Time-Dependent Changes in Learning Audiovisual Associations: A Single-Trial fMRI Study

Désirée Gonzalo,* Tim Shallice,† and Raymond Dolan*

*The Wellcome Department of Cognitive Neurology, 12 Queen Square, and †Institute of Cognitive Neuroscience, 17 Queen Square, Institute of Neurology, University College London, London WC1N 3BG, United Kingdom

Functional imaging studies of learning and memory have primarily focused on stimulus material presented within a single modality (see review by Gabrieli, 1998, *Annu. Rev. Psychol.* **49:** 87–115). In the present study we investigated mechanisms for learning material presented in visual and auditory modalities, using single-trial functional magnetic resonance imaging. We evaluated time-dependent learning effects under two conditions involving presentation of consistent (repeatedly paired in the same combination) or inconsistent (items presented randomly paired) pairs. We also evaluated time-dependent changes for bimodal (auditory and visual) presentations relative to a condition in which auditory stimuli were repeatedly presented alone. Using a time by condition analysis to compare neural responses to consistent versus inconsistent audiovisual pairs, we found significant time-dependent learning effects in medial parietal and right dorsolateral prefrontal cortices. In contrast, time-dependent effects were seen in left angular gyrus, bilateral anterior cingulate gyrus, and occipital areas bilaterally. A comparison of paired (bimodal) versus unpaired (unimodal) conditions was associated with time-dependent changes in posterior hippocampal and superior frontal regions for both consistent and inconsistent pairs. The results provide evidence that associative learning for stimuli presented in different sensory modalities is supported by neural mechanisms similar to those described for other kinds of memory processes. The involvement of posterior hippocampus and superior frontal gyrus in bimodal learning for both consistent and inconsistent pairs supports a putative function for these regions in associative learning independent of sensory modality. © 2000 Academic Press

Key Words: cross modal associations; learning; hippocampus

[1] To whom correspondence should be addressed. E-mail: d.gonzalo@fil.ion.ucl.ac.uk.

INTRODUCTION

The neural consequences of learning have been studied principally at the molecular and cellular levels (e.g., long-term potentiation, Bliss and Lømo, 1973; Bliss and Collingridge, 1993) where a fundamental principle is change in neural connection strength (Hebb, 1949). Behaviorally, learning is defined as the acquisition of new information that is manifest as increased fluency on a range of psychological measures. Irrespective of the frame of reference, learning involves change over time and consequently time is critical to the definition of learning. In this experiment we used functional neuroimaging to characterize, at the systems level, time-dependent neural changes associated with associative learning across sensory modalities.

Neuroimaging studies of memory have used multiple psychological paradigms to study different aspects of learning (see review by Cabeza and Nyberg, 1997). However, few studies have examined time-dependent aspects of learning. Notable exceptions are studies that investigated practice effects or patterns of brain activation (Friston *et al.,* 1992; Jenkins *et al.,* 1994; Raichle *et al.,* 1994; Jueptner *et al.,* 1996; Petersen *et al.,* 1998) that distinguished between brain areas involved in unskilled effortful performance and those related to skilled effortless performance. In cognitive studies, Raichle *et al.* (1994) used a verbal response selection task to compare a naive with a practiced state. Certain brain regions, namely anterior cingulate, left prefrontal and left posterior temporal cortices, and right cerebellum, became less active with practice, with a reverse pattern being observed in bilateral sylvian-insular cortex and left medial extrastriate cortex (Raichle *et al.,* 1994). A more recent study, also using a verbal response selection task, provides evidence of a practice-related shift in neuronal response (Petersen *et al.,* 1998). Initial activations in left prefrontal, anterior cingulate, and right cerebellar hemisphere were distinct from activations observed after practice that were principally located in sylvian-insular cortex.

A major assumption in memory research is that psychologically distinct memory systems have different neuroanatomical components (Baddeley, 1996; McCarthy and Warrington, 1990). The components of these systems differ according to type of material and the operations necessary for effective learning. Hence, verbal learning is believed to involve primarily bilateral prefrontal cortex, posterior cingulate, precuneus, and superior temporal gyrus (Grasby et al., 1993, 1994) while visual learning involves extrastriate and higher order visual areas in the occipital and temporal cortices (Tovee et al., 1996; Eacott and Heywood, 1995). The studies have all investigated learning of associations between stimuli and responses in conceptually related domains. Neuropsychologically, however, Petrides (1985) has argued that where stimuli and responses have no relationship to each other, other systems come into play in the learning process. The process of learning associations between visual and nonverbal auditory stimuli is of this type. We are interested in the time-dependent changes involved in this type of associative learning. These questions provide the focus of the present study, which investigated the time-dependent changes associated with learning to discriminate consistent and inconsistent audiovisual pairs, as well as with the process of making associations versus not making associations. We used a hard discrimination task where stimulus pairs were composed of an auditory stimulus and a visual stimulus.

We predicted that learning pairs of stimuli, regardless of sensory modality, would activate left medial temporal and right frontal regions. The left medial temporal lobe has been extensively associated with learning and memory processes (Petrides, 1985; Wallenstein et al., 1998; Gabrieli, 1998; Eichenbaum, 1992; Fletcher et al., 1995a; Henke et al., 1997; Kelley et al., 1998; Rombouts et al., 1997; Squire and Zola, 1996), and right prefrontal activations have been implicated in incremental learning during repeated trials (Kopelman et al., 1998). Furthermore, we hypothesized that there would be differences between the networks activated in response to presentation of consistent and inconsistent audiovisual pairs. Consistent pairs are more likely to activate regions involved in retrieving new memories, such as previously reported right frontal and precuneal areas (Tulving et al., 1994; Shallice et al., 1994; Fletcher et al., 1995a; Buckner et al., 1996; Kopelman et al., 1998). However, when attentional demands increase, such as for inconsistent audiovisual pairings where a greater number of audiovisual combinations are presented, the involvement of visual and cingulate areas is also likely to increase (Pardo et al., 1990; Posner and Petersen, 1990; Kosslyn et al., 1996; Benedict et al., 1998; Carter et al., 1998). Most notably, activations in superior right frontal and medial temporal cortices were predicted under conditions where two unrelated stimuli had to be associated, since these areas have previously been found to play a critical role in associational tasks (Petrides, 1985; Wallenstein et al., 1998).

METHODS

Subjects

Fourteen healthy volunteer subjects (five females) whose ages ranged from 22 to 29 years participated in the study. All were free of neurological or psychiatric illness. Nine subjects took part in a behavioral study while the remaining five (all right-handed) took part in the functional neuroimaging. All subjects provided informed consent. This study was approved by the Joint Ethics Committee of the National Hospital for Neurology and Neurosurgery and the Institute of Neurology, London.

Cognitive Task

Subjects were studied in an event-related paradigm during three separate training sessions. During these periods, subjects were exposed to four different conditions: consistent audiovisual pairs, inconsistent audiovisual pairs, single auditory stimuli, and single visual stimuli. Behavioral data were acquired outside the scanner where nine behavioral subjects underwent exactly the same experimental procedure as those in the scanner save for three interleaving test sessions, after each training session, from which behavioral performance was inferred using a Wilcoxon paired test.

Subjects were informed that they would receive auditory and visual stimuli either paired or unpaired. They were instructed to learn which audiovisual pairs were consistent and which were inconsistent and to not attend to the unpaired stimuli. No behavioral response was required during learning. During the behavioral study performed outside the scanner, subjects were informed that three test sessions would appear interleaved with learning sessions. The task requirement was to press a key when they thought a presented pair belonged to the consistent condition and another key if it belonged to the inconsistent condition.

Experimental Materials

The auditory stimuli consisted of six distinct musical chords. The visual stimuli were six distinct and novel Chinese ideograms. Half of the auditory and the visual stimuli were presented as consistent audiovisual pairs, i.e., always the same two items in the pair. The other half were presented as inconsistent audiovisual pairs, i.e., the same two items in the pair were never presented paired twice during a single training session. The auditory and visual stimuli were presented in pairs, either consistent or inconsistent in two-thirds of trials while in the remaining third of trials the same 12 auditory and visual stimuli were presented unpaired.

The six musical chords were constructed with a sound effects program by adding together the single components according to the frequencies assigned to single notes (http://www.afx.com/Wlips/hm000026.htm). The Chinese ideograms were obtained from http:// www.wg.omron.co.jp/cgi-bin/j e/nocolor/dict.

Experimental Procedure

Subjects lay in a MRI scanner. The auditory stimuli were presented through headphones at a comfortable volume regulated individually to each subject. Visual stimuli were presented in white, against a black background, and centered on a screen located approximately 20 cm away from the eyes, subtending a visual angle of 7°.

The total presentation duration for each pair was 3 s. For presentation of audiovisual pairs, the auditory stimulus was presented for 3 s and the visual stimulus overlapped the last second (see Fig. 1). The three stimulus types were randomly presented in triplets; i.e., each event type was repeated three times. The interstimulus interval, which was randomly jittered ($\pm\frac{1}{3}$ and $\pm\frac{2}{3}$ TR), was 12.9 s on average. There were 9 repetitions of each paired event and 12 repetitions of the single events, per session. There were differences in presentation for the 9 events belonging to each paired condition at each session. Specifically, there were three consistent audiovisual pairs repeated three times, amounting to a total of 9 presentations; there were 9 different inconsistent pairs presented once but which were rearrangements of three different pairings of items. This arrangement was necessary to keep the number of event presentations equal while creating a pattern of consistency by repetition under the consistent condition and a pattern of inconsistency, by not repeating the same pairs, under the inconsistent condition. In total, over the entire experiment, there were 27 repetitions for events corresponding to the consistent condition, 27 repetitions for events corresponding to the inconsistent condition, 18 repetitions of single auditory stimuli, and 18 repetitions of visual stimuli. The duration of the whole experiment was approximately 30 min.

Magnetic Resonance Procedures

The data were acquired from a 2-T Magnetom-VISION whole-body MRI system (Siemens, Erlangen) equipped with a head volume coil. Multislice T2* weighted fMRI images were obtained with a gradient echo planar sequence and axial slice orientation (echo time TE = 40 s, repetition time TR = 4.3 s, $64 \times 64 \times 48$ voxels). Data were acquired in three sessions. A total of 279 contiguous volume images were acquired. Each echoplanar image comprised 48 1.8-mm slices taken every 3 mm, positioned to cover the whole of the cerebrum. Each session comprised 93 volume images.

Structural images were obtained in the same orientation, yielding T1 weighted images with $1 \times 1 \times 1.5$ mm voxel size (matrix size: $256 \times 256 \times 108$). An MPRAGE (magnetized prepared rapidly acquired gradient echo) sequence was used, with 108 partitions acquired in the sagittal plane, with a flip angle of 12°, and TE = 4 ms, TR = 9.7 ms, TI = 600 ms. Structural images were acquired before the start of the experiment, followed by one or two scout scans and five dummy scans, before the first functional scan, to allow for T2 equilibration effects. The functional scans were acquired in a single-trial fashion. The scanner was synchronized with the jittered onset of events, and the ratio of interscan to interstimulus interval ensured that voxels were sampled at different phases relative to stimulus onset (a total of three scans were taken per event). This allows the gathering of data to take place at different stages of the hemodynamic response and also minimizes expectation effects.

Data Analysis

The data were analyzed using the general linear model, in this instance linear regression on condition-specific waveforms convolved with a hemodynamic response function (Friston *et al.,* 1995a) as implemented in SPM98 and SPM97devel. More specifically, the time series were realigned and resliced using a sinc interpolation, adjusting for residual motion-related changes. A mean image, created from the realigned
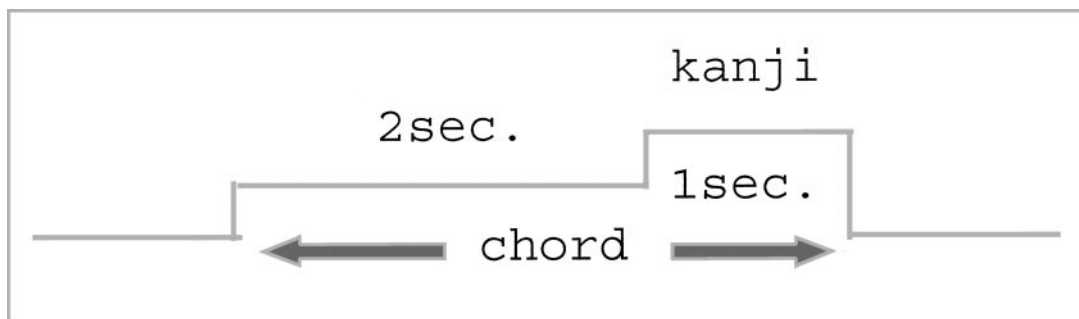

FIG. 1. Diagram indicating the temporal relation of presentation of the auditory and visual stimuli. The auditory stimulus was presented for 3 s. The visual stimulus was presented during the last second.
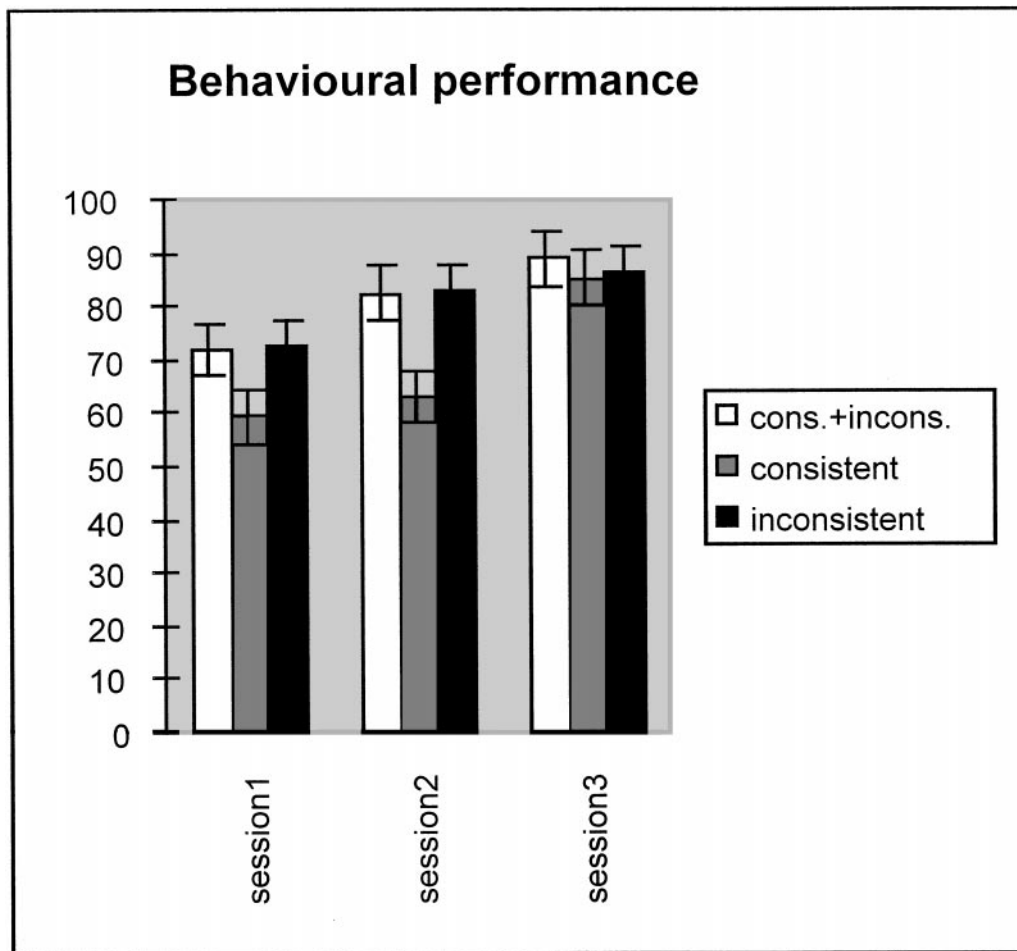
FIG. 2. Graphic plot of the behavioral test performed. The first, second, and third columns (white, gray, and black) indicate the mean group performance for each session when tested for both paired conditions (consistent and inconsistent; white); consistent audiovisual pairs only (gray); and inconsistent audiovisual pairs only (black). The difference in mean group performance between the first and the third sessions is significant (Wilcoxon paired test $P < 0.01$).

volumes, was coregistered with the T1 structural volumes (Friston *et al.,* 1995b, 1996) and spatially normalized to a standard template in the space of Talairach and Tournoux (1988). The data were smoothed spatially with a 10-mm FWHM isotropic Gaussian kernel and temporally with a 4-s Gaussian kernel and a high-pass filter providing a basis set of cosine functions with a cut-off period of 61 s to remove low-frequency drifts in the BOLD signal (Holmes *et al.,* 1997).

A group analysis was performed using a fixed effects model. The effects included in the design matrix are 15 sessions each corresponding to one session per subject (3 sessions, 5 subjects) including separate columns for main effects, and for time by condition interactions and a high-pass filter for each of the sessions. The covariates of interest included four conditions: consistent audiovisual pairs, inconsistent audiovisual pairs, auditory unpaired stimuli, and visual unpaired stimuli. We focused on the interpretation of time by condition interactions whereby the effects of learning could be

characterized as time-dependent changes in neural response within each session. Our contrasts tested for greater learning effects in one versus another condition.

A priori activations were predicted on the basis of previous studies of encoding and retrieval in episodic memory. The regions considered significant were clusters of at least 40 contiguous voxels that survived thresholding at $P < 0.001$ uncorrected ($Z > 3.09$).

RESULTS

Behavioral Data

Nine subjects were tested outside the scanner following three exposures to the material emulating the scanning conditions (see Fig. 2). The procedure was similar to that followed during scanning, with the addition of alternate tests after each learning session for the purpose of measuring learning rates. The test
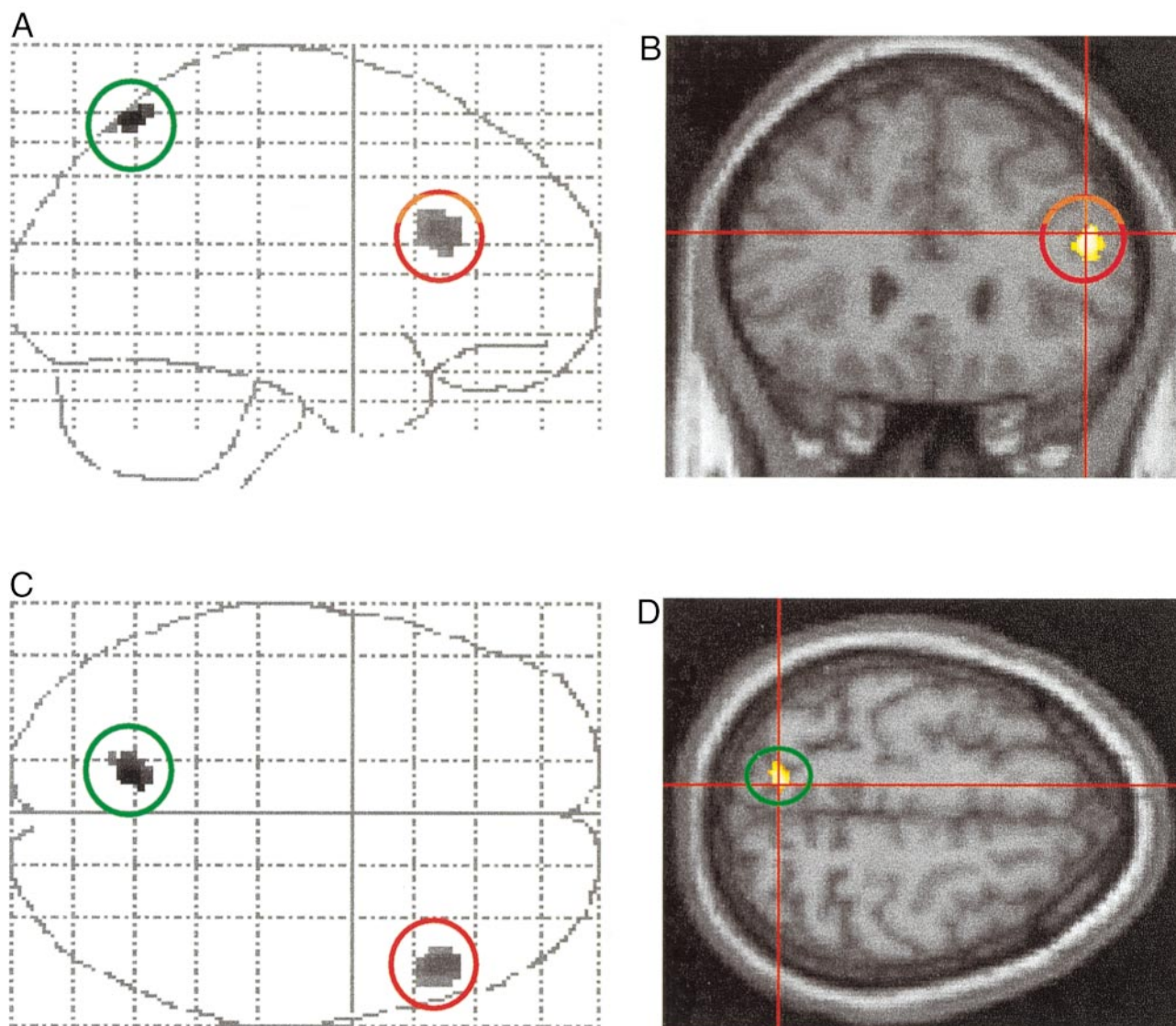
FIG. 3. (a) Lateral and (b) transverse views of the brain regions precuneus and the right dorsolateral prefrontal cortex activations when comparing time by condition interactions between the consistent and the inconsistent audiovisual pairs. (c) Coronal view of the right dorsolateral prefrontal cortex (48, 28, 24). (d) Transverse view of the medial parietal cortex (−10, −64, 60).

consisted of presentations of audiovisual pairs where the subjects had to press the right button on a keypad if they identified the pair as consistent or the left button if they identified it as inconsistent.

The fact that 12 pairs were presented at test, 3 of which were consistent and 9 of which were inconsistent, will incur a difference in chance levels for the two conditions, with chance in the consistent condition being 25%, whereas in the inconsistent condition it is 75%. However, two-sample t tests showed a significant increase above chance in correct responses between the first and the third testing sessions (Wilcoxon paired $P < 0.01$) both for the consistent and the inconsistent conditions separately.

Functional Neuroimaging Data

The effects of interest in this study are changes in neural response for consistent (where behavioral data indicate significant learning) compared to inconsistent audiovisual pairs across time, i.e., in association with learning, as well as changes in neural response for the paired conditions versus the auditory unpaired condition.

Time-Dependent Changes for Consistent versus Inconsistent Audiovisual Pairs

This analysis contrasted neural responses to consistent versus inconsistent audiovisual pair presentations
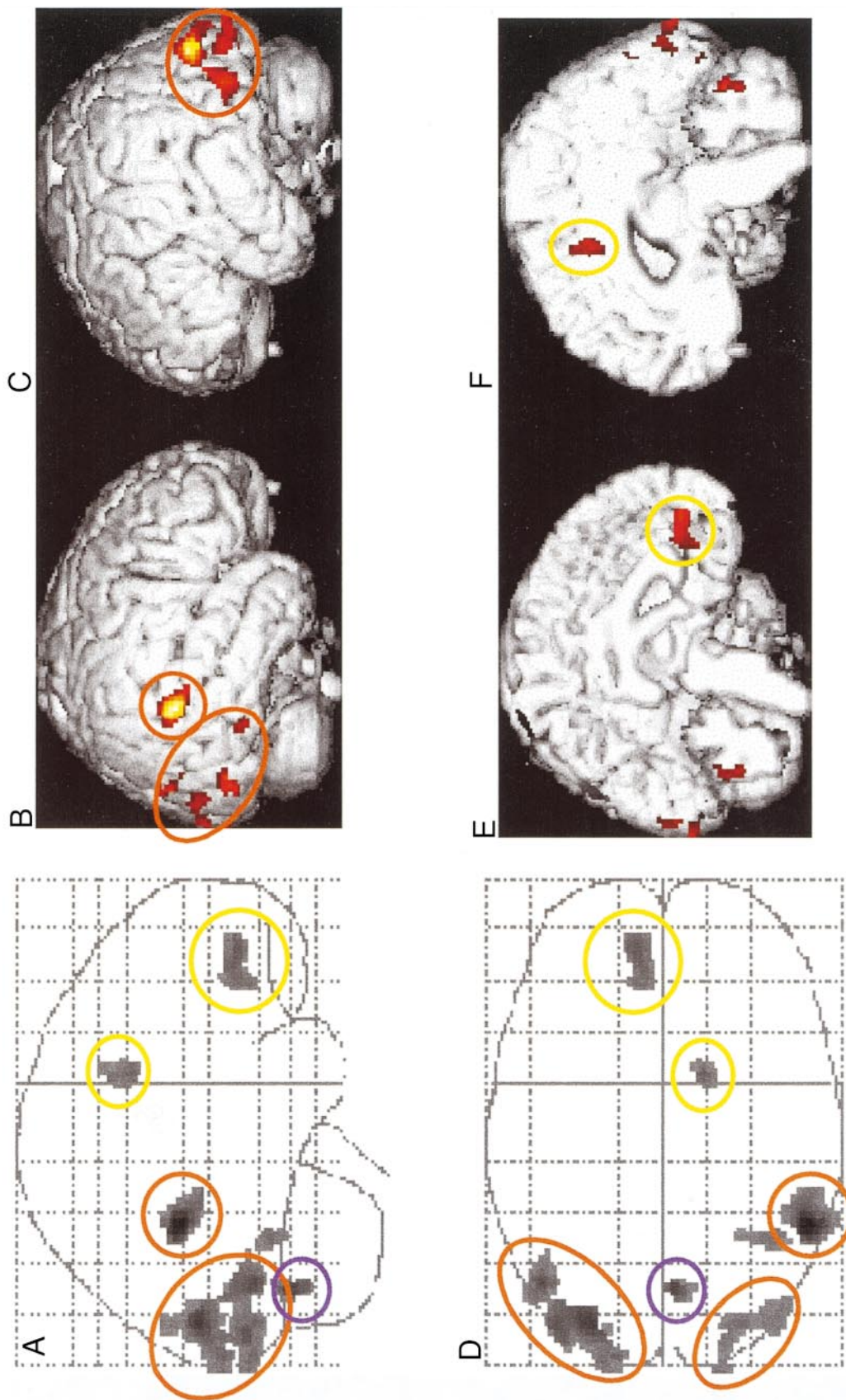
FIG. 4. (a) Lateral and (b) transverse views of the brain regions activated when comparing time by condition interactions between the inconsistent and the consistent audiovisual pairs. Yellow, cingulate; orange, extrastriate cortex; purple, cerebellum. (c) Lateral rendered views of the angular gyrus and extrastriate cortex activations. (d) Sagittal views of the cingulate activations.

TABLE 1

Local Maxima for Areas whose Activation Increased over Time for the Consistent versus the Inconsistent Audiovisual Pairs

Brain location	Coordinates (x, y, z)	Z score
Left precuneus (BA7)	−10, −64, 60	4.30
Right dorsolateral prefrontal (BA46)	48, 28, 24	3.34

as a function of time. This contrast revealed significant time-dependent differential responses in left medial parietal cortex (precuneus) and right dorsolateral prefrontal cortex (see Table 1 and Fig. 3). In other words, these are regions where there was a significant relative increase in the differential response between consistent and inconsistent pairs as a function of repeated presentations.

Time-Dependent Changes for Inconsistent versus Consistent Audiovisual Pairs

The opposite contrast revealed significant time-dependent differential responses for inconsistent relative to consistent audiovisual pairs in right angular gyrus, left middle occipital gyrus (BA18), right cerebellum, and bilateral anterior cingulate (see Table 2 and Fig. 4).

Time-Dependent Changes for Bimodal versus Unimodal Processing

Time-dependent effects of processing associations between stimuli in different sensory modalities was compared to processing within a single modality, namely unpaired auditory stimuli. Comparing consistent pairs with the auditory stimulus alone revealed significant differential time-dependent effects in right superior frontal gyrus and the left posterior hippocampus (Fig. 5, Table 3). Comparing the inconsistent condition with the single auditory stimuli revealed significant differential time-dependent responses in left posterior hippo-

TABLE 2

Local Maxima for Areas whose Activation Increased over Time for the Inconsistent versus the Consistent Audiovisual Pairs

Brain location	Coordinates (x, y, z)	Z score
Right angular gyrus	56, −50, 22	4.24
Left middle occipital gyrus (BA18)	−28, −84, 12	3.67
Left inferior occipital gyrus (BA18)	−38, −88, −4	3.14
Right cuneus (BA18)	22, −104, 4	3.05
Left middle occipital gyrus (BA19/37)	−46, −70, −8	3.27
Right cerebellum	4, −74, −22	3.50
Left anterior cingulate	−10, 48, 0	3.17
Right anterior cingulate	14, 4, 40	3.09

TABLE 3

Local Maxima for Areas whose Activation Increased over Time for the Consistent Audiovisual Pairs versus the Single Auditory Stimuli

Brain location	Coordinates (x, y, z)	Z score
Left hippocampus	−28, −44, 2	3.20
Right superior frontal gyrus	10, 26, 56	3.59

campus and right superior frontal gyrus. Additional differential responses were also observed in middle occipital gyrus bilaterally and left inferior occipital gyrus in the latter comparison (see Fig. 5 and Table 4).

The differential responses observed in posterior hippocampus and superior frontal cortex at each session for each paired condition versus the auditory control are illustrated graphically (Fig. 6). A general pattern can be observed whereby the greatest changes (increases or decreases) took place during the first experimental session and then became smaller over the subsequent sessions, possibly due to habituation effects of repeated presentations of the material. It can also be seen that the differential response originates mainly from an overall decrease in the auditory control condition relative to the paired conditions.

DISCUSSION

In this study learning to associate nonverbal auditory and visual stimuli involves operations such as sensory processing as well as establishing whether auditory–visual stimuli are consistently or inconsistently paired. On one hand, repeated presentations of specific pairs will lead to automatic recognition where the auditory stimulus, presented first, predicts subsequent presentation of a specific visual stimulus for consistent but not inconsistent pairs. An additional component involves monitoring with respect to whether a given pair belongs to a consistent or an inconsistent category.

TABLE 4

Local Maxima for Areas whose Activation Increased over Time for the Inconsistent Audiovisual Pairs versus the Single Auditory Stimuli

Brain location	Coordinates (x, y, z)	Z score
Left hippocampus	−28, −42, 0	4.26
Right superior frontal gyrus	14, 6, 50	3.42
Right middle occipital gyrus	40, −78, 28	3.35
Left inferior occipital gyrus	−16, −100, −4	3.18
Left middle occipital gyrus	−22, −30, 32	3.13

Time-Dependent Changes for Consistent versus Inconsistent Audiovisual Pairs

Learning to discriminate consistent from inconsistent pairs requires formation of specific associations between sounds and visual stimuli. As subjects learn to discriminate consistent pairs, presentation of preceding sounds provides a context for retrieval of the associated visual stimuli. The learning process involved in our experiment will comprise both encoding and retrieval operations, with encoding being more dominant in the initial stages and retrieval becoming more prevalent in later stages. In this regard it is striking that the brain regions activated over time included medial parietal and right dorsolateral prefrontal cortex, regions widely reported as part of a retrieval network for either auditory or visual stimuli alone (Tulving *et al.,* 1994; Shallice *et al.,* 1994; Fletcher *et al.,* 1995a; Buckner *et al.,* 1996). In our view, retrieval strategies operate concurrently with encoding operations. This suggestion is supported by findings that when previously presented material is re-presented, the repeated rehearsal of already learned material is associated with the activation of right prefrontal and precuneus, a well-described retrieval system (Kopelman *et al.,* 1998).

In terms of functional specification, the right prefrontal cortex is implicated in retrieval processes involving success, effort, or attempt (Wagner *et al.,* 1998; Schacter and Buckner, 1998; Buckner *et al.,* 1998a, b; Rugg *et al.,* 1996). In our experiment, the success hypothesis is a possible explanation since time-related changes in performance correlate with the amount of material retrieved. More specifically, Rugg *et al.*'s (1998) suggestion that right anterior prefrontal cortex is involved in evaluating whether retrieved information in response to a test item represents an appropriate prior episode is consistent with our finding. In addition, the right dorsolateral prefrontal cortex has been linked to monitoring of information retrieved from episodic memory (Shallice *et al.,* 1994; Henson *et al.,* in press) and to mediating monitoring processes necessary for optimal recall (Fletcher *et al.,* 1998). This putative monitoring role for the dorsolateral prefrontal cortex applies particularly to our paradigm where subjects assess response success over time given that each repetition of consistent pairs allows confirmation of prior correct or incorrect responses.

Kelley *et al.* (1998) reported lateralization within the dorsal prefrontal and the medial temporal cortices according to the nature of the material encoded, such that right prefrontal activations were found during encoding of nonverbal material. Our results might support this suggestion insofar as the material we used was nonverbal and prefrontal activation was right-sided. However, a critical difference is that in our study right prefrontal activation reflects learning-related changes rather than steady state conditions.

The other time-dependent activation was in left medial parietal cortex. The medial parietal cortex is associated with visual imagery, and more specifically with cued recall, rather than encoding, of imageable verbal paired associates (Grasby *et al.,* 1993, 1994; Fletcher *et al.,* 1995b; Mellet *et al.,* 1996; Platel *et al.,* 1997). In our medial parietal cortex activation increased over time paralleling an increasing familiarity for the consistent pairs. With time there is an increase in the degree to which presentation of auditory stimuli automatically cues the associated visual stimulus, which, due to its very nature (Chinese ideogram), cannot be verbalized. Thus, we conclude that this activation reflects processes necessary for generating a visual image of ideograms under the consistent condition. Note that under the inconsistent condition the cue, i.e., the sound, cannot generate a visual image as there is no consistent visual stimulus pairing.

Time-Dependent Changes for Inconsistent Compared with Consistent Audiovisual Pairs

During learning, subjects increasingly distinguish inconsistent pairs from consistent pairs, a task that requires generation of hypotheses in relation to whether the pairs were consistent. Such hypotheses could be subsequently revised with practice and learning. It is likely that attentional demands under the inconsistent condition were greater than under the consistent condition solely on the grounds that a greater number of audiovisual combinations are presented. Notably, we found increasing time-dependent relative activations in two anterior cingulate foci, ventral and dorsal, for the inconsistent condition, relative to the consistent condition. It has been argued that the anterior cingulate is a component of a network responsible for selective attention regardless of sensory modality (Posner and Petersen, 1990; Benedict *et al.,* 1998), as is the right angular gyrus, the other principle focus of activation (Kosslyn *et al.,* 1996). We suggest that these activations in conjunction with the observed bilateral extrastriate activations over time fit with greater attentional demands under this condition.

The more dorsal cingulate region we observed has in the past been implicated in selection between competing alternatives (Pardo *et al.,* 1990; Elliot and Dolan, 1998) and in error detection (Carter *et al.,* 1998). In our study, subjects must make a prediction of whether the pairs are consistent or inconsistent, and this process is more demanding for inconsistent trials. We can only speculate in relation to the more ventral cingulate activation. Animal studies show that lesions in a region corresponding to our observed ventral focus of the cingulate lead to a failure of learned response extinction (Roberts *et al.,* 1992). We suggest that our activa-
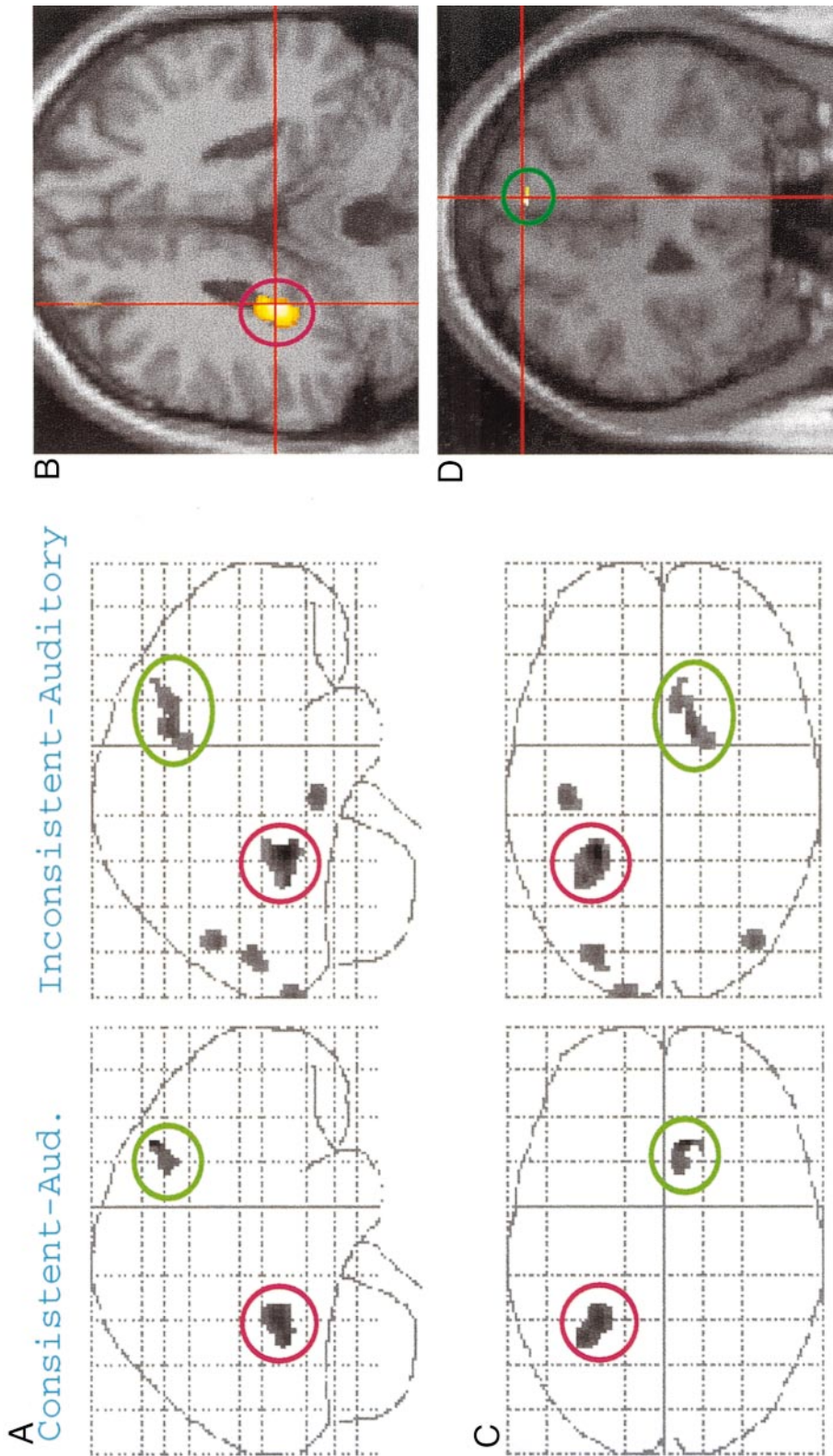
FIG. 5. (a, c) Lateral and (b, d) transverse views of the brain regions activated when comparing time by condition interactions between the paired conditions [(a, b) consistent and (c, d) inconsistent] and the auditory unpaired stimuli. Pink, left posterior hippocampus; green, superior frontal gyrus (BA6). (e) Coronal section of the left posterior hippocampal activation (−28, −42, 0). (f) Coronal section of the superior frontal cortex activation (10, 26, 56).
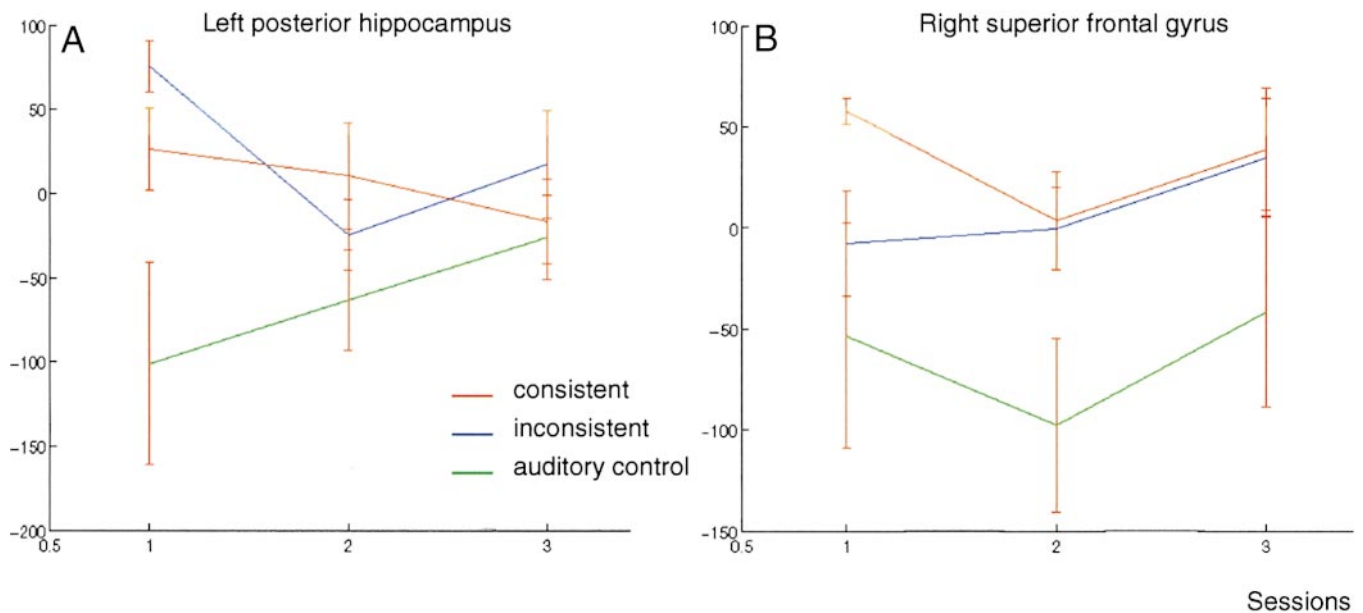
Within session time effects

FIG. 6. Within-session time effects in (a) left posterior hippocampus and (b) right superior frontal gyrus for the three sessions. Positive ordinate values represent event-related responses that increase over time within sessions (in arbitrary units); negative values represent responses that decrease over time within a session. Red lines represent consistent audiovisual pairs, blue lines represent inconsistent audiovisual pairs, and green lines represent auditory controls.

tion might reflect a similar process whereby initial hypotheses regarding pairs are subsequently rejected.

Our analysis of time-dependent changes controls for nonspecific effects by the use of a reference state (consistent versus inconsistent and vice versa). The legitimacy of these comparisons rests on an assumption that nonspecific effects, such as time-dependent effects, are equally expressed in both states. This possibility of differential interaction between task and nonspecific factors cannot be ruled out and should caution interpretation of our findings. Furthermore, other nonspecific interaction effects may be found underlying the difference of frequency in item presentation for consistent versus inconsistent conditions during each session. These possible interactions are not likely to override the main results presented and we note that evidence for differential interactions between task and nonspecific factors is limited (Petersson *et al.,* 1999).

Time-Dependent Changes for Consistent and Inconsistent Audiovisual Pairs Contrasted with Single Auditory Stimuli

These contrasts highlight differential time-dependent within-session responses for bimodal associative, as opposed to single-item, processing. The main differential responses for paired conditions, compared to unpaired conditions, were in left posterior hippocampus and right superior frontal gyrus. The fact that both

kinds of pairs show similar time-dependent responses reflect commonalities specific to associative processing across two sensory modalities. An extensive animal and human literature highlights the critical role of the hippocampus in associative learning (Bunsey and Eichenbaum, 1996; Eichenbaum, 1992; Henke *et al.,* 1997; Rombouts *et al.,* 1997). A more recent proposal is that the hippocampus is related to associative processes for discontiguous items in space or time (Levy and Steward, 1983; Levy, 1996; Wallenstein *et al.,* 1998). The hippocampal responses we observed are consistent with both these suggestions. Note for both consistent and inconsistent conditions, our task involves not only items presented in different sensory modalities but also a temporal dimension where auditory and visual stimuli were presented at different points in time but always separated by the same interval. From this perspective, the differential response for bimodal relative to unimodal conditions fits with neuropsychological findings that left hippocampal lesions lead to impairments in nonspatial conditional association tasks (Petrides, 1985), where subjects must make associations between two items presented one shortly after the other. Our task has many similarities to this task. In Petrides's study, frontal-lobe and left temporal-lobe patients (incorporating the left hippocampal region) were impaired in a nonspatial associational task involving learning associations between six colored-light caps and six

hand postures matched in unique one-to-one combinations. These patients were able to discriminate between single stimuli but experimented difficulty when having to select the correct movement to a given visual stimulus.

The hippocampal response we observed was located in posterior hippocampus. A functional rostrocaudal gradient along the hippocampus has been suggested in a meta-analysis of Lepage *et al.* (1998), who proposed a model (the HIPER model) claiming that encoding is reflected primarily rostrally and retrieval caudally, in the hippocampus. Recent empirical data have also provided support for this suggestion (Strange *et al.,* 1999). On the other hand, Schacter and Wagner (1999), from a review of fMRI studies, suggest that the posterior medial temporal lobe is associated with episodic encoding, whereas the anterior part might play a role in relational encoding. Our posterior hippocampal differential response reflecting changes with time would seem to support a retrieval-based proposal, insofar as retrieval processes are increasingly engaged with learning. Moreover, the left medial temporal region has been found activated during incremental learning, i.e., repetitive trials, a process that might be involved in the consolidation of new memories (Kopelman *et al.,* 1998).

Finally, the other activation, namely in superior prefrontal cortex, is intriguing in light of the fact that patients with lesions to the frontal lobe, encroaching on our observed area, are also impaired on nonspatial associative tasks (Petrides, 1985). The excisions performed in these patients are rarely restricted to particular anatomically defined areas of the frontal cortex, thus making it impossible to disentangle which specific area(s) is responsible for the deficit. However, this study shares commonalities with our study in that two different stimuli were associated in a consistent fashion, spaced in time so that one becomes predictive of its match over time, and the results show distinctively that left temporal-lobe patients with lesions extending to the left hippocampal region were impaired, as were frontal-lobe patients.

Conclusion

In this experiment we investigated the time-dependent changes in neural response during learning where associations were made between auditory and visual stimuli. Strikingly, areas identified in bimodal learning are part of a well-documented retrieval memory system for items within a single modality. Posterior hippocampal time-dependent responses during associative processing are consistent with a suggested role for this region in forming links across sensory modality and time. Our findings highlight the flexibility afforded by single-trial functional magnetic resonance imaging in studies of memory that allow mixed trial stimulus presentation and assessment of neural responses as a function of time.

REFERENCES

Baddeley, A. 1996. The fractionation of human memory. *Proc. Natl. Acad. Sci. USA* **93:**13468–13472.

Banati, *et al.* 1997. The functional anatomy of sensory integration in man. *Neuroimage* **5:**S213.

Benedict, R. H. M., Lockwood, A. H., Shucard, J. L., Shucard, D. W., Wack, D., and Murphy, B. W. 1998. Functional neuroimaging of attention in the auditory modality. *NeuroReport* **9**(1):121–126.

Bliss, T. V. P., and Collingridge. 1993. A synaptic model of memory—Long-term potentiation in the hippocampus. *Nature* **361**(6407): 31–39.

Bliss, T. V. P., and Lømo, T. J. 1973. Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetised rabbit following stimulation of the perforant path. *J. Physiol.* **232:**331–356.

Buckner, R. L., Raichle, M. E., Miezin, F. M., and Petersen, S. E. 1996. Functional anatomic studies of memory retrieval for auditory words and visual pictures. *J. Neurosci.* **16:**6219–6235.

Buckner, R. L., Koutstaal, W., Schacter, D. L., Dale, A. M., Rotte, M., and Rosen, B. R. 1998a. Functional anatomic study of episodic retrieval. II. Selective averaging of event-related fMRI trials to test the retrieval success hypothesis. *Neuroimage* **7:**163–175.

Buckner, R. L., Koutstaal, W., Schacter, D. L., Wagner, A. D., and Rosen, B. R. 1998b. Functional anatomic study of episodic retrieval using fMRI. I. Retrieval effort versus retrieval success. *Neuroimage* **7:**151–162.

Bunsey, M., and Eichenbaum, H. 1996. Conservation of hippocampal memory function in rats and humans. *Nature* **379:**255–257.

Cabeza, R., and Nyberg, L. 1997. Imaging cognition: An empirical review of PET studies with normal subjects. *J. Cogn. Neurosci.* **9**(1):1–26.

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., and Cohen, J. D. 1998. Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science* **280**(5364):747–749.

Chee, M. W., *et al.* 1997. Auditory and visual word processing studied with fMRI. *Neuroimage* **5:**S557.

Davenport, R. K., and Rogers, C. M. 1970. Intermodal equivalence of stimuli in apes. *Science* **168:**279–280.

Davenport, R. K., Rogers, C. M., and Russell, I. S. 1973. Cross modal perception in apes. *Neuropsychologia* **11:**21–28.

Davenport, R. K., Rogers, C. M., and Russell, I. S. 1975. Cross-modal perception in apes: Altered visual cues and delay. *Neuropsychologia* **13:**229–235.

Eacott, M. J., and Heywood, C. A. 1995. Perception and memory: Action and interaction. *Crit. Rev. Neurobiol.* **9**(4):311–320.

Eichenbaum, H. 1992. The hippocampal system and declarative memory in animals. *J. Cogn. Neurosci.* **4**(3):217–231.

Elliot, R., and Dolan, R. J. 1998. Activation of different anterior cingulate foci in association with hypothesis testing and response selection. *Neuroimage* **8:**17–29.

Fletcher, P. C., Frith, C. D., Grasby, P. M., Shallice, T., Frackowiak, R. S. J., and Dolan, R. J. 1995a. Brain systems for encoding and retrieval of auditory–verbal memory. *Brain* **118:**401–416.

Fletcher, P. C., Frith, C. D., Baker, S. C., Shallice, T., Frackowiak, R. S. J., and Dolan, R. J. 1995b. The mind's eye—Precuneus activation in memory-related imagery. *Neuroimage* **2:**195–200.

Fletcher, P. C., Shallice, T., Frith, C. D., Frackowiak, R. S. J., and Dolan, R. J. 1998. The functional role of prefrontal cortex in episodic memory. *Brain* **121**:1249–1256.

Friston, K. J., Frith, C. D., Passingham, R. E., Liddle, P. F., and Frackowiak, R. S. J. 1992. Motor practice and neurophysiological adaptation in the cerebellum: A positron tomography study. *Proc. R. Soc. London B* **248**:223–228.

Friston, K. J., Holmes, A. P., Poline, J. B., Grasby, P. J., Williams, S. C. R., Frackowiak, R. S. J., and Turner, R. 1995a. Analysis of fMRI time-series revisited. *Neuroimage* **2**:45–53.

Friston, K. J., Ashburner, J., Frith, C. D., Poline, J. B., Heather, J. D., and Frackowiak, R. S. J. 1995b. Spatial registration and normalisation of images. *Hum. Brain Mapping* **3**(3):165–189.

Friston, K. J., Williams, S., Howard, R., Frackowiak, R. S. J., and Turner, R. 1996. Movement-related effects in fMRI time-series. *Magn. Reson. Med.* **35**(3):346–355.

Gabrieli, J. D. E. 1998. Cognitive neuroscience of human memory. *Annu. Rev. Psychol.* **49**:87–115.

Grasby, P. M., Frith, C. D., Friston, K. J., Bench, C., Frackowiak, R. S., and Dolan, R. J. 1993. Functional mapping of brain areas implicated in auditory–verbal memory function. *Brain* **116**:1–20.

Grasby, P. M., Frith, C. D., Friston, K. J., Simpson, J., Fletcher, P. C., and Frackowiak, R. S. 1994. A graded task approach to the functional mapping of brain areas implicated in auditory–verbal memory. *Brain* **117**:1271–82.

Hebb, D. O. 1949. *The Organization of Behaviour.* Wiley, New York.

Henke, K., Buck, A., Weber, B., and Wieser, H. G. 1997. Human hippocampus establishes associations in memory. *Hippocampus* **7**(3):249–256.

Henson, R. N. A., Shallice, T., and Dolan, R. J. 1999. Right prefrontal cortex and episodic memory retrieval: A functional MRI test of the monitoring hypothesis. *Brain,* **122**(7):1367–1381.

Holmes, A. P., Josephs, O., Büchel, C., and Friston, K. J. 1997. Statistical modeling of low frequency confounds in fMRI. *Neuroimage* **5**:S480.

Jenkins, I. H., Brooks, D. J., Nixon, P. D., Frackowiak, R. S. J., and Passingham, R. E. 1994. Motor sequence learning—A study with positron emission tomography. *J. Neurosci.* **14**(6):3775–3790.

Jueptner, M., Stephan, K. M., Frith, C. D., Brooks, D. J., Frackowiak, R. S. J., and Passingham, R. E. 1996. The anatomy of motor learning. I. The frontal cortex and attention to action. *J. Neurophysiol.* **77**:1313–1337.

Kelley, W. M., Miezin, F. M., McDermott, K. B., Buckner, R. L., Raichle, M. E., Cohen, N. J., Ollinger, J. M., Akbudak, E., Conturo, T. E., Snyder, A. Z., and Petersen, S. E. 1998. Hemispheric specialisation in human dorsal frontal cortex and medial temporal lobe for verbal and nonverbal memory encoding. *Neuron* **20**:927–936.

King, A. J., and Palmer, A. R. 1985. Integration of visual and auditory information in bimodal neurones in the guinea-pig superior culliculus. *Exp. Brain Res.* **60**:492–500.

Kopelman, M. D., Stevens, T. G., Foli, S., and Grasby, P. 1998. PET activation of the medial temporal lobe in learning. *Brain* **121**:875–887.

Kosslyn, S. M., Shin, L. M., Thompson, W. L., McNally, R. J., Rauch, S. L., Pitman, R. K., and Alpert, N. M. 1996. Neural effects of visualising and perceiving aversive stimuli: A PET investigation. *NeuroReport* **7**(10): 1569–1576.

Lepage, M., Habib, R., and Tulving, E. 1998. Hippocampal PET activations of memory encoding and retrieval: The HIPER model. *Hippocampus* **8**:313–322.

Levy, W. B. 1996. A sequence predicting CA3 is a flexible associator that learns and uses context to solve hippocampal-like tasks. *Hippocampus* **6**(6):579–590.

Levy, W. B., and Steward, O. 1983. Temporal contiguity requirements for long-term associative potentiation/depression in the hippocampus. *Neuroscience* **8**:791–797.

McCarthy, R., and Warrington, E. K. 1990. The dissolution of semantics. *Nature* **343**(6259):599.

Mellet, E., Tzourio, N., Crivello, F., Joliot, M., Denis, M., and Mazoyer, B. 1996. Functional anatomy of spatial mental imagery generated from verbal instructions. *J. Neurosci.* **16**(20):6504–6512.

Meredith, M. A., Wallace, M. T., and Stein, B. E. 1992. Visual, auditory and somatosensory convergence in output neurons of the cat superior colliculus: Multisensory properties of the tecto-reticulo-spinal projection. *Exp. Brain Res.* **88**:181–186.

Nolde, S. F., Johnson, M. K., and Raye, C. L. 1998. The role of prefrontal cortex during tests of episodic memory. *Trends Cogn. Sci.* **2**(19):399–406.

Pardo, J. V., Pardo, P. J., Janer, K. W., and Raichle, M. E. 1990. The anterior cingulate cortex mediates processing selection in the Stroop attentional conflict paradigm. *Proc. Natl. Acad. Sci. USA* **87**:256–259.

Petersen, S. E., van Mier, H., Fiez, J. A., and Raichle, M. E. 1998. The effects of practice on the functional anatomy of task performance. *Proc. Natl. Acad. Sci. USA* **95**:853–860.

Petersson, K. M., Gisselgård, J., and Ingvar, M. 1999. Free recall of pseudowords modulated by practice. *Neuroimage* **9**:S967.

Petrides, M. 1985. Deficits on conditional associative-learning tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia* **23**(5):601–614.

Petrides, M. 1991. Learning impairments following excisions of the primate frontal cortex. In *Frontal Lobe Function and Dysfunction* (H. S. Levin and H. M. Eisenberg, Eds.), pp. 256–272. Oxford Univ. Press, New York.

Platel, H., Price, C., Baron, J. C., Wise, R., Lambert, J., Frackowiak, R. S. J., Lechevalier, B., and Eustache, F. 1997. The structural components of music perception—A functional anatomical study. *Brain* **120**(2):229–243.

Posner, M. I., and Petersen, S. E. 1990. The attention system of the human brain. *Annu. Rev. Neurosci.* **13**:25–42.

Raichle, M. E., Fiez, J. A., Videen, T. O., MacLeod, A. M. K., Pardo, J. V., Fox, P. T., and Petersen, S. E. 1994. Practice-related changes in human brain functional anatomy during nonmotor learning. *Cereb. Cortex* **4**(1):8–26.

Roberts, A. C., Robbins, T. W., Everitt, B. J., and Muir, J. L. 1992. A specific form of cognitive rigidity following excitotoxic lesions of the basal forebrain in marmosets. *Neuroscience* **47**:251–64.

Rombouts, S. A. R. B., Machielsen, W. C. M., Witter, M. P., Barkhof, F., and Lindeboom, J. 1997. Visual association encoding activates the medial temporal lobe: A functional magnetic resonance imaging study. *Hippocampus* **7**(6):594–601.

Rugg, M. D., Fletcher, P. C., Frith, C. D., Frackowiak, R. S. J., and Dolan, R. J. 1996. Differential activation of the prefrontal cortex in successful and unsuccessful memory retrieval. *Brain* **119**:2073–2083.

Rugg, M. D., Fletcher, P. C., Allan, K., Frith, C. D., Frackowiak, R. S. J., and Dolan, R. J. 1998. Neural correlates of memory retrieval during recognition memory and cued recall. *Neuroimage* **8**:262–273.

Schacter, D. L., and Buckner, R. L. 1998. On the relations among priming, conscious recollection, and intentional retrieval: Evidence from neuroimaging research. *Neurobiol. Learning Memory* **70**:284–303.

Schacter, D. L., and Wagner, A. D. 1999. Medial temporal lobe activations in fMRI and PET studies of episodic encoding and retrieval. *Hippocampus* **9**(1):7–24.

Shallice, T., Fletcher, P., Frith, C. D., Grasby, P., Fracknowiak, R. S. J., and Dolan, R. J. 1994. Brain regions associated with acquisition and retrieval of verbal episodic memory. *Nature* **368:** 633–635.

Squire, L. R., and Zola, S. M. 1996. Memory, memory impairment, and the medial temporal lobe. *Cold Spring Harbor Symp. Quant. Biol.* **61:**185–195.

Stein, B. E., Meredith, M. A., Huneycutt, W. S., and McDade, L. 1989. Behavioural indices of multisensory integration: Orientation to visual cues is affected by auditory stimuli. *J. Cogn. Neurosci.* **1:**12–24.

Strange, B. A., Fletcher, P. C., Henson, R. N. A., Friston, K. J., and Dolan, R. J. 1999. Segregating the functions of human hippocampus. *Proc. Natl. Acad. Sci. USA* **96**(17):4034–4039.

Streicher, M., and Ettlinger, G. 1987. Crossmodal recognition of familiar and unfamiliar objects by the monkey: The effects of ablation of polysensory neocortex or of the amygdaloid complex. *Behav. Brain Res.* **23:**95–107.

Tailarach, J., and Tournoux, P. 1988. *Co-planar Stereotaxic Atlas of the Human Brain.* Thieme, Stuttgart.

Tovee, M. J., Rolls, E. T., and Ramachandran, V. S. 1996. Rapid visual learning in neurones of the primate temporal visual cortex. *NeuroReport* **7**(15–17):2757–2760.

Tulving, E., Kapur, S., Markowitsch, H. J., Craik, F. I. M., and Habib, R. 1994. Neuroanatomical correlates of retrieval in episodic memory. *Proc. Natl. Acad. Sci. USA* **91:**2012–2015.

Wagner, A. D., Desmond, J. E., Glover, G. H., and Gabrieli, D. E. 1998. Prefrontal cortex and recognition memory. Functional-MRI evidence for context-dependent retrieval processes. *Brain* **121:**1985–2002.

Wallenstein, G. V., Eichenbaum, H., and Hasselmo, M. E. 1998. The hippocampus as an associator of discontiguous events. *Trends Neurosci.* **21**(8):317–323.

Wilson, W. A., and Shaffer, O. 1963. Intermodality transfer of specific discriminations in the monkey. *Nature* **4862:**107.