registration with automated algorithm. *J Comput Assist Tomogr* 17:536–546.

Worsley KJ, Evans AC, Marrett S, Neelin P (1992): A three-dimensional statistical analysis for CBF activation studies in human brain. *J Cereb Blood Flow Metab* 12:900–918.

Zatorre RJ, Evans AC, Meyer E, Gjedde A (1992): Lateralization of phonetic and pitch discrimination in speech processing. *Science* 256:846–849.

Zatorre RJ, Evans AC, Meyer E (1994): Neural mechanisms underlying melodic perception and memory for pitch. *J Neuroscience* 14:1908–1919.