Puzzling Thoughts for H.M.: Can New Semantic Information Be Anchored to Old Semantic Memories?

Brian G. Skotko
Duke University

Elizabeth A. Kensinger and Joseph J. Locascio
Massachusetts Institute of Technology

Gillian Einstein and David C. Rubin
Duke University

Larry A. Tupler
Duke University Medical Center

Anne Krendl and Suzanne Corkin
Massachusetts Institute of Technology

Researchers currently debate whether new semantic knowledge can be learned and retrieved despite extensive damage to medial temporal lobe (MTL) structures. The authors explored whether H.M., a patient with amnesia, could acquire new semantic information in the context of his lifelong hobby of solving crossword puzzles. First, H.M. was tested on a series of word-skills tests believed important in solving crosswords. He also completed 3 new crosswords: 1 puzzle testing pre-1953 knowledge, another testing post-1953 knowledge, and another combining the 2 by giving postoperative semantic clues for preoperative answers. From the results, the authors concluded that H.M. can acquire new semantic knowledge, at least temporarily, when he can anchor it to mental representations established preoperatively.

Memory is not a unitary capacity but, instead, can be divided into anatomically and functionally independent systems. Prominent among these divisions is the distinction between declarative (explicit) memory, the conscious recollection of facts and past events, and nondeclarative (implicit) memory, the influence of previous experience on task performance without conscious referral to stored information (N. J. Cohen & Squire, 1980; Graf & Schacter, 1985; Schacter, Chio, & Ochsner, 1993). Declarative memory is often further subdivided into semantic memory, an individual’s lifetime accumulation of universal factual knowledge, and episodic memory, an individual’s record of personal events (Hayman, MacDonald, & Tulving, 1993; Tulving, 1972, 1984, 1985, 1987; Tulving, Hayman, & MacDonald, 1991).

It is well established that damage to medial temporal lobe (MTL) structures (e.g., the hippocampus, the parahippocampal gyrus, the perirhinal cortex, and the entorhinal cortex) leaves the formation of declarative memory impaired (Gabrieli, Cohen, & Corkin, 1988; Nadel & Moscovitch, 1997; Parkin & Leng, 1993; Scoville & Milner, 1957; Squire & Zola, 1998). Much debate exists, however, as to whether some new semantic information can...
be learned despite extensive damage to MTL structures. Recent experiments have intimated that fragments of semantic information can be acquired even with extensive MTL damage (Bayley & Squire, 2002; Corkin, 1984, 2002; Glisky & Schacter, 1988; Glisky, Schacter, & Tulving, 1986a, 1986b; Hamann & Squire, 1995; Hayman et al., 1993; Hirst, Phelps, Johnson, & Volpe, 1988; Kitchener, Hodges, & McCarthy, 1998; Kovel, Mattis, & Goldmeier, 1983; Mattis & Kovner, 1984; McAndrews, Glisky, & Schacter, 1987; O’Kane, Kensinger, & Corkin, 2004; Schacter, Harbluk, & McLachlan, 1984; Shimamura & Squire, 1987; Tulving et al., 1991; Van der Linden, Brédart, Depoorter, & Coyette, 1996; Van der Linden et al., 2001; Van der Linden & Coyette, 1995; Verfaellie, Koseff, & Alexander, 2000; Westmacott & Moscovitch, 2001). Few of these studies, however, have tested patients with near-total damage to MTL structures. Our purpose in this current study was to explore whether the patient H.M., who has extensive damage to MTL structures bilaterally, could show any ability to learn postoperative semantic information.

H.M. was one of the first patients to clarify the role of MTL structures in semantic memory (Corkin, 1984, 2002; Corkin, Amaral, González, Johnson, & Hyman, 1997; Milner, Corkin, & Teuber, 1968; Scoville & Milner, 1957). Following a large bilateral MTL excision, he exhibited massive anterograde amnesia (Scoville, 1954, 1968; Scoville, Dunsmore, Liberson, Henry, & Pepe, 1953). He has displayed normal memory for semantic knowledge acquired before 1953 (Kensinger, Ullman, & Corkin, 2001), but he has been profoundly impaired in the acquisition of new semantic knowledge (Gabrieli et al., 1988; Postle & Corkin, 1998).

Of interest, though, is that H.M. has maintained a lifelong love of crossword puzzles. According to his own reports, at approximately age 15 he began solving crosswords that were printed in his local newspaper. He continues to work on two or more puzzles each day, and those who are close to him substantiate that he even engages in challenging puzzles featured in books published by the New York Times and Merriam-Webster. Skotko et al. (2001) examined H.M.’s leisure-time crossword puzzles for their linguistic content. H.M. responded with valid words that were rated as appropriate answers by well-educated judges. He responded correctly to different clue types (i.e., synonyms, definitions, categories, and fill in the blank), recognized the need for a proper noun or proper adjective, and answered with an appropriate part of speech. In terms of language skills, H.M. is a good puzzle solver. But to the extent that crossword puzzles rely heavily on current events and popular fads, H.M.’s sustained interest in these puzzles is noteworthy.

Crossword puzzles have been studied previously for their linguistic and cognitive content. Nickerson (1977) observed that crossword puzzles are cued retrieval tasks, and other studies have further suggested that crossword proficiency depends more on the manipulation of letters and word fragments than on knowledge of a large number of words (Underwood, Diehim, & Batt, 1994). Goldblum and Frost (1988) showed that any cluster of three adjacent letters facilitated retrieval better than did dispersed letters. In addition, syllabic clusters facilitated more recall than did non-syllabic clusters or unpronounceable clusters. For expert crossword solvers, puzzle proficiency could be predicted by word generation from a string of letters, performance on anagrams, sensitivity to a suffix of a word, sensitivity to a pseudosuffix in a word, and vocabulary scores (Underwood et al., 1994). Witte and Freund (1995) further showed that performance on anagrams—another independent test of word retrieval skills—was positively correlated with previous crossword puzzle experience. In a more recent study, general knowledge, in addition to word retrieval skills and crossword puzzle experience, was a significant predictor of crossword puzzle proficiency (Hambrick, Salthouse, & Meinz, 1999). Together, these variables accounted for approximately 85% of the variance in puzzle success, whereas reasoning skill tests were not significant predictors of puzzle-solving proficiency.

We were interested in testing H.M. on these general knowledge and word retrieval skills so that we could characterize his crossword-puzzle-solving abilities. In Experiment 1, we measured several components of H.M.’s memory and cognition, using the series of tests created by Hambrick et al. (1999) and Witte and Freund (1995). Here we asked the following: How does H.M. compare with age- and education-matched healthy volunteers on these predictive measures of crossword puzzle proficiency? Does he show normal retrieval of semantic knowledge?

After describing his puzzle skills, we asked in Experiment 2 whether H.M. was able to retrieve any factual information encountered postoperatively in the context of crossword puzzles. We anticipated that he would be capable of effectively solving clues with knowledge prior to 1953. The critical question was whether he would be able to show any ability to solve clue–answer pairs referencing information after 1953. In short, we used H.M.’s affinity for crossword puzzles to create a more ecologically valid way to examine his efforts to retrieve already learned semantic knowledge (Experiment 1) and his ability to learn new semantic information (Experiment 2).

**Experiment 1**

**Method**

**Participants**

H.M., a patient with amnesia who has 12 years of education, was 72–74 years old during this 3-year study. He has had profound amnesia since 1953, when he underwent an experimental resection of MTL structures to relieve medically intractable epilepsy (Scoville, 1954, 1968; Scoville et al., 1953; Scoville & Milner, 1957). The resected tissue included all MTL structures except, approximately, the caudal 2 cm of the hippocampus and parahippocampal gyrus (Corkin et al., 1997).

We compared H.M.’s performance to that of age- and education-matched healthy volunteers from previously published studies on crossword proficiency (Hambrick et al., 1999; see Table 3). Hambrick et al.’s (1999) participants were over 70 years old (M = 73.38, SD = 3.39) and had 12 years of education (see Table 1). These participants averaged 2.47 (SD = 3.22) crossword puzzles each week and spent 2.34 hr (SD = 3.17) each week solving them.

We also compared H.M.’s results to those of a group of “old frequent” puzzle solvers reported in a study by Witte and Freund (1995). These participants solved crossword puzzles at least once a week. H.M., however, was slightly older and less educated than these participants (see Table 1).

---

1 Although definitions of semantic information vary among investigators, for our purposes we define semantic information as factual information that is retrieved without regard to the context in which the information was learned. Thus, we include facts associated with proper names (e.g., that Judy Garland was an actress) within the domain of semantic information.
Table 1
Age and Education Level for Participants in Experiment 1:
Means and Standard Deviations

<table>
<thead>
<tr>
<th>Participants</th>
<th>Age (years)</th>
<th>Education (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>M</td>
</tr>
<tr>
<td>H.M.</td>
<td>1</td>
<td>74.05</td>
</tr>
<tr>
<td>Hambrick et al. (1999)</td>
<td>16</td>
<td>75.38</td>
</tr>
<tr>
<td>Witte and Freund (1995)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 1</td>
<td>9</td>
<td>67.33</td>
</tr>
<tr>
<td>Experiment 2</td>
<td>19</td>
<td>69.74</td>
</tr>
</tbody>
</table>

Materials and Procedure

We administered a series of tests measuring various forms of general knowledge, word retrieval skills, and transformation efficiency. These tests, obtained from Hambrick et al. (1999) and Witte and Freund (1995), are briefly described below. The time limits were extended to permit H.M. to complete each test.

Anagram tests (Hayslip & Sterns, 1979; Witte & Freund, 1995). H.M. performed two tests. In the first, 30 five-letter anagrams, plus 2 practice anagrams, were printed in uppercase letters on separate 3 × 5-in. (7.6 × 12.7-cm) index cards. The anagrams ranged in difficulty and included 10 easy, 10 medium, and 10 hard items. H.M. viewed each card for 60 s and was asked to solve the anagram during that time. In the second test, he unscrambled 12 five-letter anagrams, plus 1 practice anagram, that were initially presented in a word condition (e.g., FORTH from FROTH) and in a nonword condition (e.g., FORTH from FRTHO). Upon his first verbal response, time was recorded and the correct answer was provided, if necessary. He performed these experiments on at least 4 consecutive days during Visits 4 and 5, which were separated by approximately 6 months (see Table 4).

General knowledge tests (Hambrick et al., 1999, Study 1). H.M. performed three general knowledge tests during Visit 5 (see Table 4). The first consisted of a 40-item (four-alternative, forced-choice) test of general information, with 5 items each on American history, American literature, art and architecture, geography, music, mythology, world history, and world literature. We also administered two 10-item tests, each lasting 5 min, that assessed knowledge of synonyms and antonyms in a five-alternative, forced-choice format.

Word retrieval and transformation efficiency (Hambrick et al., 1999, Study 1). H.M. performed four word retrieval tests during Visit 5 (see Table 4). The first was a fluency test in which he was asked to write as many words as possible beginning with the letters F and S (1 min each). The second test consisted of two lists of letters (BFHILNORW and ACDBGKMTU) from which he was asked to generate as many words as possible (1 min each). The third test consisted of 25 five-letter anagrams that he had 10 min to unscramble. Finally, the word switch test consisted of 15 pairs of four-letter words and required him to morph the first word into the second word by changing one letter at a time (5 min).

Results

Witte and Freund (1995) found that anagram performance is positively correlated with crossword puzzle experience. When given these same anagram tests, H.M. performed between 0.92 and 2.17 standard deviations below the means of healthy volunteers (see Table 2). On Test 1, his score was closest to that of the healthy volunteers when he was given the easy anagrams (H.M.: $M = 6.30$; healthy volunteers: $M = 8.11$, $SD = 1.96$). On the medium and hard anagrams, he performed substantially poorer. On Test 2, his score was closest to that of the healthy volunteers when he was given anagrams in the word condition. Like the healthy volunteers, however, he solved more anagrams from the nonword condition than from the word condition. This result is consistent with Ekstrand and Dominowski’s (1968) report that an anagram is more difficult to solve if it is a word than if it is a nonsense string of letters. Overall, H.M. performed approximately 1–2 standard deviations below the means of healthy volunteers on anagram word tests.

Hambrick et al. (1999) showed that general knowledge, word retrieval skills, and crossword puzzle experience are major predictors of crossword puzzle proficiency. When H.M. completed the same tests used in that study, he performed between 0.56 and 2.54 standard deviations below the means of healthy volunteers on all tests. On measures of word retrieval, he scored 0.49–1.65 standard deviations below the means of healthy volunteers, with the exception of the fluency test, on which he scored 2.54 standard deviations below the means of healthy volunteers. This result is consistent with the findings in two other reports (Kensinger et al., 2001; Schmolck, Kensinger, Corkin, & Squire, 2002).

Discussion

On various measures of cognitive function believed to underlie skill in solving crossword puzzles, H.M. scored within 1 or 2 standard deviations below the means of healthy volunteers, with the exception of the fluency test, on which he scored 2.54 standard deviations below the means of the healthy volunteers. Lesions of the lateral temporal neocortex disproportionately affect category fluency, whereas frontal lobe lesions affect letter fluency (Gurovitch et al., 2000; Martin, Wiggs, Lalonde, & Mack, 1994; Mummery, Patterson, Hodges, & Wise, 1996; Newcombe, 1969; B. Milner, personal communication, May 24, 2001). The deficit in verbal fluency could be related to minimal damage to the lateral temporal neocortex, to H.M.’s lower socioeconomic status and

Table 2
H.M.’s Performance on Anagram Tests Compared With That of Healthy Volunteers

<table>
<thead>
<tr>
<th>Test</th>
<th>Healthy volunteers$^a$</th>
<th>H.M.$^b$</th>
<th>SD between $M$s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test 1 (30 items)</td>
<td>18.78</td>
<td>6.65</td>
<td>9.30</td>
</tr>
<tr>
<td>Easy (10 items)</td>
<td>8.11</td>
<td>1.96</td>
<td>6.30</td>
</tr>
<tr>
<td>Medium (10 items)</td>
<td>5.67</td>
<td>2.50</td>
<td>1.90</td>
</tr>
<tr>
<td>Hard (10 items)</td>
<td>5.00</td>
<td>2.69</td>
<td>1.10</td>
</tr>
<tr>
<td>Test 2 (24 items)</td>
<td>9.58</td>
<td>1.46</td>
<td>7.17</td>
</tr>
<tr>
<td>Word condition (12 items)</td>
<td>8.78</td>
<td>1.48</td>
<td>6.33</td>
</tr>
<tr>
<td>Nonword condition (12 items)</td>
<td>10.30</td>
<td>1.06</td>
<td>8.00</td>
</tr>
</tbody>
</table>

Note. Values represent numbers of correctly solved anagrams. Values for the healthy volunteers were obtained from Witte and Freund (1995).

$^a$ For Test 1, $N = 9$; for Test 2, $N = 19$. $^b$ Because H.M. showed no improvement on the anagram tests between days or visits, an average of his scores was calculated to measure his performance more accurately (for Test 1, $N = 10$ days; for Test 2, $N = 9$ days).
Compared With Healthy Volunteers

Still another possibility is that H.M., who underwent a bilateral temporal lobectomy (left temporal lobe removed) in 1953 to control his partial epileptic seizures, had poor education, or to both (Kensinger et al., 2001; Schmolck et al., 2002). It is also possible that people with amnesia are limited in their ability to rely on strategies used by healthy volunteers (e.g., visualizing an aisle at the grocery store when asked to generate their ability to rely on strategies used by healthy volunteers (e.g., poor education, or to both (Kensinger et al., 2001; Schmolck et al., 1998)).

On the measures of general knowledge and word retrieval skills, H.M. performed between 0.49 and 1.65 standard deviations below the means of healthy volunteers. Similarly, on the anagrams used by Witte and Freund (1995), he performed between 0.92 and 2.17 standard deviations below the means of the healthy volunteers. Given these predictive variables, we would expect that with memory deficits aside, H.M. would be less proficient at crossword puzzle solving than would be healthy volunteers. If there were a one-to-one correlation between these predictive variables and puzzle proficiency, we would expect H.M.’s optimal performance on actual puzzles to be within 1–2 standard deviations of the means of healthy volunteers.

In Experiment 2, we analyzed H.M.’s performance on crossword puzzles created specifically to test his semantic memory. Here we were interested in answering multiple questions: How well does he do on crossword puzzles in comparison with healthy volunteers? Can he solve clues referencing postoperative events? After repeated exposure to the same puzzles, can he improve his accuracy and/or speed?

Experiment 2

Method

Participants

H.M.’s performance was compared with that of 10 healthy volunteers recruited from the Harvard Cooperative on Aging, Cambridge, MA. They included 7 women and 3 men who gave informed consent and were paid $10/hr. As a group, the healthy volunteers were matched to H.M. in age (M = 74, SD = 4.0 years) and education (12 years for all participants). The healthy volunteers were screened to exclude those with a history of cancer, heart disease, and neurological or psychiatric disorders. All volunteers were native English speakers. Only 2 participants solved crossword puzzles more than twice a week; the rest solved crossword puzzles on an infrequent basis and, in that respect, were at a disadvantage relative to H.M.

Materials

Five different crossword puzzles of 20 clues each were created on a 20 × 20 square (5.5- × 5.5-in.; 14.0- × 14.0-cm) grid with Crossword Construction Kit 98 (Version 3.1a; Insight, 1998). Each puzzle grid was printed on the top half of a separate page (8.5 × 11 in.; 21.6 × 27.9 cm), with the corresponding clues printed directly below in a two-column format. The five puzzles tested semantic knowledge from different time periods.

Experiment 2A: Pre-1953 puzzle. The pre-1953 puzzle (see Appendix A on the Web at http://dx.doi.org/10.1037/0894-4105.18.4.756.supp) consisted of 20 clues that H.M. was expected to solve. These clues referred to historical figures and events prior to 1953 (e.g., clue: “baseball player in the 1930s who captured the home run record”) or to facts generally known by people with 12 years of education (e.g., clue: “country just north of the U.S.”).

Experiment 2B: Post-1953 puzzle. The post-1953 puzzle (see Appendix B on the Web at http://dx.doi.org/10.1037/0894-4105.18.4.756.supp) consisted of 15 clues that H.M. was not expected to solve and 5 clues that he was expected to solve. The unsolvable clues consisted of historical figures and events popularized after 1953 (e.g., clue: “husband of Jackie Onassis, assassinated while President of the U.S.”) or of objects that were introduced to the English language after 1953 (e.g., clue: “portable, personal computer”). The solvable clues consisted of clues with the same constraints as those given in the pre-1953 puzzle. These clues were randomly interspersed among the other clues to (a) decrease the level of frustration inherent in the task and (b) confirm that H.M. had read all of the clues in this puzzle and had attempted to solve them.

Experiment 2C: Pre–post puzzle. This experiment combined the previous two by giving postoperative semantic clues for preoperative answers. Pre–post puzzle 1 (see Appendix C on the Web at http://dx.doi.org/10.1037/0894-4105.18.4.756.supp) consisted of 15 clues that (a) referenced events or story lines that were particularly newsworthy after H.M.’s operation and (b) had answers that related to knowledge gained before his operation (e.g., clue: “childhood disease successfully treated by Salk vaccine” [postoperative knowledge]; answer: “polio” [preoperative knowledge]). As before, the remaining 5 clues in this puzzle consisted of items that H.M. was expected to solve.

Experiment 2D: Effects of priming. Our goal in this experiment was to examine the effects of priming on H.M.’s performance. For the first priming test, we asked the following: After completing a puzzle for 5 consecutive days, would H.M. be able to complete a grid without the clues? To answer this question, we administered two new puzzles for 5 consecutive days (see Appendices C and D on the Web at http://dx.doi.org/10.1037/0894-4105.18.4.756.supp). The first puzzle—the reordered puzzle—was identical to pre–post puzzle 1, except that the clues were randomly rescrambled and reordered by the computer program. The second puzzle—pre–post puzzle 2—consisted of a new set of clues with the same constraints as those given in Experiment 2C. On Day 6, we showed H.M. crossword grids without the accompanying clues. For each answer over four letters long, the first three letters were preprinted on the grid in capital letters. For answers that were only four letters long, only the first two letters were printed. Because of the crisscrossing nature of puzzles, six of the answers from the reordered puzzle and nine of the answers from pre–post puzzle 2 also had additional letters positioned later in a word. All

<table>
<thead>
<tr>
<th>Test</th>
<th>Healthy volunteers (N = 16)</th>
<th>H.M.</th>
<th>SD below M</th>
</tr>
</thead>
<tbody>
<tr>
<td>General knowledge</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General information test (40 items)</td>
<td>15.50 8.10 11.00 0.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antonyms (10 items)</td>
<td>4.81 3.71 1.00 1.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Synonyms (10 items)</td>
<td>5.88 3.30 4.00 0.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word retrieval</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fluency test</td>
<td>10.16 2.82 3.00 2.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Make words test</td>
<td>4.44 3.57 0.50 1.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anagram test (25 items)</td>
<td>14.13 8.57 0.00 1.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word switch test (15 items)</td>
<td>0.98 1.98 0.00 0.49</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Values represent numbers of correct items for each test. Values for the healthy volunteers were calculated from a database kindly provided by Hambrick et al. (1999). The fluency test score is an average of the number of words generated for the letters F and S. The make words test score is an average of the number of words produced from the first and second strings of letters.

<table>
<thead>
<tr>
<th>Test</th>
<th>Healthy volunteers (N = 16)</th>
<th>H.M.</th>
<th>SD below M</th>
</tr>
</thead>
<tbody>
<tr>
<td>General knowledge</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General information test (40 items)</td>
<td>15.50 8.10 11.00 0.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antonyms (10 items)</td>
<td>4.81 3.71 1.00 1.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Synonyms (10 items)</td>
<td>5.88 3.30 4.00 0.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word retrieval</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fluency test</td>
<td>10.16 2.82 3.00 2.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Make words test</td>
<td>4.44 3.57 0.50 1.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anagram test (25 items)</td>
<td>14.13 8.57 0.00 1.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word switch test (15 items)</td>
<td>0.98 1.98 0.00 0.49</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Values represent numbers of correct items for each test. Values for the healthy volunteers were calculated from a database kindly provided by Hambrick et al. (1999). The fluency test score is an average of the number of words generated for the letters F and S. The make words test score is an average of the number of words produced from the first and second strings of letters.
of the boxes could still fit more than one possible answer from the English lexicon, and therefore, H.M. could not generate an answer by default with the orthographic clues and number of crossword boxes. He was asked to complete the word stems with whatever came to his mind.

For the second priming test, we asked the following: After completing a puzzle for 5 consecutive days, would H.M. be able to complete a grid with only the pre-1953 fragments of the clues? Each of our clues contained some preoperative information, so we wanted to test the possibility that H.M.’s improvement on the reordered puzzle and pre–post puzzle 2 could be attributed to his recognition of the preoperative clue fragments alone. For example, with the clue “The National Academy of Recording Arts and Sciences recognized this male dancer with a Lifetime Achievement Award in 1989,” was H.M. merely focusing on the fragment “male dancer” to come up with the answer, “Astaire”? To answer this question, we gave H.M. the crossword puzzles described in the first priming test; this time the preoperative fragment of each clue accompanied the now-blank puzzle grids. If he scored well on this test, it would indicate that he could generate the correct answer using just the preoperative facts of each clue and was not necessarily learning the postoperative information. If he scored poorly on this test, it would indicate that the preoperative facts of each clue were not sufficient for him to generate the answers. This result would suggest that H.M. had learned some of the postoperative facts in the puzzles.

Procedure

Participants were told that they could complete each puzzle in any manner they liked and that they could erase answers, if necessary. They were also asked to indicate when they had finished—that is, when they had completed the entire puzzle or could no longer solve any more clues. The experimenter was positioned behind the participant to minimize any opportunities for extraneous conversations. There were no time limits, and all testing sessions were videotaped for subsequent analysis.

After the participants had completed the puzzle or indicated that they could not solve any more clues, the experimenter sat in front of them, presented the answer key, and asked them to verify their answers. Specifically, they were asked to change any misspelled, wrong, or unanswered responses so that they matched the correct ones. They were also asked to read aloud the corresponding clues while making the corrections to ensure that they were making the appropriate connections between the clues and the answers. If participants did not correct their puzzles completely, the experimenter guided them through the process.

Participants had a break of at least 5 min before beginning another puzzle; each puzzle was completed only once on a given day. H.M. repeated each puzzle for at least 4 consecutive days. The healthy volunteers repeated the procedure at least twice. The entire procedure was repeated for each of the puzzles at approximately 6-month intervals (see Table 4). A maximum of three crosswords were administered during any one visit; the project was completed in 3 years.

Data Analysis

Each puzzle was scored for accuracy, difficulty, and completion time. Accuracy was determined by dividing the number of boxes correctly answered by the total number of boxes in a given puzzle. Responses that were clearly misspelled were scored box by box. For example, a misspelled response of “Gettysberg” would receive 9 out of a possible 10 points. Responses that were clearly incorrect—not misspellings—were scored 0, despite any coincidental overlap of letters. In addition, we tallied the whole number of correctly solved clues for each puzzle. Instances of clear misspellings were counted as a correctly solved clue. The lengths of each puzzle were standardized as best as possible: In total, the pre-1953 puzzle had 121 boxes; the post-1953 puzzle had 122; the pre–post puzzle 1 and the reordered puzzle had 115; and the pre–post puzzle 2 had 120.

We further asked this question: Did H.M.’s accuracy improve significantly as each puzzle was repeated? To answer this, we used a permutation test (Edgington, 1995; Good, 1994), suitable because of the unusual nature of our single-participant data. The number of observations for analysis is usually small in studies of single participants, and violation of assumptions of parametric tests cannot be checked. Because the permutation test is a nonparametric test, it has fewer assumptions. For this test, H.M.’s scores across days were converted to ranks; then, all permutations of these ranks across days were enumerated, and improvement was considered significant at p ≤ .05 if the actual rank ordering was among the 5% most extreme in the direction of improvement. For example, H.M. solved pre–post puzzle 1 for 5 consecutive days during Visit 2. Therefore, there were 5! (= 120) different patterns of change possible. If he consistently answered more boxes than he did in his previous day’s work, that would constitute just 1 pattern among the 120 possibilities, giving a p value of 1/120 or .008. If, in this case, the pattern was not among the 6 most extreme cases, the results would have been nonsignificant. This analysis was repeated for each puzzle during each visit.

We next asked the question: Did H.M.’s accuracy differ significantly for each puzzle between visits separated by 6 months? To answer this, we used the Mann–Whitney test. Although the assumptions of independent observations were violated, the observations were far enough apart in time (especially in the context of H.M.’s amnesia) to lessen the likelihood of correlated error. We also used Mann–Whitney tests to compare his performance to that of healthy volunteers during each visit.

In an effort to assess the difficulty of each puzzle, we asked 4 of the healthy volunteers to rank their performance on each clue after the 3rd day of testing using a scale of 1 (they immediately knew the answer as soon as

### Table 4

**Administration of Crossword Puzzles and Information-Processing Tasks to H.M. Over a 3-Year Period**

<table>
<thead>
<tr>
<th>Experiment and puzzle</th>
<th>Visit 1</th>
<th>Visit 2</th>
<th>Visit 3</th>
<th>Visit 4</th>
<th>Visit 5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experiment 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anagram tests</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crossword skills</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Experiment 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-1953 puzzle</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Post-1953 puzzle</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Pre–post puzzle 1</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Reordered puzzle</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Note. The crossword skills category comprises the three general knowledge tests and the four word retrieval tests.
they finished the clue) to 7 (the answer was completely unfamiliar and was solvable only because they remembered it from a prior day).

The time taken to complete each puzzle extended from the time the participants turned over the puzzle to begin to the time they began to write their last response. Because we were concerned only with the time spent solving the puzzles, the stopwatch was paused when participants engaged in conversation or when they printed their responses in the puzzle grid. Participants naturally spent more time on the puzzles when they solved more clues. To standardize the completion times, we calculated a time-per-box score by dividing the total completion time by the total number of answers for each puzzle. Standard t tests were used to evaluate the differences between time-per-box scores across visits for each puzzle.

Results

Experiment 2A: Pre-1953 Puzzle

H.M. responded with high accuracy (M = 86%, SD = 5%), scoring within 1 standard deviation of the healthy volunteers’ mean on Day 1 (M = 90%, SD = 9%). He did not improve across the 6 days of testing but performed consistently well (see Figure 1 and Table 5). In fact, he missed only two clues, “Chaplin” and “Gershwin,” on a regular basis. Both of these clues were ranked as the two most difficult in this puzzle by the healthy volunteers (“Chaplin”: M = 3.5, SD = 2.5; “Gershwin”: M = 3.0, SD = 2.4).

Approximately 6 months later, H.M. repeated the experiment. Although he was slightly more accurate on Day 1 of testing but performed consistently well (Visit 1: M = 4.16 s, SD = 1.05 s; Visit 2: M = 5.56 s, SD = 0.84 s; p < .07; see Figure 4). Overall, with pre-1953 clues, H.M. responded with high accuracy and a rapid response time.

Experiment 2B: Post-1953 Puzzle

H.M. responded with low accuracy during both visits (M = 22%, SD = 6%; see Figure 2 and Table 5), scoring approximately 4 standard deviations lower than the mean of the healthy volunteers (M = 66%, SD = 11%). During Visits 1 and 2, H.M.’s scores did not improve significantly across 6 days. Also, his performance during Visit 1 did not differ significantly from his performance during Visit 2 (p = .41).

The clues that H.M. answered correctly were always from the five pre-1953 clues that were embedded in the puzzle. However, he only once answered “Olympics,” probably because he was unfamiliar with this term in the context of the clue (“group of modern international athletic contests held every 4 years in a different city”). If we exclude “Olympics,” the solvable clues account for 25% of the total boxes. His mean score of 22% correct responses fell below this conservative expectation.

For the healthy volunteers, the post-1953 puzzle was inherently more difficult than the pre-1953 puzzle on Day 1 (pre-1953 puzzle: M = 90%, SD = 9%; post-1953 puzzle: M = 66%, SD = 11%). The time-per-box scores were slightly faster on Visit 1 (M = 12.91 s, SD = 4.45 s) than on Visit 2 (M = 17.95 s, SD = 6.30 s), but the difference was not significant (p < .41; see Figure 4). In summary, H.M. performed poorly on postoperative clues for which he had no prior semantic knowledge, and his accuracy and response times did not change significantly over repeated exposures.

Experiment 2C: Pre–Post Puzzle

On Day 1 of testing, H.M. answered approximately 20% of the clues correctly, placing him almost 2.70 standard deviations below the mean of the healthy volunteers (Day 1: M = 71%, SD = 19%). On this puzzle, he showed significant improvement across 5 days (Visit 1: M = 54%, SD = 12%; Visit 2: M = 71%, SD = 9%; p < .008). In fact, he showed a 226% increase, rising from 20% correct responses on Day 1 to roughly 45% correct responses on Day 5 (see Figure 3 and Table 5). His performance on Day 5 placed him within 1.40 standard deviations of the healthy volunteers’ mean on Day 1. Also, H.M.’s performance during Visit 2 did not differ significantly from his performance during Visit 3 (p = .56). The similar data from the two visits provide strong evidence that his day-to-day improvement within visits was real; however, the memory traces established during Visit 2 were not sufficiently robust to carry over to Visit 3, which occurred 6 months later.

Five solvable clues were embedded in pre–post puzzle 1. One of them was “fall holiday,” but H.M. never responded with “Halloween.” When asked afterward whether he was familiar with the term, he answered “yes” but suggested that he never thought of it as a fall holiday. He did insert portions of “Thanksgiving” in the allotted boxes on several occasions. Clearly, this clue did not serve its purpose of decreasing frustration and, therefore, was omitted from the baseline tally. The remaining four solvable clues totaled 23% of the total boxes, H.M. performed well beyond this mark on Day 5 (roughly 45%), providing a remarkable and novel instance of his ability to associate new semantic postoperative knowledge with preoperative information. Specifically, he correctly learned to associate the postoperative knowledge for “polio,” “Hiss,” “Gone with the Wind,” “Ike,” “St. Louis,” and “Warsaw.”

In solving these clues, H.M. answered the two toughest clues of the puzzle, as ranked by healthy volunteers (“Warsaw”: M = 4.8,
2.9, rank 1, tie; "Hiss": M = 4.8, SD = 1.5, rank = 1, tie; "St. Louis": M = 1.8, SD = 1.2, rank = 7; "polio," "Ike," and "Gone with the Wind": M = 1.0, SD = 0, rank = last). Overall, healthy volunteers found the pre–post puzzle 1 easier than the post-1953 puzzle but more challenging than the pre-1953 puzzle on Day 1 (pre-1953 puzzle: M = 90%, SD = 9%; post-1953 puzzle: M = 66%, SD = 11%; pre–post puzzle 1: M = 71%, SD = 19%).

H.M.'s time-per-box scores for Visit 2 decreased markedly from Day 1 to Day 2 and then remained constant until Day 5. During Visit 3, his scores remained relatively constant from Day 1 to Day 4 and then increased markedly on Day 5 (see Figure 4). In short, H.M.'s accuracy improved consistently on a task requiring him to link postoperative knowledge with preoperative semantic memory traces. We found no evidence of procedural learning, however, as his speed of performance did not decrease across days.

Experiment 2D: Effects of Priming

H.M. completed both grids containing two- or three-letter word stems with high accuracy (reordered puzzle = 90%; pre–post puzzle 2 = 71%). When pre-1953 clue fragments accompanied a blank grid, H.M. was unable to fill in the responses (reordered puzzle = 17%; pre–post puzzle 2 = 10%).

Table 5
Number of Clues Correctly Answered, Out of 20, Per Puzzle

<table>
<thead>
<tr>
<th>Puzzle Type</th>
<th>Visit 1</th>
<th>Visit 2</th>
<th>Visit 3</th>
<th>Visit 4</th>
<th>Visit 5</th>
<th>Visit 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-1953 puzzle</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H.M.</td>
<td>18</td>
<td>15</td>
<td>18</td>
<td>18</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Healthy volunteers</td>
<td>18.2 (1.5)</td>
<td>19.0 (1.5)</td>
<td>19.8 (0.4)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Post-1953 puzzle</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H.M.</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Healthy volunteers</td>
<td>4.7 (2.5)</td>
<td>16.1 (3.0)</td>
<td>17.7 (2.9)</td>
<td>18.0 (2.8)</td>
<td>19.5 (0.7)</td>
<td>—</td>
</tr>
<tr>
<td>Pre–post puzzle 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H.M.</td>
<td>4</td>
<td>5</td>
<td>8</td>
<td>10</td>
<td>10</td>
<td>—</td>
</tr>
<tr>
<td>Healthy volunteers</td>
<td>14.4 (2.4)</td>
<td>17.3 (3.9)</td>
<td>18.9 (1.7)</td>
<td>18.5 (2.1)</td>
<td>20.0 (0.0)</td>
<td>—</td>
</tr>
</tbody>
</table>

Note. Dashes indicate that puzzles were not completed on the corresponding day.

* N = 10; values for healthy volunteers are reported as means (± SDs).

Figure 2. H.M. had a low number of correct responses on the post-1953 puzzle (M = 22.31%, SD = 5.89%). The dotted line represents the score he would have achieved if he had correctly solved the clues that were based on preoperative knowledge. He rarely surpassed this baseline, indicating that he performed poorly on clues for which he had no prior semantic knowledge. His performance over 4 consecutive days during Visit 1 did not differ significantly from his performance over 6 consecutive days during Visit 2 (p = .41). (The data from Day 3 of Visit 2 were incomplete because H.M. fell asleep during testing.) Healthy volunteers (N = 10) completed the puzzle for 2–5 days.

Figure 3. H.M. showed a 226% increase on pre–post puzzle 1, rising from 20% correct on Day 1 to about 45% correct on Day 5. The dotted line represents the score he would have achieved if he had correctly solved the clues that were based on preoperative knowledge. On this puzzle, he showed significant improvement across 5 days (Visit 1: p = .042; Visit 2: p = .008), progressing well beyond the baseline mark. In fact, his performance on Day 5 placed him within 1.4 standard deviations of the healthy volunteers' mean on Day 1. His performance during Visit 2 did not differ significantly from his performance during Visit 3 (p = .56), providing strong evidence that his improvement was real and repeatable. (The data from Day 3 of Visit 2 and Day 6 of Visit 3 were incomplete because H.M. fell asleep during testing.) Healthy volunteers (N = 10) completed the puzzle for 2–5 days.
Discussion

Preoperative Semantic Knowledge

H.M. performed near optimally on a puzzle that solely measured pre-1953 semantic knowledge. It should be noted, however, that although H.M. had the potential to know all trivia answers prior to 1953, his actual knowledge of those current events was largely dependent on his interests. The reason for his omissions of "Chaplin" and "Gershwin" after 10 repetitions of the test is unknown, although these two clues were ranked by healthy volunteers as the hardest ones on the pre-1953 puzzle. Perhaps the clues were too cryptic for H.M. to generate the correct answers, or maybe H.M. was not familiar with these performers in a substantial way before his operation. Regardless, H.M.’s ability to solve successfully the majority of the clues on the pre-1953 puzzle is consistent with his relatively preserved semantic retrieval ability, as demonstrated in Experiment 1 of this study as well as in prior studies (Kensinger et al., 2001; Schmolck et al., 2002).

Postoperative Semantic Knowledge

On the puzzle that measured post-1953 semantic knowledge, H.M. performed poorly and showed no improvement after repeated exposures. This result is consistent with his profound global amnesia (Corkin, 1984; Gabrieli et al., 1988; Scoville & Milner, 1957) and is reminiscent of his inability to learn a seven-digit series despite 25 consecutive presentations (Drachman & Arbib, 1966), a tactual stylus maze after 80 trials (Corkin, 1965), or a visual stylus maze after 215 trials (Milner, 1965). His failure to increase his performance speed on this puzzle across 6 days suggests that the sensorimotor skill underlying puzzle completion did not benefit from repeatedly inserting the same items in the grid. Given his lifelong preoccupation with crossword puzzles, we believe it is likely that the motor component had reached asymptote many years ago. The post-1953 puzzle also appeared to be the most difficult among the three puzzles for healthy volunteers. Thus, in a difficult learning task, when knowledge could not easily be anchored to already acquired information, H.M. showed no evidence of learning following a limited number of repetitions. That he never generated a correct answer also attests to the fact that he has been unable since the onset of his amnesia to acquire semantic knowledge that would benefit his performance on this particular type of task.

Acquisition of New Postoperative Semantic Knowledge

We find it interesting that on a puzzle supplying post-1953 clues for pre-1953 answers, H.M. improved significantly. On this task, his accuracy increased consistently across 5 days of testing. He showed a 226% increase in performance, to the point that on Day 5 he performed just below the level of the healthy volunteers on Day 1. Thus, in marked contrast to his inability to show any learning on the post-1953 puzzle, H.M. did appear able to show some learning when information could be linked to mental representations established preoperatively.

The results from the first priming experiment suggest that after repeated exposures to the answer keys during the 5 consecutive days of testing, the representations of the pre-1953 answers were primed. Critically, then, he had read and processed all of the words presented to him over the 5 days, and this exposure to the information had resulted in nondeclarative learning. This learning was not sufficient, however, to support his performance on a puzzle, because he showed priming for a far greater number of answers than he was able to provide explicitly on the original crossword puzzles. His performance on the second priming test, which provided only preoperative clue fragments, was significantly worse than his performance on both the original puzzles and the first priming test. For example, he was able to answer “Astaire” on the 5th consecutive day of testing for pre–post puzzle 2 and on the first
priming puzzle. When given the partial clue ("male American dancer"), however, he was unable to solve that particular clue. These results can be interpreted in three ways. First, H.M. may actually have benefited from the postoperative semantic information provided in the clue. Second, the preoperative information might have been too general or too cryptic to support his learning, and last, H.M.’s learning might have been sufficiently inflexible such that it required exactly the same puzzle clues (see Bayley & Squire, 2002, for a discussion about patient E.P., who showed similar inflexible learning).

Charness and Bieman-Copland (1992) suggested that there may be a snowball effect in knowledge acquisition, in that the “more you know, the easier it is to add new knowledge” (p. 322). Tulving et al. (1991) further suggested that “the meaningfulness of materials may be an important determinant of new semantic learning in amnesia” (p. 611). Semantic knowledge becomes meaningful, then, if a person already knows the information or if the new facts are consistent with preexisting concepts. We propose that H.M. can acquire new semantic knowledge, at least temporarily, when he can attach it to mental representations established preoperatively—that is, when he encounters semantic information that is meaningful to him.²

This conclusion is consistent with anecdotal evidence described by Milner et al. (1968). They asked H.M. whether President Kennedy was alive or dead; he answered, without hesitation, that Kennedy had been assassinated. Given that H.M.’s operation was in 1953, we were surprised that H.M. was cognizant of both facts. John F. Kennedy (JFK) was inaugurated in 1961 and was assassinated in 1963. Nevertheless, the Kennedy family was a dynasty in politics, and H.M. was likely familiar with JFK long before 1953. For example, JFK served three terms in the House of Representatives (1946, 1948, and 1950) and was elected senator in 1952. During all of this time, H.M. had a chance to establish preoperative memories of JFK. We propose, then, that he was able to attach the new information about Kennedy’s presidency and assassination because of his previously established memories.

Similar anecdotal evidence was described by Corkin (1984), who reported that H.M. correctly knew that “Raymond Burr played the part of a detective on television.” However, the show Perry Mason—the title character was the detective-like role for which Burr is best remembered—was first broadcast on September 21, 1957, nearly 4 years after H.M.’s operation. In light of our findings, we suggest that H.M. was likely to remember this new fact because he already had an established knowledge base about Burr. In fact, before starring as Perry Mason, Burr appeared in the 1948 film Pitfall and worked on 90 films in the next 11 years. During our 2003 visit with H.M., he remembered Burr as a "movie star." In short, there were numerous opportunities for H.M. to have developed semantic memories about Burr before H.M.’s surgery in 1953.

Recently, Skotko et al. (2001) obtained additional evidence to support this view. Investigators spent 2 days with H.M. in his residential setting, studying his oral language production and comprehension. During this time, H.M. repeated that Kennedy was president and had been assassinated. He further mentioned that "Jackie Onassis" was the wife of JFK. (It was not until 1968 that Jackie took the name "Onassis," when she married for a second time.) H.M. also mentioned that "[Raymond Burr] played more of a detective, in a way," and H.M. also knew that "Joe DiMaggio" was the husband of Marilyn Monroe. Monroe was first photographed in 1944 and, within a year, appeared on the cover of more than 33 magazines. In 1947, she starred in her first film, The Shocking Miss Pilgrim, and continued to act thereafter. Monroe and DiMaggio were married in 1954. Thus, it is likely that H.M. had established a representation of Monroe and/or DiMaggio before his operation and was able to link the two. During a 2003 visit with H.M., he stated that Joe DiMaggio was a "baseball player."

There have also been some instances in which H.M. was surprisingly able to make semantic connections with previously unencountered material. As reported by Ogden and Corkin (1991) and Skotko et al. (2001), H.M. appeared to have a vague recollection of the Challenger mission nearly 2 weeks after the disaster. The space shuttle dominated the news because it was destined to carry Christa McAuliffe, a high school teacher, into space. The space shuttle exploded in 1986, however, and H.M. is not likely to have any previous memories of either the Challenger or Christa McAuliffe (neither was newsworthy prior to 1953). Nonetheless, H.M. was able to retain some memory of the explosion—albeit, a fragmented and incomplete one. According to Skotko et al., it is possible that H.M. falsely anchored his semantic facts about the Challenger to his memories of the Titanic disaster, Similarly, as reported by Milner, Corkin, and Teuber (1968), H.M. was able to recall the death of Pope John XXIII in 1963. Again, this memory is noteworthy because the pope did not take the name John XXIII until he was elected in 1958. (H.M. is Catholic, however, and the possibility exists that Angelo Giuseppe Roncalli, the pope’s birth name, was mentioned during H.M.’s religious education.) Needless to say, these two instances are even more impressive in light of H.M.’s profound global amnesia. Other examples were described by Corkin (2002).

Remembering Proper Names

Previous studies have shown that proper names are more difficult to remember than other types of semantic information, such as that tested in crossword puzzles (G. Cohen & Burke, 1993). Several theoretical models have been developed to explain this difficulty (sequential stage model, Bruce & Young, 1986; McWeeny, Young, Hay, & Ellis, 1987; Young, Hay, & Ellis, 1985; interactive activation and competition model, Burton & Bruce, 1992; representational model, G. Cohen, 1990; node structure theory, MacKay, 1987; token reference model, Semenza & Zettin, ² In October and November of 2003, long after H.M. should have forgotten the postoperative facts acquired from pre–post puzzle 1, we visited him to probe his lasting representations of our puzzle topics. For nearly all of the clues in which he had learned to associate postoperative facts, H.M. demonstrated knowledge (sometimes specific and other times general) about the subject material: Polio is “an illness” that “does something to the nerves.” Gone With the Wind is a “book and a movie written by a woman” and is important “because it was about the people after the Civil War” and features “Clark Gable.” Ike was a “politician” and a “general during WWII” who later became president. St. Louis is in “Missouri,” Raymond Burr was a “movie star,” and Joe DiMaggio was a “baseball player.” H.M. also associated the Warsaw Pact with “Europe during WWII” between the “U.S. and Germany” that involved “no trade or contact with us.” These recent descriptions from H.M. suggest that he is, at least, familiar with the topics tested in our pre–post puzzle 1.
Despite their differences, most of the models concur on one point: Proper name recall is disadvantaged because names have low meaningfulness, and as a result, only a single connection is often made between an individual and his or her name. Memory of proper names also appears to decline as a person grows older (Burke, MacKay, Worthley, & Wade, 1991; G. Cohen & Burke, 1988). Despite their differences, most of the models concur on one point: Proper name recall is disadvantaged because names have low meaningfulness, and as a result, only a single connection is often made between an individual and his or her name. Memory of proper names also appears to decline as a person grows older (Burke, MacKay, Worthley, & Wade, 1991; G. Cohen & Burke, 1988).

These results differ markedly from those of prior studies documenting H.M.’s inability to learn new vocabulary (Gabrieli et al., 1988; Postle & Corkin, 1998). Our results, therefore, raise two questions: What kind of learning supports H.M.’s performance on the crossword puzzles, and what is the neural substrate for this learning?

### What Kind of Learning Supports H.M.’s Improvement on the Pre–Post Puzzle?

Tulving et al. (1991) postulated that an errorless training method is paramount for learning to occur in patients with profound anterograde amnesia. Mere repetition with the opportunity to make mistakes did not seem to be effective for their patient. Similarly, our crossword puzzle experiments were designed with errorless training methods, and H.M. was forced to correct all mistakes at the end of each puzzle. In previous experiments by Postle and Corkin (1998), repetition was included as part of the study. Yet, in previous experiments by Gabrieli et al. (1988), in which H.M. did not demonstrate semantic learning for new vocabulary, an errorless training method was also used. Therefore, the discrepancy between the learning demonstrated in our experiments and those of previous experiments (Gabrieli et al., 1988) cannot be attributed wholly to our strict training method.

H.M. probably performed poorly in experiments by Gabrieli et al. (1988) because he did not have preoperative mental establishments on which he could build new associations. Specifically, he likely had no preoperative knowledge for words such as “egress,” “welkin,” and “tyro” and, therefore, could not establish new memory traces, even after 60 trials with the words. Similarly, in our crossword puzzle experiment, H.M. never learned to solve “the current President of the United States,” despite repeated attempts on this puzzle item. As President Clinton was not in the national newspapers prior to 1953, H.M. did not have an opportunity to form any prior memories of the president. Therefore, in the instances in which H.M. had no preoperative semantic knowledge of an event, he presumably had no mental foundation upon which to construct new learning.

The possibility also exists that the type of learning needed in our crossword paradigm is fundamentally different from that needed in vocabulary acquisition (Gabrieli et al., 1988). We are hesitant to refer to the learning demonstrated in our experiments as semantic learning. The learning was qualitatively different from the general conceptualization of semantic learning. It was not permanent, and it was relatively inflexible. In this way, the present findings are similar to those of a prior investigation with a patient with amnesia, E.P. (Bayley & Squire, 2002), which showed that he was capable of learning words that were part of three-word sentences. However, the information was inflexible (i.e., it did not generalize to sentences in which one word had been replaced with a synonym) and did not appear to be accessible to E.P.’s conscious awareness. Thus, the authors postulated that factual knowledge can, in some instances, be acquired using a nondeclarative mechanism that is fundamentally different from the mechanism that is typically used to learn factual knowledge.

Our present experiments differ from the study with E.P. in that the number of item repetitions were far fewer (5 days compared with over 30 test sessions with E.P.). H.M. also seemed to have
more conscious awareness of his accuracy with the responses than did E.P. H.M.’s learning is also difficult to attribute solely to perceptual priming (e.g., the priming control study indicated that H.M. did not simply believe that “polio” fit into certain boxes, regardless of the clues) or semantic priming (e.g., the priming control study showed that H.M.’s learning was not simply due to “polio” being primed when he saw the word “disease”). Nevertheless, H.M.’s learning was inflexible enough that it may have been supported by some form of nondeclarative memory. Or, perhaps, it was supported by a fragmentary declarative memory system that was capable of temporarily acquiring, but not permanently storing, the new knowledge. Future research must answer the question: Is there a type of transient acquisition of episodic memory that can occur with extensive damage to MTL structures, or is some form of nondeclarative learning the only option?

What Are the Neural Substrates Supporting Learning in Amnesia?

A number of other studies have described adult patients with amnesia who were capable of learning new semantic information, albeit more slowly and less successfully than did healthy participants. The precise anatomical lesions of these patients varied, and in most cases, their damage to MTL structures was not as complete as that of H.M.’s, but it is still significant to note the acquisition of postoperative semantic memory in these cases. Persons with profound anterograde amnesia have remembered common words at the ends of meaningful sentences (Shimamura & Squire, 1988); common words embedded in the ends of meaningful sentences (Shimamura & Squire, 1988); postoperative semantic memory in these cases. Persons with pro-

One patient who seems to be an exception to these results is S.S., a 50-year-old man who, after contracting herpes simplex encephalitis (HSE) in 1971, became densely amnesic (Cermak & O’Connor, 1983). Despite having a master’s degree in physics and being the president of an optical physics firm, S.S. could not recall after 1 min the contents of a newspaper article on a recent laser discovery. Even after repeated exposures, S.S. “continued to fail to incorporate this new information into his memory and was unable to recall the article or even recognize its more salient features a few moments after each exposure” (Cermak & O’Connor, 1983, p. 231). These results demonstrate that S.S. was unable to build upon intensely familiar preoperative semantic knowledge. However, S.S.’s lesion was not confined to MTL structures and included surrounding cortical areas, including the left frontal lobe and orbitofrontal regions, findings that are consistent with the etiology of HSE. (O’Connor, Cermak, & Seidman, 1995; Verfaellie et al., 2000). This additional damage likely contributed to his failure to recall the new information.

How, then, was H.M. able to form new semantic associations for our pre–post puzzle? He is a unique patient because of the symmetry and extent of his MTL resection and because his profound amnesia resulted entirely from the MTL lesion and not from neocortical damage. The learning demonstrated in the present study cannot be attributed to residual episodic memory, and the neural substrate for this learning must be localized to structures other than the hippocampus. The preserved 2 cm of the caudal hippocampus have been deafferented by the removal of the entorhinal cortex and, therefore, are an unlikely substrate for any cognitive process. Rather, the neural processes that support the new semantic learning are probably localized to the preserved perirhinal and parahippocampal cortices (Corkin et al., 1997), which receive input from neocortical areas and may have residual function (Corkin, 2002).

The perirhinal and parahippocampal structures have been implicated in the ability to encode information (Brewer, Zhao, Desmond, Glover, & Gabrieli, 1998; Kensinger, Clarke, & Corkin, 2003; Kirchhoff, Wagner, Mariel, & Stern, 2000) and to retrieve information (Davachi, Mitchell, & Wagner, 2003; Strange, Otten, Josephs, Rugg, & Dolan, 2002). Prior behavioral evidence suggests that H.M. may retain some function of his posterior parahippocampal gyrus: He was able to draw an accurate floor plan of the house he moved into after onset of his amnesia (Corkin, 2002) and to show some learning on a navigation task (Bobbot et al., 1998). Experiments with animals have also revealed that the parahippocampal and perirhinal cortices are essential for the formation of long-term memories (Squire & Zola-Morgan, 1991). Further, a functional MRI study by Henke, Buck, Weber, & Wieser (1997) revealed that associative learning activated hippocampal and parahippocampal regions. Perhaps, then, under specific circumstances (e.g., when information can be anchored to already acquired information), this region is capable of supporting some learning of factual knowledge. Thus, H.M.’s preserved MTL cortex may allow him to acquire, at least temporarily, some knowledge that is rather sparse and inflexible.
It is also possible that regions outside the MTL contribute to the type of learning demonstrated by H.M. Slow neocortical learning may supplement residual function in the perirhinal or parahippocampal cortices or both to allow for this type of learning. After this short-term acquisition of knowledge, registration of that information into permanent semantic stores may require MTL structures, because H.M.'s performance regressed to his baseline after the initial visit. Thus, he was unable to register his newly acquired postoperative semantic facts into his sustained long-term memory.

Recent evidence suggests that some neocortical regions may be capable of such limited neocortical semantic learning (Brown, Hagoort, & Kutas, 2000; Henke, Weber, Kneifel, Wieser, & Buck, 1999; Vargha-Khadem et al., 1997). Encoding of new words has been associated with increased activity in the left inferior frontal gyrus (Heun et al., 1999; Wagner, Paré-Blagoev, Clark, & Poldrack, 2001; Wiggs, Weisberg, & Martin, 1999), left superior parietal cortex (Heun et al., 1999; Wiggs et al., 1999), and left prefrontal cortices (Wiggs et al., 1999).

In conclusion, these experiments demonstrate that H.M. is capable of learning some new factual information when it can be fixed to already acquired knowledge. The cognitive and neural processes contributing to this learning remain unspecified. We suggest three possibilities: (a) Some residual declarative, semantic learning may be mediated by intact caudal, perirhinal, and parahippocampal cortices, (b) H.M.'s learning may be an instance of relatively inflexible learning via nondeclarative memory in neocortical structures supporting the acquisition of factual knowledge, or (c) the two processes may both contribute. Our results suggest that, despite a lack of episodic learning, patients with dense amnesia may be able to use cortical structures outside of the MTL to acquire new semantic facts temporarily. Nevertheless, MTL structures may be needed to register that factual information into a permanent memory store.

References
H.M.’s Puzzle Memries

769


Received June 20, 2002
Revision received December 17, 2003
Accepted December 22, 2003

Call for Nominations

The Publications and Communications (P&C) Board has opened nominations for the editorships of Clinician’s Research Digest, Emotion, JEP: Learning, Memory, and Cognition, Professional Psychology: Research and Practice, and Psychology, Public Policy, and Law for the years 2007–2012. Elizabeth M. Altmaier, PhD; Richard J. Davidson, PhD, and Klaus R. Scherer, PhD; Thomas O. Nelson, PhD; Mary Beth Kenkel, PhD; and Jane Goodman-Delahunty, PhD, respectively, are the incumbent editors.

Candidates should be members of APA and should be available to start receiving manuscripts in early 2006 to prepare for issues published in 2007. Please note that the P&C Board encourages participation by members of underrepresented groups in the publication process and would particularly welcome such nominees. Self-nominations also are encouraged.

Search chairs have been appointed as follows:

• Clinician’s Research Digest: William C. Howell, PhD
• Emotion: David C. Funder, PhD
• JEP: Learning, Memory, and Cognition: Linda P. Spear, PhD, and Peter Ornstein, PhD
• Professional Psychology: Susan H. McDaniel, PhD, and J. Gilbert Benedict, PhD
• Psychology, Public Policy, and Law: Mark Appelbaum, PhD, and Gary R. VandenBos, PhD

Candidates should be nominated by accessing APA’s EditorQuest site on the Web. Using your Web browser, go to http://editorquest.apa.org. On the Home menu on the left, find Guests. Next, click on the link “Submit a Nomination,” enter your nominee’s information, and click “Submit.” Preparad statements of one page or less in support of a nominee can also be submitted by e-mail to Karen Sellman, P&C Board Search Liaison, at ksellman@apa.org.

The deadline for accepting nominations is December 10, 2004, when reviews will begin.