Visual long-term memory has a massive storage capacity for object details

Timothy F. Brady, Talia Konkle, George A. Alvarez, & Aude Oliva

Department of Brain and Cognitive Sciences,

Massachusetts Institute of Technology,

Cambridge, MA, 02139

Please address correspondence to:

Timothy F. Brady or Aude Oliva

Department of Brain and Cognitive Sciences, 46-4078

Massachusetts Institute of Technology

77 Massachusetts Avenue

Cambridge, MA 02139

Email : tfbrady@mit.edu / oliva@mit.edu

Classification   : Social Sciences (Psychology), Biological Sciences (Neuroscience)

Manuscript info : 41,600 characters (including Figures) and 32,060 (not including Figures).

Abstract

One of the major lessons of memory research has been that human memory is fallible, imprecise, and subject to interference. Thus, although observers can remember thousands of images, it is widely assumed that these memories lack detail. Contrary to this assumption, here we show that long-term memory is capable of storing a massive number of objects with details from the image. Participants viewed pictures of 2500 objects over the course of 5.5 hours. Afterwards, they were shown pairs of images, and indicated which of the two they had seen. The previously viewed item could be paired with either an object from a novel category, an object of the same basic level category, or the same object in a different state or pose. Performance in each of these conditions was remarkably high (92%, 88%, 87%, respectively), suggesting participants successfully maintained detailed representations of thousands of images. These results have implications for cognitive models in which capacity limitations impose a primary computational constraint (e.g., models of object recognition), and pose a challenge to neural models of memory storage and retrieval, which must be able to account for such a large and detailed storage capacity.

Keywords: visual long-term memory, memory capacity, object recognition

\body

Introduction

We have all had the experience of watching a movie trailer and having the overwhelming feeling that we can see much more than we could possibly report later. This subjective experience is consistent with research on human memory, which suggests that as information passes from sensory memory, to short-term memory, to long-term memory, the amount of perceptual detail stored decreases. For example, within a few hundred milliseconds of perceiving an image, sensory memory confers a truly photographic experience, enabling you to report any of the image details (1). Seconds later, short-term memory enables you to report only sparse details from the image (2). Days later you might only be able to report the gist of what you had seen (3).

While long-term memory is generally believed to lack detail, it is well established that long-term memory can store a massive number of items. Landmark studies in the 1970s demonstrated that after viewing 10,000 scenes for a few seconds each, people could determine which of two images had been seen with 83% accuracy (4; see also 5, 6). This level of performance indicates the existence of a large storage capacity for images.

However, remembering the gist of an image (e.g., "I saw a picture of a wedding not a beach") requires the storage of much less information than remembering the gist and specific details (e.g., "I saw that specific wedding picture"). Thus, to truly estimate the information capacity of long-term memory, it is necessary to determine both the *quantity* of items that can be remembered and the *fidelity* (amount of detail) with which each item is remembered. This point highlights an important limitation of previous large

scale memory studies (4, 5, 6): the level of detail required to succeed at the memory tests has never been systematically examined. In these studies, the stimuli were images taken from magazines, where the foil items used in the two-alternative forced-choice tests were random images drawn from the same set (4). Thus, the foil items were typically quite different from the studied images, making it impossible to conclude whether the memories for each item in these previous experiments consisted of only the 'gist' or category of the image, or whether they contained specific details about the images. It therefore remains unclear exactly how much visual information can be stored in human long-term memory.

There are reasons for thinking that the memories for each item in these large-scale experiments might have consisted of only the 'gist' or category of the image. For example, a well-known body of research has shown that human observers often fail to notice significant changes in visual scenes (for instance, if their conversation partner is switched to another person, or if large background objects suddenly disappear; 7, 8). These 'change blindness' studies suggest that the amount of information we remember about each item may be quite low (e.g., 8). In addition, it has been elegantly demonstrated that the details of visual memories can easily be interfered with by experimenter suggestion, a matter of serious concern for eyewitness testimony, as well as another indication that visual memories might be very sparse (9). Taken together, these results have led many to infer that the representations used to remember the thousands of images from the experiments of Shepard (5) and Standing (4) were in fact quite sparse, with few or no details about the images except for their basic-level categories (e.g. 8, 10, 11, 12).

However, recent work has also suggested that visual long-term memory representations can be more detailed than previously believed. Long-term memory for objects in scenes can contain more information than only the 'gist' of the object (13, 14, 15, 16). For instance Hollingworth (13) showed that, when requiring memory for a hundred or more objects, observers remain significantly above chance at remembering which exemplar of an object they have seen (e.g., 'did you see this power drill or that one?') even after seeing up to 400 objects in between studying the object and being tested on it. This suggests that memory is capable of storing fairly detailed visual representations of objects over long time periods (e.g., longer than working memory).

The current study was designed to estimate the information capacity of visual long-term memory by simultaneously pushing the system in terms of both the quantity and the fidelity of the representations that must be stored. First, we used isolated objects that were not embedded in scenes, in order to more systematically control the conceptual content of the stimulus set and prevent the role of contextual cues that may have contributed to memory performance in previous experiments. In addition, we used very subtle visual discriminations to probe the fidelity of the visual representations. Finally, we had people remember several thousand objects. Combined, these manipulations enable us to estimate a new bound on the capacity of memory to store visual information.

Results

Observers were presented with pictures of 2500 real world objects for 3 seconds each. The experiment instructions and displays were designed to optimize the encoding of object information into memory. First, observers were informed that they should try to remember all the details of the items (e.g., 17). Second, objects from mostly distinct

basic-level categories were chosen to minimize conceptual interference (e.g., 18). Finally, memory was tested with a two alternative forced choice test in which a studied item was paired with a foil and the task was to choose the studied item, allowing for recognition memory rather than recall memory (as in 4).

We varied the similarity of the studied item and the foil item in three ways (see Figure 1). In the novel condition, the old item was paired with a new item that was categorically distinct from all of the previously studied objects. In this case, remembering the category of the object, even without remembering the visual details of the object, would be sufficient to choose the appropriate item. In the exemplar condition, the old item was paired with a physically similar new object from the same basic level category. In this condition, remembering only the basic level category of the object would result in chance performance. Finally, in the state condition, the old item was paired with a new item that was exactly the same object, but appeared in a different state or pose. In this condition, memory for the category of the object, or even for the object's identity, would be insufficient to select the old item from the pair. Thus, memory for specific details from the image is required to select the appropriate object in both the exemplar and state conditions. Critically, observers did not know during the study session which items of the 2500 would be tested afterwards, nor what they would be tested against. Thus any strategic encoding of a specific detail that would distinguish between the item and the foil was not possible. To perform well on average in both the exemplar and the state conditions, observers would have to encode many specific details from each object.

Performance was remarkably high in all three of the test conditions in the two-alternative forced choice (see Figure 2). As anticipated based on previous research (4), performance was high in the novel condition, with participants correctly reporting the old

item on 92.5% (s.e.m. = 1.6%) of the trials. Surprisingly, performance was also exceptionally high in both conditions that required memory for details from the images: on average, participants responded correctly on 87.6% (s.e.m. = 1.8%) of exemplar trials, and 87.2% (s.e.m. = 1.8%) of the state trials. A one-way repeated measures ANOVA revealed a significant effect of condition, $F(2,26) = 11.3$, $p < 0.001$. Planned pairwise t-tests show that performance in the novel condition was significantly more accurate than the state and exemplar conditions (novel vs. exemplar: $t(13)=3.4$, $p<0.01$; novel vs. state: $t(13)=4.3$, $p<0.01$; exemplar vs. state, n.s. $p>0.10$). However, reaction time data was slowest in the state condition, intermediate in the exemplar condition, and fastest in the novel conditions (M=2.58s, 2.42s, 2.28s respectively; novel vs. exemplar: $t(13) = 1.81$, $p = 0.09$; novel vs. state: $t(13) = 4.05$, $p = 0.001$; exemplar vs. state: $t(13) = 2.71$, $p = 0.02$), consistent with the idea that the novel, exemplar, and state conditions required increasing detail. Participant reports afterwards indicated that they were usually explicitly aware of which item they had seen, as they expressed confidence in their performance and volunteered information about the details that enabled them to pick the correct items.

During the presentation of the 2500 objects, participants monitored for any repeat images. Unbeknownst to the participants, these repeats occurred anywhere from 1 to 1024 images previously in the sequence (on average 1 in 8 images was a repeat). This task insured that participants were actively attending to the stream of images as they were presented, and also provided an online estimate of memory storage capacity over the course of the entire study session.

Performance on the repeat-detection task also demonstrated remarkable memory. Participants rarely false-alarmed (1.3%, ± 1% s.e.m), and were highly accurate in reporting actual repeats (96% overall, ± 1% s.e.m.). Accuracy was near ceiling for repeat

images with up to 63 intervening items, and declined gradually for detecting repeat items with more intervening items (see Supporting Information, Figure S1). Even at the longest condition of 1023 intervening items (i.e. items that were initially presented ~ 2 hours previously), the repeats were detected about 80% of the time. The correlation between observers' performance at the repeat detection task and their performance in the forced-choice was high (r=0.81). The error rate as a function of number of intervening items fits well with a standard power law of forgetting (r^2=0.98). The repeat detection task also shows that this high capacity memory arises not only in two-alternative forced choice tasks, but also in ongoing old/new recognition tasks, though the repeat-detection task did not probe for more detailed representations beyond the category level. Together with the memory test, these results indicate a massive capacity memory system, in terms of both the quantity and fidelity of the visual information that can be remembered.

Discussion

We found that observers could successfully remember details about thousands of images after only a single viewing. What do these data say about the information capacity of visual long-term memory? It is known from previous research that people can remember large numbers of pictures (4, 5, 6) but it has often been assumed that they were storing only the gist of these images (8, 10, 11, 12). While some evidence suggested observers are capable of remembering details about a few hundred objects over long time periods (e.g., 13), no experiment had previously demonstrated accurate memory at the exemplar or state level on such a large scale. The present results demonstrate visual memory is a massive store that is not exhausted by a set of 2500 detailed representations of objects. Importantly, these data cannot reveal the format of these representations, and

should not be taken to suggest that observers have a photographic-like memory (19). Further work is required to understand how the details from the images are encoded and stored in visual long-term memory.

The information capacity of memory

Memory capacity cannot solely be characterized by the number of items stored: a proper capacity estimate takes into account the number of items remembered and multiplies this by the amount of information per item. In the present experiment we show that the information remembered per item is much higher than previously believed, as observers can correctly choose among visually similar foils. Therefore, any estimate of long-term memory capacity will be significantly increased by the present data. Ideally, we could quantify this increase, e.g. using information-theoretic bits, in terms of the actual visual code used to represent the objects. Unfortunately, one must know how the brain encodes visual information into memory to truly quantify capacity in this way.

However, Landauer (20) provided an alternate method for quantifying the capacity of memory by calculating the number of bits required to correctly make a *decision* about which items have been seen and which have not (see also 21). Rather than assign images codes based on visual similarity, this model assigns each image a random code regardless of its visual appearance. Memory errors happen when two images are assigned the same code. In this model the optimal code length is computed from the total number of items to remember and the percent correct achieved on a two-alternative forced choice task (see Supporting Information). Importantly, this model does not take into account the content of the remembered items: the same code length would be obtained if people remembered 80% of 100 natural scenes or 80% of 100 colored letters.

In other words, the bits in the model refer to content-independent memory addresses, and not estimated codes used by the visual system.

Given the 93% performance in the *novel* condition, the optimal code would require 13.8 bits/item, which is comparable to estimates of 10-14 bits needed for previous large-scale experiments (20). To expand on the model of Landauer, we assume a hierarchical model of memory where we first specify the category and the additional bits of information specify the exemplar and state of the item in that category (see Supporting Information). To match 88% performance in the exemplar conditions, 2.0 additional bits/item are required for each item. Similarly, 2.0 additional bits are required to achieve 87% correct in the state condition. Thus, we increase the estimated code length from 13.8 to 17.8 bits/item. This raises the lower bound on our estimate of the representational capacity of long-term memory by an order of magnitude, from ~14,000 ($2^{13.8}$) to ~228,000 ($2^{17.8}$) unique codes. This number does not tell us the true visual information capacity of the system. However, this model is a formal way of demonstrating how quickly any estimate of memory capacity grows if we increase the size of the representation of each individual object in memory.

Why examine people's capacity to remember visual information? One reason is that the evolution of more complicated cognition and behavioral repertoires involved the gradual enlargement of the long-term memory capacities of the brain (22). In particular, there are reasons to believe the capacity of our memory systems to store *perceptual* information may be a critical factor in abstract reasoning (23, 24). It has been argued, for example, that abstract conceptual knowledge that appears amodal and abstracted from actual experience (e.g., 25) may in fact be grounded in perceptual knowledge (e.g., perceptual symbol systems: 26). Under this view, abstract conceptual properties are

created on the fly by mental simulations on perceptual knowledge. This suggests an adaptive significance for the ability to encode a large amount of information in memory: storing large amounts of perceptual information allows abstraction based on all available information, rather than requiring a decision about what information might be necessary at some later point in time (24, 27).

Organization of memory

All 2500 items in our study stream were categorically distinct and thus had different high-level, conceptual representations. Long-term memory is often seen as organized by conceptual similarity (e.g., in spreading activation models; 28, 29). Thus, the conceptual distinctiveness of the objects may have reduced interference between them and helped support the remarkable memory performance we observed (18). In addition, recent work has suggested that the representation of the perceptual features of an object may often differ depending on the category the object is drawn from (e.g., 30). Taken together, these ideas suggest an important role for categories and concepts in the storage of the visual details of objects, an important area of future research.

Another possible distinction in the organization of memory is between memory for objects, memory for collections of objects, and memory for scenes. While some work has shown that it is possible to remember details from scenes drawn from the same category (15), future work is required to examine massive and detailed memory for complex scenes.

Familiarity versus recollection

The literature on long-term memory frequently distinguishes between two types of recognition memory: familiarity, the sense that you have seen something before; and recollection, specific knowledge of where you have seen it (31). However, there remains controversy over the extent to which these types of memory can be dissociated, and the extent to which forced-choice judgments tap into familiarity more than recollection or vice versa (see 32 for a review). In addition, while some have argued that perceptual information is often more associated with familiarity and conceptual information is associated more with recollection (e.g., 31), this view also remains disputed (32, 33). Thus, it is unclear the relative extent to which observers' choices in the current two-alternative forced choice tests were based on familiarity versus recollection. Given the perceptual nature of the details required to select the studied item, it is likely that familiarity plays a major role, and that recollection aids recognition performance on the subset of trials in which observers were explicitly aware of the details that were most helpful to their decision (a potentially large subset of trials, based on self-reports). Importantly, however, whether observers were depending on recollection or familiarity alone, the stored representation still requires enough detail to distinguish it from the foil at test. Our main conclusion is the same whether the memory is subserved by familiarity or recollection: observers encoded and retained many specific details about each object.

Constraints on models of object recognition and categorization

Long-term memory capacity imposes a constraint on high-level cognitive functions and on neural models of such functions. For example, approaches to object

recognition often vary by either relying on brute force online processing or a massive parallel memory (for review see 34, 35). The present data lend credence to object recognition approaches that require massive storage of multiple object viewpoints and exemplars (36-39). Similarly, in the domain of categorization, a popular class of models, so-called exemplar models (40), have suggested that human categorization can best be modeled by positing storage of each exemplar that is viewed in a category. The present results demonstrate the feasibility of models requiring such large memory capacities.

In the domain of neural models, the present results imply that visual processing stages in the brain do not, by necessity, discard visual details. Current models of visual perception posit a hierarchy of processing stages that reach more and more abstract representations in higher-level cortical areas (35, 41). Thus to maintain featural details, long-term memory representations of objects might be stored throughout the entire hierarchy of the visual processing stream, including early visual areas, possibly retrieved on demand via a feedback process (41, 42). Indeed, imagery processes, a form of representation retrieval, have been shown to activate both high-level visual cortical areas and primary visual cortex (43). In addition, fMRI studies have indicated that a relatively mid-level visual area, the right fusiform gyrus, responds more when observers are encoding objects for which they will later remember the specific exemplar, compared to objects for which they will later remember only the gist (44). Understanding the neural substrates underlying this massive and detailed storage of visual information is an important goal for future research and will inform the study of visual object recognition and categorization.

Conclusion

The information capacity of human memory plays an important role in cognitive and neural models of memory, recognition, and categorization, as models of these processes implicitly or explicitly make claims about the level of detail stored in memory. Detailed representations allow more computational flexibility because they enable processing at task-relevant levels of abstraction (24, 27), but these computational advantages tradeoff with the costs of additional storage. Therefore, establishing the bounds on the information capacity of human memory is critical to understanding the computational constraints on visual and cognitive tasks.

The upper bound on the size of visual long-term memory has not been reached, even with previous attempts to push the quantity of items (4), or the present study's attempt to push both the quantity and fidelity. Here we raise only the lower bound of what is possible, by showing that visual long-term memory representations can contain not only gist information but also details sufficient to discriminate between exemplars and states. We think that examining the fidelity of memory representations is an important addition to existing frameworks of visual long-term memory capacity. While in everyday life we may often fail to encode the details of objects or scenes (7, 8, 17), our results suggest that under conditions where we attempt to encode such details, we are capable of succeeding.

Materials and Methods

Participants

Fourteen adults (aged 20 – 35) gave informed consent and participated in the experiment. All of the participants were tested simultaneously, using computer workstations that were closely matched for monitor size and viewing distance.

Stimuli

Stimuli were gathered using both a commercially available database (Hemera Photo-Objects, Vol. I & II) and internet searches using Google Image Search. Overall, 2600 categorically distinct images were gathered for the main database, plus 200 paired exemplar images and 200 paired state images drawn from categories not represented in the main database. The experimental stimuli are available on the authors' websites. Once these images had been gathered, 200 were selected at random from the 2600 objects to serve in the novel test condition. All participants were thus tested with the same 300 pairs of novel, exemplar and state images. However, the item seen during the study session and the item used as the foil at test were randomized across participants.

Procedure

Study Blocks. The experiment was broken up into ten study blocks of approximately 20 minutes each, followed by a 30 minute testing session. Between blocks participants were given a five minute break, and were not allowed to discuss any of the images they had seen. During a block, approximately 300 images were shown, with 2896 images shown overall: 2500 new and 396 repeated images. Each image (subtending 7.5 by 7.5 degrees of visual angle) was presented for 3 seconds, followed by an 800 ms fixation cross.

Repeat Detection Task. To maintain attention and to probe online memory capacity, participants performed a repeat detection task during the ten study blocks. Repeated images were inserted into the stream such that there were between 0 and 1023 intervening items, and participants were told to respond using the spacebar anytime that an image repeated throughout the entire study period. They were not informed of the structure of the repeat conditions. Participants were given feedback only when they responded, with the fixation cross turning red if they had incorrectly pressed the space bar (false alarm) or green if they had correctly detected a repeat (hit), and were given no feedback for misses or correct rejections.

Overall, 56 images were repeated immediately (1-back), 52 were repeated with 1 intervening item (2-back), 48 were repeated with 3 intervening items (4-back), 44 were repeated with 7 intervening items (8-back), and so forth, down to 16 repeated with 1023 intervening items (1024-back). Repeat items were inserted into the stream uniformly, with the constraint that all of the lengths of n-backs (1-back, 2-back, 4-back,...1024-back) had to occur equally in the first half of the experiment and the second half. This ensured that fatigue would not differentially affect images that were repeated from further back in the stream. Due to the complexity of generating a properly counterbalanced set of repeats, all participants had repeated images appear at the same places within the stream. However, each participant saw a different order of the 2500 objects, and the specific images repeated in the n-back conditions were also different across participants. Images that would later be tested in one of the three memory conditions were never repeated during the study period.

Forced Choice Tests. After a 10 minute break following the study period, we probed the fidelity with which objects were remembered. Two items were presented on

the screen, one previously seen old item, and one new foil item. Observers reported

which item they had seen before in a two-alternative forced choice task.

Participants were allowed to proceed at their own pace, and were told to emphasize

accuracy, not speed, in making their judgments. The 300 test trials were presented in a

random order for each participant, with the three types of test trials (novel, exemplar and

state) interleaved. The images that would later be tested were distributed uniformly

throughout the study period.

Acknowledgements

References

1. Sperling, G. (1960). The information available in brief visual presentations. *Psychol. Monogr.* 74: 1-29.

2. Phillips, W.A. (1974). On the distinction between sensory storage and short-term visual memory. *Percept. Psychophys.* 16: 283-290.

3. Brainerd, C.J. & Reyna, V.F. (2005). *The Science of False Memory*. Oxford University Press.

4. Standing, L. (1973). Learning 10,000 pictures. , *Q. J. Exp. Psychol.* 25: 207-222.

5. Shepard, R. N. (1967) Recognition memory for words, sentences, and pictures. *J. Verb. Learn. Verb. Behav.* 6: 156-163.

6. Standing, L., Conezio, J., Haber, R.N. (1970). Perception and memory for pictures: Single-trial learning of 2500 visual stimuli. *Psychon. Sci.* 19: 73-74.

7. Rensink R.A., O'Regan J.K., and Clark J.J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychol. Sci.* 8: 368-373.

8. Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends Cogn. Sci.* 1:261-267.

9. Loftus, E. F. (2003). Our changeable memories: legal and practical implications. *Nature Rev. Neurosci.* 4: 231-234.

10. Chun, M. M. (2003). In *Psychology of Learning and Motivation: Advances in Research and Theory: Cognitive Vision, Vol. 42,* eds Irwin D, Ross, B (Academic Press, San Diego, CA), pp 79-108.

11. Wolfe, J.M. (1998). Visual Memory: What do you know about what you saw? *Curr. Biol.* 8: R303-R304.

12. O'Regan, J.K. & Noë, A. (2001). A sensorimotor account of vision and visual consciousness, *Behav. Brain Sci.* 24: 939-1011.

13. Hollingworth, A. (2004). Constructing visual representations of natural scenes: The roles of short- and long-term visual memory. *J Exp Psychol Hum Percept Perform.* 30: 519-537.

14. Tatler B.W., & Melcher D. (2007). Pictures in mind: Initial encoding of object properties varies with the realism of the scene stimulus. *Perception* 36: 1715 – 1729.

15. Vogt, S., & Magnussen, S. (2007). Long-term memory for 400 pictures on a common theme. *Exp. Psycol.* 54: 298-303.

16. Castelhano, M., & Henderson, J. (2005). Incidental visual memory for objects in scenes. *Vis. Cog.* 12: 1017-1040.

17. Marmie, W. R. & Healy, A. F. (2004). Memory for common objects: Brief intentional study is sufficient to overcome poor recall of US coin features. *Appl. Cognit. Psychol.* 18: 445-453.

18. Koutstaal, W. & Schacter, D.L. (1997). Gist-based false recognition of pictures in older and younger adults. *J. Mem. Lang.* 37: 555-583.

19. Intraub, H., and Hoffman, J.E. (1992). Remembering scenes that were never seen: Reading and visual memory. *Am. J. Psychol*, 105, 101-114.

20. Landauer, T. K. (1986). How much do people remember? Some estimates of the quantity of learned information in long-term memory. *Cognit. Sci.* 10: 477-493.

21. Dudai, Y. (1997). How big is human memory, or on being just useful enough. *Learn. Mem.* 3: 341-365.

22. Fagot, J., & Cook, R. G. (2006). Evidence for large long-term memory capacities in baboons and pigeons and its implications for learning and the evolution of cognition. *Proc. Natl Acad. Sci. USA*, 103: 17564-17567.

23. Fuster, J.M. (2003). More than working memory rides on long-term memory. *Behav. Brain Sci.* 26: 737-737.

24. Palmeri, T. J., & Tarr, M.J. (in press), In *Visual Memory*, eds. S. J. Luck & A. Hollingworth, (Oxford University Press: New York, NY).

25. Fodor, J. A. (1975). *The language of thought.* Harvard University Press, Cambridge.

26. Barsalou, L.W. (1999). Perceptual symbol systems. *Behav. Brain Sci.* 22: 577-660.

27. Barsalou, L. W. (1990). In *Content and process specificity in the effects of prior experiences*, eds T. K. Srull & R. S. Wyer, Jr., (Erlbaum, Hillsdale, NJ), pp. 61-88.

28. Collins, A.M. & Loftus, E.F. (1975). A spreading activation theory of semantic processing. *Psychol. Rev.* 82: 407-428.

29. Anderson, J. R. (1983). Retrieval of information from long-term memory. *Science* 220: 25-30.

30. Sigala, N., Gabbiani, F., & Logothetis, N. K. (2002). Visual categorization and object representation in monkeys and humans. *J. Cog. Neuro.* 14: 187-198.

31. Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychol. Rev.*, 87: 252-271.

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

33. Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *J. Mem. Lang.*, 30: 513–541.

34. Tarr, M. J. & Bulthoff, H. H. (1998). Image-based object recognition in man, monkey and machine. *Cognition* 67: 1-20.

35. Riesenhuber, M. & Poggio, T. (1999). Hierarchical models of object recognition in cortex, *Nature Neurosci.* 2: 1019–1025.

36. Logothetis, N.K., Pauls, J. & Poggio, T. (1995). Shape representation in the inferior temporal cortex of monkeys. *Curr. Biol.* 5**:** 552-563.

37. Tarr, M.J., Williams, P., Hayward, W.G. and Gauthier, I., (1998). Three-dimensional object recognition is viewpoint dependent. *Nature Neurosci.* 1: 275–277.

38. Bulthoff, H.H., & Edelman, S. (1992). Psychological support for a two-dimensional view interpolation theory of object recognition. *Proc. Natl Acad. Sci. USA.* 89: 60-64.

39. Torralba, A., Fergus, R., & Freeman, W. (in press). 80 million tiny images: a large dataset for non-parametric object and scene recognition. *IEEE Trans. PAMI.*

40. Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *J. Exp. Psychol. Gen.,* 115: 39-61.

41. Ahissar, M. & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning, *Trends Cogn. Sci.* 8: 457–464.

42. Wheeler, M. E., Petersen, S. E., & Buckner, R. L. (2000). Memory's echo: vivid remembering reactivates sensory-specific cortex. *Proc Natl Acad Sci USA* 97: 11125-9.

43. Kosslyn, S. M., Pascual-Leone, A., Felician, O., Camposano, S., Keenan, J. P., Thompson, W. L., Ganis, G., Sukel, K. E. & Alpert, N. M. (1999). The Role of Area 17 in Visual Imagery: Convergent Evidence from PET and rTMS. *Science*, 284: 167-170.

44. Garoff, R. J., Slotnick, S. D., & Schacter, D. L. (2005). The neural origins of specific and general memory: the role of fusiform cortex. *Neuropsychologia,* 43: 847-859.
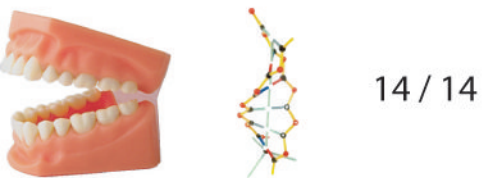
Figure Legends

Figure 1: Example test pairs presented during the two-alternative forced choice task for all three conditions (novel, exemplar, state). The number of observers reporting the correct item is shown for each of the depicted pairs. The experimental stimuli are available on the authors' websites.

Figure 2: Memory performance for each of the three test conditions (novel, exemplar, state) is shown above. Error bars represent the s.e.m. The dashed line indicates chance performance.

Novel	Exemplar	State
14 / 14	13 / 14	13 / 14
13 / 14	14 / 14	12 / 14
12 / 14	12 / 14	13 / 14
14 / 14	13 / 14	12 / 14
14 / 14	14 / 14	14 / 14
12 / 14	13 / 14	14 / 14
12 / 14	10 / 14	11 / 14
13 / 14	12 / 14	13 / 14
14 / 14	9 / 14	12 / 14
14 / 14	11 / 14	11 / 14

Memory Performance