Report

Monkeys Recall and Reproduce Simple Shapes from Memory

Benjamin M. Basile[1,*] and Robert R. Hampton[1]
[1]Department of Psychology and Yerkes National Primate Research Center, Emory University, Atlanta, GA 30322, USA

Summary

If you draw from memory a picture of the front of your childhood home, you will have demonstrated recall. You could also recognize this house upon seeing it. Unlike recognition, recall demonstrates memory for things that are not present. Recall is necessary for planning and imagining, and it can increase the flexibility of navigation, social behavior, and other cognitive skills. Without recall, memory is more limited to recognition of the immediate environment. Amnesic patients are impaired on recall tests [1, 2], and recall performance often declines with aging [3]. Despite its importance, we know relatively little about nonhuman animals' ability to recall information; we lack suitable recall tests for them and depend instead on recognition tests to measure nonhuman memory. Here we report that rhesus monkeys can recall simple shapes from memory and reproduce them on a touchscreen. As in humans [4, 5], monkeys remembered less in recall than recognition tests, and their recall performance deteriorated more slowly. Transfer tests showed that monkeys used a flexible memory mechanism rather than memorizing specific actions for each shape. Observation of recall in Old World monkeys suggests that it has been adaptive for over 30 million years [6] and does not depend on language.

Results and Discussion

Humans can freely recall information from memory, as when describing a criminal suspect. We can also recognize whether something currently experienced was experienced before, as when choosing a suspect from a line up. Recall and recognition describe two types of tests or retrieval situations. The critical distinction is whether the material to be remembered is present when you try to remember it. In humans, accurate performance in recall and recognition situations differentially recruits two types of memory: recollection and familiarity [7, 8]. Recollection often involves a deliberate retrieval of information, sometimes accompanied by additional details, such as study context. In contrast, familiarity produces a relatively vague judgment of novelty or recency, as when you know that you have met someone before but cannot remember their name or where you met.

Successful recognition of something can occur either by recollecting it or by detecting that it is familiar [9]. In contrast, for successful recall, one must bring the memory to mind through recollection; the studied material is not present to re-experience as familiar. The ability to recall information is particularly important because it frees memory from exclusive control by immediate time and place. Recognition can only

happen when we reperceive something we have perceived before. Normal tasks, like planning a meeting or giving driving directions, would be impossible if the things we needed to remember—meeting attendees or street names—had to be present for us to remember them. Accordingly, loss of recall ability drastically impairs quality of life; poor recall performance is a central deficit in amnesic patients following brain damage ([1, 2], but see [10]), and recall performance often declines during aging [3, 11]. Recall tests are critical to our understanding of the evolution of memory and other cognitive abilities and to our ability to diagnose and treat memory impairments.

Ask humans what they recall and they can tell you; give them a blank piece of paper and they can draw what they have seen. In contrast, nonhuman animals do not have language and do not naturally draw, making it difficult to create controlled conditions under which we can measure recall. Consequently, virtually all tests of memory used with nonhumans are recognition tests. Nonhumans can be trained to touch, peck, or look at a familiar image when it is re-presented after a delay, thereby reliably measuring memory in a recognition format (e.g., [12, 13]). Some investigators have devised clever tests intended to measure the distinct contributions of recollection and familiarity to memory performance in nonhumans (e.g., [14, 15]). These tests have led to important insights; however, all of them use recognition paradigms that leave the conclusions controversial [16]. Because the major barrier to testing recall in nonhumans is the lack of language as a readout of the contents of memory, it is perhaps not surprising that one of the most convincing demonstrations of recollection in a nonhuman comes from a lexigram-trained chimpanzee [17]. After having seen food or a desired object hidden in the forest outside her enclosure, Panzee spontaneously recruited help from human caretakers, pointed outside toward the forest, and touched the lexigram that corresponded to the hidden item. This test situation parallels recall tests because she reported the location and identity of the hidden food while inside; both the forest and the food were out of sight. Panzee's use of lexigrams appears to give researchers a unique tool to access her memory. However, tests of memory using lexigrams can still be reasonably characterized as recognition tests, albeit ones with hundreds of possible choices. Panzee "did not literally draw a lexigram or a map of the forest" ([18], p. 214), but her performance suggests that she had the necessary information to do so if she could draw. If an animal could draw or reproduce a previously seen image, that would provide a powerful test of recall in nonhumans.

In the current study, we trained five rhesus monkeys on a novel recall test in which they had to reproduce a simple figure on a touchscreen from memory. Our test was modeled after the Rey-Osterrieth Complex Figure Test [19], in which humans draw a complicated shape from memory. The Rey-Osterrieth Complex Figure Test is a well-established tool that has been used to diagnose recall impairments in amnesic humans [1]. During the study phase, monkeys saw a simple shape composed of two or three colored boxes located on a 5 × 5 grid on a computer touchscreen. During the test phase, one of the boxes appeared in a new location

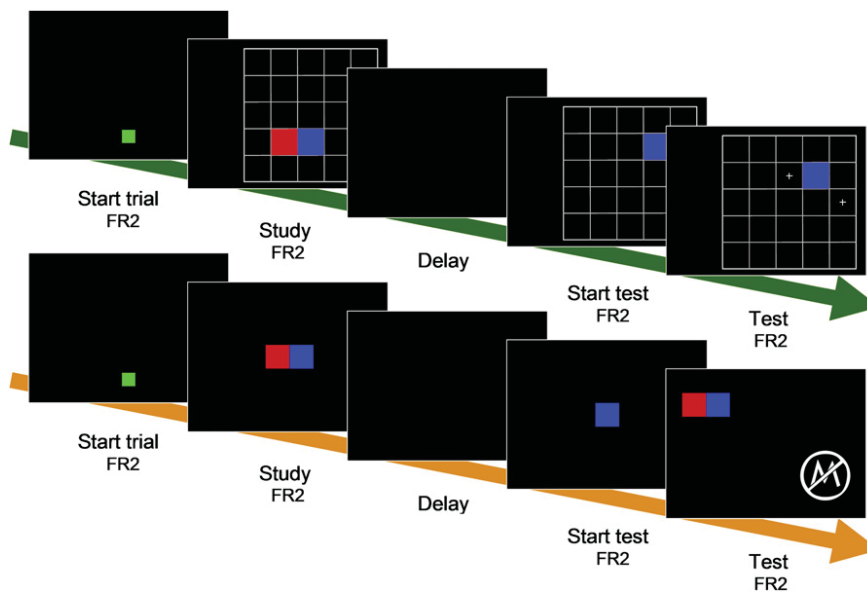*Correspondence: bbasile@emory.edu

Figure 1. Time Course of a Recall and Recognition Trial during Comparison

Schematic of the progression of a recall test (top) and matched recognition test (bottom). Monkeys started both tests by touching the green start box. An image then appeared, and they had to touch the blue box, ensuring that they had seen the sample image. After a delay, the blue box appeared in a new location, and the monkeys touched it to initiate the test phase. For the recall tests, monkeys earned food if they reproduced the studied shape by touching the appropriate grid location for the red box. For the recognition tests, monkeys earned food for touching the test stimulus if it was the same as that presented during the study phase (depicted) or for touching the nonmatch symbol if it was not. The small white crosses shown in the last panel of the recall test indicated to the monkeys which response locations were available. In the first phase of training, the white crosses were present in all the locations abutting the blue box. For the comparison with recognition, we reduced the response locations to two, which allowed us to equate the chance rate in the recall and recognition tests at 50%, permitting us to directly compare performance in the two types of test. To prevent accidental responding, each response required two consecutive touches to the same location.

on the grid. Monkeys could reproduce the absent box or boxes by touching the appropriate grid locations (Figure 1, top; see also Movie S1 available online). When successful, they earned food; errors were followed by a time out and no food. Critically, monkeys could not solve this memory test using familiarity, because the image to be remembered was not present during the test phase to experience as familiar. We hypothesized that if monkeys have recollection, they would be able to reproduce these simple shapes in this recall format.

All monkeys learned to reproduce two-box shapes after a brief delay more accurately than expected by chance (chance = 12.5%; mean accuracy = 27.6%; binomial tests; all p < 0.002). This performance parallels the way humans reproduce the Rey-Osterrieth Complex Figure. Unlike recognition tests, in which the target shape would have been present during the test phase, the monkeys had to reproduce the target from memory, making this the first pure recall test for monkeys.

Having established that monkeys can perform in a test methodologically similar to human recall tests, we further assessed the validity of our new paradigm by comparing it to a precisely matched recognition test. We tested for two performance differences diagnostic of recollection and familiarity. First, humans usually recall less information than they recognize [4]. This is because recall performance is based solely on successful recollection, whereas recognition performance is a combination of both recollection and correct familiarity judgments. Second, familiarity makes proportionally more contribution to human recognition performance at short memory delays [5, 20]. Consequently, higher accuracy on recognition tests should be most evident at short delays. To test for these patterns in our monkeys, we compared performance on our recall test to that on a precisely matched recognition test that had the same chance rate, used the same stimuli, required the same responses, and used the same study-test intervals (Figure 1, bottom).

Consistent with the hypothesis that the recall and recognition tests measure different kinds of memory, we found that recognition accuracy was higher than recall accuracy at short delays but declined more rapidly (Figure 2). The striking similarity of these patterns in monkey recall and recognition performance to human performance on comparable tests [5, 20] suggests two things. First, our shape reproduction test measures recollection, similar to human drawing tests. Second, monkey memory is similar to human memory; it likely includes two processes, recollection and familiarity, that contribute differentially to recall and recognition performance.

One common criticism of studies of nonhumans is that subjects may solve even complex tasks using relatively simple, inflexible stimulus-response rules acquired through extensive training. Monkeys might have learned a fixed response appropriate for each sample image. Such inflexible stimulus-response rules could result in performance that superficially resembled recall but would not generalize to novel images. To evaluate whether monkeys used flexible recollective memory or rigid response rules, we tested whether performance generalized to novel three-box shapes. We did this both under conditions in which the chance rate remained the same as in earlier tests (reproduce one box of a novel three-box shape) and under conditions in which the difficulty was increased (reproduce two boxes of a novel three-box shape). We hypothesized that if monkeys had learned a general reproduction rule rather than inflexible stimulus-response rules, accuracy would be significantly above chance in the first session with each of these novel test conditions.

Monkeys immediately transferred recall performance to novel three-box shapes (Figure 3). Accuracy was significantly above chance both when monkeys had to produce one box of a novel three-box shape and when they had to reproduce two boxes to complete a three-box shape. Generalization to novel shapes shows that monkeys remembered the images in a flexible way that parallels human recall. The small number of trials
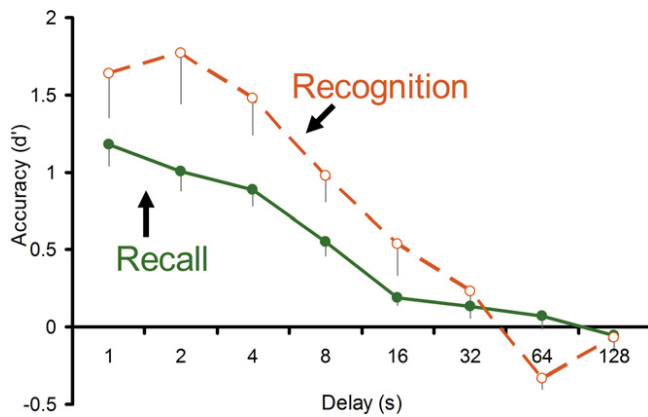
Figure 2. Comparison of Recall and Recognition Accuracy as a Function of Memory Delay

Under precisely matched testing conditions, monkeys showed greater accuracy and faster forgetting in the recognition test than in the recall test (two-factor within-subject analysis of variance [test type X delay]: main effect of test type [$F_{(1,4)}$ = 6.66, p = 0.061], main effect of delay [$F_{(7,28)}$ = 38.54, p < 0.001], and interaction [$F_{(7,28)}$ = 3.96, p = 0.004]). Accuracy in both recall and recognition is reported as d′ [26], as a function of the delay in seconds between study and test. Error bars represent 1 standard error of the mean.
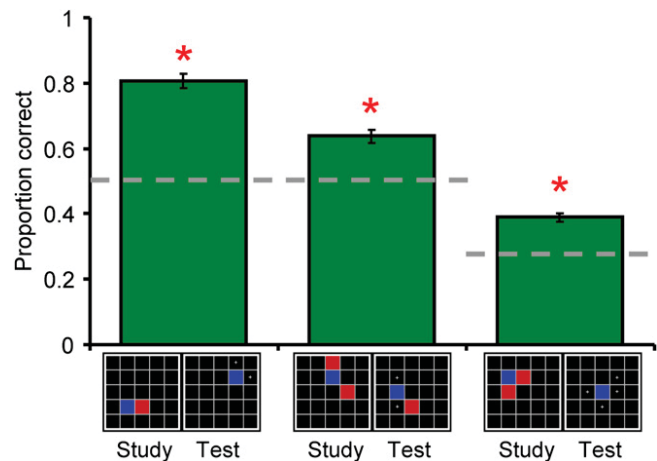

Figure 3. Recall Performance on Trained Shapes and on Novel Shapes during Transfer Tests

Monkeys successfully generalized to novel shapes. Accuracy was above chance on the first session both when monkeys had to reproduce one box of a three-box shape (middle; chance = 0.5; one-sample t test: $t_{(4)}$ = 7.03, p = 0.002) and when they had to reproduce two boxes of a three-box shape (right; chance = 0.25; one-sample t test: $t_{(4)}$ = 10.75, p < 0.001). Dashed lines represent accuracy expected by chance. Asterisks mark performance that is significantly above chance. Error bars are ± 1 standard error of the mean.

received during the generalization tests (144 trials), and the large number of novel shapes (28 three-box shapes), makes it unlikely that the monkeys learned a new set of response rules for each new shape.

Comparing recall and recognition is complicated. The two tasks usually require different types of responses and have drastically different chance rates. In the current experiment, we matched the recall and recognition tests on all critical procedural details, allowing us to attribute the observed differences in accuracy and forgetting to the types of memory used in the two tests. The design of our tests also rules out several alternative strategies. Monkeys could not have solved the recall test by constantly touching the location of the studied shape during the delay because the shape moved between study and test. They could not have solved the task by repeating a motor response made during the study phase because they were not required to touch all boxes during the study phase. Finally, they could not have solved the task using a set of inflexible stimulus-response rules because they immediately transferred performance to novel shapes. Monkeys appear to have solved the recall task by recollecting the studied shape when they could not see it.

This new recall test for nonhuman primates advances our understanding of the range of memory types present in monkeys and available for neurobiological study. Comparisons of recall and recognition performance in amnesic patients have stimulated considerable excitement and controversy about the neural substrates of memory [21, 22]. These controversies are difficult to address conclusively in humans because accidental brain damage is rarely selective or complete for a given structure. Studies of nonhumans allow for tighter experimental control over variables of interest, such as prior stimulus exposure and training, and permit methodologies that are difficult or impossible to use in humans. Use of these techniques with this new recall test promises new insights into the organization of human and nonhuman memory.

The presence of recollection in rhesus monkeys suggests that ancestors common to humans and Old World monkeys evolved under selection pressures favoring the ability to recall as well as recognize. Recollection and familiarity likely evolved because they solved functionally incompatible problems [23]. For example, familiarity does not support detailed memory for context, but it is quick [24] and resistant to distraction [25]. Recollection is slower and more vulnerable to distraction but supports a more detailed and flexible use of memory. Familiarity might better allow rapid responses to foods and predators under distracting conditions, whereas recollection might be necessary to access knowledge of distant food locations or past social interactions for planning future behavior. In this study, we have demonstrated recall performance in monkeys under limited laboratory conditions. Further work will be required to understand how this performance relates to natural behavior.

Experimental Procedures

Subjects and Apparatus

Five adult male rhesus macaque monkeys (*Macaca mulatta*) were tested 6 days a week in their home cages using portable touchscreen computer rigs (see Supplemental Experimental Procedures). All procedures were approved by the Emory University Institutional Animal Care and Use Committee and complied with United States law.

Initial Training

Monkeys first learned to touch accurately within the small boxes of the response grid to turn boxes red. Next, they learned to reproduce one box of a stationary two-box shape after a 0 s delay. Finally, they learned to reproduce the box after a 1 s delay when it appeared in different locations during the test phase and during the study phase. The final phase was identical to that in Figure 1 (top), with the exception that white crosses appeared in all eight adjacent grid locations and therefore did not limit the available response locations. See Supplemental Experimental Procedures for additional training details.

Comparison with Recognition

Monkeys learned a match/nonmatch recognition test (Figure 1, bottom), and this new test and the recall test were trained to stability (six sessions with no significant change in accuracy; see Supplemental Experimental Procedures). Identical shapes were used in the recognition and recall tests. Chance rates in the two tests were equated at 50% by providing one correct and one incorrect choice during the test phase in both tasks. For the recall test, the blue box was presented with only two possible adjacent choices, indicated by white crosshairs, rather than the eight choices used in initial training. For the recognition test, one shape was presented with a nonmatch symbol (Figure 1).

One session of each type was given per day, with testing order alternated between days, for 11 days. The delay was 1 s on slightly less than half the trials (because of constraints of counterbalancing trial types), and delays of 2, 4, 8, 16, 32, 64, and 128 s were mixed pseudorandomly among the other half of the trials. Because the addition of varied delays was novel, we excluded the first session of each test type, leaving 600 trials at the trained delay and 120 trials at each of the other delays from each monkey for analysis.

Because our recognition test was a match/nonmatch test, we used d′ values, which provide a measure of accuracy that is unbiased by any overall tendency to choose match or nonmatch [26]. Using d′ scores also allowed us to directly compare accuracy on the match/nonmatch recognition test with accuracy on the two-choice recall test by transforming the proportion of correct scores on the recall test into d′ scores ([26], Table A5.7).

Transfer of Recall Performance to Novel Shapes

In the first recall transfer test, we assessed whether monkeys would generalize performance to novel three-box shapes under conditions in which they had to reproduce one box during the test phase. During the study phase, monkeys saw a shape composed of one blue box and two red boxes. During the test phase, the blue box appeared in a new location along with one of the red boxes (chosen at random) and two possible response locations indicated by white crosses. Monkeys received a single session consisting of all 504 possible shape/location configurations in a random order. Only the first 144 trials were used to assess transfer in order to equate this test with those used in the previous experiments and to limit the opportunity for the monkeys to learn responses specific to each new stimulus.

In the second transfer test, monkeys again saw three-box shapes during the study phase, but only the blue box appeared during the test phase, along with four possible response locations indicated by white crosses (see Movie S1). Monkeys had to reproduce both red boxes correctly by touching the two correct grid locations. After one red box was added correctly, one of the two remaining incorrect response locations became unresponsive, and the corresponding white cross disappeared. Chance was 50% for each box and 25% for reproducing both boxes correctly. Again, the first 144 trials served as the critical transfer data. Proportions were arcsine transformed prior to analysis to better approximate normality [27].

Supplemental Information

Acknowledgments

References

1. Vargha-Khadem, F., Gadian, D.G., Watkins, K.E., Connelly, A., Van Paesschen, W., and Mishkin, M. (1997). Differential effects of early hippocampal pathology on episodic and semantic memory. Science 277, 376–380.

2. Mayes, A.R., Holdstock, J.S., Isaac, C.L., Hunkin, N.M., and Roberts, N. (2002). Relative sparing of item recognition memory in a patient with adult-onset damage limited to the hippocampus. Hippocampus 12, 325–340.

3. Craik, F.I.M., and McDowd, J.M. (1987). Age differences in recall and recognition. J. Exp. Psychol. Learn. Mem. Cogn. 13, 474–479.

4. Postman, L., Jenkins, W.O., and Postman, D.L. (1948). An experimental comparison of active recall and recognition. Am. J. Psychol. 61, 511–519.

5. Yonelinas, A.P., and Levy, B.J. (2002). Dissociating familiarity from recollection in human recognition memory: Different rates of forgetting over short retention intervals. Psychon. Bull. Rev. 9, 575–582.

6. Steiper, M.E., and Young, N.M. (2006). Primate molecular divergence dates. Mol. Phylogenet. Evol. 41, 384–394.

7. Yonelinas, A.P. (2002). The nature of recollection and familiarity: A review of 30 years of research. J. Mem. Lang. 46, 441–517.

8. Kelley, C.M., and Jacoby, L.L. (2000). Recollection and familiarity. In The Oxford Handbook of Memory, E. Tulving and F.I.M. Craik, eds. (Oxford: Oxford University Press), pp. 215–228.

9. Yonelinas, A.P. (1997). Recognition memory ROCs for item and associative information: The contribution of recollection and familiarity. Mem. Cognit. 25, 747–763.

10. Haist, F., Shimamura, A.P., and Squire, L.R. (1992). On the relationship between recall and recognition memory. J. Exp. Psychol. Learn. Mem. Cogn. 18, 691–702.

11. Prull, M.W., Dawes, L.L.C., Martin, A.M., 3rd, Rosenberg, H.F., and Light, L.L. (2006). Recollection and familiarity in recognition memory: Adult age differences and neuropsychological test correlates. Psychol. Aging 21, 107–118.

12. Murray, E.A., and Mishkin, M. (1998). Object recognition and location memory in monkeys with excitotoxic lesions of the amygdala and hippocampus. J. Neurosci. 18, 6568–6582.

13. Sutton, J.E., and Shettleworth, S.J. (2008). Memory without awareness: Pigeons do not show metamemory in delayed matching to sample. J. Exp. Psychol. Anim. Behav. Process. 34, 266–282.

14. Fortin, N.J., Wright, S.P., and Eichenbaum, H. (2004). Recollection-like memory retrieval in rats is dependent on the hippocampus. Nature 431, 188–191.

15. Sauvage, M.M., Fortin, N.J., Owens, C.B., Yonelinas, A.P., and Eichenbaum, H. (2008). Recognition memory: Opposite effects of hippocampal damage on recollection and familiarity. Nat. Neurosci. 11, 16–18.

16. Wixted, J.T., and Squire, L.R. (2008). Constructing receiver operating characteristics (ROCs) with experimental animals: Cautionary notes. Learn. Mem. 15, 687–690.

17. Menzel, C.R. (1999). Unprompted recall and reporting of hidden objects by a chimpanzee (Pan troglodytes) after extended delays. J. Comp. Psychol. 113, 426–434.

18. Menzel, C.R. (2005). Progress in the study of chimpanzee recall and episodic memory. In The Missing Link in Cognition: Origins of Self-Reflective Consciousness, H.S. Terrace and J. Metcalfe, eds. (New York: Oxford University Press), pp. 188–224.

19. Rey, A. (1941). L'examen psychologique dans les cas d'encephalopathie traumatique. Arch. Psychol. 28, 286–340.

20. Hockley, W.E. (1992). Item versus associative information: Further comparisons of forgetting rates. J. Exp. Psychol. Learn. Mem. Cogn. 18, 1321–1330.

21. Squire, L.R., Wixted, J.T., and Clark, R.E. (2007). Recognition memory and the medial temporal lobe: A new perspective. Nat. Rev. Neurosci. 8, 872–883.

22. Eichenbaum, H., Yonelinas, A.P., and Ranganath, C. (2007). The medial temporal lobe and recognition memory. Annu. Rev. Neurosci. 30, 123–152.

23. Sherry, D.F., and Schacter, D.L. (1987). The evolution of multiple memory-systems. Psychol. Rev. 94, 439–454.

24. Yonelinas, A.P., and Jacoby, L.L. (1994). Dissociations of processes in recognition memory: Effects of interference and of response speed. Can. J. Exp. Psychol. 48, 516–535.

25. Anderson, N.D., Craik, F.I.M., and Naveh-Benjamin, M. (1998). The attentional demands of encoding and retrieval in younger and older adults: 1. Evidence from divided attention costs. Psychol. Aging *13*, 405–423.
26. Macmillan, N.A., and Creelman, C.D. (2005). Detection Theory: A User's Guide, Second Edition (Mahwah, NJ: Lawrence Erlbaum).
27. Aron, A., and Aron, E. (1999). Statistics for Psychology (Upper Saddle River, NJ: Prentice Hall).