Detecting changes in real-world objects: The relationship between visual long-term memory and change blindness

[Communicative & Integrative Biology 2:1, 1-3; January/February 2009]; ©2009 Landes Bioscience

Article Addendum

Detecting changes in real-world objects

The relationship between visual long-term memory and change blindness

Timothy F. Brady,[1] Talia Konkle,[1] Aude Oliva[1] and George A. Alvarez[2,*]

[1]Department of Brain and Cognitive Sciences; Massachusetts Institute of Technology; Cambridge, Massachusetts USA; [2]Department of Psychology; Harvard University; Cambridge, Massachusetts USA

Key words: object recognition, gist, fidelity, change blindness, change detection

A large body of literature has shown that observers often fail to notice significant changes in visual scenes, even when these changes happen right in front of their eyes. For instance, people often fail to notice if their conversation partner is switched to another person, or if large background objects suddenly disappear.[1,2] These 'change blindness' studies have led to the inference that the amount of information we remember about each item in a visual scene may be quite low.[1] However, in recent work we have demonstrated that long-term memory is capable of storing a massive number of visual objects with significant detail about each item.[3] In the present paper we attempt to reconcile these findings by demonstrating that observers do not experience 'change blindness' with the real world objects used in our previous experiment if they are given sufficient time to encode each item. The results reported here suggest that one of the major causes of change blindness for real-world objects is a lack of encoding time or attention to each object (see also refs. 4 and 5).

Introduction

One of the most well known phenomena in the study of visual memory is the remarkable failure of observers to detect what should be salient changes in visual scenes if detection of those changes depends on visual memory (change blindness: reviewed in ref. 6). Studies have shown that even after very brief storage intervals, large changes to images can go undetected, and that this change blindness can occur even in real life social situations.[1,2] These results have sometimes been taken to suggest that we maintain only a sparse representation of the world in visual memory.[7-9]

However, in a recently published experiment we showed that visual long-term memory is capable of storing a massive number of items with a large amount of detail per item;[3] (for other work showing storage of a large number of items in long-term memory (see refs. 10–12). In this experiment, participants viewed pictures of 2,500 categorically distinct objects one at a time for 3 seconds each, 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 (a cup you saw versus a clock you never saw), an object of the same basic level category (a cup you saw versus a similar but different cup), or the same object in a different state (the cup you saw empty versus the same cup with juice in it). Performance in each of these conditions was remarkably high (92%, 88% and 87%, respectively), suggesting participants successfully maintained detailed representations of thousands of images.

There are many possible reasons why our recent long-term memory experiment found evidence for detailed visual memory where other researchers have not. One counter intuitive possibility is that long-term memory represents items in greater detail than short-term memory. A more likely possibility is that differences between our long-term memory paradigm and a typical change blindness paradigm led to the difference in results. Here we sought to test this directly, examining the effect of encoding time on the ability to detect changes to real-world objects.

Method and Results

We presented observers (N = 6) with six real-world objects arrayed in two rows of three (Fig. 1). The objects were taken from the test pairs used in our previous study of long-term visual memory[3] (the images can be downloaded at http://cvcl.mit.edu/MM/). On each trial, an onscreen message informed viewers of how long the objects would appear on that trial (1.2, 6 or 18 seconds). Observers then pressed a key and the six objects appeared for the specified amount of time. Then the objects disappeared for 1 second, after which a single object reappeared and observers had to indicate whether it was the same exact object that had previously occupied that location. On half of the trials the object was exactly the same, and on the other half of the trials the object changed. We manipulated the similarity of the object that reappeared to the original object in the same manner as in our long-term memory experiment (using the stimuli from[3]): the image that reappeared could either be an entirely different object (from a *novel* category), a different *exemplar* of the same category, or the same exact object in a different *state* or pose.

*Correspondence to: George A. Alvarez; Harvard University; Department of Psychology; 33 Kirkland Street; 760 William James Hall; Cambridge, Massachusetts 02138 USA; Email: alvarez@wjh.harvard.edu

Submitted: 11/04/08; Accepted: 11/04/08

Previously published online as a *Communicative & Integrative Biology* E-publication: http://www.landesbioscience.com/journals/cib/article/7349

Addendum to: Brady TF, Konkle T, Alvarez GA, Oliva A. Visual long-term memory has a massive storage capacity for object details. Proc Natl Acad Sci USA 2008; 105:14325–9; PMID: 18787113; DOI: 10.1073/pnas.0803390105.

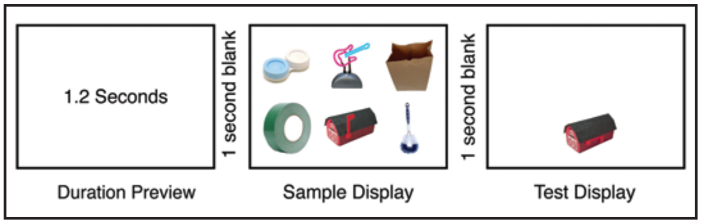

Figure 1. A screen initially displayed the duration the objects would be present for. This was followed by a brief blank, and then the six objects were presented. Afterwards, there was another brief blank and then a test object. Observers had to say whether the test object was the same or different than the object previously presented at that location.

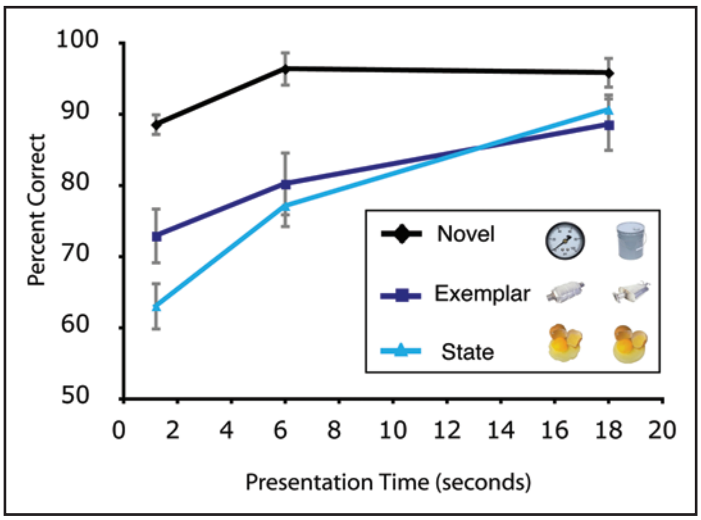

Figure 2. Percent correct for objects in the three change-type conditions (novel, exemplar-level change, state-level change) as a function of encoding durations (1.2 s, 6 s and 18 s). There were six objects presented, so the encoding durations correspond to 200 ms/object, 1 sec/object and 3 sec/object, respectively.

The three duration conditions correspond to 200 ms/item, 1 sec/item and 3 sec/item. The final condition matches the amount of time each object was presented in the long-term memory study.[3] In the present study we did not control how long people looked at each item in the display; observers could only use the duration preview at the beginning of each trial in order to allocate their encoding time for each object.

The results are plotted in Figure 2, showing percent correct for objects in the three change-type conditions (novel, exemplar, state change) at all three encoding durations (1.2 s, 6 s and 18 s). At the shortest duration, observers were able to detect changes of an object's category (*novel* condition) with almost 90% accuracy, significantly better than both exemplar-level changes at 73% ($t(5) = 4.84$, $p < 0.05$) and state-level changes at 63% ($t(5) = 8.33$, $p < 0.001$). With increasing presentation time, the performance for state and exemplar-level change detection also increases. At the longest presentation time, the novel changes are detected 96% of the time, with state and exemplar changes detected 89% and 91% of the time, respectively. (ANOVA: Main effect of condition: $F(2,10) = 44.77$, $p < 0.001$; Main effect of Encoding Time: $F(2,10) = 18.02$, $p < 0.001$; Condition x Encoding Time interaction $F(4,20) = 3.46$, $p < 0.05$). These accuracy levels match fairly well with the long-term memory study, in which the same novel, exemplar and state pairs were presented in a two-alternative forced choice (novel: 92%, exemplar and state: 87%). These results show that observers can successfully detect large changes (at the basic-level category) with only brief time to encode the display, but encoding all of the objects with sufficient detail to allow change detection at the level of exemplars or states took a considerable amount of time (as much as 3 seconds/item). Interestingly, these rates are much slower than typically found for simpler stimuli, where 100 ms of encoding time is often sufficient.[13] This suggests that the formation of detailed visual memories of real-world objects takes seconds to form, and factors other than encoding time limit change-detection for simple shapes (reviewed in ref. 14).

Discussion

Our data suggest that having sufficient time to encode information from each object is crucial to the formation of the detailed visual memories needed to support change detection. Thus, change blindness for real-world objects may sometimes occur because of a failure to sufficiently encode the details of objects in visual memory, preventing the later comparison of those objects when they reappear.

Many studies that have demonstrated poor memory for image details in the form of change blindness have either tested memory for non-salient background regions, or have tested memory for central objects when they were task irrelevant.[1,2] Thus, in line with the current results, it is possible that poor performance on such tasks does not reflect memory limitations, but the failure of observers to encode the relevant details in the first place. In addition, even if items are successfully encoded there may still be errors of comparison or persistence of the representation that cause change blindness (reviewed in ref. 6).

In further support of this encoding failure interpretation of change blindness, studies that have demonstrated better memory for image details have tested task-relevant and attended foreground objects and also have informed observers of exactly what details are relevant for the task (reviewed in refs. 5, 15 and 16). Hollingworth[16] demonstrated that when observers are required to attend to all of the items that might potentially change with the intention of remembering them, they could successfully detect changes to a large number of items from an image (i.e., change-blindness is attenuated; reviewed in ref. 4). The current results support this interpretation that change blindness may result from a failure to encode the items into visual memory in the first place.

These results also highlight an important aspect of change detection paradigms: memory performance will depend not only on the duration of encoding, but also on the difficulty of the subsequent test. For example, the memory performance data in the present experiment could be converted into capacity estimates (e.g., using Cowan's K[17]) but this analysis would give three different answers to how many objects observers could hold in memory depending on whether the novel, exemplar, or state change performance was used. The variability in K as a function of the test type suggests that the right units for quantifying the capacity of visual short term memory (or visual memory on any time scale), are not in number of objects, but in an alternate measure that takes into account both the number of items represented and also the precision of each item's representation.[18-20]

Acknowledgements

This work was partly funded by a National Science Foundation Graduate Fellowship (to Timothy F. Brady), a National Defense Science and Engineering Graduate Fellowship (to Talia Konkle), a National Research Service Award Fellowship F32-EY016982 (to George A. Alvarez), and a National Science Foundation (NSF) Career Award IIS-0546262 and NSF Grant IIS-0705677 (to Aude Oliva).

References

1. Simons DJ, Levin DT. Change blindness. Trends Cogn Sci 1997; 1:261-7.
2. Rensink RA, O'Regan JK, Clark JJ. To see or not to see: The need for attention to perceive changes in scenes. Psychol Sci 1997; 8:368-73.
3. Brady TF, Konkle T, Alvarez GA, Oliva A. Visual long-term memory has a massive storage capacity for object details. Proc Nat Acad Sci USA 2008; 105:14325-9.
4. Hollingworth A. Visual memory for natural scenes: Evidence from change detection and visual search. Vis Cogn 2006; 14:781-807.
5. Scholl BJ. Attenuated change blindness for exogenously attended items in a flicker paradigm. Vis Cogn 2000; 7:377-96.
6. Simons DJ, Rensink RA. Change blindness: Past, present and future. Trends Cogn Sci 2005; 9:16-20.
7. O'Regan JK. Solving the "real" mysteries of visual perception: the world as an outside memory. Can J Psychol 1992; 46:461-88.
8. O'Regan JK, Noë A. A sensorimotor account of vision and visual consciousness. Behav Brain Sci 2001; 24:939-1011.
9. Rensink RA. The dynamic representation of scenes. Vis Cogn 2000; 7:17-42.
10. Standing L. Learning 10,000 pictures. Q J Exp Psychol 1973; 25:207-22.
11. Shepard RN. Recognition memory for words, sentences and pictures. J Verb Learn Verb Behav 1967; 6:156-63.
12. Standing L, Conezio J, Haber RN. Perception and memory for pictures: Single-trial learning of 2,500 visual stimuli. Psychon Sci 1970; 19:73-4.
13. Luck SJ, Vogel EK. The capacity of visual working memory for features and conjunctions. Nature 1997; 390:279-81.
14. Alvarez GA, Thompson TW. Overwriting and rebinding: why feature-switch detection tasks underestimate the binding capacity of visual working memory. Visual Cognition 2008; In press. doi: 10.1080/13506280802265496
15. Hollingworth A, Henderson JM. Accurate visual memory for previously attended objects in natural scenes. J Exp Psychol: Hum Percept Perform 2002; 28:113-36.
16. Hollingworth A. Constructing visual representations of natural scenes: The roles of short- and long-term visual memory. J Exp Psychol Hum Percept Perform 2004; 30:519-37.
17. Cowan N. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. Behav Brain Sci 2001; 24:87-185.
18. Awh E, Barton B, Vogel EK. Visual working memory represents a fixed number of items, regardless of complexity. Psychol Sci 2007; 18:622-8.
19. Alvarez GA, Cavanagh P. The capacity of visual short-term memory is set both by visual information load and by number of objects. Psychol Sci 2004; 15:106-11.
20. Brady TF, Konkle T, Alvarez GA. Compression in visual working memory: Using statistical regularities to form more efficient memory representations. Submitted.