Supporting Information

Brady *et al*. 10.1073/pnas.0803390105

SI Text

Modeling Methods. Our model of memory capacity in bits per item is based on work by Landauer (1), who provided a method for estimating the number of bits required to correctly make a decision about which items have been seen and which have not. This simple model assigns each picture a random code (b bits long), rather than assign them based on visual similarity. If a foil item is assigned the same code as any old item, the old item cannot be distinguished from the foil. Because this is a two-alternative forced-choice task, an error will occur on one half of these cases. Given this model we can compute the length of code (in bits) necessary to achieve any given accuracy level for a given number of items.

For example, to achieve 88% correct with 1,000 items in memory, you must have a code at least 11.9 bits long (3,821 possible codes). The chance of a new item being assigned a code that overlaps an old item would then be 23% ($[1 - 1/3{,}821]^{1{,}000}$), and on half of trials observers would still answer correctly, resulting in 12% errors, or 88% correct performance. In general, the number of bits is related to memory performance in this model by the following equation: $b = -\log_2[1 - (2p - 1)^{1/n}]$, where p is the percentage correct, and n is the number of items in memory.

If this model is capturing a systematic property of memory, similar estimates for the number of bits should be obtained by using different numbers of pictures in memory (because performance should increase correspondingly). Supporting this, Landauer (1) found that similar calculations result in estimates of 10.0 bits with 20 pictures in memory and 99% correct judgments, 10.2 bits with 400 pictures in memory and 86% correct judgments, etc. By using this model, we find observers must have 13.8 bits of information per item based on our novel condition (92% correct with 2,500 pictures in memory). This suggests that the maximum capacity of memory, if all items were coded by using the optimal set of features (decision-level bits), would be $2^{13.8}$ (14,000) unique items.

To model the exemplar and state results, we make the assumption that memory is organized hierarchically, such that the bits for the category appear before the bits for the exemplar per state, resulting in greater similarity in the codes (and greater chance of confusion) for items within a category than items in different categories. In the exemplar condition, observers are holding one item of the category in memory and this results in ≈87.5% correct. If observers have two bits of exemplar-level information about each studied item, then the probability of the a foil exemplar overlapping with a studied item would be 25% ($1/2^2$), resulting in 12.5% error trials, thus 87.5% correct performance, matching our empirical results. The same logic holds in the state condition which also has ≈87.5% performance. Thus, our overall estimate of memory capacity is 17.8 bits per item, where 13.8 bits are required to code the category of the object, and the additional four bits per item are required to code which exemplar (two bits) and state (two bits) the object is to successfully distinguish it from the foils. This suggests that maximum number of unique items that can be put in memory (assuming an optimal feature set) is $2^{17.8}$, or 228,000.

This model, in which the exemplar bits are separate from the category bits, is more conservative than giving unique codes without regard to category, since it accounts for the idea that we are more likely to confuse two teacups than a teacup and a tractor. However, our calculations do assume that the exemplar and state conditions draw on different bits, such that the information used to perform well in the exemplar tests is not the same as the information used for the state tests. This is compatible with memory representations in which it is possible to know that you saw an open door rather than a closed door (state condition) without knowing exactly which door it was (exemplar condition). However, even if up to half of the information was shared between the exemplar and state condition, we would still obtain an estimate of 16.8 bits of information, or 114,000 unique codes, still an order of magnitude over previous estimates.

A previous study by Hollingworth (2) also examined object representations on the order of hundreds of images. Participants were shown scenes with many embedded objects and were subsequently tested with exemplar-level foil items. To quantify the capacity of memory estimated in this experiment, we used the same hierarchical model. This study did not include a novel test condition to estimate the category bits, so we assumed a generous 99% performance, giving 14.27 bits. Performance in the exemplar level tests was 65%, or an additional 0.5 bits. Thus we estimate the memory capacity demonstrated in Hollingworth (2) between 14–15 bits, approximately equal to the estimates arrived at by Landauer (1). Interestingly, this study demonstrates that memory can store hundreds of objects with exemplar-level fidelity, and even this does not guarantee an increased estimate of memory capacity.

The hierarchical decision-level model is based on the number of items studied, and importantly, on the number of questions asked about each item. Previous large-scale memory studies only tested against a novel foil (one question). Here, we ask about novel, exemplar, and state comparisons (three questions, thus three sets of bits). Our estimate of memory capacity could be increased even further if more questions were asked about what kind of information observers have about remembered items. However, there is an important caution to this modeling approach: the questions asked about each item are probably not independent. For example, we could have included a fourth kind of test asking about the orientation of the presented object, and added more bits to the estimated code of each item. However, information about the state is probably also informative about orientation. Thus, asking 1,000 different questions to show an enormous memory capacity estimate is not sufficient because such questions will likely have overlapping information when considered under the true coding model of the visual system. On a positive note, if one could ask the right set of completely independent questions, one might be approximating the visual coding scheme.

1. Landauer TK (1986) How much do people remember? Some estimates of the quantity of learned information in long-term memory. *Cognit Sci* 10:477–493.
2. 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.

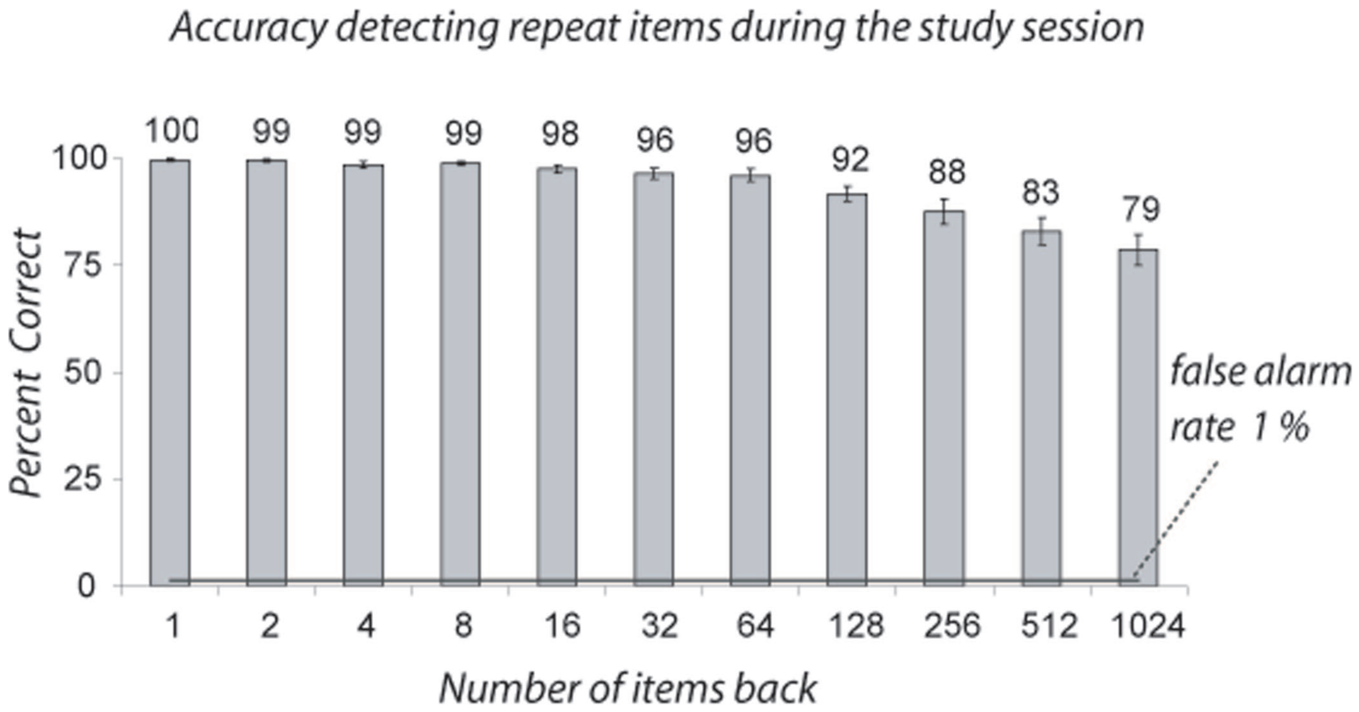

Fig. S1. Performance on detecting repeat images during the 5.5 h study session. Images were repeated with a different number of intervening items, from 0 to 1,023, by powers of 2.