Entorhinal Cortex Lesions Disrupt the Relational Organization of Memory in Monkeys

Cindy A. Buckmaster,1,3 Howard Eichenbaum,1 David G. Amaral,5 Wendy A. Suzuki,6 and Peter R. Rapp1,2

1Fishberg Department of Neuroscience and 2Kastor Neurobiology of Aging Laboratories, Department of Geriatrics and Adult Development, Mount Sinai School of Medicine, New York, New York 10029-6574, 3Department of Neurobiology and Behavior, State University of New York at Stony Brook, Stony Brook, New York 11794, 4Laboratory of Cognitive Neuroscience, Boston University, Boston, Massachusetts 02215, 5The Medical Investigation of Neurodevelopmental Disorders Institute, University of California, Davis, California 95817, and 6Center for Neural Science, New York University, New York, New York 10003

Recent accounts suggest that the hippocampal system critically supports two central characteristics of episodic memory: the ability to establish and maintain representations for the salient relationships between experienced events (relational representation) and the capacity to flexibly manipulate memory (flexible memory expression). To test this proposal in monkeys, intact controls and subjects with bilateral aspiration lesions of the entorhinal cortex were trained postoperatively on two standard memory tasks, delayed nonmatching-to-sample (DNMS) and two-choice object discrimination (OD) learning, and three procedures intended to emphasize relational representation and flexible memory expression: a paired associate (PA) task, a transitive inference (TI) test of learning and memory for hierarchical stimulus relationships, and a spatial delayed recognition span (SDRS) procedure. The latter assessments each included critical “probe” tests that asked monkeys to evaluate the relationships among previously learned stimuli presented in novel combinations. Subjects with entorhinal cortex lesions scored as accurately as controls on all phases of DNMS and OD, procedures that can be solved on the basis of memory for individual stimuli. In contrast, experimental monkeys displayed deficits relative to controls on all phases of the PA, TI, and SDRS tasks that emphasized the flexible manipulation of memory for the relationships between familiar items. Together, the findings support the conclusion that the primate hippocampal system critically enables the relational organization of declarative memory.

Key words: memory; hippocampus; relational representation; medial temporal lobe; nonhuman primate

Introduction

Normal memory depends on the hippocampal formation (widely viewed on grounds of connectivity as comprising the dentate gyrus, hippocampus proper, subicular complex, and entorhinal cortex) (Insauti et al., 1987; Amaral and Witter, 1995) and associated perirhinal and parahippocampal cortices (Milner, 1972; Zola-Morgan and Squire, 1993; Mishkin et al., 1997). The specific cognitive operations implemented by this system in support of memory, however, remain undefined. A popular view is that the hippocampus is specialized or disproportionately devoted to learning and memory for spatial information (Parkinson et al., 1988; Nadal, 1991; Gagliardo et al., 1999; Matsumura et al., 1999; Ekstrom et al., 2003). Other conceptualizations suggest that medial temporal lobe memory is more accurately characterized by the underlying processing operations of the system rather than the content of acquired information (Cohen and Squire, 1980; Graf and Schacter, 1985; Squire, 1992; Bechara et al., 1995; Vargha-Khadem et al., 1997; Chun and Phelps, 1999). Specific formulations include the proposal that the hippocampal formation subserves a broad organizational function, establishing representations of the relevant relationships between experienced events and enabling the flexible utilization of memory (Eichenbaum et al., 1999; Eichenbaum, 2000).

The relational account receives support from human and experimental animal research. In one illuminating investigation, amnesics improved as quickly as controls on a probabilistic classification test of implicit learning that involved the nonconscious, incremental acquisition of response biases. Unlike intact subjects, however, patients were unable to use memory flexibly and apply task knowledge under modified testing conditions (Reber et al., 1996). Conceptually related impairments have been reported in mice and rats with hippocampal damage (Kogan et al., 2000; Agster et al., 2002; Fortin et al., 2002), including deficits in the ability to make inferential judgments about the relationships between familiar items presented in novel combinations and poor performance on previously acquired discriminations when items are presented in reverse of their original training order ( Bunsey and Eichenbaum, 1996; Dusek and Eichenbaum, 1997).
To reduce bleeding, lidocaine (1%) with epinephrine (1:100,000) was injected along the midline of the scalp before the first surgical incision. The skin and galea were then cut and reflected anteriorly and posteriorly, exposing the zygomatic arches. The zygomatic arches were removed, providing access to the ventrolateral aspect of the skull, and the temporal muscles were cut parasagittal to the anteroposterior midline, leaving a narrow strip of muscle attached to each side of the fascia for reattachment at the end of the procedure. The muscle was gently retracted and a 2 × 3 cm craniotomies were produced to expose the anterior and ventrolateral portions of the temporal lobe. Mannitol (30 ml at 25%, i.v., over 15 min) was then provided to prevent edema and reduce brain volume in preparation for the entorhinal ablation.

After rotating the animal ~60° from supine position to permit optimal access to the medial temporal lobe, the dura mater was cut and reflected, exposing the brain. The target area, manipulated into view under a surgical microscope, was bounded laterally by the rhinal sulcus and medially by the sulcus seminulatus (rostroly) and the choroidal fissure (caudally). Lesions were intended to spare the underlying white matter and ablate all cortical tissue extending rostromedially into the temporal pole and caudally to a point ~1 mm beyond the caudal limit of the rhinal sulcus. The surface of the entorhinal cortex was cauterized, and the damaged tissue was aspirated using a glass pipette with an angled tip. After ablation, the dura mater was sutured and the craniotomy was packed with Gelfoam (Upjohn, Kalamazoo, MI). The same procedures were used to make the contralateral lesion, and the wound was closed in anatomical layers. Once awake, the animal was allowed to recover in a warming chamber and monitored closely. Cefazolin (50 mg, i.m.), dexamethasone phosphate (1–2 mg, i.m.), and ceftazidime (30 ml at 25%, i.v., over 15 min) were administered postoperatively to prevent infection, swelling, and pain. In most cases, monkeys were allowed 6 days of recovery before the behavioral testing began.

Behavioral testing

Apparatus

Behavioral assessment was conducted postoperatively in a modified Wisconsin General Test Apparatus (WGTA) (Harlow and Bromer, 1938), similar to previous descriptions (Rapp and Amaral, 1989, 1991). Briefly, the test chamber was arranged so that the animals could manipulate stimuli presented on a stationary tray that contained three evenly spaced reward wells, illuminated from above by a fluorescent light strip. The apparatus was equipped with an experimenter-controlled opaque door that blocked the animal’s view of the tray during retention and intertrial intervals (ITIs). A second door, positioned between the experimenter and the food wells of the stimulus tray, contained a one-way mirror that allowed observation of the subject but prevented unintentional experimenter cueing. A white noise generator masked background sounds throughout testing.

Pretraining

Before formal testing, monkeys were trained by successive approximation to displace objects and retrieve food rewards from the wells of the stimulus tray. Pretraining progressed at similar rates in the experimental (E) and control (C) groups, and all monkeys advanced to the test battery within 1 month.

Standard tasks

Tasks were administered in the order listed below and noncorrection procedures were followed throughout.

Delayed nonmatching-to-sample. Training was similar to previous descriptions (Rapp and Amaral, 1991). Trials consisted of two phases: a sample presentation followed by a recognition test. In the sample phase, monkeys displaced an object over the baited central well of the stimulus tray, and the opaque door was lowered to impose a delay. The sample and a novel object were then positioned over the lateral wells with only the novel, “nonmatching” stimulus covering a reward. The left–right position of the rewarded objects was balanced across trials within each test.
session (Gellermann, 1933), and the sample–novel object assignments were balanced across animals within a group. New pairs of objects were presented on each trial (Mishkin and Delacour, 1975) and drawn in a predetermined sequence from a collection of 400 perceptually distinct “junk” items (three-dimensional objects differing widely in shape, size, color, and texture). Test sessions consisted of 20 trials per day, using a 30 sec ITI. Initially, monkeys were trained with a 10 sec delay to a criterion of 90% correct over five consecutive sessions (10 or fewer errors in 100 trials). After learning the nonmatching procedure, the effects of increased memory demand were evaluated by imposing successively longer retention intervals of 15, 30, 60, and 120 sec, and 10 min between the sample and recognition phases of each trial. Animals were tested for a total of 100 trials at delays through 120 sec (20 trials per day) and for 50 trials (5 per day) with a 10 min retention interval. Monkeys remained in the WGTa during all retention intervals.

Object discrimination acquisition and retention. Monkeys were trained successively on four two-choice object discrimination problems that required learning which stimulus in a pair was consistently associated with reward (Zola-Morgan and Squire, 1985). Discrimination problems were presented for two consecutive daily sessions followed by a third session 48 hr later. Objects were presented simultaneously over the lateral wells of the stimulus tray, with the rewarded stimulus appearing equally as often on the left and right across trials (Gellermann, 1933). The positive stimulus in each pair was balanced within groups. Test sessions consisted of 20 trials per day separated by 20 sec ITIs, except that, as a result of experimenter error, the ITI for the lesion group was 30 sec. Scores for object discrimination (OD) were calculated as the average percentage correct achieved across the four problems tested.

Relational tasks
Testing continued using three procedures designed to emphasize relational information processing: the paired associate (PA), transitive inference (TI), and spatial delayed recognition span (SDRS) tasks. In overview, the instruction phases of these procedures consisted of training on multiple conditional discriminations that contained overlapping elements and encouraged learning the predictive relationships among stimuli. Once these problems were acquired, the familiar stimuli were presented in novel combinations, which challenged animals to draw inferences or generalizations on the basis of memory for the relationships between the items.

Stimuli. The present experiments took advantage of previous findings demonstrating that relative to standard object-guided tasks, learning is substantially facilitated when test items serve as both discriminative stimuli and reinforcers (Jarvik, 1953, 1956; Murphy and Miller, 1955, 1958; Miller and Murphy, 1964; Smith et al., 1976). In the present experiments, custom-made sugar cookies, distinct in shape and color but similar in taste and overall size (1.5–2.0 cm across), were used as stimuli. Just before each session, the reward value of the items was manipulated by spraying the underside of negative cookies with a bitter flavoring (Grannick’s Bitter Apple, Norwalk, CT), rendering them less palatable than untreated, positive stimuli. A control procedure, conducted between the TI and SDRS tasks, confirmed that monkeys were unable to distinguish positive and negative items on the basis of unintended visual or olfactory cues associated with the bitter apple solution. Specifically, over the course of 100 trials in which animals chose between positive and negative cookies of identical shape and color (i.e., an interval over which monkeys achieved scores of 90% and better on standard two-choice discrimination problems), the control and experimental groups failed to perform above chance (mean percentage of selection of positive stimuli ± SE: C = 52.8 ± 1.6%, E = 55.9 ± 2.5%; one-sample t tests, t < 2.4; p > 0.10). Before formal training, monkeys were allowed to retrieve and consume cookies in both the home cage and test apparatus. The cookies provided during this brief accommodation period were different from the task stimuli in shape.

Paired associate. Overview. The PA procedure was adapted from parallel research in rats (Busney and Eichenbaum, 1996). Monkeys were trained on two series of instruction or “premise” discriminations: a “relational” series, including stimuli that overlapped and bridged the individual problems, and a “discontinuous” series of similarly configured conditional discriminations that lacked overlapping, linking elements. As illustrated in Figure 1, three items were presented concurrently on each trial, and the visual identity of the centrally positioned stimulus (e.g., A or X) signaled the correct choice among the two lateral items (B and Y). Later in training on the relational series, the choice items from the
initial problems (B and Y) were used as samples for a second set of discriminations, with items C and Z appearing as choices. The intention of this design was to encourage learning according to the associative links between stimuli, such that A predicts B and B predicts C, whereas X predicts Y and Y predicts Z. Subsequent probe tests evaluated memory for these interitem associations by asking subjects to judge the relationship between familiar stimuli presented in novel combinations (i.e., sample A or X presented with choice items C and Z) (Fig. 1). Probe tests for the discontinuous series followed the same format as for the relational probes (i.e., a sample from one set of discriminations was presented with choices from the second set), but because the premise pairs lacked overlapping elements, these trials could only be solved as novel conditional discriminations constituting a test of learning for new associations among familiar stimuli.

**Instruction discrimination training.** Training was conducted in phases, as detailed in Table 1. Animals learned the discontinuous series of premise discriminations first, followed by the relational series. In both cases, the premise problems were presented initially in alternating blocks of trials that decreased in size as accuracy improved, and training continued until the animals achieved a predetermined performance criterion when all four problems of a series appeared in an unpredictable order within sessions. A spatial cueing procedure was used early in training on the discontinuous series in an effort to emphasize the association between the sample and correct choice. Specifically, the sample stimulus appeared immediately adjacent to the positive choice item and was moved progressively closer to the central well as accuracy increased. Spatial cueing was omitted during training on the relational series of discriminations (i.e., the sample was presented in the central position, equidistant from the two choice items, from the outset). The left–right position of the positive choice was balanced across trials for each problem within a session, and the specific stimulus selected to be positive was balanced across subjects within a group. Daily sessions consisted of 20 trials separated by 30 sec ITIs.

**Probe testing.** Relational series: Probe sessions were similar to the final phase of premise pair training, except that four probe trials were interspersed among the familiar discriminations. Probe trials involved presentation of a sample stimulus from the first set of premise problems (i.e., A or X) with choice items from the second set (i.e., C and Z), and monkeys were rewarded for selecting the item predicted by the associative relationships between stimuli (e.g., choosing C over Z when A appeared as the sample) (Fig. 1). Five daily probe sessions were provided (20 trials per day; 30 sec ITI), yielding a total of 10 probe trials of each type.

**Discontinuous series:** After completing probe testing for the relational series, monkeys were refamiliarized with the discontinuous premise problems until they achieved a criterion of 80% correct, in a single session, on phases 5 and 6 of the original training protocol (Table 1). Subsequent probe testing presented subjects with novel combinations of the familiar stimuli, following the same format as for the relational series. In this case, however, because the discontinuous instruction pairs lacked shared stimuli that bridged the problems, probes were configured such that the choice item selected to be positive in the presence of each sample was determined pseudorandomly. These sample–choice reinforcement contingencies were held constant across probe trials. In this way, probe testing for the discontinuous series provided a window on the capacity for learning new conditional discriminations among familiar stimuli. Other procedural details were the same as for the relational probe tests.

**Transitive inference.** Transitive inference refers to the ability to judge from memory the relationships between nonadjacent items in a hierarchically organized series. For example, after learning that A is longer than B and B is longer than C, normal subjects can infer that A is longer than C in the absence of any explicit instruction about the relationship between these items. A practical limitation of previous object-guided versions of TI, however, is that acquisition is protracted in intact monkeys and typically requires >1000 training trials (Rapp et al., 1996). In an effort to facilitate learning, the present experiments took advantage of a modified protocol using cookie stimuli similar to but different in shape from memory the relationships between nonadjacent items in a hierarchically organized series. For example, after learning that A is longer than B and B is longer than C, normal subjects can infer that A is longer than C in the absence of any explicit instruction about the relationship between these items. A practical limitation of previous object-guided versions of TI, however, is that acquisition is protracted in intact monkeys and typically requires >1000 training trials (Rapp et al., 1996). In an effort to facilitate learning, the present experiments took advantage of a modified protocol using cookie stimuli similar to but different in shape and color from those in the PA task.

**Overview.** Monkeys initially learned four overlapping two-choice discrimination problems in which the reward value of most items depended on the stimulus with which they were paired (i.e., A+B, A+C, B+C, and C+D) (Fig. 2). By this design, the relationship between the items can be formally described by the hierarchical organization A>B>C>D>E. To determine whether the stimuli were in fact learned according to the described relationships, subsequent probe tests presented novel pairs of nonadjacent, indirectly related items (i.e., B vs D). Although multiple strategies can support inferential responding (Frank et al., 2003; Van Elzakker et al., 2003), our prediction based on previous findings (McGonigle and Chalmers, 1992; Rapp et al., 1996) was that the

### Table 1. Paired associate training protocol

<table>
<thead>
<tr>
<th>Instruction phases</th>
<th>Criterion requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase 1: PP1, PP2</td>
<td>a–h: 80% correct within one block</td>
</tr>
<tr>
<td>Phase 2: PP1, PP2</td>
<td>a–h: 80% correct within one block</td>
</tr>
<tr>
<td>Phase 3: PP1, PP2</td>
<td>a–h: 80% correct within one block</td>
</tr>
<tr>
<td>Phase 4: PP1, PP2</td>
<td>a–h: 80% correct within one block</td>
</tr>
<tr>
<td>Phase 5: PP1, PP2</td>
<td>a–h: 80% correct within one block</td>
</tr>
<tr>
<td>Phase 6: PP1, PP2</td>
<td>a–h: 80% correct within one block</td>
</tr>
</tbody>
</table>

*Phase 1, a–h, was instituted for instruction on the discontinuous series only. This spatial cueing procedure was not provided on any training phase of the relational series. PP, Premise problem; S, sample stimulus; +, positive choice stimulus.*
flexible manipulation of relational memory would enable intact monkeys to make appropriate inferential judgments, choosing B over D on these trials. Probe testing also included so-called "end-anchored" tests, composed of stimuli that had consistent reward values throughout training (i.e., A vs E) and therefore were soluble on the basis of reinforcement history alone, without reference to memory for the relationships between items. For all phases of testing, trials consisted of presenting two stimuli simultaneously, one behind each of the lateral wells of the stimulus tray (30 sec ITI). The left–right position of positive choice varied pseudorandomly and was balanced across trials within a session. Training was conducted in phases, as described below.

Detailed TI training procedures. Phase 1: individual instruction discriminations: Animals were trained on the A+B− discrimination until they chose correctly on 9 of 10 consecutive trials. Sessions included a maximum of 32 trials. Testing continued in an identical manner for each of the remaining discriminations (B+C−, C+D−, and D+E−) in sequence.

Phase 2: eight trial blocks of instruction pairs: Instruction problems were presented in blocks of eight trials each, in sequence (i.e., eight A+B− trials followed by eight B+C− trials, eight C+D− trials, and eight D+E− trials). Five sessions were provided.

Phase 3: sequential training: This phase consisted of one session in which the four instruction problems were presented once each in sequence, repeated across eight blocks of four trials (i.e., eight blocks of A+B−, B+C−, C+D−, and D+E−).

Phase 4: random-order training: Training was identical to phase 3 except that the order of presentation of the four instruction problems was unpredictable within each four-trial block. Testing continued until animals scored ≥87.5% correct, with no more than one error on an individual instruction problem within one session.

Phase 5: probe testing: Testing continued in an identical manner except that two BD and two AE probe trials per session were randomly interspersed among the instruction discriminations. As noted previously, the end-anchored AE pair could be solved on the basis of the reward history of the individual items (i.e., A was positive throughout training and E was never rewarded) (Fig. 2). In contrast, because items B and D were positive on 50% of the instruction trials, transitive responding on BD probes (i.e., selecting B over D) was interpreted as evidence for the flexible manipulation of memory for the hierarchical relationships among items (Dusek and Eichenbaum, 1997). Probe trials were reinforced in the direction predicted by the hierarchical organization among items (i.e., A+E− and B+D−) (Fig. 2). Ten probe sessions were provided, for a total of 320 instruction problems and 20 probes of each type.

Spatial delayed recognition span. The final assessment in the battery was a delayed recognition span task modified from previous descriptions (Moss et al., 1997). This procedure required monkeys to select the item occupying the novel position in an increasing list of locations. A custom-made stimulus tray was used, measuring 26.3 × 68.5 cm, angled at approximately an 8° slope toward the subject. Sugar cookies, identical in shape and color, but novel with respect to those used previously in testing, served as discriminative stimuli and reinforcers.

Trials were initiated by presenting a positive stimulus in 1 of 18 locations, evenly distributed in a 3 × 6 array (8.5 and 8.0 cm spacing in the vertical and horizontal dimensions, respectively) (Fig. 3). After the stimulus was retrieved, the opaque door of the WGTA was lowered for 8–10 sec, a negative cookie was placed in the previously rewarded location, and a positive cookie was positioned at 1 of the 17 remaining locations. This process was repeated, with previously rewarded positions occupied by negative stimuli, to a maximum list of four locations or until an error was committed. Training was conducted in the phases described below, to a predetermined performance criterion on lists of two, three, and finally four locations. Daily sessions consisted of 20 trials for two-item lists and 10 trials for three- and four-item lists (60 sec ITI). A final phase of testing incorporated probe trials intended to identify the strategies animals used to solve the task (described below).

Phase 1: 10-trial blocks of directionally oriented two-item lists. Sessions consisted of two 10-trial blocks of two-item lists, the first oriented from left to right on the stimulus tray and the second oriented from front to back. The baited locations were separated by no more than one unoccupied matrix point and balanced with respect to left–right and front–back positions within a session. Testing continued until animals achieved a criterion of ≥80% correct per block (i.e., 8 of 10 two-item lists completed correctly) across two consecutive sessions.

Phase 2: alternating five-trial blocks of directionally oriented two-item lists. Sessions were identical to phase 1 except stimuli were presented in alternating five-trial blocks of two-item lists oriented either left–right or front–back. Testing continued until choice accuracy was 80% correct across the 10 trials of each orientation for two consecutive sessions.

Phase 3: pseudorandom-order two-item lists. Sessions comprised two 10-trial blocks with left–right, front–back, and diagonally oriented two-item lists presented in pseudorandom order, balanced across days. Spacing between diagonally arranged stimuli was constrained to span no more than two unoccupied target locations of the spatial array. Testing continued until a criterion of 80% correct per 10-trial block was met for two consecutive sessions.

Phase 4: three-item lists. Testing was provided on three-item lists con-
strained by the spacing rules described previously, excluding linear configurations (i.e., all three items aligned in a single orientation). This phase was completed when accuracy was ≥80% across four consecutive sessions.

Phase 5: four-item lists. Testing continued in an identical manner using lists of four items until subjects met a criterion of 80% correct across four consecutive sessions.

Phase 6: probe tests. Multiple spatial cues were available to guide SDRS performance, including the local configuration of the test stimuli relative to each other and their locations in relation to the broader testing environment in which they appeared. Two types of probe trials, designed to identify the information animals used to solve SDRS, were provided in this final phase of testing.

Type-1 probes. Sessions consisted of standard four-item lists alternated with probe trials. The two trial types were identical except that before presenting the last item on probe tests, the previously completed three-stimulus list was shifted to occupy locations not used previously in the same trial while maintaining their local positions relative to each other (Fig. 3). As on standard trials, the fourth item also appeared at a previously unused location. Accordingly, for the final choice on type-1 probes, all four of the stimuli occupied novel locations relative to the overall configuration of the test apparatus and other ambient cues. Accuracy therefore depended on memory for the local spatial relationships among the test items themselves. Five sessions were provided, including a total of 25 type-1 probes.

Type-2 probes. Task strategies were explored further using a variant of probe testing that imposed an ambiguity between the local spatial relationships among the test stimuli versus the positions of these items in relation to the broader testing environment in which they appeared. Following the same format as type-1 probes, sessions consisted of standard trials alternated with probe tests in which the first three items of the list were shifted to novel locations before the presentation of the final choice. In this case, however, the fourth item appeared in one of the three positions occupied previously in the same trial (Fig. 3). Accordingly, the location of the fourth item was novel with respect to the local configuration of the test stimuli, but in contrast to type-1 probes, its position was familiar in relation to the reference frame provided by the experimental apparatus and environmental cues (e.g., uncontrolled auditory stimuli). Thus, although performance guided by the distribution of test stimuli would lead to selection of the fourth item on type-2 probes, the same item might be avoided if its position were defined in relation to the configuration of cues in the broader task setting. To avoid promoting a particular strategy, all four of the test stimuli were positive on the final phase of type-2 probe trials. Five sessions were provided, yielding a total of 25 trials each for the standard four-item lists and probes.

Perfusion and histological preparation

Animals were given ketamine hydrochloride (25 mg/kg, i.m.), deeply anesthetized with an overdose of sodium pentobarbital (20–35 mg/kg, i.v.), intubated, and mechanically ventilated. The heart was exposed and a vasodilator was injected into the left ventricle (1.5 mg of 0.1% sodium nitrite) to expedite subsequent fixation. The descending aorta was clamped, and the monkey was perfused transcardially with cold 1% (1 ml of 0.1% sodium paraformaldehyde (12–14 min at 230–260 ml/min) followed by 4% paraformaldehyde (12–14 min at 230–260 ml/min) in 0.1 M phosphate buffer, pH 7.2. Brains were postfixed in buffered fixative for an additional 6–9 hr and cryoprotected with a vasodilator was injected into the left ventricle (1.5 ml of 0.1% sodium paraformaldehyde (12–14 min at 230–260 ml/min) followed by 4% paraformaldehyde (12–14 min at 230–260 ml/min) in 0.1 M phosphate buffer, pH 7.2. Brains were postfixed in buffered fixative for an additional 6–9 hr and cryoprotected in 20% glycerol in 0.1 M phosphate buffer for 1 d. Whole brains were then blocked in the coronal plane. Cryoprotection continued in 20% glycerol in 0.1 M phosphate buffer for an additional 2 d before the tissue was frozen in isopentane (chilled in a dry ice–ethanol bath) and stored at −70°C. Brains were sectioned at 40 μm on a freezing sliding microtome, and a one in five series of sections (200 μm spacing) was mounted on gelatin-coated slides, stained with thionin (0.25%), and coverslipped.

Lesion analysis

The extent of the lesions was assessed qualitatively in a series of histological sections (200 μm spacing) through the medial temporal lobe examined at low magnification (10–50×) on a Zeiss Stemi SV8 stereomicroscope (Zeiss, Oberkochen, Germany). Subsequent quantitative analysis was conducted in more widely spaced sections (800 μm) from four of the controls and all of the experimental monkeys. The regional boundaries of the entorhinal cortex (Amaral et al., 1987) were digitized in each section at a final magnification of 80×, using a Leitz Medilux microscope (Leitz, Wetzlar, Germany) interfaced with a computer-aided morphometry system (StereoInvestigator; Microbrightfield, Colchester, VT). The volume of the entorhinal cortex was calculated for the control brains as the sum of the cross-sectional area measurements multiplied by the intersection interval (800 μm). Parallel analysis was performed to determine the percentage of the entorhinal cortex that was spared in the lesion brains. The cytoarchitectonic borders of the perirhinal and parahippocampal cortices and area TE were also digitized according to the description by Suzuki and Amaral (1994b, 2003), and the extent of unintended damage to these regions was quantified. Following the convention of a previous lesion study (Buffalo et al., 1999), the caudal limit of area TE in the present analysis was assumed to coincide with the rostral limit of the posterior medial temporal sulcus.

Statistical analysis

The behavioral results were analyzed by parametric statistics (e.g., factorial and repeated-measures ANOVA), except in those cases in which ceiling or floor effects dictated the use of nonparametric tests (e.g., Mann–Whitney U test). Planned follow-up comparisons were conducted using two-tailed unpaired or paired t tests, adopting the strategy of a previous study (Leonard et al., 1995). All statistical analyses were performed using StatView 5.0 (SAS Institute, Cary, NC) or SuperANOVA 1.11 (Abacus Concepts, Berkeley, CA) on a Macintosh computer (Apple, Cupertino, CA).

Results

Histological findings

Figures 4 and 5 schematically represent the intended lesion and the extent of observed medial temporal lobe damage, respectively. Photomicrographs of coronal sections through multiple rostrocaudal levels of the entorhinal cortex are shown in Figure 6.
for all of the experimental animals. All of the experimental monkeys sustained substantial bilateral damage to the entorhinal cortex, as intended. The volume of the entorhinal cortex in the control and lesion groups, respectively, averaged 206.4 and 27.8 mm³, and across the experimental subjects, 76–97% of the target region was ablated (mean, 86.5%). As described in the following sections, the lesion was restricted almost entirely to the entorhinal cortex in one animal and included varying degrees of unintended damage to other regions in the remaining subjects.

**Monkey E1**
The ablation in monkey E1 was confined almost entirely to the entorhinal cortex and most closely approximated the intended lesion (Fig. 6.1). Experimental damage to the entorhinal cortex involved all cell layers, constituting 78.8 and 73.1% of the structure in the left and right hemispheres, respectively. In general, the lesion was bilaterally symmetric throughout its rostrocaudal extent, with minor sparing observed in the most medial and lateral aspects of the entorhinal cortex. In the right hemisphere, the tissue removal began ~2.4 mm anterior to the target region, encroaching on superficial layers of the periamygdaloid cortex. The periamygdaloid cortex in the left hemisphere remained mostly intact. No damage was observed in the hippocampus, perilhinal or parahippocampal cortices, or area TE.

**Monkey E2**
Monkey E2 sustained the most complete entorhinal cortex lesion, involving 100% of the region in the left hemisphere and 94.7% of the region in the right (Fig. 6.2). Sparing was restricted to rostral levels along the medial bank of the rhinal sulcus. The lesion began ~4.0 mm rostral to the entorhinal cortex in the left hemisphere. At this level, unintended damage was restricted to superficial layers of the periamygdaloid cortex but expanded to include all cell layers caudally. Superficial damage was also observed in the periamygdaloid cortex in the right hemisphere for a distance of 3.2 mm anterior to the target region.

Aside from limited involvement of the periamygdaloid cortex, unintended damage in this animal was restricted primarily to the left hemisphere. An apparent infarct damaged ventral portions of the amygdala along most of its mediolateral and rostrocaudal extent and all cell fields of the rostral hippocampus, involving ~35% of its total rostrocaudal length. A narrow band of infarct-related necrosis was also observed in area TE, measuring ~3 mm at its widest mediolateral extent and 8.6 mm in the rostrocaudal axis. In addition, the experimental ablation extended caudally into parahippocampal area TH for a distance of 4–5 mm. This aspect of the lesion included all of the cell layers throughout the mediolateral extent of area TH and encroached upon the most medial aspect of the adjoining area TF at rostral levels. Notably, the perirhinal cortex was mostly spared in both hemispheres.

**Monkey E3**
Damage to the intended region was extensive throughout its rostrocaudal extent, involving 91.4 and 81.7% of the left and right entorhinal cortex, respectively (Fig. 6.3). Sparing in the left hemisphere was confined to a small segment at the anterior medial limit of the entorhinal cortex. Minor sparing was also observed in this region on the right side, and in this hemisphere, the medial bank of the rhinal sulcus was intact throughout its length. The lesion in animal E3 began 2.4 and 4.0 mm anterior to the rostral limit of the entorhinal cortex in the left and right hemispheres, respectively, extending through deep layers of the periamygdaloid cortex on both sides.
An infarct in the right hemisphere of animal E3 produced significant damage in area TE that extended from the posterior temporal pole to the TE–TEO border. At anterior levels, the damage was confined to area TE and spared ~35% of this region, which comprises the ventral bank of the superior temporal sulcus (Seltzer and Pandya, 1978). Immediately posterior to this level, the damage expanded medially, involving much of perirhinal area 36 for a rostrocaudal distance of 4.8 mm. Area 36 damage was confined to these levels, leaving ~72% of this cortical field intact. Involvement of the superior temporal gyrus emerged ~4.6 mm anterior to the rostral border of the occipitotemporal sulcus and extended caudally for a distance of 11.4 mm. Although the majority of the gyrus was anteriorly damaged for a short distance of 1.4 mm, the mediolateral extent of involvement narrowed caudally. Across these levels, dorsal aspects of the superior temporal gyrus were spared, and ultimately, only superficial damage restricted to the lateral surface was observed.

Unintended damage in the left hemisphere was less extensive. Infarct-related necrosis was apparent beginning at the posterior limit of the temporal pole, involving most of areas 36 and TE medial to the ventral bank of the superior temporal sulcus over a rostrocaudal distance of 7.2 mm. The mediolateral extent of damage narrowed near the rostral limit of the amygdala, appearing as a thin band of necrosis (5 mm at its widest) in deep cell layers throughout the length of area TE, restricted to the lateral convexity of the medial temporal gyrus. A small white matter abnormality was also observed subjacent to this region, in the caudal 8.0 mm of area TE. Otherwise, damage to areas TE and 36 resolved near the rostral limit of the hippocampus, sparing ~60% of the rostrocaudal length of both regions. Finally, cortical thinning was noted in the region of the craniotomy, involving the lateral surface of the superior temporal gyrus for a rostrocaudal distance of 17.2 mm.

To summarize, bilateral lesions of the entorhinal cortex were extensive in all experimental animals. Although additional, unintended damage was observed in two monkeys, both displayed substantial unilateral sparing. Specifically, infarct-related necrosis observed in the hippocampus and areas TE and TH in monkey E2 was confined to the left hemisphere. In monkey E3, although area TE and the superior temporal gyrus displayed significant damage in the right hemisphere, these regions were relatively spared on the opposite side. Notably, the perirhinal cortex was entirely intact in monkeys E1 and E2 and damaged only moderately bilaterally in monkey E3. As detailed later, the profile of behavioral deficits documented in experimental subjects appeared unrelated to the involvement of brain regions outside the entorhinal cortex.

**Behavioral findings**

**Standard tasks**

Subjects were tested first on two well characterized assessments that are known to be sensitive to large medial temporal lobe lesions [i.e., delayed nonmatching-to-sample (DNMS) and a series of two-choice object discriminations]. As described below, entorhinal cortex lesions caused no impairment at any point in training on these tasks.

**DNMS**

Experimental animals learned the DNMS procedure with a 10 sec delay as quickly as controls [mean trials to criterion ± SE: C = 182.2 ± 80.7; E = 165.3 ± 84.5; unpaired t test(7) = 0.1; p > 0.10] (Fig. 7). Previous studies reported that trials early in DNMS training are especially sensitive to memory impairment and can reveal reliable deficits even in cases in which the overall rate of acquisition appears normal (Alvarez et al., 1995). A similar ana-
lytic approach in the present experiment demonstrated that percentage correct responding was comparable in the control and experimental groups. A repeated-measures ANOVA revealed a significant main effect of test day on percentage correct responding (F(2,14) = 33.8; p < 0.001) but no group effect (F(1,7) = 0.2; p > 0.10) or group x day interaction (F(2,14) = 0.5; p > 0.10). Previous studies of OD learning in monkeys with selective hippocampal formation lesions reported deficits restricted to training trials early in testing (Teng et al., 2000). Parallel analyses in the present experiment failed to reveal group differences in the average percentage correct achieved across the first 3 (mean percentage correct ± SE: C = 73.6 ± 6.6; E = 66.7 ± 8.3), first 5 (C = 70.8 ± 7.1; E = 66.7 ± 6.0), or first 10 trials of testing (C = 74.6 ± 5.0; E = 71.7 ± 4.6; unpaired t tests, t < 0.7; p > 0.10).

Relational tasks

Paired associate

In this procedure, monkeys learned multiple conditional discriminations and were subsequently tested on probe problems that called for inferences or generalization across the familiar items presented in novel combinations.

Acquisition. PA training was conducted in phases (Table 1), beginning with the discontinuous series of conditional discriminations. Although overall acquisition of these problems was numerically slower in the E group relative to controls, this effect was not statistically reliable (mean trials to criterion ± SE: C = 371.8 ± 61.6; E = 623.3 ± 136.4; unpaired t test(7) = −2.0; p > 0.05) (Fig. 9). With one exception, performance also failed to differ between groups on each of the individual phases of testing (unpaired t tests, t < 2.4; p ≥ 0.90) (Fig. 9). The exception was that monkeys with entorhinal cortex lesions required significantly more trials than controls to achieve the performance criterion on phase 2 of training for the discontinuous discriminations (mean ± SE: C = 6.7 ± 4.2; E = 40.0 ± 20.0; Mann-Whitney U test = 2.0; p < 0.05; the presence of a floor effect for four of the six controls dictated the use of a nonparametric statistic for this comparison). Although previous test sessions presented individual problems that could be solved by simply learning to select one choice item over the other, phase 2 was the first in which each of the choices was positive and negative equally as often across trials, conditioned on the stimulus that appeared as the sample. Accordingly, it is tempting to speculate that the difficulty exhibited by the E group on phase 2 reflects the demands...
of testing on evaluating the discriminative stimuli in relation to each other.

Acquisition of the relational discriminations proceeded more slowly in experimental animals than in intact controls (mean cumulative trials to criterion ± SE: C = 220.5 ± 49.6; E = 513.0 ± 122.0; unpaired t-test(7) = −2.7; p < 0.05) (Fig. 9). This effect appeared to be attributable partly to the differential influence of previous training on subsequent performance. Specifically, although control monkeys learned the relational discriminations more quickly than the discontinuous pairs (paired t-test(5) = 3.2; p < 0.05), no facilitation was observed in the lesion group (paired t-test(2) = 1.4; p > 0.10). These findings suggest that over the course of training, intact subjects acquired a general strategy for solving conditional discrimination problems, and perhaps monkeys with entorhinal cortex damage relied on a different approach.

The pattern of results for the individual instruction phases of the relational series mirrored findings from the discontinuous series. Specifically, learning progressed at comparable rates on phases 1 and 3–6 (unpaired t-tests, t < 0.8; p > 0.10) (Fig. 9), but lesion monkeys were significantly impaired, relative to controls, in reaching the phase 2 criterion (mean trials to criterion: C = 16.7 ± 9.5; E = 105.0 ± 27.5; Mann–Whitney U test = 0.5; p < 0.05). As noted for the relational series, this pattern of impairment is interesting in that phase 2 was the first to emphasize learning according to predictive relationships among the discriminative stimuli. With sufficiently extensive training, however, monkeys in both groups achieved similar high levels of accuracy before progressing to the critical probe phase of the task.

**Relational series probe tests.** Probe problems tested whether monkeys were able to infer the relationship between sample and choice items presented in novel combinations on the basis of memory for the associative links between stimuli from different premise problems (e.g., AC and XZ) (Fig. 1). As shown in Figure 10, the experimental and control groups scored comparably on the familiar instruction discriminations during probe testing, maintaining accuracy at or near criterion levels (mean percentage correct ± SE: C = 86.4 ± 2.8; E = 79.6 ± 4.4; unpaired t-test(7) = 1.4; p > 0.10). Averaged across the critical probe tests, inferential responding among intact monkeys (i.e., selecting C over Z when A was the sample, and Z over C when X was the sample) was substantially above chance (80.8 ± 4.7; one-sample t-test(5) = 6.5; p < 0.01) and as accurate as their performance on the instruction problems (paired t-test(5) = 1.5; p > 0.10) (Fig. 10). Subjects with entorhinal cortex lesions, in comparison, performed significantly worse than group C on the probes (unpaired t-test(7) = 2.4; p < 0.05) and failed to score reliably above chance (63.3 ± 3.3; one-sample t-test(2) = 4.0; p > 0.05) (Fig. 10). Notably, there was no evidence of improvement over the course of testing on the relational probe problems for either group (single-group repeated-measures ANOVA: C, F(5,20) = 1.9; p > 0.10; E, F(2,8) = 0.2; p > 0.50), which counts against the interpretation that control subjects learned the probes as novel conditional discriminations.

**Discontinuous series probe tests: new associative learning.** In contrast to the relational series, the discontinuous instruction discriminations shared no bridging, overlapping stimuli to provide a basis for drawing inferences across pairs. Accordingly, the infrequently presented probe problems derived from this series permitted an evaluation of learning for new conditional problems composed of familiar items. The control and experimental groups performed comparably during refamiliarization with the instruction discriminations (see Materials and Methods) (mean trials to criterion ± SE: C = 49.7 ± 16.0; E = 0; p > 0.05), and they maintained criterion levels of accuracy on these pairs across probe sessions (mean percentage correct ± SE: C = 88.3 ± 2.3; E = 82.9 ± 2.1; unpaired t-test(7) = 1.5; p > 0.10) (Fig. 10). The overall result from the probe tests was that control animals...
learned the novel conditional discriminations rapidly (mean percentage correct $\pm$ SE across trials $= 71.7 \pm 5.7$; one-sample $t$ test $(5) = 3.8; p = 0.01$ relative to chance) over a period during which experimental animals showed no acquisition (mean $\pm$ SE across trials $= 55.0 \pm 5.0$%; one-sample $t$ test $(2) = 1.0; p > 0.10$ relative to chance). A repeated-measures ANOVA confirmed that performance among intact monkeys improved significantly over the course of probe testing (repeated-measures ANOVA; $F(5,20) = 2.9; p < 0.05$), but no learning occurred in the experimental group ($F(2,8) = 0.4; p > 0.50$). Importantly, the control results were qualitatively different on the relational probes in which, as noted previously, peak accuracy was displayed from the outset of training at a level equivalent to the premise pairs from which the probes were derived. In contrast, on probes from the discontinuous series, intact subjects performed significantly worse than on corresponding instruction problems (paired $t$ test $(5) = 3.0; p < 0.05$). Together, these findings suggest that although normal monkeys with substantial previous experience are capable of rapidly learning novel conditional discriminations, this mechanism is not sufficient to account for their success in solving the relational probe tests.

In summary, findings from the PA procedure demonstrate that normal monkeys readily solve problems in which indirectly related items are presented in novel configurations and are capable of robust inferential responding on the basis of previous learning. In contrast, animals with entorhinal cortex lesions display modest impairment in the initial acquisition of conditional discriminations and substantial deficits under conditions that call for the flexible manipulation of memory for the salient relationships between experienced items.

Transitive inference
In this procedure, monkeys initially learned five two-choice conditional discriminations with overlapping elements (e.g., $A + B^-$, $B - C^-$, $C + A^-$, $D + E^-$, $E - D^-$), and then were tested on a series of transitive inference problems. Performance on these tasks was assessed using a variety of measures, including the percentage of correct responses, the number of trials to criterion, and the number of errors made. The results showed that normal monkeys were able to transfer their knowledge of the conditional discriminations to new situations, whereas experimental animals with entorhinal cortex lesions performed significantly worse on these tasks. These findings support the hypothesis that transitive inference is a complex cognitive skill that requires the integration of information across multiple trials and across different stimuli, and that this ability is impaired in monkeys with entorhinal cortex lesions.
B+C−, C+D−, and D+E−). Subsequent probe testing (e.g., BD) (Fig. 2) was designed to determine whether animals learned the stimuli according to the hierarchical relationships between them and whether memory was represented in a manner accessible to flexible manipulation. Before testing on the TI task, one control subject was killed for health reasons unrelated to the experimental procedures.

**Acquisition.** Training progressed in phases, ultimately requiring criterion performance on all four conditional discriminations presented in pseudorandom order (phase 4). The control group learned the task rapidly and completed all phases of acquisition in a mean cumulative total of 284.6 ± 38.1 trials. Although the corresponding score for the experimental group was greater (466.3 ± 170.9), the difference was not statistically reliable (unpaired t test(6) = −1.3; p > 0.10) and resulted from slow acquisition in one subject on phase 4 of testing. Otherwise, performance on the individual acquisition phases was similar across groups (phase 1 mean trials to criterion ± SE: C = 11.4 ± 2.3; E = 8.0 ± 5.5; phase 2 mean percentage correct ± SE: C = 86.5 ± 1.6%; E = 88.1 ± 1.6%; phase 3 mean percentage correct ± SE: C = 80.5 ± 5.4%; E = 79.2 ± 9.9%; phase 4 mean trials to criterion ± SE: C = 83.0 ± 38.5; E = 266.3 ± 174.3; unpaired t tests, t < 1.3; p > 0.10).

**Probe tests.** Probe sessions were similar to the final phase of acquisition but included two presentations of each probe problem, BD and AE. In comparison with the AE pair, which could be solved simply by the differential reinforcement value of the stimuli, inferring the correct choice on the BD probe placed greater emphasis on the flexible manipulation of memory for the hierarchical relationships between the items. One control monkey, unlike any of the other intact or experimental animals, failed to maintain the criterion level of accuracy on the C+D− instruction pair during probe testing (77.5% compared with the lowest scores for the remaining control and lesion subjects, 92.5 and 86.2%, respectively). The same monkey also required nearly four times as many trials as the remaining controls to complete phase 4 of training. On the basis of the view that BD probe performance critically depends on having learned the instruction discriminations from which the probe was derived (i.e., pairs B+C− and C+D−), results from the control subject with abnormally poor instruction pair performance were excluded from the probe analysis.

Among the remaining monkeys, accuracy on the instruction discriminations that contained the probe stimuli was similar across groups (mean percentage correct ± SE during probe testing for the end-anchored discriminations A+B− and D+E−: C = 90.3 ± 2.2%; E = 94.6 ± 2.9%; mean percentage correct ± SE for middle problems B+C− and C+D−: C = 92.6 ± 1.2%; E = 90.6 ± 4.2%; unpaired t tests, t < 1.2; p > 0.10). The intact and lesion monkeys averaged 98.8 ± 1.2 and 100% correct, respectively, on the end-anchored probe pair AE. This preserved capacity is informative, demonstrating that the presentation of familiar items in novel combinations is not itself sufficient to disrupt performance in animals with entorhinal cortex damage. On the critical BD trials, in contrast, only the intact controls displayed reliable inferential responding. Indeed, although normal subjects selected B over D on 100% of the probes, inferential responding in the experimental group was significantly lower (Mann–Whitney U test = 0; p < 0.05) and failed to exceed chance (mean percentage correct ± SE = 73.3 ± 14.8%; one-sample t test(2) = 1.6; p > 0.10). These findings demonstrate that although monkeys with lesions of the entorhinal cortex are able to learn multiple conditional discriminations concurrently, they display severe impairment, relative to controls, under conditions that call for the flexible manipulation of memory for the hierarchical relationships among familiar items.

**Spatial delayed recognition span**

This procedure required monkeys to recognize novel positions in an increasing list and tested memory for the spatial relations that defined the rewarded locations.

**Acquisition.** Control subjects required an average ± SE of 164.8 ± 67.8 trials to solve the two-item problems of phase 1. Normative acquisition was even more rapid in subsequent phases of training, despite the requirement to remember longer lists of locations (mean trials to criterion ± SE: phase 2, 19.5 ± 11.3; phase 3, 10.0 ± 5.8; phase 4, 12.5 ± 4.8; phase 5, 27.5 ± 13.8). In contrast, the experimental group was severely impaired, and none of the animals with entorhinal cortex lesions mastered two-item lists within the limits of testing. Specifically, one monkey failed to reach criterion after 1000 trials on phase 1. The remaining two required an average of 495 and 370.5 trials to complete phases 1 and 2, respectively, but both failed to achieve the performance criterion on phase 3 after 500 trials. The results from probe testing for the control group, described below, illuminate the likely basis of this pronounced deficit.

**Probe tests.** Sessions consisted of standard four-item lists alternated with probe tests that were designed to identify the information animals used to solve SDRS. For type-1 probes, before the final item was presented, the previously completed three-item list was shifted to novel locations on the stimulus tray while the positions of the test stimuli were maintained relative to each other (Fig. 3). The fourth item was also presented at a position not used previously in the same trial. By this arrangement, all four stimuli occupied new locations with respect to the reference frame defined by the test apparatus and other ambient cues, and the position of the last item was distinctive only in relation to the local configuration of the three test stimuli presented previously in the trial. Across probe sessions, control animals demonstrated stable retention for procedural aspects of the task, averaging 84.6 ± 7.2% correct on the standard four-item lists (one-sample t test relative to chance, t test(3) = 8.2; p < 0.005). Errors involving the first three locations presented on probe trials were infrequent, and overall, control subjects completed a mean of 22.2 ± 0.8 type-1 probes. Mean probe accuracy substantially exceeded chance (79.4 ± 3.2%; one-sample t test, t(3) = 17.2; p < 0.001) and was statistically equivalent to performance on the standard four-item lists (paired t test(3) = 0.7; p > 0.50). The conclusion from these findings is that in normal monkeys, memory for the spatial relationships among the test stimuli is sufficient to support accurate SDRS performance.

Type-2 probes were procedurally similar, with the critical exception that in this case the fourth list item appeared at a position that was occupied previously in the same trial (Fig. 3). Accordingly, the location of the last item was only novel with respect to the local configuration of test stimuli, and accuracy was expected to suffer if performance is also guided by memory for the spatial positions of the stimuli in relation to the reference frame of the test apparatus and other elements of the experimental setting (e.g., lighting, fixed auditory cues, etc.). Consistent with this prediction, performance on the type-2 probes failed to exceed chance in the control group (mean percentage correct ± SE = 23.4 ± 5.5%; one-sample t test(3) = −0.3; p > 0.50). This effect was not the result of a generalized decline in accuracy because scores for the interleaved standard trials averaged 94.4 ± 0.9% (one-sample t test relative to chance, t(3) = 74.3; p < 0.0001). Together, these findings suggest that normal, intact
monkeys encode the full scope of information available to solve SDRS, including local spatial cues provided both by the test stimuli themselves and by the configuration of these items in relation to the broader setting in which they appear. The observation that monkeys with entorhinal cortex lesions were unable to master even the simplest version of the task supports the conclusion that entorhinal integrity is critical for this capacity.

Discussion

Declarative memory involves the recollection of information embedded in a rich context of related events, engaging a densely interwoven fabric of representations that comprises the mental record of our experience. The findings reported here demonstrate that this capacity critically requires the integrity of the entorhinal cortex in nonhuman primates. Monkeys with entorhinal lesions scored normally on standard testing procedures that can be solved on the basis of memory for individual stimuli (i.e., DNMS and OD). Experimental animals also succeeded in acquiring the more difficult instruction discriminations of the PA and TI tasks, although in this case acquisition tended to be slower than normal, particularly under conditions that emphasized evaluating multiple stimuli in relation to each other. These findings suggest the possibility that the lesion group took advantage of alternate task strategies, perhaps guided by configurational cue arrangements such as choosing “B” when presented with the configuration “B–A–Y,” but selecting “Y” when given “B–X–Y” (Gallagher and Holland, 1992; Murray et al., 1993; Bussey and Eichenbaum, 1996; Bussey et al., 1998, 2000). Although the specific cognitive mechanisms that support successful acquisition of these instruction discriminations remain to be investigated directly, the results of probe testing provide compelling evidence that entorhinal cortex damage profoundly disrupts the representational structure of previous experience. Specifically, monkeys with entorhinal cortex lesions displayed severe impairment under all testing conditions that involved using memory flexibly to determine the relationships between familiar items presented in novel combinations or spatial arrangements. These observations and related findings in rats and humans (Davachi and Wagner, 2002; Fortin et al., 2002; Kesner et al., 2002) support the conclusion that the mammalian hippocampal system plays a critical role in the relational organization of declarative memory.

On average, nearly 90% of the entorhinal cortex was ablated bilaterally among experimental subjects in the present study. Additional incidental damage observed in two monkeys (E2 and E3) was predominantly unilateral, and in both cases there was substantial sparing of memory-related temporal lobe structures outside the entorhinal cortex. Notably, like the monkey that sustained the most selective ablation (E1), subjects E2 and E3 performed as accurately as intact controls on DNMS and OD. Thus, unintended damage failed to affect performance on two well characterized tests that are known to be exquisitely sensitive to extensive medial temporal lobe lesions (Zola-Morgan and Squire, 1985; Meunier et al., 1993, 1996). An additional implication is that the deficits observed after removal of the entorhinal cortex are unlikely to reflect a secondary disruption of memory processing mediated by the heavily interconnected perirhinal and parahippocampal cortices, in which direct damage severely impairs DNMS performance (Suzuki et al., 1993). The findings from standard assessments thereby established a baseline of preserved function for gauging the relative sensitivity of other processing capacities to damage involving the entorhinal cortical component of the hippocampal formation. A long-standing view derived from multiple lines of evidence is that hippocampal integrity is critical for learning and remembering spatial information (for review, see Eichenbaum et al., 1999; Burgess et al., 2002). Our results are consistent with this conclusion and demonstrate that entorhinal cortex lesions essentially prevent acquisition in a spatial recognition span task. Although previous studies examined the effects of hippocampal lesions using various related testing procedures (Parkinson et al., 1988), this is the first to document such severe impairment under conditions in which we directly confirmed that intact control performance is supported by memory for the spatial relationships between task stimuli.

The deficits caused by entorhinal cortex ablation were not selective for spatial information, however, and impairment of approximately equivalent magnitude was observed on additional tasks that lacked any explicit demand on spatial learning and memory. Probe tests in the PA and TI procedures asked monkeys to make inferential judgments about the relationship between familiar items that were associated only indirectly by virtue of their shared relationship with a linking or overlapping element. Consistent with previous findings in rats and nonhuman primates (McGonigle and Chalmers, 1992; Bunsey and Eichenbaum, 1996; Rapp et al., 1996; Dusek and Eichenbaum, 1997), intact rhesus monkeys scored reliably above chance on probe tests of this sort, reflecting flexible access to representations of associations acquired during initial training. Thus, rather than affecting memory for a particular class or modality of stimulus materials, the result of entorhinal cortex damage is more accurately understood as an effect on an underlying information-processing capacity. Specifically, our findings imply that intact hippocampal processing is critical for encoding experienced items and events according to the relevant relationships between them and for enabling an organization of representations that can be flexibly accessed, navigated, and expressed.

Because greater consensus has emerged concerning the anatomical structures critical for declarative memory, interest has turned increasingly to second-generation questions about the functional organization of this system. A central issue is whether the component structures of the medial temporal lobe mediate distinct capacities in support of normal memory. Functional specialization might be predicted on the basis of striking differences in cytoarchitectonic and connectional characteristics, and indeed, a number of studies reported dissociations in the behavioral effects of disrupting hippocampal function compared with the consequences of damage to the perirhinal and parahippocampal cortices (Gaffan, 1994; Meunier et al., 1996). Other data, however, suggest that the severity of impairment after medial temporal lobe damage varies primarily as a function of the extent of involvement and that the fundamental nature of the deficit is similar regardless of the specific site of damage (Zola-Morgan et al., 1994). In the present experiments, entorhinal cortex ablation failed to affect performance on two tasks with well documented sensitivity to larger lesions (Zola-Morgan and Squire, 1985; Meunier et al., 1993, 1996). Supporting the conclusion that accuracy was fully intact, monkeys in the lesion group scored normally on those specific phases of testing that were thought to provide especially sensitive measures of memory impairment (i.e., trials early in the course of DNMS acquisition and initial trials during the first few sessions of object discrimination learning) (Alvarez et al., 1995; Teng et al., 2000). This pattern of results is compatible with previous reports demonstrating that entorhinal cortex lesions cause, at most, modest deficits in recognition memory that completely resolve with continued testing (Meunier et al., 1993; Leonard et al., 1995). Against this background of preservation, experimental damage in the current
study led to severe and selective impairment under all conditions in which intact animals displayed memory for the relevant relationships between familiar items and the ability to access this representational organization flexibly when confronted with novel stimulus arrangements. Together, these findings suggest that the hippocampal formation selectively or disproportionately mediates relational representation and that the broader profile of impairment observed in amnesia reflects the contributions to declarative memory of neocortical components of the medial temporal lobe system. To confirm this proposal in monkeys, future studies will need to compare directly the effects of hippocampal and parahippocampal region lesions using testing procedures that aside from manipulating demands on relational memory are matched along other relevant dimensions, including overall task difficulty.

Although impairment in establishing a consciously accessible record of day-to-day episodes and events is a central feature of amnesia, the specific processing capacities that enable episodic memory remain to be defined. One class of theories focuses on the content of encoded information, suggesting that memory is organized according to the "what, when, and where" dimensions of experience (for review, see Clayton et al., 2003). By comparison, our findings suggest that the primate hippocampal system mediates a more general processing function, abstracting the salient relationships between items and events rather than encoding memory for the items themselves. Results from recent imaging studies in humans converge on a similar conclusion, demonstrating that although the parahippocampal region is activated during recognition judgments for individual stimuli, the hippocampal formation is selectively engaged when subjects survey memory according to the associations among familiar items (Davachi and Wagner, 2002; Heckers et al., 2004; Preston et al., 2004). Of particular note, two investigations found that the hippocampus is selectively activated during transitive inference judgments of the sort that the present results indicate require the integrity of the hippocampal system (Heckers et al., 2004; Preston et al., 2004). What remains to be determined is whether this capacity comprises a fundamental mechanism enabling other central properties of memory. Experiments aimed at defining the contribution of relational representation to the temporal and autobiographical organisation of remembered information, for example, will be an important step toward understanding the structure of episodic memory in the primate brain.

References
McCabe et al. • Relational Memory in Monkeys