



Published in final edited form as:

*J Cogn Neurosci.* 2006 July ; 18(7): 1156–1173.

## An fMRI Analysis of the Human Hippocampus: Inference, Context, and Task Awareness

Anthony J. Greene<sup>1</sup>, William L. Gross<sup>1</sup>, Catherine L. Elsinger<sup>2</sup>, and Stephen M. Rao<sup>2</sup>

<sup>1</sup> *University of Wisconsin, Milwaukee*

<sup>2</sup> *Medical College of Wisconsin*

### Abstract

The hippocampus is critical for encoding and retrieving semantic and episodic memories. Animal studies indicate that the hippocampus is also required for relational learning tasks. A prototypical relational learning task, and the one investigated in this experiment, using event-related functional magnetic resonance imaging, is the transitive inference (TI) task. In the TI task, participants were to choose between A and B (A?B) and learned by trial and error to choose A (A > B). There were four such premise pairs during a training (A > B, B > C, C > D, D > E). These can be acquired distinctly or can be organized into a superordinate hierarchy (A > B > C > D > E), which would efficiently represent all the learned relations and allow inferences (e.g., B > D). At test there was no reinforcement: In addition to premise pairs, untrained pairings were introduced (e.g., A?E, B?D). Correctly inferring that B > D is taken as evidence for the formation of a superordinate hierarchy; several alternatives to the superordinate hierarchy hypothesis are considered. Awareness of the formation of this hierarchy was measured by a postscan questionnaire. Four main findings are reported: (1) Inferential performance and task awareness dissociated behaviorally and at the level of hemodynamic response; (2) As expected, performance on the inferred relation, B > D, corresponded to the ability to simultaneously acquire B > C and C > D premise pairs during training; (3) Interestingly, acquiring these “inner pairs” corresponded to greater hippocampal activation than the “outer pairs” (A > B, D > E) for all participants. However, a distinct pattern of hippocampal activity for these inner pairs differentiated those able to perform the inferential discrimination, B > D, at test. Because these inner premise pairs require contextual discrimination (e.g., C is incorrect in the context of B but correct in the context of D), we argue that the TI task is hippocampal-dependent because the premise pair acquisition necessary for inference is hippocampal-dependent; (4) We found B > D related hippocampal activity at test that is anatomically consistent with pre-consolidation recall effects shown in other studies.

### INTRODUCTION

The modern understanding of memory formation began to take shape with the discovery that bilateral damage to the medial-temporal lobes (MTLs), which include the hippocampus and parahippocampal regions, results in dense anterograde amnesia and temporally graded retrograde amnesia for conscious memory (Scoville & Milner, 1957). Conscious memory is generally thought to consist of episodic (autobiographical) and semantic (world knowledge) memory, which are often collectively referred to as declarative memory (DM; Cohen & Squire, 1980). The pattern of amnesia observed in MTL-damaged patients indicates that the MTL is critical for acquisition, temporary access, and the consolidation of new conscious memories

Reprint requests should be sent to Anthony Greene, Department of Psychology, University of Wisconsin, Milwaukee, 2441 East Hartford Avenue, Milwaukee, WI 53211, or via e-mail: ag@uwm.edu..

The data reported in this experiment have been deposited with the fMRI Data Center ([www.fmridc.org](http://www.fmridc.org)). The accession number is 2-2005-120MH.

but is not the locus of permanent storage (Milner, 1972). In addition, certain learning tasks are impaired in MTL amnesics (for a review, see Clark, Manns, & Squire, 2002) and cannot be performed in normal-memory participants without explicit awareness of the task contingencies (Clark, Manns, & Squire, 2001). Conversely, some forms of nonconscious learning and memory (procedural learning, perceptual priming, as well as simple forms of classical and operant conditioning) are spared in MTL amnesia (e.g., Squire, Zola-Morgan, & Chen, 1988; Warrington & Weiskrantz, 1974; Milner, 1972). Accordingly, the MTL is necessary for the formation of conscious forms of learning and memory (Clark & Squire, 1998; Schacter, 1998; Squire & Zola, 1996).

Animal studies reveal that hippocampal ablation results in learning deficits specific to the acquisition of tasks involving complex contingencies but spares simple contingency learning (e.g., Rudy & Sutherland, 1989). The type of tasks that show impairment in hippocampal-lesioned animals are sometimes referred to as “relational learning” (RL) tasks (Cohen & Eichenbaum, 1993) and may be characterized as belonging to two categories: tasks that require learning contextual relations (e.g., negative patterning, transverse patterning, and multiple-contingency conditioning) or tasks that require the associative reorganization of elemental relations into more global representations such as path integration (i.e., learning a path from Point A to B, B to C, C to D, and D to E allows an animal to transverse untrained routes), transitive inference (TI, described below), chaining (e.g., if Stimulus A predicts reward and Stimulus B predicts A, then B will become a conditioned stimulus for reward), and external-cue maze navigation (for a review, see Eichenbaum & Cohen, 2002).

We consider a prototypical hippocampal-dependent task to be the TI task for several reasons: First, successful acquisition of the learning trials entails acquisition of contextual relations. In the most common version of TI, given a choice between Stimulus “A” and Stimulus “B,” one learns by trial and error to select “A” ( $A > B$ ), then  $B > C$ ,  $C > D$ ,  $D > E$ . In the TI task, end items ( $A > B$  and  $D > E$ ) may be acquired as atomic stimuli (i.e., A is always correct and E is always incorrect) but inner pairs ( $B > C$  and  $C > D$ ) must be acquired as context-dependent pairs. That is, Item C is correct in the context of D but incorrect in the context of B (Dusek & Eichenbaum, 1997). Second, successful performance at test requires reorganization of stimuli into a global representation. Testing includes the unreinforced pairing “B” and “D” (B?D) which tests the capacity for inference. Correctly inferring “B” ( $B > D$ ) is prima facie evidence that premise pairs have been reordered into a superordinate hierarchy which supports inference (i.e.,  $A > B > C > D > E$ ). Inference is operationally defined as the capacity to make novel decisions on the basis of relevant prior experience, whether by syllogism (e.g., chaining) or by comparative value (e.g., taller, further east). Successful inference in TI is evidence of a set of multiple associations which will here be referred to as a hierarchy. The capacity for inference has been shown in rats (Davis, 1992); pigeons and crows (Lazareva, Smirnova, Rayevsky, & Zorina, 2000; Weaver, Steirn, & Zentall, 1997); primates (Treichler & Van Tilburg, 1996; Gillan, 1981); and humans (Acuna, Sanes, & Donoghue, 2002; Nagode & Pardo, 2002; Greene, Spellman, Dusek, Eichenbaum, & Levy, 2001). The inferential process is highly adaptive in that a sample of data may be extrapolated into a more globally applicable framework of contingencies or ideas (e.g., Macphail, 1996), which therefore supports inference. Third, the TI task may be performed with or without conscious awareness of the task contingencies (Greene et al., 2001), which allows a comparison to DM tasks. TI and DM share several properties: Episodic and semantic memory are not optimally encoded as atomic learning events but rather are organized into representations of self and world knowledge (for this argument, see Clark & Squire 1998) that are both context-dependent (e.g., Bustamante, Jordan, Vila, Gonzalez, & Insua, 1970), associatively organized (Kumaran & Maguire, 2005; Craik & Tulving, 1975; Bransford & Franks, 1971), and capable of supporting the flexibility of predictive inference under novel conditions (e.g., Hartley, Maguire, Spiers, & Burgess, 2003; Nagode et al., 2002; Carlson, 1992). The capacity for context-dependent and inferential

expression of associative learning confers a flexibility that is likely the principal benefit of experience (for this argument, see Cohen, Poldrack, & Eichenbaum, 1997). Thus, RL and DM processes have comparable task demands (i.e., contextual, conditional, and/or superordinate organization) and are both mediated by the hippocampus.

Two general approaches have been taken to understand the role of the hippocampus: First, it has been proposed that RL tasks all require conscious awareness of the learned contingencies (e.g., Manns, Hopkins, Reed, Kitchener, & Squire, 2003; Manns & Squire, 2001; Clark & Squire, 1998; Rempel-Clower, Zola, Squire, & Amaral, 1996). This approach emphasizes that higher-order conscious processes serve to actively test hypotheses and organize contingencies. The other approach has been to assert that DM is an RL task (e.g., Cohen & Eichenbaum, 1993). By this approach, DM is construed as a contextual learning task given that context is fundamental to acquisition and expression (e.g., Craik & Tulving, 1975). One implication of this perspective may be that conscious aspects of DM are a second-order property of hippocampal function in that the context is always invoked or implied to initiate declarative retrieval (e.g., “Recall as many items as you can from the study session” or “What did you do last summer?”). The distinction between these two views is essentially that the latter proposes that DM is a subset of RL, and the former proposes either that RL is a subset of DM or that RL and DM are interchangeable concepts. The most common way to test between these two hypotheses has been to test the dependency of RL on task awareness. Recent studies (Chun & Phelps, 1999; Schacter, Church, & Bolton, 1995) have shown that some priming tasks show MTL dependency with little or no awareness of task contingencies. This evidence suggests that MTL-dependent tasks exist that do not necessarily entail task awareness and therefore supports the hypothesis that MTL learning is fundamentally relational.

To date, evidence that implicit tasks involve the MTL has largely been limited to observed impairments in amnesics (Chun & Phelps, 1999; Schacter et al., 1995). However, the amnesics involved in these studies come from varied etiologies that include degrees of hippocampal sparing, damage concomitant with hippocampal impairment, and the possibility of nonhippocampal amnesia such as diencephalic amnesia. It may be the case, for example, that the priming impairments observed in such amnesics result from parahippocampal damage, whereas conscious memory impairments may be due to concomitant hippocampal disconnect or damage (e.g., Manns & Squire, 2001). One approach for addressing this ambiguity is to obtain converging functional imaging evidence to examine hippocampal and parahippocampal activity during implicit tasks thought to be MTL mediated. If task-related activation were to be observed in the hippocampus *without* conscious awareness, it would provide further evidence that the hippocampus serves a broader role than DM.

Three recent imaging studies suggest that MTL structures may be functionally active during tasks even when there is no evident task awareness. In a serial reaction time (RT) task, MTL (nonhippocampal) activations were observed in both implicit and explicit task conditions (Schendan, Searl, Melrose, & Stern, 2003). Using a semantic association task, Henke et al., (2003) observed hippocampal activity during both priming of semantic associations and explicit retrieval of those associations; however, conscious awareness of facilitated associations was not ruled out. Finally, in a conditional discrimination task, MTL activation was observed in both aware and unaware participants, but only those participants who were aware of the contingencies could perform the discrimination (McIntosh, Rajah, & Lobaugh, 2003).

The TI task is presently the only established RL task that does not require task awareness of the critical inferential discrimination (Greene et al., 2001). Three recently published functional imaging studies have examined patterns of activation during performance of a TI task. In a functional magnetic resonance imaging (fMRI) study using simple geometric shapes as stimuli,

activations during test were observed in the frontal and posterior temporal lobes (Acuna, Eliassen, Donoghue, & Sanes, 2002). In a positron emission tomography study using faces as stimuli, bilateral posterior hippocampal activation was observed at study, providing evidence that TI is a hippocampal-dependent task in humans (Nagode et al., 2002). An fMRI study using faces and houses as stimuli showed bilateral middle hippocampal activation at test corresponding to inference (Preston, Shrager, Dudukovic, & Gabrieli, 2004). Finally, using novel visual stimuli, functional activations were observed at test in the right anterior hippocampus for inferred pairs, and recognition of nonoverlapping pairs elicited bilateral MTL activation (Heckers, Zalesak, Weiss, Ditman, & Titone, 2004). However, in all four studies, participants were informed at the outset as to the existence of a superordinate hierarchy which supports inference. The present experiment explores the role of the hippocampus and the MTL in RL with or without task awareness using the TI task.

## METHODS

### Participants

Participants were 21 right-handed undergraduate and graduate students (7 men and 14 women) from the University of Wisconsin Milwaukee and the Medical College of Wisconsin with ages ranging from 18 to 34, participating for course extra credit or financial reimbursement. All participants were neurologically normal and naive as to the purpose of the study. Participants were unfamiliar with the Hiragana script used in this study.

### Materials and Task Procedures

Stimulus items were five characters selected from the Japanese Hiragana script. Characters were initially selected as those judged to be highly discriminable. Of those, pilot testing revealed that the five characters shown in Figure 1 were the most easily discriminable (Greene et al., 2001).

Visual stimuli were computer-generated and rear-projected on an opaque screen located at the participant's feet. Participants viewed the screen through prism glasses, which could be fitted with corrective lenses if necessary. The Hiragana characters were 30 cm tall and the viewing distance was 230 cm. Responses were recorded with a custom-made, nonferrous, two-button piano-keypress device, which rested on the participants' right thigh and was occluded from view.

The training and test tasks used in the current experiment closely follow a previous study (Greene et al., 2001, Experiment 1). Pairs of characters were visually presented on the screen. Participants were instructed to select one character from each display, using their right hand, by pressing either the left piano key with their index finger (for the left character) or the right piano key with their middle finger (for the right character). Correct and incorrect responses were recorded along with stimulus-to-response latency, and no-response trials were counted as incorrect.

**Training Phase**—For each training pair, one character was always correct and the other always incorrect. The position (left–right) of the correct stimulus was randomized. Specific instructions to the participants were as follows: “In this experiment, two symbols will appear simultaneously on the computer screen. You are to select the correct symbol. At first, this will be by trial and error; however, with practice you will find that the correct figure is easily learned” (see Appendix A for the full instructions). The character pairs were presented for 3 sec, followed by feedback for 1.5 sec. Feedback consisted of the display of either “correct” (in blue), or “incorrect” (in red). Participants were trained to learn which of each pair was correct from a set of four overlapping premise pairs ( $A > B$ ,  $B > C$ ,  $C > D$ ,  $D > E$ , where “ $>$ ” indicates

that the correct choice is on the left), which could be encoded as distinct pairs or as a sequential hierarchy ( $A > B > C > D > E$ ) where all learned and unlearned relations are encoded (for evidence and discussion of a linear hierarchy, see Nakamaru & Sasaki, 2003; Dusek & Eichenbaum, 1997; Treichler & Van Tilburg, 1996). In each session, one in three trials was a fixation trial (appearance of a black dot in the center of the screen) randomly interspersed throughout the experimental run. Participants were instructed not to respond during these fixation periods. Training was divided into four blocks, each block consisting of 40 training trials. In Block 1, pairs were trained out of sequence with no adjacent pairs sharing an item (i.e., either BC, DE, AB, CD or CD, AB, DE, BC), as sequential training may serve as a clue that a hierarchy exists (i.e., when presenting AB then BC then CD, only one stimulus changes at a time). Each pair was presented five times in a row, and the list was repeated once (e.g.,  $5 \times BC, 5 \times DE, 5 \times AB, 5 \times CD, 5 \times BC, 5 \times DE, 5 \times AB, 5 \times CD$ ). In Block 2, pairs were trained in the same sequence, each pair was presented twice in a row, and the list was repeated an additional five times. In Block 3, pairs were again presented in this sequence, one presentation each, and the list was repeated an additional nine times. In Block 4, pairs were presented in random order, 10 times each. As mentioned earlier, there were 20 fixation trials in each block of 40 training trials. Having a variable delay period between trials accomplishes two goals: The introduction of the delay period minimizes participant expectations by reducing temporal regularity in the task and also introduces a random interstimulus interval (ISI) that aids in deconvolution analysis (described in Functional Image Generation section below).

Participants had an opportunity to become familiar with the task prior to entering the scanner. In contrast to the TI task, in the practice task the stimulus pairs did not overlap (i.e.,  $L > M, N > O, P > Q$ , etc.). The practice pair items were selected from the standard keyboard symbols (e.g., “#,” “&,” etc.). Participants performed as many trials as necessary to learn that a given item from each pair was correct.

**Test Phase**—Following completion of all four training blocks, participants were tested without feedback. Test items included all four premise pairs, the transitive pair BD to assess capacity for TI, and pairs containing end items (e.g., AE, CE, AD, etc.). As in Training, one of every three trials was randomly designated as a fixation trial. Stimuli were presented for 2.25 sec, corresponding to one fMRI image (the trial length changed because we omitted feedback). Three test blocks were conducted, each block consisting of 20 randomized presentations of each pair, interspersed within a random presentation of 90 fixation trials (as in training), for a total of 270 trials per block.

**Postscan Questionnaire**—After the scan session, participants were given as much time as they needed to complete a postscan questionnaire. The questionnaire was designed to assess participants' awareness of the hierarchical nature of the BD task without confounding BD performance (for questionnaire and scoring instructions, see Appendix B). Six raters independently evaluated postscan questionnaires after training on proper scoring criteria. Raters were trained on the scoring procedure using a practice questionnaire. Criteria for scoring are as follows: Awareness scores range from 1 (*no evidence of awareness of the hierarchical structure*) to 5 (*definite indications of awareness*). Appendix B provides a complete explanation of the scoring procedure. Final awareness scores were calculated by using the average of all six raters' scores. Interrater reliability was very high: Cronbach's  $\alpha^2 = .95$  (Bland & Altman, 1997); lowest pairwise  $r^2 = .69$ . Pilot testing revealed that participants who become aware of the hierarchy do so gradually (see also Greene et al., 2001, Experiment 2), and aware participants have no recollection of where in the experiment they began to figure it out, or how long they entertained the idea before concluding that it was true.

**Behavioral Analyses**—Dependent measures of interest were: (1) Proportion of correct trials on premise pair performance from each training block; (2) Proportion of premise pairs correct

from each testing block; (3) Proportion of transitive pairs (BD) correct for each testing block; (4) Proportion of outer pairs correct (AB, DE) at test; (5) RT as measured from the time of stimulus onset for all measures listed above; and (6) a posttest assessment of hierarchical awareness. MRC analyses were used to assess the effects of pair type, learning over time, and awareness, on test performance. When appropriate, follow-up *t* tests were used to compare performers with nonperformers on early learning measures. We also used an MRC analysis to examine the effect of performance accuracy and awareness on RTs to stimuli during test.

**Functional Image Acquisition**—Event-related, whole-brain imaging was performed using a research dedicated GE Signa 1.5-Tesla scanner equipped with a custom 30.5-cm i.d. three-axes local gradient coil and an end-capped bird-cage radio-frequency coil. Echo-planar (EP) images were collected using a single-shot, blipped, gradient-echo EP pulse sequence; echo time (TE) = 40 msec, with 40 msec of image acquisition time. The interscan period (TR) was 2250 msec. Image resolution was  $64 \times 64$  voxels with a 24-cm field of view (FOV). Nineteen contiguous sagittal 7-mm-thick slices were selected in order to provide coverage of the entire brain ( $3.75 \times 3.75 \times 7$  mm typical voxel size). Although distinguishing MTL regions is sometimes a challenge in functional imaging, past research conducted on this 1.5-T magnet, using these parameters, has repeatedly detected robust and reliable hippocampal activations (Douville et al., 2005; Cabeza, Rao, Wagner, Mayer, & Schacter, 2001; Leveroni et al., 2000).

During training, scanning was synchronized with trial onset, with two images (4.5 sec) acquired during each trial, for a total of 120 images per run (60 trials per run). An additional six images were added to the beginning and eight images were added to end of the run to accommodate the delayed rise and fall of the hemodynamic response. This resulted in a total of 134 images (approximately 5 min).

During test, scanning was synchronized with trial onset, with one image acquired during each trial, for a total of 270 images per run. Again, an additional six images were added to the beginning of the run and eight images were added to the end of the run, for a total of 284 images (approximately 10 min).

Prior to functional imaging, 124 high-resolution spoiled gradient-recalled at steady-state (GRASS) sagittal anatomic images [TE = 5 msec; TR (repetition time) = 24 msec,  $40^\circ$  flip angle, NEX (number of excitations) = 1, slice thickness = 1.2, FOV = 24 cm, matrix size =  $256 \times 128$ ] were acquired for each participant. These images served as the high-resolution anatomic images that allow precise localization of functional activity and coregistration.

**Image Generation**—Each image time series was spatially registered in-plane to reduce the effects of head motion using an iterative linear least squares method (Analysis of Functional Neuroimages, or AFNI software; Cox, 1996). A deconvolution analysis was used to generate IRFs of the fMRI signal on a voxelwise basis. This IRF was modeled over seven TRs, with the first time point coincident with the onset of the stimulus. This analysis produced an estimate of the hemodynamic response for each training and test condition relative to a baseline state (rest) without making a priori assumptions regarding the shape, delay, or magnitude of the IRF. Anatomical and functional images were then interpolated to volumes with  $1 \text{ mm}^3$  voxels, coregistered, converted to Talairach stereotaxic coordinate space (Talairach & Tournoux, 1988), and blurred using a 4-mm Gaussian full-width half-maximum filter to compensate for intersubject variability in anatomic and functional anatomy.

In order to examine the effect of awareness, pair type (inner vs. outer), and block (1–4) on the change in MR signal intensity during training, a multiple regression analysis was performed across all time points (TR1–TR7), with performance status at test (performers/non-performers)

as a between-subjects factor. The change in MR signal intensity represents the AUC for all time points in the IRF. For the test phase, a multiple regression analysis was performed to examine the effect of awareness and test block (1, 2, and 3) on the change in MR signal intensity, again with performance status as a between-subjects factor. A statistical threshold of  $p < .001$  and a minimum cluster size threshold of 200  $\mu\text{l}$  (using a voxel size of 1  $\mu\text{l}$  or 1  $\text{mm}^3$ ) were established based on 100,000 Monte Carlo simulations demonstrating that the chance probability of obtaining a significant activation cluster for an entire volume (Type I error) was less than  $3.5 \times 10^{-4}$ . Areas of activation were specified using the Talairach and Tournoux (1988) atlas.

Functional image analyses were performed using MRC techniques to allow for the most flexibility and power by maintaining continuous variables (i.e., awareness, accuracy) as ratio scales rather than nominal scales. During training we examined the effect of training block (1–4), pair type (inner vs. outer), awareness, and performance (BD performers vs. BD nonperformers) on patterns of functional activation. Based on previous behavioral findings (Greene et al., 2001) and pilot data, our main contrast of interest was Pair type  $\times$  Block. Because inner and outer pairs differ in their context dependency, we expected to see differences in functional activation patterns over time. At test, we examined the relationship between those who were BD performers and those who were nonperformers to determine whether patterns of functional activation differed. Similarly, we examined the extent to which task awareness affects patterns of functional activity.

## RESULTS

### Behavioral Results

The principal behavioral result is that training performance for the context-dependent pairs predicts the capacity to do inference at test. For simplicity, we first present the training and test results that do not pertain directly to the principal behavioral finding.

**Training Performance**—The results of training are summarized in Figure 2A. Performance on premise pairs follows the typical hammock shape reported in previous studies (e.g., Greene et al., 2001). Participants performed significantly worse on the inner pairs relative to the outer pairs [ $F(1,17) = 19.183, p = 4.09 \times 10^{-4}$ ]. This is likely due to the inner items being context-dependent, whereas the outer items are noncontextual and can be solved through simple elemental strategies (i.e., A is always correct and E is always incorrect; Dusek & Eichenbaum, 1997). During the first block of training, performance on the inner versus outer pairs was not significantly different [ $F(1,57) = 0.768, ns$ ]. By the final block of training, however, outer pair performance was significantly better than inner pair performance [ $F(1,57) = 22.408, p = 1.50 \times 10^{-5}$ ]. This is described by the interaction of block number and pair type [ $F(3,57) = 5.731, p = .002$ ].

Latency data (see Table 1) reveal that inner pairs were more difficult than outer pairs, and that this effect is more pronounced for those who would become performers, and that the effect increases as training continues. Overall, it took longer to respond to inner than outer pairs, and this effect was larger for those who would later be BD performers than for nonperformers. Additionally, as training continued, there was a general increase in latency, presumably because of the increased difficulty of each training block. However, this main effect of training is overshadowed by the interaction of training block and pair type. Latencies increased over training block for the inner pairs but not significantly for the outer pairs. Recall that training stages began with five sequential presentations of a given pair, which decreased in subsequent training blocks until pair presentation was random in the last training block. The latency data, with respect to the training stages, suggest that in early blocks it was possible to consider inner pairs in isolation, whereas in later training blocks inner items (B, C, or D) required evaluation

with respect to both neighboring items because a given pair was no longer presented several times in sequence.

**Test Performance**—Test results are summarized in Figure 2B. Overall performance on BD pairs averaged 63% ( $SD$  0.05, range 0.0–1.0), which is well above chance [ $t(62) = 2.555, p = .013$ ]. Although BD performance was somewhat lower in this experiment than previously reported (Greene et al., 2001), this is probably due to two factors: (1) the scanning environment often reduces task performance, and (2) in the previously published study, participants repeated training blocks if their performance did not reach a set threshold. In the present experiment, the number of training blocks remained fixed so that we could meaningfully combine imaging data across participants. Participants were classified as BD performers if they scored over 0.715 (corresponding to a binomial probability of less than .01) on the BD trials. Using these criteria, we classified 13 participants as performers and 8 participants as nonperformers.

BD performers almost uniformly responded correctly on the very first BD presentation [85%;  $t(13) = 2.51, p < .03$ ] and did so in an average time of 1389 msec. Performance on BD as a function of test trial is depicted in Figure 3. During test all participants were slower to respond to the inner pairs [mean difference of 112 msec;  $F(1,16) = 21.664, p < 2.64 \times 10^{-4}$ ] than to the outer pairs. Performers and nonperformers did not differ on RT to the first presentation of B?D [155 msec;  $t(13) = 1.11, ns$ ] nor throughout the first block of test [mean difference = 70 msec;  $F(1,17) = 0.637, ns$ ]. Although RTs to B?D decreased over time [mean decrease = 237 msec;  $F(19,323) = 3.431, p < 2.77 \times 10^{-6}$ ], these variables did not interact as a function of performance [ $F(19,323) = 0.891, ns$ ].

**Context Performance at Training Predicts Inference Performance at Test**—The most important behavioral finding is that training performance on the context-dependent inner pairs predicts inferential performance at test. In order to assess the inner context-dependent pairs ( $B > C, C > D$ ) on later BD performance, we performed a mixed multiple regression with BC and CD performance across training blocks as within-subjects factors, and BD performance at test as a between-subjects factor. The three-way interaction was significant [ $F(1,17) = 12.106, p < .003$ ], and analysis of simple effects showed that the effect was limited to the fourth training block. Nonperformers were generally able to respond correctly to either BC or CD but not consistently to both, whereas BD performers were able to perform well on both pairs. BD performance at test was also well correlated with performance on BC and CD during training ( $r = .520, p < .005$ ). Both nonperformers and performers were able to learn the end items at well above chance (.89 and .94, respectively), tending to rule out the possibility that the difference between performers and nonperformers is simply a difference in learning ability or compliance. Note also that if some version of “value transfer” (the association of B to A gives it a high value, whereas the association of D to E gives it a low value) was responsible for BD performance (e.g., Wynne, 1995), then acquisition of the end items would be sufficient for the correct  $B > D$  at test, and this is clearly not the case. Thus, the ability to acquire both of the context-dependent pairs is an excellent predictor of later ability to perform TI successfully.

**Awareness Is Distinct from Inferential Performance**—Awareness scores are shown in Figure 2C. Awareness scores ranged from 1 to 5, ( $M = 2.15, SD = 0.28$ ). Most participants scored at or near the lowest awareness score. Because the awareness scores were not close to normally distributed (bimodal with a positive skew), participants were classified as aware or unaware by a median split of the awareness scores (median = 1.50). Using this criterion, 10 participants were classified as aware and 11 as unaware. The correlation between performance and awareness was not significant ( $r^2 = 0.157, ns$ ) with the majority of subjects scoring high on BD performance but very low on inferential awareness. This is consistent with the finding from Greene et al. (2001), showing near-zero correlations between awareness and performance (Experiment 1:  $r^2 = .01, ns$ ; Experiment 2:  $r^2 = .01, ns$ ). Another principal finding from Greene

et al. was that when training was interrupted as soon as participants reached criteria on all premise pairs, BD performance was at or near its asymptote but performance was quite low, whereas complete training did not improve performance but did improve awareness; this finding thus suggests that BD performance preceded any inferential awareness on the TI task.

To provide further evidence of the dissociation between BD performance and awareness, we performed an analysis of response latency as a function of BD performance and awareness. Interestingly, participants who were categorized as performers responded at longer latencies to the premise pairs at test [mean difference = 114 msec;  $F(1,16) = 5.048, p < .039$ ], whereas participants categorized as aware tended to respond faster [mean difference = 129 msec;  $F(1,16) = 13.586, p < .002$ ], presumably because they explicitly recognized the BD pair and had a solution readily in mind (see also McIntosh et al., 2003). There was no significant interaction of awareness and performance on RT [ $F(1,7) < 1, ns$ ]. The confluence of dissociations suggests that awareness and inferential performance are either distinct processes or may often be distinct processes because inferential performance does not require task awareness (for additional discussion, see Frank, Rudy, & O'Reilly, 2003).

### Functional Imaging Results

The functional imaging analysis yielded four principal findings. First, for all participants at study, context-dependent inner pairs showed greater hippocampal activation late in training than did the outer pairs. Second, at test, BD performers showed greater hippocampal activation to the inferential pair BD than did the BD nonperformers. Third, during study, those who would be BD performers showed greater hippocampal activity at study attributable to the processing of context-dependent inner pairs. Fourth, at test, those participants who were classified as unaware BD performers showed greater parahippocampal activation to the BD pairs than did the aware BD performers.

**The Hippocampus and Contextual Processing at Study**—We examined the relationship between pair type (inner vs. outer) and training block (1–4) on functional activation. The regions of interest (ROIs) are shown in Figure 4 (all activated regions are listed in Table 2). Both of these activations appear primarily within the hippocampus proper. The area under the curve (AUC) of the impulse response functions (IRFs) for the region of activation in the anterior hippocampus shows that this effect is due to a selective increase in activation for the inner pairs relative to the outer pairs. The inner pairs elicit significantly higher activation by the third block of training than the outer pairs. This pattern is the same for the region of activation in the posterior hippocampus.

**The Hippocampus and Inference at Test**—Importantly, we observed hippocampal activation to the B?D pair at test. A posterior area of the left hippocampus, shown with accompanying AUC of the IRFs in Figure 5, exhibited a main effect of performance on B > D (other regions listed in Table 3). This area has previously been shown to be involved in recall of items (e.g., Dolan & Fletcher, 1999; Strange, Fletcher, Henson, Friston, & Dolan, 1999). Although multiple regression and correlation (MRC) analyses revealed no hippocampal activations related to awareness or the interaction of awareness and performance, a central aim of this study is to examine the role of awareness in inference. It is therefore important to demonstrate not just that this performance-related hippocampal activation is unrelated to awareness, but also that the activation is present in analyses constrained to unaware participants. We therefore treated the active hippocampal region as an ROI and constrained the analyses to unaware participants only ( $n = 11$ ), with  $\alpha = .01$ . This hippocampal region showed significant activation in the unaware participants ( $F = 10.580$ ). Together with the behavioral findings, this activation suggests that success on this inference task involves the hippocampus but does not depend upon task awareness.

On initial presentations, activation to the B>D pair could be due to transitivity or novelty. To test this possibility, we contrasted activation observed during B?D with activation observed during novel end-item pairs (A?E, A?C, C?E), and with premise pairs' activations containing an end item (A?B, D?E; presented at training). We did not observe differential activation patterns in any of these combinations. In addition, we observed no repetition-related changes in hemodynamic activity for any of the novel pairs (B?D, A?E, A?C, C?E) that one would expect of a novelty activation. Taken together, these findings suggest that the solutions to the novel pairings (B > D, A > E, A > C, C > E) are already encoded as part of the hierarchy that supports all choices and are therefore not treated as novel.

### **Hippocampal Activation for Context at Study Predicts BD Performance at Test**

—The principal result of this set of experiments was that there was differential bilateral hippocampal activation for contextual pairs late in training, which predicted inferential performance at test (see Figure 6). During the final block of training, activation was greater to the inner items for BD performers only (Pair type  $\times$  Block  $\times$  Performance; additional regions listed in Table 4).

**The MTL and Inference without Task Awareness**—The analysis at test focused on postexperimental awareness and BD performance and their effect on change in MR signal on BD trials only. We found a significant two-way interaction of BD performance and awareness in the left middle parahippocampus (see Figure 7; additional regions are listed in Table 5). This parahippocampal activation was only observed in unaware performers. It seems likely that aware performers recruit different systems after the initial hippocampal encoding, or perhaps they are more efficient and thus do not show observable activation. This activation is consistent with human and animal findings that show hippocampal regions are involved early in discovering associations and parahippocampal regions are involved subsequently in communicating (e.g., Lavenex & Amaral, 2000) and consolidating (e.g., Jarrard, 2001) those associations to cortical areas (for a review, see Eichenbaum & Cohen, 2001). This parahippocampal activity may be regarded as related to a process subsequent to the discovery of the associations supporting inference but prior to task awareness, which is consistent with the findings of Greene et al. (2001), wherein the capacity to correctly select B > D preceded task awareness. It may be the case, for example, that this activation is involved in the processing that leads from correct B > D performance to task awareness. If so, this activation may correspond to a stage of consolidation that would predict later task awareness in a study designed specifically to examine that issue. It is also important to note that awareness and performance showed no significant interaction within the hippocampus proper.

## **DISCUSSION**

The behavioral and functional imaging findings reported here provide the basis for several insights into the role of the hippocampus in human TI. Behaviorally, test performance on the transitive B > D pair depends on the simultaneous solution to the premise pairs B > C and C > D during training. Although nonperformers were generally able to solve one of these pairs as well as the end items, they could not solve them simultaneously. In addition, when first presented with the inferential pair B?D at test, performers solved it quickly and accurately. Together these findings suggest that the contextual conflict (C is correct in the presence of D but incorrect in the presence of B) is resolved efficiently as B > C > D, which serendipitously allowed for the correct solution of B > D at test.

Our principal finding is that hippocampal activation during training predicts inferential task performance at test. This activation is present in both aware and unaware participants. Behavioral data show no significant correlation between task awareness and performance. A further dissociation between task awareness and inferential performance is that latency data

show that BD performers are slower to respond at to B?D at test, whereas aware participants are faster to respond to B?D at test. Finally, the only MTL activation corresponding to task awareness was for the absence of awareness and then only in participants who showed good BD performance. These findings support our claim that BD performance is hippocampus-dependent and does not require awareness of the inference.

This pattern of functional activity and behavioral data supports the hypothesis that hippocampal regions are involved in inferential processes irrespective of task awareness. These functional imaging results converge with neuropsychological findings that participants may perform a hippocampal-dependent task without awareness (e.g., see Chun & Phelps, 1999).

Although there is a growing literature suggesting that the hippocampus may indeed serve a broader role than DM, no single study has shown correct task solution on an implicit task that involves the hippocampus. Previous studies have shown that amnesics cannot perform certain priming tasks (e.g., Manns & Squire, 2001; Chun & Phelps, 1999; Schacter et al., 1995; Schacter, Church, & Treadwell, 1994). The existing amnesic studies cannot entirely establish that these priming tasks are hippocampal-dependent because the participants include both MTL and diencephalic amnesics, as well as amnesics with varied degrees of hippocampal sparing. Existing functional imaging studies of hippocampal tasks have either informed participants of the task demands (Henke et al., 2003; e.g., Preston et al., 2004; Nagode & Pardo, 2002), or have not shown successful task performance in the absence of awareness (McIntosh et al., 2003). Nevertheless, the evidence to date shows that certain implicit tasks are impaired in amnesics (Chun & Phelps, 1999; Schacter & Church, 1995), the hippocampus may be active without task awareness (McIntosh et al., 2003), TI is a hippocampal-dependent task in animals (Dusek & Eichenbaum, 1997), TI may be performed without task awareness in humans (Greene et al., 2001), and the hippocampus is differentially active in successful explicit TI processing both at study (Nagode & Pardo, 2002) and at test (Preston et al., 2004; Henke et al., 2003). The findings presented here allow many of these disparate demonstrations to be considered simultaneously because we show that hippocampal activation associated with successful TI performance (for both context at study and inference at test) does not depend upon task awareness. The confluence of imaging and amnesic studies provides a compelling case that the hippocampus may be necessary for tasks without task awareness. It is becoming increasingly clear that the hippocampus may serve a broader role than DM. In addition, no previous study has shown hippocampal activation corresponding to inference, although inference is clearly necessary for experience to be useful in predicting novel outcomes.

If the hippocampus has an identifiable unitive role in learning and memory, the question then is what might that role be? One difficulty in understanding the associative role of the hippocampus is that there appear to be at least two distinct categories of relational tasks: The first of these is context-dependent tasks, such as transverse patterning or the exclusive-OR task (negative patterning), where context—rather than atomic stimuli—determines the correct choice. The second class of tasks is inferential, such as TI and path integration, where the integration of previous learning allows decisions under novelty. The present findings show a relationship between contextual and inferential processing, which suggests an extant view of hippocampal function. To correctly infer that  $B > D$ , participants first correctly resolve the context-dependent premise pairs  $B > C$  and  $C > D$ . One may correctly assert that we have shown only that one cannot infer  $B > D$  without first properly encoding  $B > C$  and  $C > D$ . Indeed, learning these inner pairs during training requires the hippocampus because these pairs require context to be solved correctly (Dusek & Eichenbaum, 1997; Greene et al., 2001; Wynne, 1995). The role of the hippocampus for contextual learning (e.g., Nadel, Samsonovich, Ryan, & Moscovitch, 2000; Ryan, Althoff, Whitlow, & Cohen, 2000) is affirmed by differential hippocampal activity for the context-dependent training pairs.

We propose that the existing evidence on TI in both animals and humans suggests that one may correctly solve both  $B > C$  and  $C > D$  by associatively linking the overlapping pairs into a superordinate hierarchy that simultaneously allows contextual and inferential decisions by representing a more generalized level of associative relationships, viz.  $A > B > C > D > E$  (Nakamaru & Sasaki, 2003; Treichler & Van Tilburg, 1996). Accordingly, this hierarchy (1) efficiently represents the premise pairs with the logically minimal items and relations in such a way that resolves the context-dependency and (2) automatically supports all other logical relations, including the inference  $B > D$ . By this view, the two faculties of RL—contextual discrimination and inference—are facets of a single process, formation of a superordinate representation, integrated from distinct events, that resolves contextual relations which would otherwise be in conflict.

These perspectives on RL also permit possible insights into the nature of DM. Autobiographical and world knowledge, as with TI, are obtained from discrete experiences and organized associatively into a global representation that simultaneously allows context-dependent relations and inference under novelty. Accordingly, both episodic and semantic memory require, as part of their formation, this superordinate associative organization of general and autobiographical information, which is context-dependent, and permit the flexibility that memory systems clearly evolved to support.

#### Acknowledgements

This study was supported in part by a grant from the National Institute of Aging (R01 AG022304 to S. M. R.), the Medical College of Wisconsin General Clinical Research Center (M01 RR00058), and the W. M. Keck Foundation.

#### References

- Acuna BD, Eliassen JC, Donoghue JP, Sanes JN. Frontal and parietal lobe activation during transitive inference in humans. *Cerebral Cortex* 2002;12:1312–1321. [PubMed: 12427681]
- Acuna BD, Sanes JN, Donoghue JP. Cognitive mechanisms of transitive inference. *Experimental Brain Research* 2002;146:1–10.
- Bland JM, Altman DG. Cronbach's alpha. *British Medical Journal* 1997;314:572. [PubMed: 9055718]
- Bransford JD, Franks JJ. The abstraction of linguistic ideas. *Cognitive Psychology* 1971;2:331–350.
- Bustamante JA, Jordan A, Vila M, Gonzalez A, Insua A. State dependent learning in humans. *Physiology and Behavior* 1970;5:793–796. [PubMed: 5522495]
- Cabeza R, Rao SM, Wagner AD, Mayer AR, Schacter DL. Can medial temporal lobe regions distinguish true from false? An event-related functional MRI study of veridical and illusory recognition memory. *Proceedings of the National Academy of Sciences, USA* 2001;98:4805–4810.
- Carlson RA. Syllogistic inference chains in meaningful text. *American Journal of Psychology* 1992;105:75–99.
- Chun MM, Phelps EA. Memory deficits for implicit contextual information in amnesic subjects with hippocampal damage. *Nature Neuroscience* 1999;2:844–847.
- Clark RE, Manns JR, Squire LR. Trace and delay eyeblink conditioning: Contrasting phenomena of declarative and nondeclarative memory. *Psychological Science* 2001;12:304–308. [PubMed: 11476097]
- Clark RE, Manns JR, Squire LR. Classical conditioning, awareness, and brain systems. *Trends in Cognitive Sciences* 2002;6:524–531. [PubMed: 12475713]
- Clark RE, Squire LR. Classical conditioning and brain systems: The role of awareness. *Science* 1998;280:77–81. [PubMed: 9525860]
- Cohen, NJ.; Eichenbaum, H. *Memory, amnesia, and the hippocampal system*. Cambridge: MIT Press; 1993.
- Cohen NJ, Poldrack RA, Eichenbaum H. Memory for items and memory for relations in the procedural/declarative memory framework. *Memory* 1997;5:131–178. [PubMed: 9156097]

- Cohen NJ, Squire LR. Preserved learning and retention of pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science* 1980;210:207–210. [PubMed: 7414331]
- Cox RW. AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research* 1996;29:162–173. [PubMed: 8812068]
- Craik FIM, Tulving E. Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General* 1975;104:268–294.
- Davis H. Transitive inference in rats (*Rattus norvegicus*). *Journal of Comparative Psychology* 1992;106:342–349. [PubMed: 1451416]
- Dolan RJ, Fletcher PF. Encoding and retrieval in human medial temporal lobes: An empirical investigation using functional magnetic resonance imaging (fMRI). *Hippocampus* 1999;9:25–34. [PubMed: 10088897]
- Douville K, Woodard JL, Seidenberg M, Miller SK, Leveroni CL, Nielson KA, Kranczak M, Antuono P, Rao SM. Medial temporal lobe activity for recognition of recent and remote famous names: An event-related fMRI study. *Neuropsychologia* 2005;43:693–703. [PubMed: 15721182]
- Dusek JA, Eichenbaum H. The hippocampus and memory for orderly stimulus relations. *Proceedings of the National Academy of Sciences, USA* 1997;94:7109–7114.
- Eichenbaum, H.; Cohen, NJ. *From conditioning to conscious recollection: Memory systems of the brain.* New York: Oxford University Press; 2001.
- Eichenbaum, H.; Cohen, NJ. *From conditioning to conscious recollection: Memory systems of the brain.* Oxford: Oxford University Press; 2002.
- Frank MJ, Rudy JW, O'Reilly RC. Transitivity, flexibility, conjunctive representations, and the hippocampus: II. A computational analysis. *Hippocampus* 2003;13:341–354. [PubMed: 12722975]
- Gillan DJ. Reasoning in the chimpanzee: II. Transitive inference. *Journal of Experimental Psychology: Animal Behavior Processes* 1981;7:150–164.
- Greene AJ, Spellman B, Dusek JA, Eichenbaum HB, Levy WB. Relational learning with and without awareness: Transitive inference using nonverbal stimuli in humans. *Memory and Cognition* 2001;29:893–902.
- Hartley T, Maguire EA, Spiers HJ, Burgess N. The well-worn route and the path less traveled: Distinct neural bases of route following and wayfinding in humans. *Neuron* 2003;37:877–888. [PubMed: 12628177]
- Heckers S, Zalesak M, Weiss AP, Ditman T, Titone D. Hippocampal activation during transitive inference in humans. *Hippocampus* 2004;14:153–162. [PubMed: 15098721]
- Henke K, Mondadori CR, Treyer V, Nitsch RM, Buck A, Hock C. Nonconscious formation and reactivation of semantic associations by way of the medial temporal lobe. *Neuropsychologia* 2003;41:863–876. [PubMed: 12667523]
- Jarrard LE. Retrograde amnesia and consolidation: Anatomical and lesion considerations. *Hippocampus* 2001;11:43–49. [PubMed: 11261771]
- Kumaran D, Maguire EA. The human hippocampus: Cognitive maps or relational memory? *Journal of Neuroscience* 2005;25:7254–7259. [PubMed: 16079407]
- Lavenex P, Amaral DG. Hippocampal–neocortical interaction: A hierarchy of associativity. *Hippocampus* 2000;10:420–430. [PubMed: 10985281]
- Lazareva OF, Smirnova AA, Rayevsky VV, Zorina ZA. Transitive inference in hooded crows: Preliminary data. *Doklady Biological Sciences* 2000;370:30–32. [PubMed: 10781324]
- Leveroni CL, Seidenberg M, Mayer AR, Mead LA, Binder JR, Rao SM. Neural systems underlying the recognition of familiar and newly learned faces. *Journal of Neuroscience* 2000;20:878–886. [PubMed: 10632617]
- Macphail EM. Cognitive function in mammals: The evolutionary perspective. *Brain Research, Cognitive Brain Research* 1996;3:279–290. [PubMed: 8806029]
- Manns JR, Hopkins RO, Reed JM, Kitchener EG, Squire LR. Recognition memory and the human hippocampus. *Neuron* 2003;37:171–180. [PubMed: 12526782]
- Manns JR, Squire LR. Perceptual learning, awareness, and the hippocampus. *Hippocampus* 2001;11:776–782. [PubMed: 11811672]

- McIntosh AR, Rajah MN, Lobaugh NJ. Functional connectivity of the medial temporal lobe relates to learning and awareness. *Journal of Neuroscience* 2003;23:6520–6528. [PubMed: 12878693]
- Milner B. Disorders of learning and memory after temporal lobe lesions in man. *Clinical Neurosurgery* 1972;19:421–446. [PubMed: 4637561]
- Nadel L, Samsonovich A, Ryan L, Moscovitch M. Multiple trace theory of human memory: Computational, neuroimaging, and neuropsychological results. *Hippocampus* 2000;10:352–368. [PubMed: 10985275]
- Nagode JC, Pardo JV. Human hippocampal activation during transitive inference. *NeuroReport* 2002;13:939–944. [PubMed: 12004195]
- Nakamaru M, Sasaki A. Can transitive inference evolve in animals playing the hawk–dove game? *Journal of Theoretical Biology* 2003;222:461–470. [PubMed: 12781744]
- Preston AR, Shrager Y, Dudukovic NM, Gabrieli JD. Hippocampal contribution to the novel use of relational information in declarative memory. *Hippocampus* 2004;14:148–152. [PubMed: 15098720]
- Rempel-Clower NL, Zola SM, Squire LR, Amaral DG. Three cases of enduring memory impairment after bilateral damage limited to the hippocampal formation. *Journal of Neuroscience* 1996;16:5233–5255. [PubMed: 8756452]
- Rudy JW, Sutherland RJ. The hippocampal formation is necessary for rats to learn and remember configural discriminations. *Behavioral Brain Research* 1989;34:97–109.
- Ryan JD, Althoff RR, Whitlow S, Cohen NJ. Amnesia is a deficit in relational memory. *Psychological Science* 2000;11:454–461. [PubMed: 11202489]
- Schacter DL. Memory and awareness. *Science* 1998;280:59–60. [PubMed: 9556454]
- Schacter DL, Church B. Implicit memory in amnesic patients: When is auditory priming spared? *Journal of the International Neuropsychological Society* 1995;1:434–442. [PubMed: 9375229]
- Schacter DL, Church B, Bolton E. Implicit memory in amnesic patients: Impairment of voice-specific priming. *Psychological Science* 1995;6:20–25.
- Schacter DL, Church B, Treadwell J. Implicit memory in amnesic patients: Evidence for spared auditory priming. *Psychological Science* 1994;5:20–25.
- Schendan HE, Searl MM, Melrose RJ, Stern CE. An fMRI study of the role of the medial temporal lobe in implicit and explicit sequence learning. *Neuron* 2003;37:1013–1025. [PubMed: 12670429]
- Scoville WB, Milner B. Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery and Psychiatry* 1957;20:11–21.
- Squire LR, Zola S. Memory, memory impairment, and the medial temporal lobe. *Cold Spring Harbor Symposium on Quantitative Biology* 1996;61:185–195.
- Squire LR, Zola-Morgan S, Chen KS. Human amnesia and animal models of amnesia: Performance of amnesic patients on tests designed for the monkey. *Behavioral Neuroscience* 1988;102:210–221. [PubMed: 3130073]
- Strange BA, Fletcher PC, Henson RN, Friston KJ, Dolan RJ. Segregating the functions of human hippocampus. *Proceedings of the National Academy of Sciences, USA* 1999;96:4034–4039.
- Talairach, J.; Tournoux, P. Co-planar stereotaxic atlas of the human brain. New York: Thieme; 1988.
- Treichler FR, Van Tilburg D. Concurrent conditional discrimination tests of transitive inference by macaque monkeys: List linking. *Journal of Experimental Psychology: Animal Behavior Processes* 1996;22:105–117. [PubMed: 8568492]
- Warrington EK, Weiskrantz L. The effect of prior learning on subsequent retention in amnesic patients. *Neuropsychologia* 1974;12:419–428. [PubMed: 4437740]
- Weaver JE, Steirn JN, Zentall TR. Transitive inference in pigeons: Control for differential value transfer. *Psychonomic Bulletin & Review* 1997;4:113–117.
- Wynne CDL. Reinforcement accounts for transitive inference performance. *Animal Learning and Behavior* 1995;23:207–217.

## APPENDIX A: PRE-EXPERIMENT INSTRUCTIONS (FOR EXPERIMENTERS)

“In this experiment, two figures will appear simultaneously on the computer screen. You are to select the ‘correct’ figure. At first, this will be by trial and error, however, with practice, you

will find that the correct figure is easily learned. You will need to respond quickly, however, accuracy is more important. In this practice experiment you will use the right mouse button to indicate that the figure on the right is 'correct' and the left mouse button to indicate the figure on the left is 'correct'. In the scanner you will have a box with similar buttons and you will respond in the same manner."

## APPENDIX B: AWARENESS RATINGS

*Postscan Questionnaire* (given to participants)

(Page breaks, with instructions not to return to a prior page follow questions 3, 6, and 7)

1. What did you think we were trying to find out with this experiment?
2. What did you think was the point of the trials where you were not told if you were right or wrong (the no feedback condition)?
3. Regarding question 2, were all the pairs of shapes in the no feedback condition the same as pairs you had already learned when feedback was given?

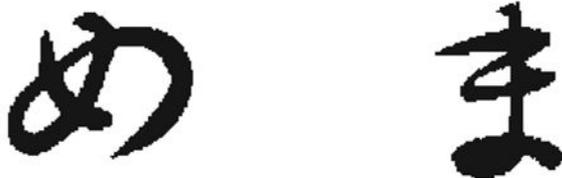
Yes  Not Sure  No

If no, do you think there was a correct answer?

Yes  Not Sure  No

If you believe there was a correct answer, explain why:

4. You were given the following pair of shapes several times, but never told an answer. Circle the shape you believe is correct (guess if necessary):



5. Regarding question 4, what reason (if any) did you have for your choice (check one):

There is a logically correct choice because (explain):

One just seemed right but I can't explain why.

I guessed: There may be a correct shape, but I don't know what it is.

I made a random choice because there is no correct choice.

Other: Explain \_\_\_\_\_

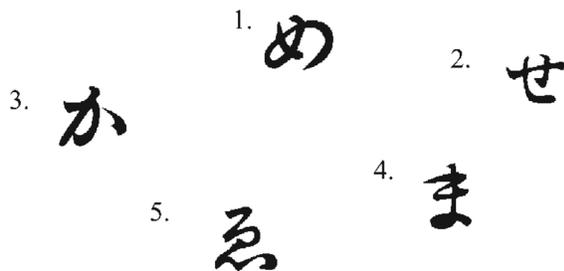
6. What strategy (if any) did you use to learn the shapes (check one):

I already knew the shapes: If so, from where? \_\_\_\_\_

I gave them names.

- I memorized each shape.
- I memorized part of the each shape.
- I just watched and eventually got it.
- I used their similarity to familiar shapes.
- No strategy.
- Other strategy (please describe in the space below)

7. Based on your understanding of the relationships among these shapes, arrange the shapes appropriately. Use the numbers provided to stand for the shapes.



8. If applicable, please indicate how much knowledge you have of formal logic, syllogism, or transitive inference:

*Postscan Questionnaire Scoring Procedure* (for experimenters)

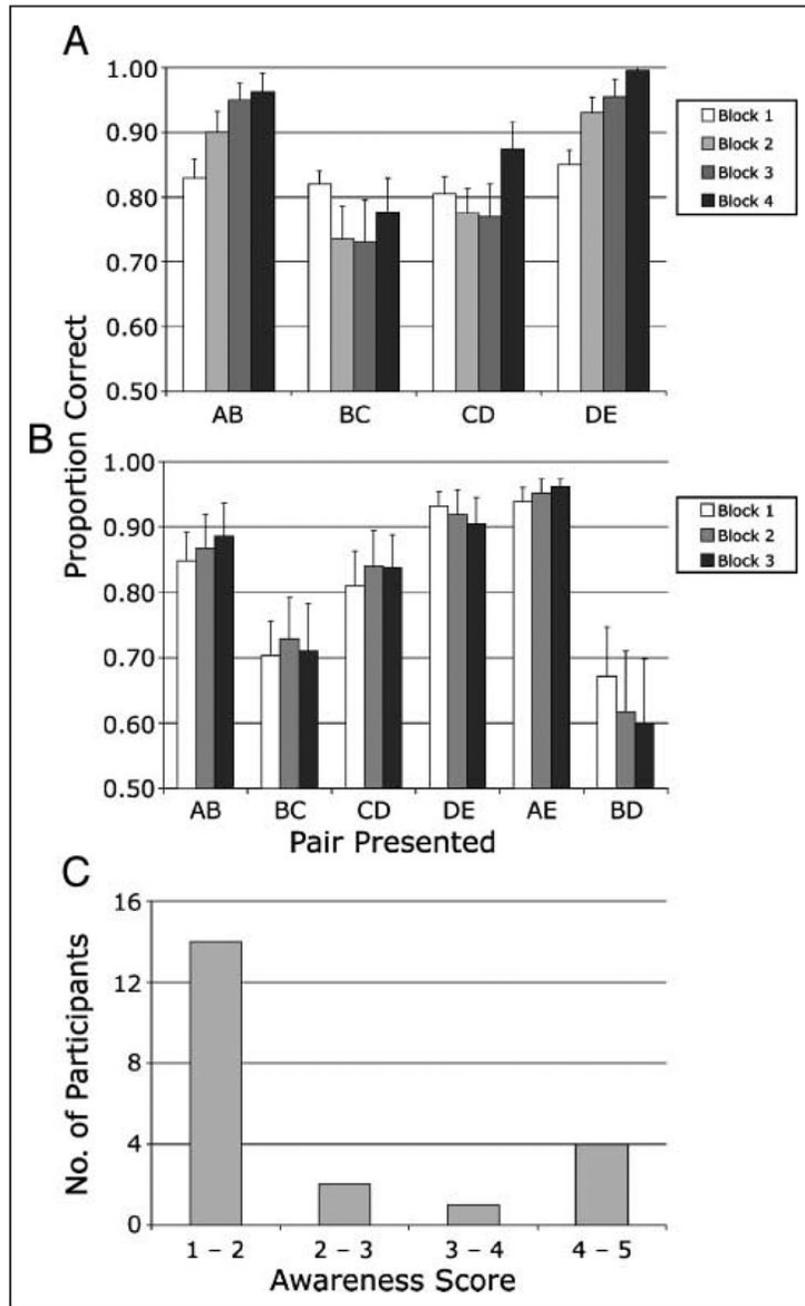
The questionnaire was designed to give increasingly leading questions in order to assess how explicitly the hierarchy among characters was understood.

Ratings on a scale of 1–5 were agreed to represent the following: 5 = knowledge of hierarchy, knowledge of logical BD choice; 4 = some knowledge of hierarchy, possible knowledge of logical BD choice; 3 = possible knowledge of hierarchy, possible or vague knowledge of logical BD choice; 2 = highly questionable knowledge of hierarchy, confusion or incorrect assertion about logical BD choice; 1 = no knowledge of hierarchy, no indication of any logical processes governing BD choice.

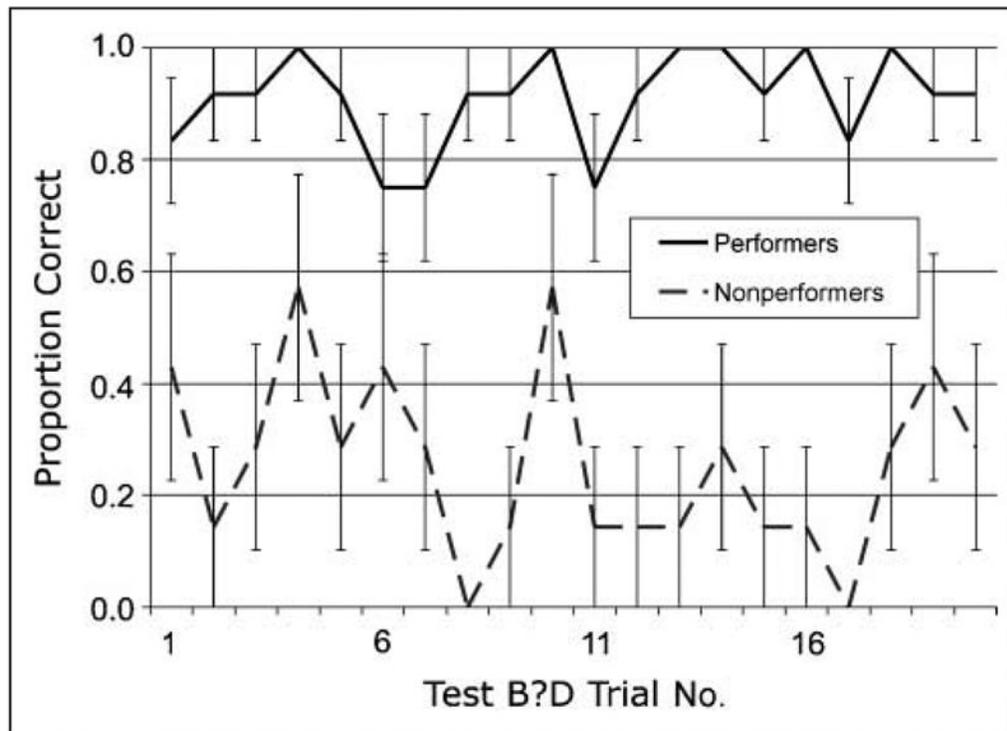
Ratings on the basis of question responses were assigned as follows: Questions 1–3: some assertion that there was a hierarchy or serial ordering resulted in a rating of 4 or 5 depending upon the clarity of the assertion. Question 5: some assertion that the given pair (BD) had a logical solution resulted in a rating of 3 or 4, depending upon the degree to which the logic participants provided corresponded to the logic of transitive inference. Question 8: if participants arranged the stimuli in some sort of sequence, a score of 2 was assigned irrespective of whether or not they were in the correct order. If the stimuli were arranged in some other way (e.g., as separate pairs), then a score of 1 was assigned.



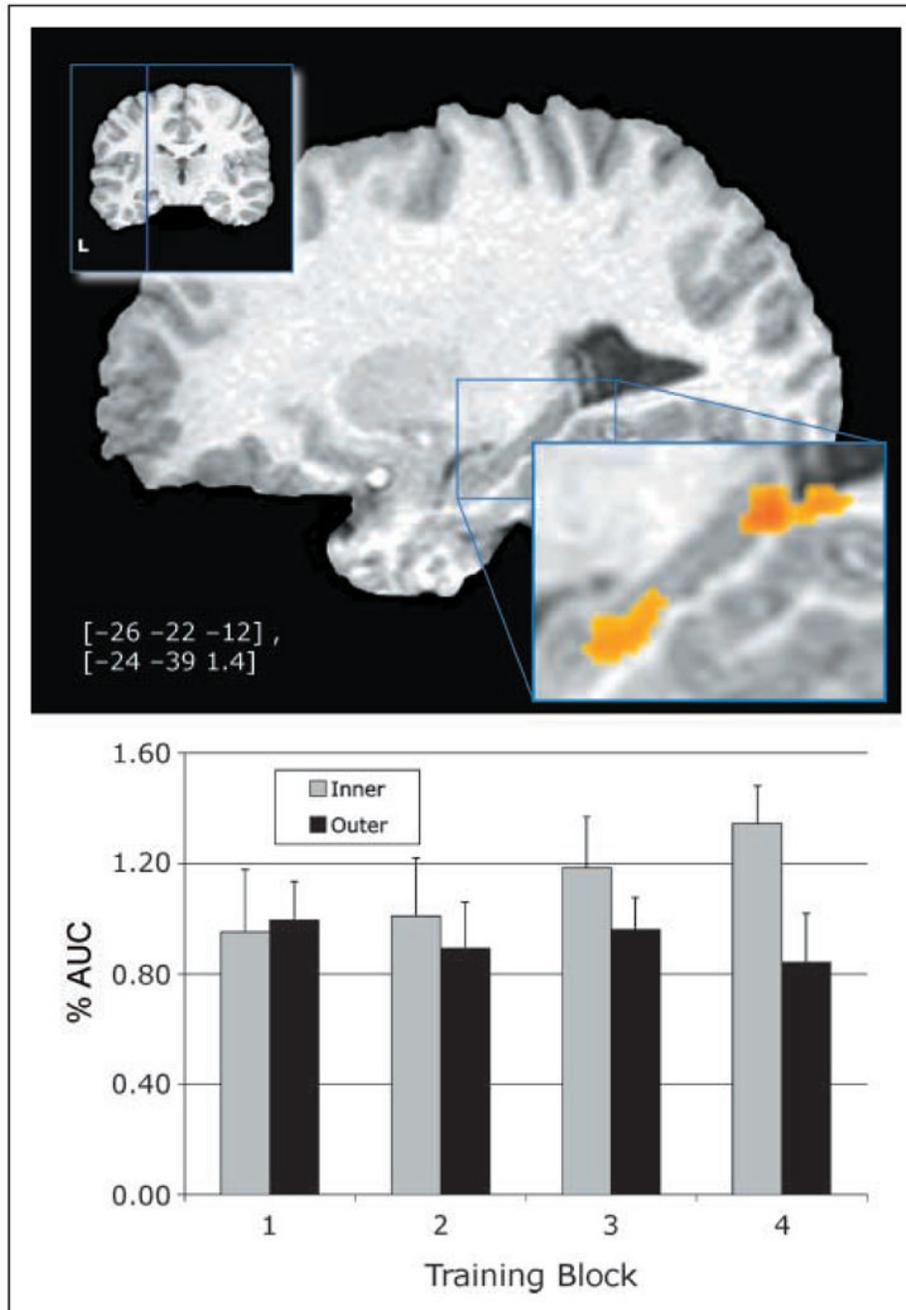
**Figure 1.**  
An example of the five Hiragana characters used as stimuli in the TI task.



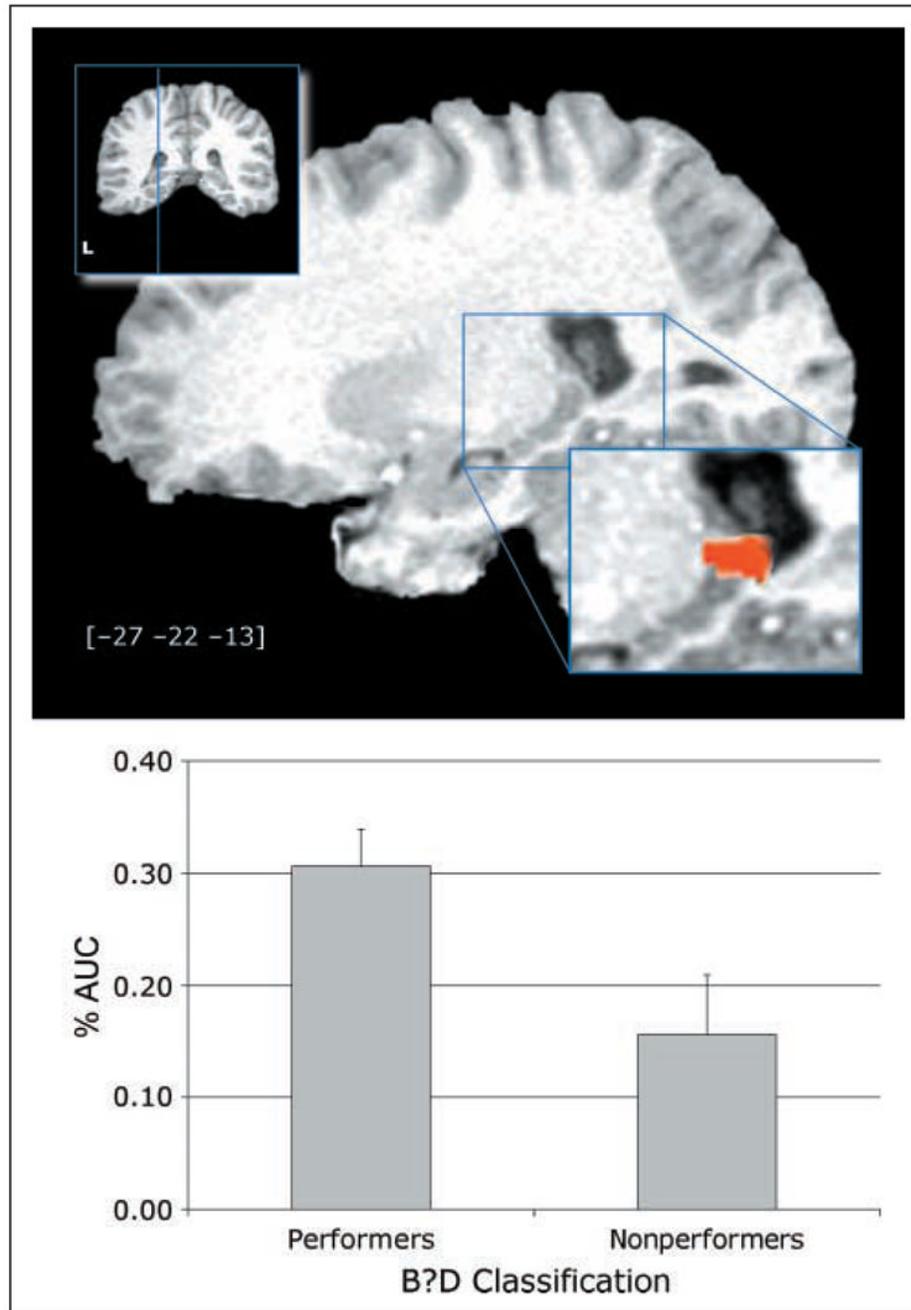
**Figure 2.** Behavioral results: accuracy during training (A) and test (B) as a function of pair type and block number. Frequency distribution of awareness scores (C).



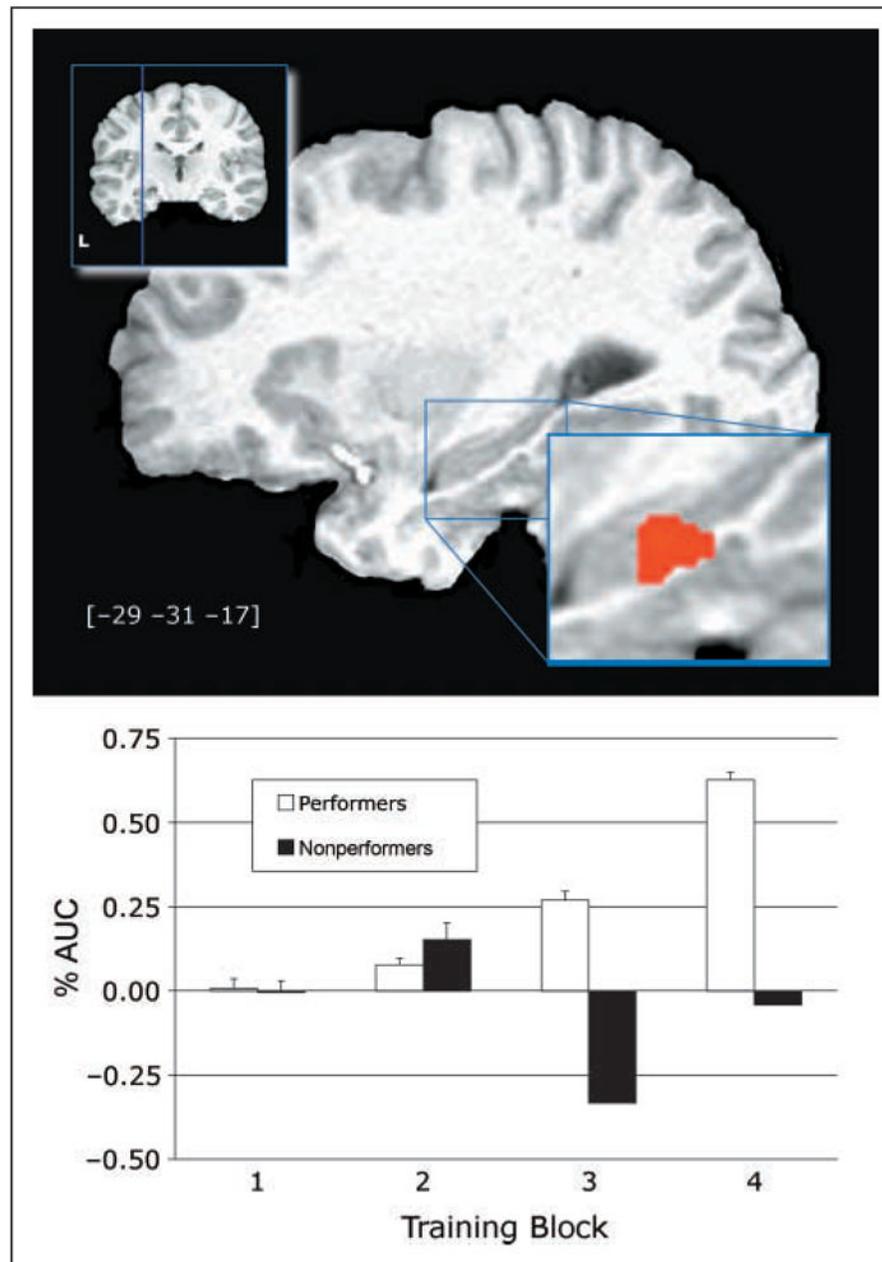
**Figure 3.** Mean accuracy on the BD pair as a function of trial presentation number. Coordinates (Talairach & Tournoux, 1988), in brackets, are the center of mass for the active region.



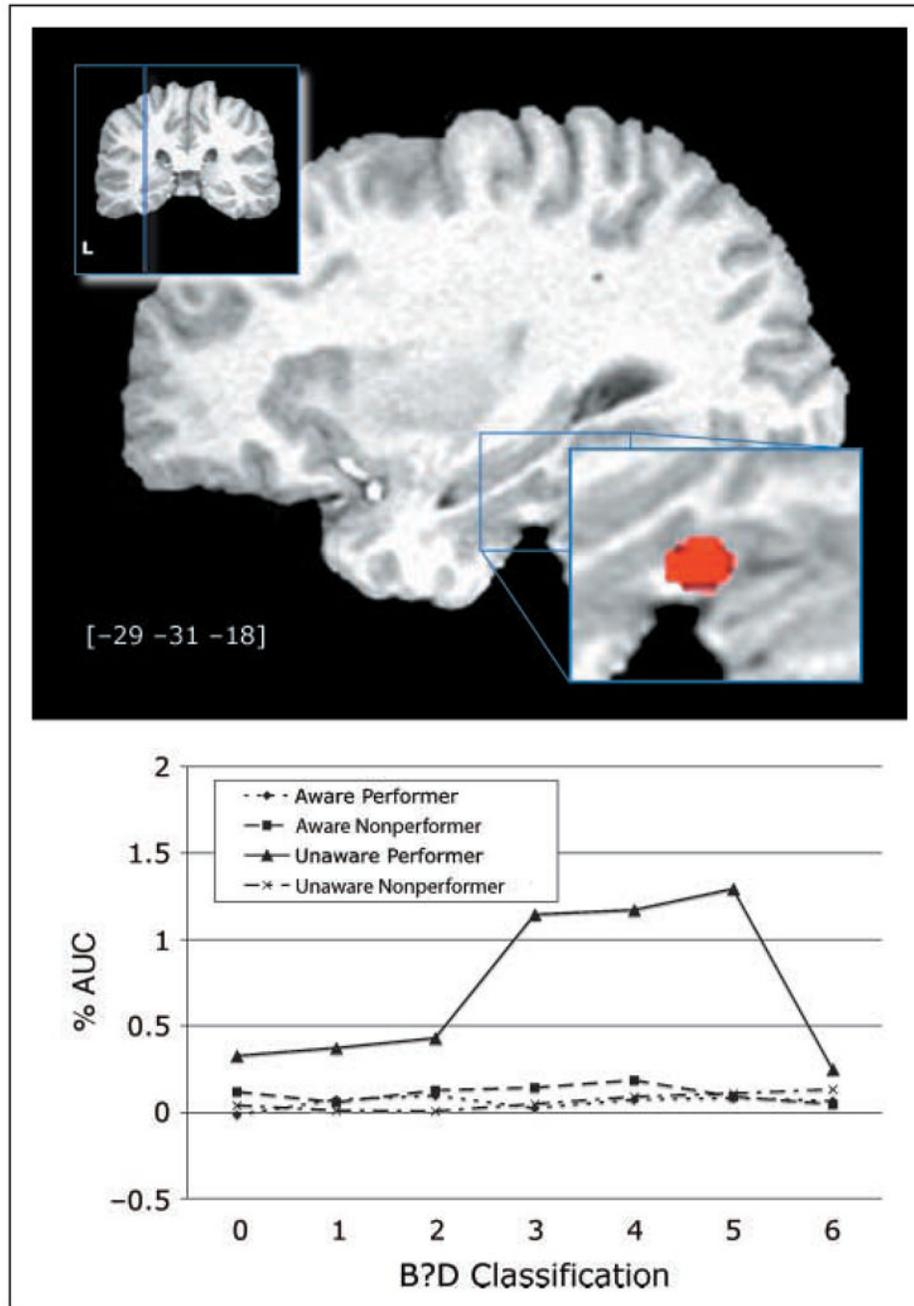
**Figure 4.** Training activation by pair type (inner vs. outer) over time in the left hippocampus and accompanying AUC graph. Coordinates (Talairach & Tournoux, 1988), in brackets, are the center of mass for the active region.



**Figure 5.** Test activation for B?D pair presentation over all participants and accompanying AUC graph. Coordinates (Talairach & Tournoux, 1988), in brackets, are the center of mass for the active region.



**Figure 6.** Performance-dependent training interaction contrast: The interaction of inner and outer pairs by time over levels of performance in the left hippocampus. The simplified AUC graph displays activation to the outer pairs subtracted from the inner pairs. Coordinates (Talairach & Tournoux, 1988), in brackets, are the center of mass for the active region.



**Figure 7.** Activation during the B?D pair at test related to both performance and awareness. Coordinates (Talairach & Tournoux, 1988), in brackets, are the center of mass for the active region.

**Table 1**  
Effects of Reaction Time for Training Summarized

Effect	Mean Difference (msec)	Statistic	p
1 Main Effect of Pair Type (Inner vs. Outer)	169	$F(1,18) = 35.69$	$1.18 \times 10^{-5}$
2 Interaction of Pair Type and BD Performance: Performers vs.			
a. Performers	235		
b. Nonperformers	103	$F(1,18) = 4.77$	.04
3 Main Effect of Training	86	$F(3,54) = 3.462$	.02
4 Interaction of Training and Pair Type			
a. Inner	171		
b. Outer	2	$F(3,54) = 6.045$	.001

**Table 2**  
Regions Differentially Activated to the Training Effect of Pair Type (Inner vs. Outer) × Training Block (Collapsed Over Levels of Performance and Awareness)

Region	BA	Volume (mm <sup>3</sup> )	Talairach Coordinates			z
			x	y	z	
<i>Frontal</i>						
L. Middle Frontal Gyrus	6	16,297	-23	5	41	
R. Middle Frontal Gyrus	9	14,086	33	10	31	
L. Inferior Frontal Gyrus	13	3140	-44	27	3	
L. Medial Frontal Gyrus	11	1188	-19	50	-9	
R. Paracentral Frontal Gyrus	5	635	6	-45	60	
L. Medial Frontal Gyrus	11	355	-1	60	-14	
L. Superior Frontal Gyrus	10	351	-21	61	2	
R. Medial Frontal Gyrus	6	327	6	-12	58	
L. Medial Frontal Gyrus	6	324	-10	-18	60	
L. Middle Frontal Gyrus	6	303	-34	0	58	
L. Precentral Gyrus	6	292	-14	-19	70	
R. Inferior Frontal Gyrus	47	290	37	30	-4	
L. Superior Frontal Gyrus	6	241	-17	-9	70	
R. Inferior Frontal Gyrus	10	231	38	45	3	
R. Precentral Gyrus	4	220	25	-24	68	
<i>Cingulate</i>						
R. Anterior	32	2284	22	45	9	
L. Anterior	23	366	-1	34	27	
<i>Parietal</i>						
R. Precuneus	7	14,578	25	-60	37	
L. Precuneus	7	11,639	-24	-66	34	
L. Supramarginal Gyrus	40	4572	-36	-38	34	
L. Precuneus	7	2743	-11	-53	57	
R. Inferior Parietal Lobule	40	481	38	-51	59	
R. Superior Parietal Lobule	7	395	15	-66	55	
L. Postcentral Gyrus	43	286	-61	-13	18	
R. Superior Parietal Lobule	7	248	28	-59	51	
<i>Temporal</i>						
L. Fusiform Gyrus	19	3865	-26	-54	-4	
R. Superior Temporal Gyrus	13	3127	39	-45	11	
L. Inferior Temporal Gyrus	20	1206	-35	-10	-37	
L. Middle Temporal Gyrus	21	806	-56	-26	-2	
L. Hippocampus	28	472	-20	-25	-7	
L. Hippocampus	30	396	-22	-36	1	
R. Parahippocampal Gyrus	36	303	32	-32	-5	
L. Parahippocampal Gyrus	21	276	-32	-16	-11	
L. Middle Temporal Gyrus	38	267	-40	3	-38	
R. Middle Temporal Gyrus	37	266	47	-51	-5	
L. Middle Temporal Gyrus	22	258	-46	-37	3	
<i>Occipital</i>						
R. Lingual Gyrus	18	1247	10	-82	-11	
R. Cuneus	23	1000	7	11	11	
R. Middle Occipital Gyrus	19	1790	28	-83	10	
R. Middle Occipital Gyrus	18	205	20	-97	7	
L. Cuneus	18	866	-6	-85	17	
<i>Subcortical</i>						
R. Putamen		3312	19	8	5	
R. Thalamus		1192	7	-20	14	
L. Thalamus		511	-11	-8	6	

Region	BA	Volume (mm <sup>3</sup> )	Talairach Coordinates		
			x	y	z
L. Claustrum		234	-27	12	11

**Table 3**  
Regions Differentially Activated to the B?D Pair at Test across Levels of B?D Performance

Region	BA	Volume (mm <sup>3</sup> )	Talairach Coordinates			z
			x	y	z	
<i>Performers &gt; Nonperformers</i>						
Frontal						
L. Middle Frontal Gyrus	11	993	-20	48	-8	
L. Medial Frontal Gyrus	11	765	-7	61	-13	
R. Inferior Frontal Gyrus	47	488	28	22	-19	
L. Medial Frontal Gyrus	10	328	-7	49	13	
Cingulate						
R. Anterior	32	285	9	24	21	
L. Cingulate Gyrus	23	385	0	-31	27	
Parietal						
R. Angular Gyrus	39	569	47	-58	33	
Temporal						
L. Inferior Temporal Gyrus	20	1712	-34	-11	-39	
<b>L. Hippocampus</b>	<b>35</b>	<b>239</b>	<b>-27</b>	<b>-22</b>	<b>-13</b>	
Occipital						
L. Inferior Occipital Gyrus	18	306	-26	-87	-15	
Subcortical						
R. Caudate		332	10	-5	17	
<i>Nonperformers &gt; Performers</i>						
Frontal						
L. Superior Frontal Gyrus	6	424	-15	-11	62	
L. Middle Frontal Gyrus	10	235	-44	47	20	
L. Precentral Gyrus	6	200	-43	0	36	
Parietal						
L. Superior Parietal Lobule	7	237	-5	-63	57	

**Table 4** Regions Differentially Activated to the Training Effect of Pair Type (Inner vs. Outer) × Test B?D Performance Level (Collapsed over Levels of Awareness)

Region	BA	Volume (mm <sup>3</sup> )	Talairach Coordinates			z
			x	y	z	
<i>Frontal</i>						
R. Middle Frontal Gyrus	10	1321	31	52	8	
L. Middle Frontal Gyrus	10	1173	-27	58	18	
L. Middle Frontal Gyrus	11	696	-25	49	-12	
R. Medial Frontal Gyrus	11	570	8	56	-14	
L. Medial Frontal Gyrus	10	527	-4	67	10	
L. Middle Frontal Gyrus	10	468	-33	48	3	
L. Medial Frontal Gyrus	11	461	-4	41	-17	
L. Inferior Frontal Gyrus	44	436	-56	8	14	
R. Middle Frontal Gyrus	11	302	28	47	-9	
L. Superior Frontal Gyrus	9	297	-2	55	27	
L. Inferior Frontal Gyrus	47	284	-49	21	-3	
R. Medial Frontal Gyrus	11	239	14	47	-12	
L. Precentral Gyrus	6	216	-50	3	45	
R. Superior Frontal Gyrus	6	202	8	15	61	
L. Superior Frontal Gyrus	6	207	-5	27	57	
<i>Cingulate</i>						
L. Anterior	32	348	-11	31	-7	
<i>Parietal</i>						
R. Angular Gyrus	39	571	51	-64	33	
L. Superior Parietal Lobule	7	392	-22	-68	56	
R. Precuneus	7	346	8	-55	60	
L. Precuneus	7	316	-6	-77	46	
L. Cuneus	18	200	-15	-98	19	
<i>Temporal</i>						
<b>L. Hippocampal Gyrus</b>	<b>36</b>	<b>548</b>	<b>-29</b>	<b>-31</b>	<b>-17</b>	
R. Inferior Temporal Gyrus	20	367	53	-25	-19	
L. Superior Temporal Gyrus	38	336	-56	6	-6	
<i>Occipital</i>						
R. Inferior Occipital Gyrus	19	330	32	-78	-5	
R. Middle Occipital Gyrus	19	204	44	-80	2	

**Table 5**  
Regions Differentially Activated to the B?D Pair during the Analysis of B?D Performance  $\times$  Inferential Awareness

Region	BA	Volume (mm <sup>3</sup> )	Talairach Coordinates			z
			x	y	z	
<i>Frontal</i>						
L. Superior Frontal Gyrus	10	1830	-14	62	14	14
L. Inferior Frontal Gyrus	11	1150	-22	36	-22	-22
R. Middle Frontal Gyrus	10	608	32	55	4	4
R. Superior Frontal Gyrus	6	595	19	15	45	45
R. Superior Frontal Gyrus	6	506	16	28	55	55
L. Inferior Frontal Gyrus	47	465	-48	20	-3	-3
L. Superior Frontal Gyrus	11	394	-22	50	-16	-16
L. Inferior Frontal Gyrus	44	387	-52	12	13	13
L. Inferior Frontal Gyrus	47	341	-39	32	-2	-2
L. Medial Frontal Gyrus	11	340	-1	46	-16	-16
R. Superior Frontal Gyrus	10	313	9	57	-3	-3
R. Superior Frontal Gyrus	11	250	21	61	-14	-14
R. Rectal Gyrus	11	225	7	18	-20	-20
L. Medial Frontal Gyrus	10	222	-3	51	-4	-4
L. Superior Frontal Gyrus	9	218	-12	53	32	32
<i>Cingulate</i>						
L. Anterior	32	4262	-4	38	9	9
<i>Parietal</i>						
R. Cuneus	17	675	6	-80	10	10
L. Superior Parietal Lobule	7	577	-19	-49	61	61
L. Cuneus	18	429	-16	-100	5	5
R. Superior Parietal Lobule	7	297	19	-68	54	54
L. Superior Parietal Lobule	7	275	-20	-68	54	54
L. Inferior Parietal Lobule	40	224	-46	-38	48	48
L. Precuneus	7	208	-9	-76	46	46
<i>Temporal</i>						
R. Middle Temporal Gyrus	21	1704	50	0	-19	-19
L. Middle Temporal Gyrus	39	523	-36	-73	20	20
<b>L. Parahippocampal Gyrus</b>	<b>36</b>	<b>475</b>	<b>-29</b>	<b>-31</b>	<b>-18</b>	<b>-18</b>
L. Fusiform Gyrus	37	414	-45	-57	-17	-17
L. Superior Temporal Gyrus	13	412	-35	11	-16	-16
<i>Occipital</i>						
L. Lingual Gyrus	19	457	-24	-72	3	3