

LETTERS

Robust habit learning in the absence of awareness and independent of the medial temporal lobe

Peter J. Bayley¹, Jennifer C. Frascino¹ & Larry R. Squire^{1,2,3,4}

Habit memory is thought to involve slowly acquired associations between stimuli and responses and to depend on the basal ganglia¹. Habit memory has been well studied in experimental animals but is poorly understood in humans because of their strong tendency to acquire information as conscious (declarative) knowledge. Here we show that humans have a robust capacity for gradual trial-and-error learning that operates outside awareness for what is learned and independently of the medial temporal lobe. We tested two patients with large medial temporal lobe lesions and no capacity for declarative memory. Both patients gradually acquired a standard eight-pair object discrimination task over many weeks but at the start of each session could not describe the task, the instructions or the objects. The acquired knowledge was rigidly organized, and performance collapsed when the task format was altered.

Declarative memory affords the capacity for conscious recollection of facts and events and depends on the integrity of the memory system based in the medial temporal lobe^{2,3}. Declarative memory can be contrasted with a collection of phylogenetically early, nondeclarative (and unconscious) memory abilities, including skills and habits, simple forms of conditioning, and other forms of experience-dependent behaviour that are expressed through performance rather than conscious recollection. A large body of literature involving humans, monkeys and rodents can be understood by recognizing that tasks requiring declarative memory depend on the integrity of the hippocampus and related structures, whereas other tasks can be learned nondeclaratively and are supported by other brain systems^{4–6}.

A continuing uncertainty about this taxonomy, and about the organization of mammalian memory systems, follows from the observation that some memory tasks, which humans acquire as declarative knowledge through memorization, can be learned nondeclaratively by experimental animals⁷. In this instance, the same task that is performed poorly by amnesic patients with hippocampal lesions is acquired at a normal rate by monkeys with hippocampal lesions. These findings raise questions about the extent to which one memory system can substitute for another, and about whether humans have the same capacity for nondeclarative memory that nonhuman animals have. In humans, perhaps some forms of nondeclarative memory are not well developed or are so overridden by the tendency to engage conscious memory strategies that nondeclarative memory is ineffective.

The problem is well illustrated by concurrent discrimination learning, a standard task that has been used to study mammalian memory for more than 50 years^{8,9}. In a common version of the task, eight pairs of objects are presented five times each day, one pair at a time in a mixed order, for a total of 40 trials. One object in each pair is always correct, and a choice of the correct object yields a reward. Humans readily learn this task, performing at about 90% correct after only one or two days of training. That the task ordinarily

depends on declarative memory, and on memorizing which object is correct in each pair, is demonstrated by the fact that task performance was correlated highly with the ability to describe the objects and by the fact that amnesic patients exhibited little learning during the period that their controls mastered the task^{10,11}.

In contrast to the findings from humans, monkeys learned the same concurrent discrimination task gradually in several hundred trials, and after hippocampal lesions they learned this task (or a related version) at a normal rate^{7,12,13}. A standard interpretation of the monkey data is that monkeys acquire the concurrent discrimination task by trial-and-error (habit) learning with the support of the basal ganglia^{1,13,14}. That is, they acquire a disposition to respond appropriately to each object pair. Habit memory is proposed to involve slowly acquired associations between stimuli and responses that develop outside awareness and are rigidly organized, with the result that what is learned is not readily expressed except when the task is presented just as it was during training.

One possibility is that habit learning is only weakly developed in humans, and some amount of declarative memory must be available to guide the learning. Indeed, successful learning of habit-like tasks has been reported only for moderately impaired amnesic patients who retain a considerable capacity for declarative memory^{15,16}. Moreover, some amnesic patients do not learn such tasks¹⁷. A related view is that the concept of an independent habit system is unnecessary, because habit learning and other kinds of learning depend on similar mechanisms¹⁸. Alternatively, if the capacity for habit learning is as well developed in humans as it appears to be, for example, in the monkey, then patients with large medial temporal lobe lesions and no capacity for declarative memory should be able to acquire the concurrent learning task gradually and to a high level of proficiency in the same manner that monkeys learn the task.

We tested two males (patients EP and GP) who have large medial temporal lobe lesions¹⁹ (see also Supplementary Information) and profound memory impairment as a result of herpes simplex encephalitis (Fig. 1). Four controls were also tested. Both patients fail altogether at formal tests of new learning ability, for example recall or recognition of word lists, stories and diagrams^{20,21}. Further, they have acquired little, if any, new declarative knowledge about the world since the onset of their amnesia, and neither of them can draw an accurate floor plan of his current residence²². Their capacity for habit learning has not been explored.

Figure 2 shows that the controls learned the task easily during three days of testing. EP and GP learned gradually during 36 and 28 sessions, respectively. EP improved at a nearly linear rate from 45.0% correct to 85.0% correct (linear trend: $F_{1,35} = 54.7$, $P < 0.001$). GP improved from 55.0% to 92.5% correct (linear trend: $F_{1,27} = 13.6$, $P = 0.001$). During training, participants gave verbal labels to some of the objects.

The successful learning exhibited by the patients across weeks was

¹Department of Psychiatry, ²Department of Neurosciences, and ³Department of Psychology, University of California, San Diego, La Jolla, California 92093, USA. ⁴Veterans Affairs Medical Center (116A), 3350 La Jolla Village Drive, San Diego, California 92161, USA.

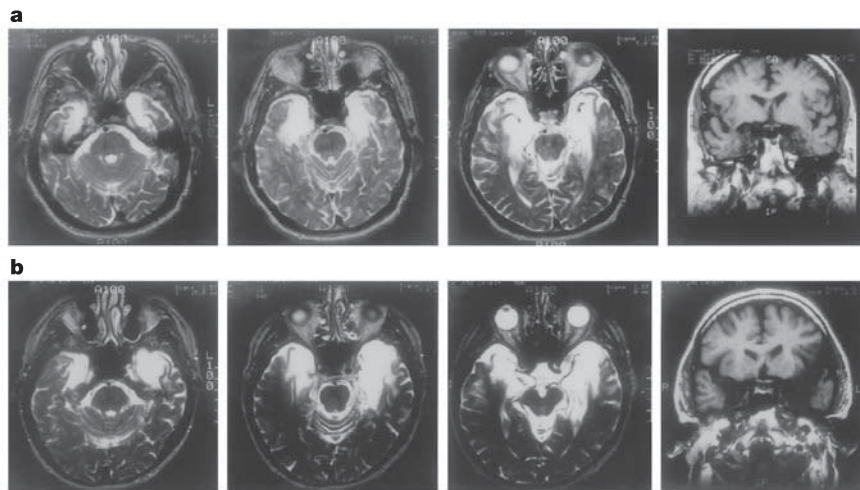


Figure 1 | Magnetic resonance images showing the extent of damage to the temporal lobe in amnesic patients. Images are shown for patient EP (**a**) and patient GP (**b**). The first three images in each row are T₂-weighted axial images through the temporal lobe. The images are continuous 5-mm sections (with 2.5-mm gaps) and are arranged from ventral to dorsal (left to

right). Damaged tissue is indicated by a bright signal. The last image in each row is a coronal T₁-weighted image at the level of the amygdala. Damaged tissue is indicated by a dark signal. See ref. 19 and Supplementary Information for a detailed description of the lesions.

not accompanied by reportable knowledge about the nature of the task. Even during testing, they never recognized that they had encountered the same task in previous sessions. By the fourth session both patients, without specific instruction, began to pick up the objects and turn them over (to look for the word “correct”), but they never acquired declarative memory of what they did each day. Before each session, both patients were told that they had been participating in testing on a regular basis, and they were asked to describe the testing. EP consistently replied that he would be asked questions and have discussions, but he never mentioned objects or described a learning task. When he was then told that the testing involved objects, EP regularly stated that he was to describe the objects and tell what they were used for. GP was able to report a little more about the testing. Before every session except one, he stated that he would look at photographs and answer questions such as “What’s missing?” or “What doesn’t belong?” On 11 occasions he said that he was to “pick the correct one”. On 10 occasions he also mentioned objects, but usually misdescribed what he was to do (for example, “choose the one that doesn’t belong”; “point out what’s missing”). On four different occasions he correctly stated that he was to pick the correct object, but he was unable to describe any of them.

Comments offered by the patients during and at the end of each test day described an automatic kind of responding. For example, EP was regularly asked at the conclusion of testing whether he had a strategy for choosing the correct object. After session 34, he said “No. It’s just up here [pointing to head] from the memory. It seems like it’s up there, and it comes down and out.” [Examiner: “Do you say to yourself, ‘I remember seeing that one?’”]. “No. It just seems that’s the one. It’s here (pointing to head) somehow or another and the hand goes for it.” After session 32, when asked if he had a strategy, EP replied, “No. It seems [pause] like from memory, or I’m just recalling this is the one and I picked it up. I can’t say memory. I just feel this is the one.” During session 23, GP stated, “I’m missing the same one.” [Examiner: “Why are you selecting that one?”] “It’s just jumping out at me. I’m the one. I’m the one. I keep wanting to pick it.”

To test whether the knowledge that had been acquired was rigidly organized, as is thought to occur for habit learning^{1,23}, or whether it could be used flexibly, we administered a sorting task 3–6 days after the conclusion of formal training. Controls succeeded with no difficulty, scoring 95.3% correct (Fig. 2a). In contrast, EP and GP failed the sorting task altogether (56.3% and 50.0% correct,

respectively; chance performance = 50%; Fig. 2b, c). EP placed nine objects in the group meant to contain correct objects and seven objects in the other group. GP placed eight objects in each group. Both patients expressed uncertainty and made exclamations when they first saw the array of 16 objects. EP said, “Gosh sakes. How to remember this?” When he moved to look underneath one of the objects, and was stopped by the examiner, he said “That’s just a habit, I think. It’s just, what’s underneath?”

Immediately after the sorting task, the familiar task format was reinstated (pairs of objects one at a time for 40 trials) but now participants were instructed to verbalize their choices (“left” or “right”) rather than pick up the objects. All participants scored 90% correct or better (Fig. 2). Immediately thereafter, each correct object was paired with an incorrect object, different from the one it had been paired with during training, and 40 additional test trials were given (re-pairing). The performance of EP and GP declined only a little, from 90.0% to 85.0% correct and from 92.5% to 75.0% correct, respectively (controls = 99.4% correct). These findings agree with the results of a recent study²⁴ in which rats with hippocampal lesions, trained on three visual discrimination problems, exhibited perfect transfer when the stimuli were recombined into novel pairs. It seems that re-pairing of stimuli in such tasks does not effectively challenge the ability to express knowledge flexibly, and simple associative processes can guide choice behaviour (for example, choose the object in each pair with the greatest positive value or least negative value).

Three days later, both patients were tested again with the standard 40-trial procedure but now no instructions were given. The first pair of objects was presented, and the examiner simply said that participants should do what comes naturally. EP picked up an object, turned it over to reveal the word “correct”, and said, “that’s the natural thing to me.” Similarly, GP said, “the correct one”, then picked up the correct object and turned it over. EP and GP went on to obtain scores of 95% and 100% correct, respectively. After the test was concluded, EP was asked how he knew what to do. “It seems to be automatic [pointing to head and then to table]. My mind just seemed to tell me, ‘just pick it up, it’s the right one.’” Finally 17 days after its first administration, each patient was again given the sorting task, followed immediately by 40 trials of the standard task (Fig. 2). The results were the same as before. EP and GP failed the sorting task (56.3% and 50.0% correct, respectively), but succeeded at the

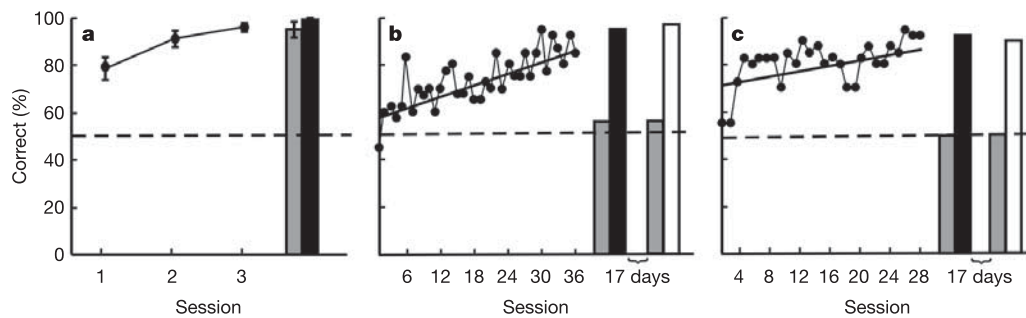


Figure 2 | Performance on the concurrent discrimination task. **a**, Controls ($n = 4$) learned the task easily within three sessions and performed well on the sorting task 3–6 days later (grey bar). The black bar shows performance immediately afterwards when participants were asked to verbalize their choices rather than reach for objects. Results are means \pm s.e.m. **b**, EP gradually learned the object pairs across 18 weeks. Five days later, he failed the sorting task (grey bar) but then, immediately afterwards, performed well in the standard task format while verbalizing his responses (black bar).

Seventeen days later, EP again failed the sorting task (grey bar) but performed perfectly when 40 trials were given exactly as in original training (white bar). **c**, GP learned the object pairs gradually during 14 weeks. Like EP, he failed the sorting task on two different occasions, 5 days after training and again 17 days later. In both instances he performed well immediately afterwards when the original task format was reinstated (black bar, verbalizing; white bar, standard task). The dashed line indicates the results expected by chance.

standard task (97.5% and 90.0% correct, respectively). EP placed 11 objects in the group meant to contain correct objects and 5 objects in the other group. GP placed 8 objects in each group. Whereas the two patients performed well on the standard task and at chance on the sorting task, the four controls performed nearly perfectly on both tasks. Control scores were also similar on the two tasks in a second study when performance was reduced by testing 4–7 days after only 1 day of training (40 trials) on a new set of eight object pairs (standard task 87.5% correct; sorting task 79.7% correct; $t_3 = 1.7$, $P = 0.19$).

These findings show directly the existence of a robust capacity for habit learning that can operate outside awareness and independently of declarative memory. The information acquired by both patients was rigidly organized and most accessible when the task was structured just as it was during training. Neither the instructions nor the response requirements (verbalizing versus reaching for objects) were critical, but performance did depend on the presentation of pairs of objects and on the requirement to choose one object in each pair. Thus, performance collapsed entirely in the sorting task when all the objects were encountered at once.

If the patients did learn the concurrent discrimination task as a habit in the manner that monkeys learn the task (that is, very gradually through reinforcement and with the support of the basal ganglia rather than the medial temporal lobe), then one might expect that they should have learned in about the same number of trials as monkeys with large medial temporal lobe lesions. The data support this expectation. Four monkeys with large medial temporal lobe lesions learned this same concurrent discrimination task (five sessions/week rather than two) in a mean of 1,100 trials⁷. EP and GP first reached a score of 95% correct after 1,200 and 1,040 trials, respectively. This comparison is not straightforward because normal monkeys acquired the task in about 500 trials, and monkeys that required more trials, including the four animals just mentioned, had damage that extended laterally to include some lateral temporal cortex⁷. Still, both patients also have some lateral temporal damage (in the fusiform and insular cortices) that could have slowed their learning. Accordingly, it seems reasonable to suppose that the rate of learning shown by the patients approximated the learning rate of monkeys with similar lesions.

The present results provide particularly useful evidence for the distinction between conscious (declarative) and unconscious (non-declarative) learning systems. The biological reality of this distinction has sometimes been challenged^{25,26}. For example, amnesic patients might have conscious knowledge about a successfully performed task but forget it by the time they are queried. Further, task information that is acquired successfully might be easier to acquire

or be presented more often than task information that is not acquired. The present findings sweep aside these objections, because towards the end of training both patients regularly performed at high levels immediately after failing to describe the task, the instructions and the objects. The findings therefore affirm the validity of unconscious learning and show that humans possess a robust capacity for gradual, trial-and-error (habit) learning that can operate outside awareness for what is learned and independently of the medial temporal lobe. It seems likely that many tasks, including concurrent discrimination, which humans ordinarily acquire rapidly as declarative knowledge (through memorization) can also be acquired more slowly as habit memory. However, in such cases, what is acquired is rigidly organized and altogether different from declarative memory.

METHODS

Subjects. EP was born in 1922, had 12 years of education, and developed amnesia in 1992. GP was born in 1946, had 16 years of education, and developed amnesia in 1987. Estimates of brain damage were based on a quantitative analysis of magnetic resonance images¹⁹ (also see Supplementary Information), following published procedures for segmenting the medial temporal lobe^{27–29}. Volume estimates were also calculated for the lateral temporal, frontal, parietal and occipital lobes. Four controls averaged 70.2 years of age and had 13.2 years of education.

The concurrent discrimination task. Eight pairs of junk objects (miscellaneous pieces of plastic or metal that were not readily nameable, each mounted on 8.9 cm \times 12.7 cm cardboard) were first presented one at a time. Participants were asked to choose one of the objects in each pair, and the object not chosen was then designated the ‘correct’ object for the duration of testing. Formal testing began 6 or 7 days later with the presentation of all eight pairs of objects five times each (session 1). Participants were told that the same object in each pair would always be correct and that they should try to learn the correct objects by trial and error. Specifically, they were told to make a choice between each pair of objects and then pick up the object, turn it over, and discover whether the word ‘correct’ appeared under the cardboard base. Testing continued in this manner twice each week on nonconsecutive days (40 trials per session). The order in which the object pairs were presented varied each day with the constraint that all eight pairs were presented within each block of eight trials. The left–right position of the correct object also varied pseudorandomly.

The sorting task. All 16 objects were placed together on a table, and participants were told that some of the objects had been consistently designated as correct. They were then asked to sort the objects, placing the correct objects to one side of the table and the other objects to the other side.

Received 24 March; accepted 31 May 2005.

1. Mishkin, M. & Petri, H. L. in *Neuropsychology of Memory* (eds Squire, L. R. & Butters, N.) 287–296 (Guilford, New York, 1984).
2. Squire, L. R. & Zola-Morgan, S. The medial temporal lobe memory system. *Science* 253, 1380–1386 (1991).

3. Eichenbaum, H. & Cohen, N. J. *From Conditioning to Conscious Recollection: Memory Systems of the Brain* (Oxford Univ. Press, New York, 2001).
4. Squire, L. R. Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychol. Rev.* **99**, 195–231 (1992).
5. Squire, L. R. Memory systems of the brain: a brief history and current perspective. *Neurobiol. Learn. Mem.* **82**, 171–177 (2004).
6. Schacter, D. L. & Tulving, E. *Memory Systems 1994* (MIT Press, Cambridge, Massachusetts, 1994).
7. Buffalo, E. A., Stefanacci, L., Squire, L. R. & Zola, S. M. A reexamination of the concurrent discrimination learning task: the importance of anterior inferotemporal cortex, area TE. *Behav. Neurosci.* **112**, 3–14 (1998).
8. Hayes, K. J., Thompson, R. & Hayes, C. Concurrent discrimination learning in chimpanzees. *J. Comp. Physiol. Psychol.* **46**, 105–107 (1953).
9. Correll, R. E. & Scoville, W. B. Effects of medial temporal lesions on visual discrimination performance. *J. Comp. Physiol. Psychol.* **60**, 175–181 (1965).
10. Squire, L. R., Zola-Morgan, S. & Chen, K. Human amnesia and animal models of amnesia: performance of amnesic patients on tests designed for the monkey. *Behav. Neurosci.* **11**, 210–221 (1988).
11. Hood, K. L., Postle, B. R. & Corkin, S. An evaluation of the concurrent discrimination task as a measure of habit learning: performance of amnesic subjects. *Neuropsychologia* **37**, 1375–1386 (1999).
12. Malamut, B. L., Saunders, R. C. & Mishkin, M. Monkeys with combined amygdalo-hippocampal lesions succeed in object discrimination learning despite 24-hour intertrial intervals. *Behav. Neurosci.* **98**, 759–769 (1984).
13. Teng, E., Stefanacci, L., Squire, L. R. & Zola, S. M. Contrasting effects on discrimination learning following hippocampal lesions or conjoint hippocampal-caudate lesions in monkeys. *J. Neurosci.* **20**, 3853–3863 (2000).
14. Fernandez-Ruiz, J., Wang, J., Aigner, T. G. & Mishkin, M. Visual habit formation in monkeys with neurotoxic lesions of the ventrocaudal neostriatum. *Proc. Natl Acad. Sci. USA* **98**, 4196–4201 (2001).
15. Knowlton, B. J., Squire, L. R. & Gluck, M. Probabilistic classification learning in amnesia. *Learn. Mem.* **1**, 106–120 (1994).
16. Myers, C. E. *et al.* Dissociating hippocampal versus basal ganglia contributions to learning and transfer. *J. Cogn. Neurosci.* **15**, 185–193 (2003).
17. Hopkins, R. O., Myers, C. E., Shohamy, D., Grossman, S. & Gluck, M. A. Impaired probabilistic category learning in hypoxic subjects with hippocampal damage. *Neuropsychologia* **42**, 524–535 (2004).
18. Gaffan, D. Memory, action and the corpus striatum: current developments in the memory-habit distinction. *Semin. Neurosci.* **8**, 33–38 (1996).
19. Bayley, P. J., Gold, J. J., Hopkins, R. O. & Squire, L. R. The neuroanatomy of remote memory. *Neuron* **46**, 799–810 (2005).
20. Stefanacci, L., Buffalo, E. A., Schmolck, H. & Squire, L. R. Profound amnesia following damage to the medial temporal lobe: a neuroanatomical and neuropsychological profile of patient E.P. *J. Neurosci.* **20**, 7024–7036 (2000).
21. Levy, D. A., Bayley, P. J. & Squire, L. R. The anatomy of semantic knowledge: medial vs. lateral temporal lobe. *Proc. Natl Acad. Sci. USA* **101**, 6710–6715 (2004).
22. Bayley, P. J. & Squire, L. R. Failure to acquire new semantic knowledge in patients with large medial temporal lobe lesions. *Hippocampus* **15**, 273–280 (2005).
23. Wise, S. P. The role of the basal ganglia in procedural memory. *Sem Neurosci.* **8**, 39–46 (1996).
24. Driscoll, I., Sutherland, R. J., Prusky, G. T. & Rudy, J. W. Damage to the hippocampal formation does not disrupt representational flexibility as measured by a novelty transfer test. *Behav. Neurosci.* **118**, 1427–1432 (2004).
25. Shanks, D. R. & StJohn, M. F. Characteristics of dissociable human learning systems. *Behav. Brain Sci.* **17**, 367–447 (1994).
26. Nosofsky, R. M. & Zaki, S. R. Math modelling, neuropsychology, and category learning: Response to B. Knowlton (1999). *Trends Cogn. Sci.* **3**, 125–126 (1999).
27. Insausti, R. *et al.* MR volumetric analysis of the human entorhinal, perirhinal, and temporopolar cortices. *AJNR Am. J. Neuroradiol.* **19**, 659–671 (1998).
28. Amaral, D. G. & Insausti, R. in *The Human Nervous System* (ed. Paxinos, G.) 711–755 (Academic, San Diego, 1990).
29. Gold, J. J. & Squire, L. R. Quantifying medial temporal lobe damage in memory-impaired patients. *Hippocampus* **15**, 79–85 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank B. Suchan for assistance, and R. Clark and J. Wixted for discussion. This work was supported by the Medical Research Service of the Department of Veterans Affairs, an NIMH grant, and the Metropolitan Life Foundation.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to L.S. (Lsquire@ucsd.edu).