Hippocampal Atrophy Disrupts Transfer Generalization in Nondemented Elderly

Catherine E. Myers, PhD, Alan Kluger, PhD, James Golomb, MD, Steven Ferris, PhD, Mony J. de Leon, EdD, Geoffrey Schnirman, PhD, and Mark A. Gluck, PhD

ABSTRACT

Specific reductions in hippocampal volume in nondemented elderly individuals with mild cognitive impairment have been shown to correlate with future development of Alzheimer's disease (AD). Hippocampal atrophy (HA) is also correlated with cognitive impairments, leading to the promise of behavioral markers for early AD. Prior theoretical work has suggested that hippocampal dysfunction may selectively impair generalization involving novel recombinations of familiar stimuli. In this study, nondemented elderly individuals were trained on a series of concurrent visual discriminations and were then tested for transfer when stimulus features were recombined in new ways. Presence or absence of HA, revealed by neuroimaging, was not correlated with concurrent discrimination performance; however, individuals with mild HA showed significant decreases in transfer performance relative to nonatrophied participants. These preliminary results suggest that even very mild degrees of hippocampal atrophy may be associated with subtle behavioral impairments. (*J Geriatr Psychiatry Neurol* 2002; 15:82–90).

Cognitive impairment, especially memory dysfunction, occurs in the course of human aging and precedes the development of Alzheimer's disease (AD), a degenerative neurologic illness.[1,2] Older individuals with subjective memory complaints frequently test within age-adjusted means on standardized neuropsychological tests; however, recent studies have shown a correlation between mild memory impairments and increased risk for developing AD.[1–3]

The hippocampus and associated structures within the medial temporal lobe are critical substrates of memory formation,[4] and progressive hippocampal dysfunction has been proposed as a possible neuroanatomic basis of AD.[5–7] Neuroradiographic evidence for hippocampal formation tissue loss has been observed in patients with mild AD[8–12] and in elderly individuals exhibiting cognitive impairment suggestive of possible AD[13–15]; further, the hippocampal formation was the only temporal lobe subvolume sensitive to these small cognitive differences.[16] In fact, for older adults with mild cognitive impairment, hippocampal atrophy (HA) can predict whether an individual with mild cognitive impairment is at short-term risk of subsequently developing dementia.[8] Such early prediction is critical, given that all existing AD therapies work to arrest the progress of the disease rather than reverse its effects.

Several previous studies with cognitively normal older adults have shown that HA correlates significantly with performance on memory tests,[17–20] although other studies have found that hippocampal volume does not correlate strongly with memory performance.[21,22] Recall of verbal information may be especially sensitive to hippocampal volume.[21,23] In one study considering nondemented elderly individuals, HA was shown to predict longitudinal decline on tests of delayed paragraph recall.[23] Paragraph delayed recall (PDR) tasks are also particularly sensitive to hippocampal region damage resulting from other etiologies.[4] Therefore, it seems plausible that other tasks that are also sensitive to hippocampal region damage might similarly be disrupted in individuals with mild HA, and such hippocampal-sensitive tasks might be disrupted in early AD before more general cognitive deficits are evident.

Received December 7, 2000. Received revised July 17, 2001. Accepted for publication August 1, 2001.

From the Department of Psychology (Dr. Myers), Rutgers University, Newark, New Jersey; the Department of Psychology (Dr. Kluger), Lehman College/CUNY, New York; the Departments of Psychiatry (Drs. Kluger, Ferris, and de Leon) and Neurology (Dr. Golomb), New York University Medical Center, New York, New York; Nathan Kline Institute (Dr. de Leon), New York, New York; Massachusetts Mental Health Center (Dr. Schnirman), Boston, Massachusetts; and the Center for Molecular and Behavioral Neuroscience (Dr. Gluck), Rutgers University, Newark, New Jersey.

This work was partially funded as a pilot subproject under National Institute on Aging (NIA) grant #P30 AG08051 to the NYU Aging and Dementia Research Center (to Dr. Ferris), by NIA grants RO1 AG12101 and RO1 AG13616 (to Dr. de Leon), and by a grant from the Alzheimer's Association (to Dr. Gluck).

Reprint requests: Dr. Catherine E. Myers, Department of Psychology, Rutgers University, Newark, NJ 07102.

A large literature on animal models has suggested that particular classes of memory task are especially sensitive to hippocampal region damage. One theme that unifies many of the hippocampal-sensitive tasks is the trade-off between generalization and specificity. For example, normal rats trained to choose the rewarded odor from a pair of odors (e.g., A+B–, C+D–, etc.) would transfer well to novel recombinations of familiar odors (e.g., A+D–, C+B–).[24] However, animals with hippocampal region dysfunction subsequent to fornix transection performed at chance on these novel recombinations. This effect has been interpreted as indicating that hippocampal-lesioned animals overcompress odors: perceiving an AB compound rather than its component odors A and B; presentation of AD therefore represents a novel compound rather than a recombination of familiar components.[25,26] Humans with hippocampal region damage are also often characterized as displaying a similar "hyperspecificity": they are able to retrieve studied information when study and test conditions are identical but not when test conditions are varied.[27,28] We therefore hypothesized that individuals with HA might show a similar impairment when challenged to respond to familiar cues in novel recombinations and that this impairment might be evident before the appearance of more generalized cognitive and memory deficits.

In the experiments presented here, nondemented elderly participants were trained on a series of eight concurrent discriminations. Each discrimination pair consisted of two objects that varied in color or shape but not both. Thus, for each discrimination pair, one stimulus feature — color or shape — was relevant and one was irrelevant. For example, in one pair, a red triangle might be rewarded but a yellow triangle might not; thus, color was relevant but shape was irrelevant. Once the discrimination pairs were mastered, a transfer phase followed in which the relevant features remained the same but the irrelevant features were altered. For example, the discrimination pair might now involve a red circle and a yellow circle; the red object was still rewarded even though its shape had changed.

We expected that individuals without HA would learn associations involving the relevant stimulus features but not the irrelevant features. In the example above, this would mean learning to choose red over yellow, regardless of shape. In the transfer phase, these individuals should continue to perform well since the relevant features remain the same. On the other hand, theoretical work[25] and animal studies[26] predict that selective hippocampal damage or dysfunction should impair the ability to selectively learn about relevant features; instead, individuals with HA should solve the discrimination task by learning about entire stimulus items — for example, learning to prefer a red triangle over a yellow triangle. Such individuals should perceive the transfer items red circle and yellow circle as novel and show little or no transfer of prior learning. Thus, our prediction was that although individuals with mild HA might perform well on the original discriminations, they would show significantly less transfer than participants without HA.

METHODS

Participants

Thirty-four individuals, including 16 females and 18 males, aged 45 to 80 years, participated in the current study. Participants were recruited through the New York University (NYU) Medical Center Aging and Dementia Research Center, where they were given neuroradiographic and neuropsychological workups as part of ongoing clinical and research projects. To be eligible to participate in the current study, participants were required to be medically healthy. Exclusion criterion included clinical or radiographic evidence for structural or metabolic central nervous system abnormalities (including cerebral infarction), more than borderline hypertension (> 160/90 mm Hg), history of excessive alcohol intake, or significant cardiovascular, rheumatologic, endocrinologic, hematologic, neoplastic, pulmonary, or psychiatric disorders (including depression). Participants were also required to be native English speakers and to be taking no medication affecting cognition. A further inclusion criterion was nondemented status, as revealed in clinical assessment (detailed below). Informed consent was obtained from all participants prior to the start of behavioral testing.

Neuroimaging

Each participant received magnetic resonance imaging (MR), using a 1.5-T GE Advantage system (General Electric, Milwaukee, Wisconsin). Diagnostic screening was completed to identify and exclude participants with MRI evidence of infarct, hydrocephalus, intracranial mass, or moderate to severe white matter lesions.

Scans were obtained from a GE 1.5-T MRI scanner using a three-dimension spoiled gradient recalled acquisition (SPGR) sequence. Data were acquired in the coronal plane, including using a T_1-weighted sequence: TR 35 ms, TE 9 ms, 60-degree flip angle, 256 × 128 acquisition matrix, 1.3-mm section thickness, 18-cm field of view, for a total acquisition time of 9 minutes. The SPGR study provided data for reformatted (coronal, axial, sagittal, and oblique) images.

For the purposes of qualitative HA ratings, 3-mm axial reformats were created from the coronal SPGR scan, parallel to the long axis of the hippocampus. Detailed examination of two to four axial slices parallel to the plane of the hippocampus was made by a highly experienced observer (Dr. de Leon), who was blind to the participant's clinical status and behavioral performance. Evidence of enlargement of the hippocampal fissures (transverse, choroidal, and hippocampal) was subjectively determined using a 4-point rating scale.[8,29] Assess-

Table 1. Breakdown of Neuroimaging Results for the 18 Subjects in Group HA According to HA Rating to the Left and Right Hippocampus

HA Rating		
Right Side	*Left Side*	*Number of Subjects*
0	1	5
0	2	1
1	0	3
1	1	1
1	2	4
2	1	1
2	2	2
3	3	1

HA = hippocampal atrophy.

ments ranged from rating of 0 = no atrophy through 1 = questionable or mild atrophy, 2 = mild to moderate atrophy, and 3 = moderate to severe atrophy. Such subjective evaluation of scans has proven to be highly correlated with objective (quantitative) volumetric measures.[13,29]

Normally, ratings of 2 and higher are considered to indicate definite evidence of hippocampal atrophy and risk for subsequent cognitive decline and AD risk. In the current study, however, we were primarily concerned with elderly individuals who show minimal evidence of cognitive decline and correspondingly little evidence of HA. Accordingly, for the purposes of the current analysis, participants were assigned to the no HA group if there was no evidence of any HA (rating = 0 bilaterally); they were assigned to the HA group if there was at least some evidence of HA (rating > 0 on at least one side). Thus, our definition of HA includes even very mild degrees of atrophy, which would not normally be considered significant to indicate risk of cognitive decline.

According to this measure, 18 participants were determined to have at least questionable or mild HA, whereas 16 participants had no visible evidence of HA. Table 1 summarizes these findings. Within the HA group, there were more participants with atrophy on the left than the right side; 9 subjects received a rating of at least 2 on at least one hippocampus.

Clinical and Neuropsychological Assessment

All participants were assessed for cognitive impairments using the Global Deterioration Scale (GDS),[30] which ranks individuals according to a 7-point scale. The GDS 1 rating is given to an individual with no memory impairment. The GDS 2 rating is given to an individual who is functionally unimpaired but with subjective complaints of mild forgetfulness that is not recognized by family members or coworkers and for which there is no clinical evidence. The GDS 3 rating is given to an individual with subtle functional deficits and mild cognitive impairment, revealed with extensive clinical interview. Whereas GDS 3 rating does not indicate dementia, individuals with GDS 3 ratings are at heightened risk to subsequently develop AD, compared with individuals given GDS ratings of 1 and 2.[1,3] Global Deterioration Scale ratings of 4 and higher indicate dementia with increasingly severe cognitive and functional impairments; GDS 4 is often considered indicative of mild AD.

To be included in the current study, individuals were required to have ratings of GDS 3 or lower, indicative of nondemented clinical status. Overall, participants in the current study had an average GDS rating of 2.25 (SD = 0.43).

In addition to GDS rating, all participants were administered the PDR component of the NYU Paragraph Recall Test (derived from the Guild Memory Test[31,32]). In this test, the experimenter reads a brief paragraph and immediately asks the participant to repeat as much of the content as possible. The paragraph is then read a second time, followed by a delay of 5 to 10 minutes during which other tasks are administered. Finally, the participant is again asked to recall as much content as possible. Participants are given a PDR score reflecting how many elements are recalled verbatim, averaged over two paragraphs (maximum score = 21.5).

This test has been shown to be sensitive to the effects of aging, mild cognitive impairment (MCI), and early AD.[1,33,34] Additionally, the PDR score has been shown to be correlated with medial temporal lobe (especially hippocampal) size, revealed through neuroimaging in nondemented elderly.[17,18] For example, Golomb et al[17] found that a sample of nondemented (GDS 1–2) elderly subjects (mean age = 67.9 years) with no HA received an average PDR score of 10.5, whereas a similar group of nondemented individuals with HA averaged 8.08. Thus, poor PDR scores correlated with HA. Additionally, PDR score has been used to classify which individuals among a group of nondemented elderly will or will not show subsequent cognitive decline, particularly from MCI to AD.[3] Participants in the current study received an average score of 8.07 on this test (SD = 0.51, range 2–12).

Additionally, most participants were given the Finger Tapping Test,[35,36] a test of manual dexterity and gross motor speed that may also be predictive of daily living skills in elderly patients with possible dementia. It involves five 10-second trials with each hand. Bornstein[37] provides normative data for individuals aged 60 to 69 years with at least a grade 12 education: for males, a score of 43.0 (SD = 4.7) with the dominant hand and 39.3 (SD = 6.2) with the nondominant hand; for females, 35.2 (SD = 6.0) with the dominant hand and 32.0 (SD = 4.9) with the nondominant hand. In the current study, scores were averaged across the two hands; male participants received an average score of 39.7 (range = 27.7–52.1), whereas female participants averaged 31.62 (range = 15.5–40.9).

Summary demographic, clinical, and neuropsychological data for the HA and no HA groups are given in Table 2. There were more males than females in the

Table 2. Summary of Demographic, Clinical, and Neuropsychological Profile for the HA and No HA Groups

	HA Group (n = 16)	*No HA Group (n = 18)*
Gender	5 female, 13 male	11 female, 5 male
Age (yr)	70.60 (9.26)	65.56 (9.61)
GDS rating	2.22 (0.43)	2.25 (0.45)
PDR score	8.17 (3.35)	7.97 (2.59)
Finger tapping	36.99 (7.80)	31.64 (7.50)

Standard deviations given in parentheses.
HA = hippocampal atrophy; GDS = Global Deterioration Scale; PDR = paragraph delayed recall.

HA group, whereas the reverse was true for the no HA group; this was a significant difference in gender distribution (Yates's corrected χ^2 test: $\chi^2 = 5.71$, df = 1, $P < .05$). Otherwise, there were no significant differences between the HA and no HA groups in age (independent-samples t-test; $t(32) = 1.39$, $P > .05$), GDS rating ($t(32) = 0.19$, $P > .05$), or PDR score ($t(32) = 0.19$, $P > .05$). On the finger tapping measure, which is generally sensitive to subject gender, there was no significant difference between the HA and the no HA group, analysis of variance (ANOVA): $F(1, 16) = 0.27$, $P > .5$, no effect of subject gender, $F(1, 16) = 2.36$, $P = .144$, and no group–gender interaction, $F(1, 16) = 0.01$, $P > .5$.

Behavioral Task

Apparatus

Behavioral experiments were automated on a Macintosh LC, Powerbook, or equivalent computer, with a color screen, programmed in the SuperCard language (Allegiant Technologies, San Diego, CA). Testing took place in a quiet room, with the participant seated in front of the computer at a comfortable viewing distance. The keyboard was masked except for two keys, labeled "LEFT" and "RIGHT," which the participant could press to record a response.

Stimuli

Phase 1 of the experiment was a concurrent discrimination. Stimuli consisted of 16 colored shapes, organized into the eight discrimination pairs shown in Figure 1. Four of the pairs (C1–C4) differed in color (relevant feature) but not in shape (irrelevant feature); four pairs (S1–S4) differed in shape (relevant feature) but not in color (irrelevant feature). Within each discrimination pair, one stimulus was designated as rewarded (+) and one was designated as nonrewarded (–). Assignments of particular color, shape, and reward to discrimination pairs were made according to a pseudorandom procedure but were held constant across the experiment. Both color and shape features were selected to be highly distinguishable within and across stimulus pairs. Each stimulus appeared to be about 2.5 cm square on the computer screen, with approximately 7.5 cm between members of a discrimination pair.

Phase 2 of the experiment was a transfer test. Stimuli consisted of 16 colored shapes, which were recombinations of the shape and color features in phase 1. Each of the eight discrimination pairs was organized around the same relevant features as in phase 1; only the irrelevant features were changed. Thus, whereas C1 was a red/gold discrimination (shape irrelevant), C1-T was also a red/gold discrimination but with different shapes. Similarly, whereas S1 was a shape discrimination (color irrelevant), S1–T involved the same shape discrimination but different colors. The features that were rewarded in phase 1 were also rewarded in phase 2. Thus, a set of response rules that emphasized the relevant features in phase 1 would perfectly predict the rewarded stimuli in phase 2. Alternatively, a set of response rules that empha-

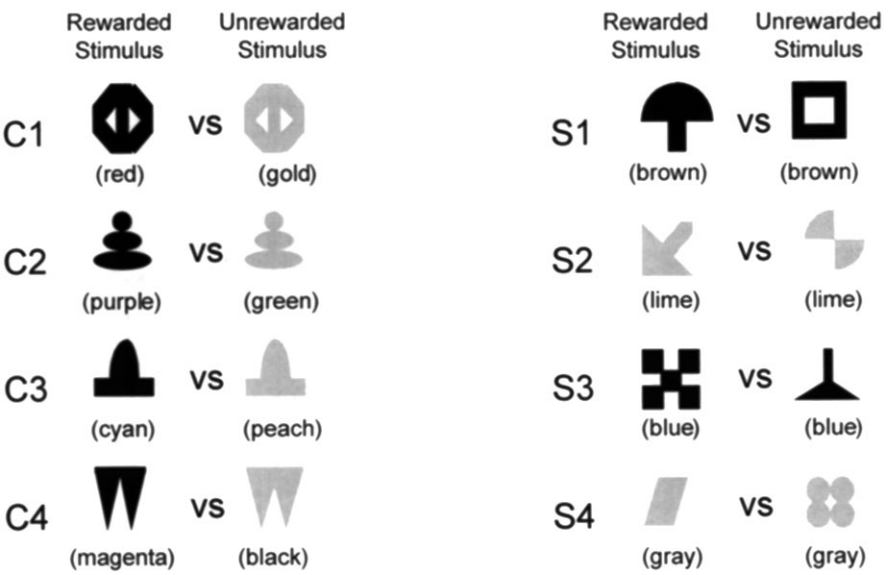

Figure 1. Stimuli used in phase 1 included eight discrimination pairs, S1–S4 and C1–C4. Each S1–S4 pair had two objects that differed in shape but not color. Each C1–C4 pair had two objects that differed in color but not shape. One object in each discrimination pair was arbitrarily designated as rewarded (+) and one as nonrewarded (–). No individual color or shape appeared in more than one discrimination pair in phase 1. Phase 2 stimuli were constructed from the phase 1 discrimination pairs by swapping irrelevant features.

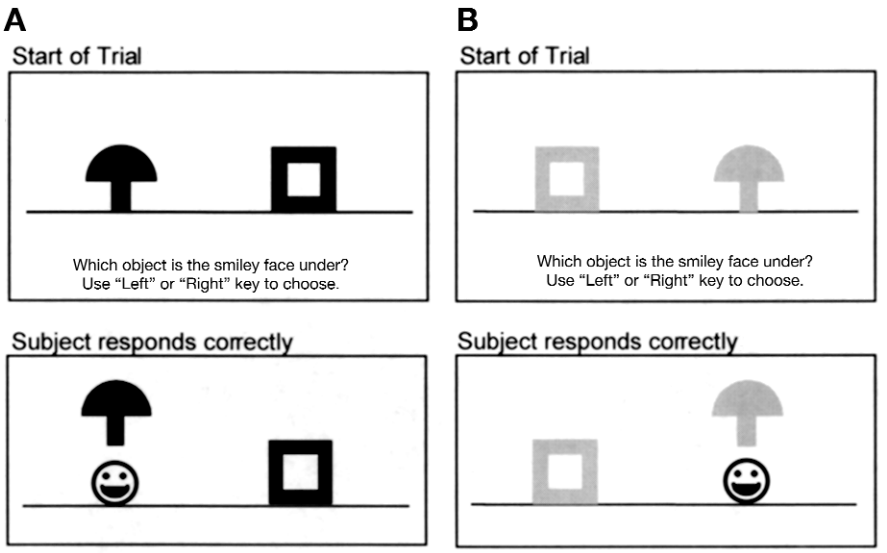

Figure 2. *A,* Screen events on a sample trial of phase 1. *Top:* On each trial, the discrimination pair is presented in either left-right order and a prompt appears. *Bottom:* If the participant responds correctly (in this case, choosing the mushroom shape), the chosen object is raised to reveal a smiley face icon underneath. *B,* Screen events on a sample trial of phase 2: events are similar to phase 1, but the objects are changed so that the relevant dimension (here the shape) is the same, whereas the irrelevant dimension (here the color approximated by grayscale) is novel.

sized the entire stimulus (including relevant and irrelevant features) in phase 1 would not transfer well to the new feature recombinations in phase 2.

Procedure

At the start of the experiment, the following instructions appeared on the screen: "Welcome to the experiment. You will see pairs of objects. Each time, there is a smiley face hidden under one of the two objects. It looks like this: J. Find as many as you can." The experimenter read these instructions aloud and then clicked the computer mouse button to begin phase 1 of the experiment.

On each trial of phase 1, participants saw one of the discrimination pairs shown in Figure 1. Trials were organized into blocks, each containing 16 trials: one presentation of each discrimination pair in each possible left-right ordering. Trials in a block occurred in a pseudorandom but fixed order. Figure 2A shows screen events in a typical trial. Below the stimuli, a prompt appeared: "Which object is the smiley face under? Use the "LEFT" or "RIGHT" key to choose." Participants then responded by pressing one of the two labeled keys. The selected stimulus then rose 2.5 cm on the screen. If it was the rewarded stimulus, a smiley face icon was revealed underneath and displayed for approximately 1 second. The object then returned to its original position, obscuring the smiley face icon below. The objects were then removed and a new trial initiated. There was no limit on response time. Phase 1 continued until the participant completed 16 consecutive trials correctly or for a maximum of 96 trials (6 blocks).

As soon as phase 1 terminated, phase 2 began without any warning that task demands had shifted. The screen events were identical to phase 1 (Figure 2B) except that the discrimination pairs were altered as described above. Again, trials were organized into blocks of 16 trials, one with each discrimination pair in each possible left-right ordering, in a pseudorandom but fixed order. Phase 2 continued until the participant completed 16 consecutive trials correctly or to a maximum of 48 trials (3 blocks).

The entire procedure, including phase 1 and phase 2, took approximately 15 to 20 minutes to complete.

Data Collection

On each trial, the computer recorded the discrimination pair, its left-right ordering, the desired response, and the participant's response. For both phases, the total trials to criterion were recorded. This total did not include the final 16 consecutive correct trials. Additionally, the total errors in each phase were recorded.

RESULTS

Concurrent Discrimination (Phase 1)

Overall, most participants finished phase 1 within the 96 trial maximum. Three participants in group HA failed to reach the performance criterion in phase 1, responding at or near chance levels even in the final block of 16 trials. Four participants in the no HA group also failed to reach the performance criterion in phase 1, but all were making 15 of 16 correct responses by the final block. Figure 3A shows the mean total errors for each group in phase 1. An ANOVA with phase 1 errors as the dependent variable revealed no significant difference in phase 1 performance between the HA and no HA groups, $F(1, 28) = 0.66$, $P = .424$, with no effects of subject gender, $F(1, 28) = 0.02$, $P > .5$; age, $F(1, 28) = 0.74$, $P = .397$; PDR score, $F(1, 28) = 1.20$, $P = .28$; or GDS rating, $F(1, 28) = 0.23$, $P > .5$.

There was no significant difference in learning to the color versus shape discriminations in terms of percent errors on each kind of discrimination (repeated-measures t-test, $t(33) = 0.79$, $P = .434$).

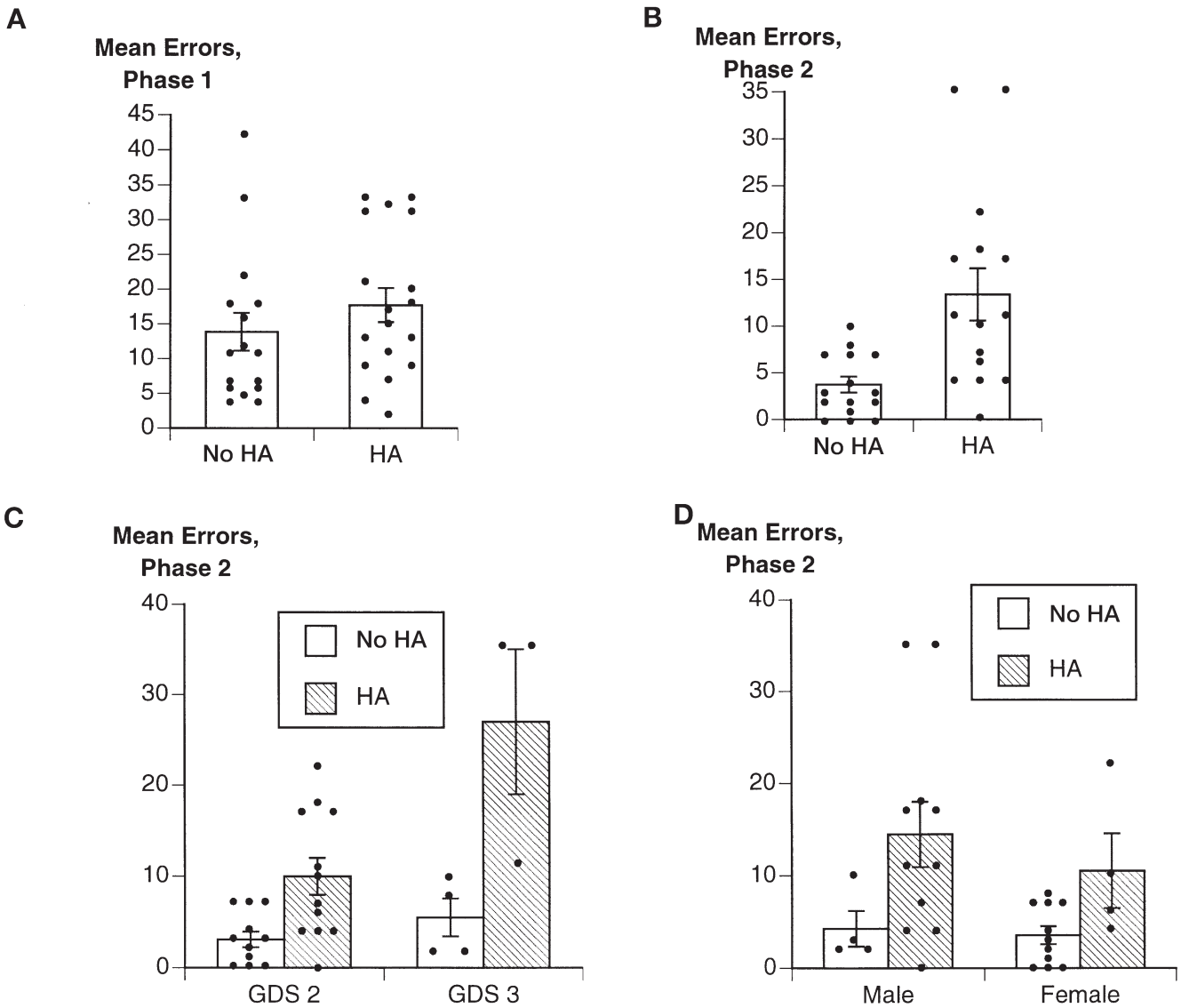

Figure 3. *A,* Performance on the concurrent discrimination task (phase 1) did not differ in the hippocampal atrophy (HA) and no HA groups. *B,* The HA group averaged significantly more errors on the transfer task (phase 2). *C,* In phase 2, there was also an effect of Global Deterioration Scale (GDS) rating, with GDS 2 participants averaging more errors than GDS 3 participants. *D,* Both male and female subjects showed a similar effect of HA. *Dots* represent individual subject scores.

Transfer Task (Phase 2)

One of the participants in the no HA group was interrupted before finishing phase 2; his phase 2 data were discarded. Phase 2 data from the three HA subjects who failed to reach criterion performance in phase 1 were likewise excluded from phase 2 analysis.

Among the remaining subjects, two subjects in group HA failed to reach criterion performance in phase 2. Figure 3B shows that mean phase 2 errors were higher in the HA than in the no HA participants; an ANOVA with phase 1 errors as the dependent variable confirmed a significant effect of group, $F(1, 23) = 13.67$, $P = .001$, with no effect of subject gender, $F(1, 23) = 0.85$, $P = .366$; age, $F(1, 23) = 0.19$, $P > .5$; PDR score, $F(1, 23) = 1.53$, $P = .229$; or phase 1 performance in terms of total errors, $F(1, 23) = 1.85$, $P = .187$; there was a significant effect of GDS rating, $F(1, 23) = 5.39$, $P = .030$, as shown in Figure 3C, with GDS 2 participants outperforming GDS 3 participants in both the HA and the no HA groups.

Although the ANOVA revealed no significant gender effect on phase 2 performance, it is notable that the HA group was disproportionately male, whereas the no HA group was disproportionately female; accordingly, gender imbalance could have contributed to the group differences. Figure 3D shows, however, that the males and females in the HA and no HA groups both showed the same general pattern of performance, with the no HA subjects making fewer errors than the HA subjects.

Considering the subset of HA participants who had an atrophy rating of at least 2 on at least one hippocampus (n = 5), there continues to be a significant difference in phase 2 errors in this HA group compared with

the 16 participants who had no visible atrophy to either hippocampus, ANOVA: $F(1, 19) = 4.64$, $P = .044$; no other variables (gender, age, PDR score, GDS rating, phase 1 performance) had statistically significant effects (all $P > .05$).

DISCUSSION

The current study showed that performance on a simple transfer task was disrupted in nondemented elderly participants with very small degrees of HA. By contrast, participants with and without HA performed equivalently on an eight-pair concurrent visual discrimination. Overall, participants without HA were generally able to transfer well when familiar stimuli or stimulus features were presented in novel recombinations. Participants with very mild HA were less able to transfer information to the novel recombinations. This pattern is consistent with prior animal and theoretical work suggesting that the hippocampus (and related medial temporal structures) is critical in forming stimulus representations that emphasize relevant stimulus features and de-emphasize irrelevant stimulus features.[24,25] Thus, hippocampal damage (including HA) may result in stimulus overcompression, or hyperspecificity, meaning that learned associations do not generalize well when familiar items are presented in novel recombinations.

There was no significant effect of subject age on either the original concurrent discrimination or the transfer task, although subjects with MCI (GDS 3) showed a tendency to make more errors than subjects with no objective evidence of cognitive impairment (GDS 2). Although age is often correlated with both hippocampal atrophy and cognitive impairment,[1,2] age did not appear to be a critical factor in the present study.

Similarly, there was no significant correlation between memory ability, as indexed by PDR score, and either discrimination or transfer performance. Likewise, there was no difference in PDR score between the HA and the no HA groups in the current study. Golomb et al[17] have shown that PDR score is highly correlated with HA rating; however, in that study, HA was defined as the presence of definite atrophy (rating = 2 or higher) in at least one hippocampus. Thus, Golomb et al's participants generally had greater degrees of HA than many of the participants classified as HA in the current study, which only required a rating of HA > 0. Since the PDR alone did not successfully distinguish individuals with and without mild HA in the current study, one possibility is that the transfer task is a more sensitive indicator of very mild HA, whereas the PDR score may be more useful in identifying individuals with more advanced HA than the participants in the current study. Further studies with a larger population size are clearly indicated to examine this issue further.

It is important to iterate that the individuals comprising the HA group in this study had minimal levels of HA. In many studies, a qualitative HA rating of 1 is considered to be only very mild or questionable atrophy, whereas a rating of 2 or higher is required for the individual to be judged as having definite atrophy. However, even restricting the HA group to those subjects with a rating of at least 2 on at least one hippocampus, there was still a significant difference between the HA and the no HA participants on phase 2 performance.

Conversely, ratings of hippocampal atrophy were made simultaneously for the left and right hippocampi. This leaves open the possibility that if both hippocampi were mildly atrophied, the degree of atrophy would be difficult to determine. However, even if the subjects rated as "nonatrophied" in the current study did, indeed, have very mild degrees of atrophy, they were still significantly less impaired on phase 2 than subjects with more visible atrophy. In either case, based on the current results, the transfer task seems sensitive to very mild HA before behavioral abnormalities would show up in more general performance in daily living; this is consistent with the nondemented (GDS 2–3) ranking of the participants.

One important limitation of the current work is the relatively small sample size. In addition, the participants were all individuals who presented themselves at the NYU Aging and Dementia Research Center and agreed to participate as volunteers. They were generally well educated, highly motivated, and informed about AD and memory function. This is probably less true of the elderly community at large, so any generalizations from this population must be made with caution. Additionally, there was a gender imbalance in our study: the HA group was predominantly male, whereas the no HA group was predominantly female. Both the current results (Figure 3D) and our recent normative study in 175 healthy individuals (Myers C, Houk V, Berlapsch K, et al. Age effects in concurrent discrimination learning. [In preparation]) revealed no significant difference between male and female subjects. However, future work will need to be done to establish the robustness of the finding that very mild HA disrupts the transfer test, across a larger cross-section of the elderly population, in both males and females.

Another important limitation of the current work is the use of qualitative atrophy assessments rather than volumetric measurements of the hippocampus. Previous studies have shown that HA ratings correlate well with hippocampal volume assessed through quantitative MRI[13,29]; however, quantitive assessment would allow direct calculation of hippocampal volume normalized with respect to head volume. Additionally, the mere presence of HA says little about what other brain regions may be atrophied or dysfunctional. For example, recent studies have demonstrated that prefrontal damage can disrupt performance on tasks that involve a shift among relevant stimulus dimensions.[38] Although there was no a priori reason to expect frontal damage

in the current subject population, this was not explicitly ruled out radiographically. Other brain areas such as the amygdala, important for registration of reinforcement, and the entorhinal cortex, implicated in early AD, may also be involved in the current task; damage to the fornix, which causes impairments in animals given a related task, may likewise be sufficient to disrupt performance in humans. Age-, injury-, or disease-associated damage to any or all of these structures may contribute to (or even cause) poor task performance. Future work should include assessment of frontal volumes, as well as other medial temporal structures such as the entorhinal cortex, to determine the specificity of HA in causing the observed pattern of behavior. In ongoing studies, we are considering the performance on this task of various patient populations including medial temporal amnesics, basal forebrain amnesics, and patients with prefrontal damage. Based solely on the current findings, though, it appears that HA correlates with impaired transfer performance, but it is unclear whether HA is truly causative.

Most individuals rated as GDS 2 are not considered at risk to develop AD within the time frame of a few years; individuals rated as GDS 3 have a relatively higher risk of declining to AD.[1,3,23,39] On the other hand, HA is known to correlate with increased risk for subsequent cognitive decline and AD.[8] Another open question is therefore whether the transfer task can be used to predict HA and therefore to predict risk of subsequent AD. An important future direction is to track the HA participants in this study and determine which, if any, show cognitive decline over the next several years. The outcome of such a longitudinal study could demonstrate whether the transfer test is predictive of AD risk. If so, this form of transfer test may have some utility as a general screening tool to identify individuals with possible HA; subsequent neuroradiographic, neuropsychological, and/or neurochemical investigation may then be warranted to verify HA and risk for decline to AD.

References

1. Flicker C, Ferris S, Reisberg B. Mild cognitive impairment in the elderly: predictors of dementia. *Neurology* 1991; 41: 1006–1009.
2. Masur D, Sliwinski M, Lipton R, et al. Neuropsychological prediction of dementia and the absence of dementia in healthy elderly persons. *Neurology* 1994; 44:1427–1432.
3. Kluger A, Ferris S, Golomb J, et al. Neurophysiological prediction of decline to dementia in nondemented elderly. *J Geriatr Psychiatry Neurol* 1999; 12:168–179.
4. Squire L, Zola-Morgan S. The medial temporal lobe memory system. *Science* 1991; 253:1380–1386.
5. Ball M. Neuronal loss, neurofibrillary tangles and granulovacuolar degeneration in the hippocampus with ageing and dementia. A quantitative study. *Acta Neuropathol (Berl)* 1977; 37:111–118.
6. Hyman B, Van Horsen G, Damasio A, Barnes C. Alzheimer's disease: cell-specific pathology isolates the hippocampal formation. *Science* 1984; 225:1168–1170.
7. Moscovitch M, Winocur G. The neuropsychology of memory and aging. In: Craik F, Salthouse T, eds. *The handbook of aging and cognition.* Hillsdale, NJ: Lawrence Erlbaum Associates, 1992:315–372.
8. de Leon M, George A, Stylopoulos L, et al. Early marker for Alzheimer's disease: the atrophic hippocampus. *Lancet* 1989; 2:672–673.
9. George A, de Leon M, Stylopoulos L, et al. CT diagnostic features of Alzheimer's disease: importance of the choroidal/hippocampal fissure complex. *AJNR Am J Neuroradiol* 1990; 11:101–107.
10. Kesslak J, Nalcioglu O, Cotman C. Quantification of magnetic resonance scans for hippocampal and parahippocampal atrophy in Alzheimer's disease. *Neurology* 1991; 41:51–54.
11. Jack C, Petersen R, O'Brien P, Tangalos E. MR-based hippocampal volumetry in the diagnosis of Alzheimer's disease. *Neurology* 1992; 42:183–188.
12. Jobst K, Smith A, Szatmari M, et al. Detection in life of confirmed Alzheimer's disease using a simple measurement of medial temporal lobe atrophy by computed tomography. *Lancet* 1992; 340:1179–1183.
13. Convit A, de Leon M, Golomb J, et al. Hippocampal atrophy in early Alzheimer's disease, anatomic specificity and validation. *Psychiatr Q* 1993; 64:371–387.
14. de Leon M, Golomb J, George A, et al. Hippocampal formation atrophy: prognostic significance for Alzheimer's disease. In: Corain B, Iqbal K, Nicolini M, et al, eds. *Alzheimer's disease: advances in clinical and brain research*. New York: John Wiley & Sons, 1993:35–46.
15. Killiany R, Moss M, Albert M, et al. Temporal lobe regions on magnetic resonance imaging identify patients with early Alzheimer's disease. *Arch Neurol* 1993; 50:949–954.
16. Convit A, de Leon M, Tarshish C, et al. Hippocampal volume losses in minimally impaired elderly. *Lancet* 1995; 345:266.
17. Golomb J, de Leon M, Kluger A, et al. Hippocampal atrophy in normal aging: an association with recent memory impairment. *Arch Neurol* 1993; 50:967–973.
18. Golomb J, Kluger A, de Leon M, et al. Hippocampal formation size in normal human aging: a correlate of delayed secondary memory performance. *Learn Mem* 1994; 1:45–54.
19. O'Brien J, Desmond P, Ames D, et al. Magnetic resonance imaging correlates of memory impairment in the healthy elderly: association with medial temporal lobe atrophy but not white matter lesions. *Int J Geriatr Psychiatry* 1997; 12:369–374.
20. Raz N, Gunning-Dixon F, Head D, et al. Neuroanatomical correlates of cognitive aging: evidence from structural magnetic resonance imaging. *Neuropsychology* 1998; 12:95–114.
21. Soininen H, Partanan K, Pitkänen A, et al. Volumetric MRI analysis of the amygdala and the hippocampus in subjects with age-associated memory impairment. *Neurology* 1994; 44: 1660–1668.
22. Sullivan E, Marsh L, Mathaion D, et al. Age-related decline in MRI volumes of temporal lobe gray matter but not hippocampus. *Neurobiol Aging* 1995; 16:591–606.
23. Golomb J, Kluger A, de Leon M, et al. Hippocampal formation size predicts declining memory performance in normal aging. *Neurology* 1996; 47:810–813.
24. Eichenbaum H, Mathews P, Cohen N. Further studies of hippocampal representation during odor discrimination learning. *Behav Neurosci* 1989; 103:1207–1216.
25. Myers C, Gluck M. Cortico-hippocampal representations in simultaneous odor discrimination learning: a computational

interpretation of Eichenbaum, Mathews & Cohen (1989). *Behav Neurosci* 1996; 110:685–706.

26. Eichenbaum H, Otto T, Cohen N. Two functional components of the hippocampal memory system. *Behav Brain Sci* 1994; 17:449–518.
27. Schacter D. Multiple forms of memory in humans and animals. In: Weinberger N, McGaugh J, Lynch G, eds. *Memory systems of the brain: animal and human cognitive processes*. New York: Guildford Press, 1985:351–379.
28. Winocur G, Kinsbourne M. Contextual cueing as an aid to Korsakoff amnesics. *Neuropsychologia* 1978; 16:671–682.
29. Holodny A, Waxman R, George A, et al. MR differential diagnosis of normal-pressure hydrocephalus and Alzheimer disease: significance of perihippocampal fissures. *AJNR Am J Neuroradiol* 1998; 19:813–819.
30. Reisberg B, Ferris S, de Leon M, Crook T. The Global Deterioration Scale for assessment of primary degenerative dementia. *Am J Psychiatry* 1982; 139:1136–1139.
31. Gilbert J, Levee R, Catalano F. A preliminary report on a new memory scale. *Percept Mot Skills* 1968; 27:277–278.
32. Crook T, Gilbert J, Ferris S. Operationalizing memory impairment for elderly persons: the Guild Memory Test. *Psychol Rep* 1980; 47:1315–1318.
33. Reisberg B, Ferris S, de Leon M, et al. Stage-specific behavioral, cognitive, and in vivo changes in community residing subjects with age-associated memory impairment (AAMI) and primary degenerative dementia of the Alzheimer type. *Drug Dev Res* 1988; 15:101–114.
34. Kluger A, Gianutsos J, Golomb J, et al. Patterns of motor impairment in normal aging, mild cognitive decline, and early Alzheimer's disease. *J Gerontol B Psychol Sci Soc Sci* 1997; 52B:28–39.
35. Halstead W. *Brain and intelligence*. Chicago: University of Chicago Press, 1947.
36. Reitan R, Wolfson D. *The Halstead-Reitan Neuropsychological Test Battery: Theory and clinical interpretation.* Tucson, AZ: Neuropsychology Press, 1993.
37. Bornstein R. Normative data on selected neuropsychological measures from a nonclinical sample. *J Clin Psychol* 1985; 41:651–659.
38. Owen A, Roberts A, Polkey C, et al. Extra-dimensional versus intra-dimensional set shifting performance following frontal lobe excisions, temporal lobe excisions or amygdalo-hippocampectomy in man. *Neuropsychologia* 1991; 29:993–1006.
39. de Leon M, Golomb J, George A, et al. The radiologic prediction of Alzheimer's disease: the atrophic hippocampal formation. *AJNR Am J Neuroradiol* 1993; 14:897–906.