American Psychological Association

Journal of Experimental Psychology: General

0096-3445/20/$12.00

2020, Vol. 149, No. 3, 530–549
http://dx.doi.org/10.1037/xge0000664

Independent Storage of Different Features of Real-World Objects in Long-Term Memory

Igor S. Utochkin
National Research University Higher School of Economics, Moscow, Russia

Timothy F. Brady
University of California, San Diego

People can store thousands of real-world objects in visual long-term memory with high precision. But are these objects stored as unitary, bound entities, as often assumed, or as bundles of separable features? We tested this in several experiments. In the first series of studies, participants were instructed to remember specific exemplars of real-world objects presented in a particular state (e.g., open/closed, full/empty, etc.), and then were asked to recognize either which exemplars they had seen (e.g., I saw this coffee mug), or which exemplar-state conjunctions they had seen (e.g., I saw this coffee mug and it was full). Participants had a large number of within-category confusions, for example misremembering which states went with which exemplars, while simultaneously showing strong memory for the features themselves (e.g., which states they had seen, which exemplars they had seen). In a second series of studies, we found further evidence of independence: participants were very good at remembering which exemplars they had seen independently of whether these items were presented in a new or old state, but the same did not occur for features known to be truly holistically represented. Thus, we find through 2 lines of evidence that the features of real-world objects that support exemplar discrimination and state discrimination are not bound, suggesting visual objects are not inherently unitary entities in memory.

Keywords: visual memory, long-term memory, feature binding, binding errors

Long-term memory is constructive (Bartlett, 1932): People take pieces of information and recombine them to arrive at a memory for a complex event. However, the fundamental "units" of visual memories remain unknown. That is, people will mistakenly remember a stop sign for a yield sign (Loftus, Miller, & Burns, 1978), but would they misremember a combined object with one feature of the stop sign and one of the yield sign? It seems natural to suppose that objects would be a fundamental "unit" of representation, as objects are fundamental to attention and working memory (Scholl, 2001). However, the extent to which distinct features of an object are stored in a bound, unitized object representation (e.g., I either remember or do not remember an entire object) versus different object features stored independently (I remember particular features of many objects, but not all the features of any of them) remains an important open question. This question of whether objects are a fundamental bound unit has significant repercussions for many important issues that cut across perception, working memory, and long-term memory.

For example, in the literature on long-term memory, there is often a distinction made between familiarity—a kind of holistic item memory— and recollection, or memory for the episodic details and context of an item (Diana, Yonelinas, & Ranganath, 2007; Ranganath et al., 2004). This distinction implicitly treats objects as unitary, where familiarity processes operate over object representations that do not require any binding, while recollection processes help retrieve information about how objects are bound to their contexts. Similarly, the cognitive neuroscience literature often investigates claims that hippocampus or medial temporal lobe structures are used for binding of objects to contexts, and this literature often implicitly treats object-only memory as not requiring binding (Davachi, 2006; Olsen et al., 2015). Finally, the literature on how we perceptually recognize objects—object recognition—also depends on the extent to which objects are recognized based on holistic viewpoint

This article was published Online First August 15, 2019.

Igor S. Utochkin, Department of Psychology, National Research University Higher School of Economics, Moscow, Russia; Timothy F. Brady, Department of Psychology, University of California, San Diego.

The study was conducted within the Basic Research Program of the National Research University Higher School of Economics in 2017 (TZ-59) and in 2018 (TZ-57). We thank Natalia Tiurina, Yuri Markov, and Vladislav Khvostov for their assistance in data collection. Parts of this work were previously presented at the Vision Sciences Society in 2017, and a preprint of this article was posted on PsyArXiv https://doi.org/10.31234/osf.io/8b64k before submission. Igor S. Utochkin and Timothy F. Brady designed the experiments. Igor S. Utochkin made experimental scripts and ran Experiments 1A–B, 2A, and A1. Timothy F. Brady made experimental scripts and ran Experiments 2B–C. Igor S. Utochkin and Timothy F. Brady analyzed data and wrote the article. Experimental data can be publicly accessed via Open Science Framework at https://osf.io/mgxf5/ and the object stimuli are available from Timothy F. Brady's website, https://bradylab.ucsd.edu/stimuli.html.

Correspondence concerning this article should be addressed to Igor S. Utochkin, Department of Psychology, National Research University Higher School of Economics, 101000, Armyansky per., 4, Moscow, Russia. E-mail: isutochkin@inbox.ru

images versus stored in part-based manners, which again is ultimately about the issue of whether objects are fundamentally unitized (Gauthier & Tarr, 2016; Hollingworth, 2006). Thus, the question of whether visual objects are actually represented as holistic units, or whether different properties or dimensions of objects are represented separably, is a core question about object representation and long-term memory.

Binding of Object Features in Working Memory

This question of whether objects are unitary has been the source of a large amount of research in perception (e.g., Di Lollo, 2012; Treisman, 1996; Wolfe & Cave, 1999) and visual working memory (e.g., Fougnie, Cormiea, & Alvarez, 2013; Luck & Vogel, 1997), although this work has almost exclusively focused on memory for arbitrary combinations of simple visual features (e.g., random colors and orientations). Initial evidence from visual working memory argued that people are as efficient at storing 16 features of four objects (four features per object) as at storing only four features of the same number of objects (one feature per object), and thus memories must be stored as integrated objects (Luck & Vogel, 1997)—even in cases where the conjunctions are ultimately arbitrary and change randomly trial to trial. However, later studies have shown that while there is some benefit to object-based storage in these situations (Fougnie et al., 2013; Fougnie, Asplund, & Marois, 2010; Olson & Jiang, 2002; Wheeler & Treisman, 2002), the probability of a poor memory for one feature of an object is not correlated with the probability of a poor memory for another feature of the same object (Fougnie & Alvarez, 2011), and participants in fact perform considerably worse on each feature when asked to store more features per object, contrary to the claims of Luck and Vogel (1997; e.g., Cowan, Blume, & Saults, 2013; Fougnie et al., 2010; Hardman & Cowan, 2015; Oberauer & Eichenberger, 2013). Thus, in the literature on visual working memory, focused largely on memory for simple features (e.g., color and orientation) rather than realistic objects, it has become clear that binding is an effortful process and objects are not automatically stored in a fully bound, unitized format. However, it remains largely unknown how this generalizes to long-term memory for complex and meaningful images. Are objects stored as holistic, units under these more real-world conditions?

Binding of Object Features in Long-Term Memory

Existing work on binding in long-term memory has examined memory for basic features like color (e.g., Chalfonte & Johnson, 1996; Hicks & Starns, 2015) and for words (e.g., Reinitz, Lammers, & Cochran, 1992; Underwood, Kapelak, & Malmi, 1976) suggesting the possibility of independent storage and recombination of component features even of individual objects, though in these cases the features are again arbitrarily rather than semantically related (e.g., the color of a word). A priori it seems like realistic objects would be much more likely to be holistically represented, and thus form a basic unit of memories. For example, whether a coffee mug is full or empty is a semantically meaningful distinction, rather than an arbitrary conjunction of features (like the color of a word), and we may well have existing representations of "full coffee mugs" that can be brought to bear on forming a new memory of such an object. For such meaningful objects (a full coffee mug), do people actually encode the features separately? Or do they store objects as unitized, fully bound units in memory?

Some work has suggested that more realistic objects may also be represented as independent features that can be recombined in memory (Brady, Konkle, Alvarez, & Oliva, 2013; Reinitz et al., 1992). For example, Reinitz et al. (1992) show that participants are more likely to false alarm to schematic faces made up of familiar parts, suggesting the possibility of binding errors in more realistic stimuli. However, their data is also consistent with a combination of bound memories and separate independent feature memory, and their stimuli featured fully spatially separable features, so it is unclear how well they generalize to realistic objects. Brady, Konkle, Alvarez, et al. (2013) looked at the case of semantically meaningful features of objects (e.g., which cup of orange juice you saw and whether it was full or empty), and showed that different features of the same real-world object can be forgotten at different rates, and that forgetting one feature does not necessarily entail forgetting another. These findings provided further evidence that real-world objects might be stored as a set of unbound features that are subject to independent forgetting, and thus possibly open to recombination in memory. In particular, Brady, Konkle, Alvarez, et al.'s (2013) data provide evidence for the idea that features of objects may be, at minimum, forgotten independently and thus are not completely holistically stored.

By contrast to these studies, a significant amount of work on the cognitive neuroscience of long-term memory has argued that a central feature of our long-term memory system is unitization—for example, that incoming features must be combined in such a way that similar experiences yield fundamentally distinct neural representations in long-term memory ("pattern separation"; Yassa & Stark, 2011). In the medial temporal lobe as well as more ventral visual regions, it is often found that objects (Erez, Cusack, Kendall, & Barense, 2016) or structured scenes or events (e.g., van den Honert, McCarthy, & Johnson, 2017) are represented holistically: That is, that brain responses cannot be explained by the sum of the component stimuli or features alone (Erez et al., 2016). This is particularly the focus of work arguing that parts of the medial temporal lobe—crucial for long-term memory—are specifically designed to perform pattern separation in order to make even similar remembered items as distinct as possible (e.g., Diana et al., 2007; Norman & O'Reilly, 2003; Yassa & Stark, 2011). This has been used to argue that a central feature of building more complex object and scene representations—and holding them in memory—is a holistic representation (e.g., van den Honert et al., 2017) that does not rely simply on the similarity of the underlying feature representations but goes beyond these to novel, unitized representations design to prevent confusions of similar items.

Thus, there is existing evidence suggesting that objects are not represented entirely holistically (e.g., Brady, Konkle, Alvarez, et al., 2013), and at minimum distinct features can be forgotten separately. However, there is also a long tradition of work—from object recognition to long-term memory—arguing for unitized, bound object representations, where objects differing in even a single feature from each other are stored using entirely unique and unrelated representations.

The Current Studies

In the current set of studies, we test the specific idea that different features of real-world objects are stored independently in memory. In the first set of studies, we focus on the idea that if some important object features are stored independently, rather than holistically, storing multiple items from a category should cause confusions or interference between the features of those items. In other words, if you see two coffee mugs and one is full and one is empty, this should make for a difficult binding problem—which was full and which was empty? Models of memory where objects are stored as fully unitized, predict that there is no difficulty in storing related items with different features, as these are separated into distinct holistic representations. In the second set of studies, we focus on generalization. In particular, if you are asked which cabinet you saw, and it changed states in the meantime (was open and is now closed), how does this impact your memory for the particular cabinet? If memory is holistic for all of the object features, and similar items are "pattern separated" to have new, distinct representations, generalizing over such a change should be quite difficult. If different aspects of a stimulus are stored independently, this generalization should be straightforward.

To test this set of predictions, we need to define possibly separable "properties" of real-world objects. To do so, we rely on previous work showing that the visual or semantic features people use to recognize which "state" or "pose" an object is in (e.g., was the cup full or empty; was the cabinet open or closed) tend to be forgotten independently of the features people use to distinguish which exemplar of a category they have seen (which cabinet did I see, which mug did I see; Brady, Konkle, Alvarez, et al., 2013). Note that both object "state" and "exemplar" properties are likely quite complex and different kinds of "state" changes (i.e., different ways the pose or configuration of an object could be changed) may rely on different visual or semantic features. However, distinguishing between two different states or poses of an object, and between two different exemplars of the same object category are two common and important tasks that we perform every day in the context of real-world objects, and these distinctions have frequently been used in the literature on visual long-term memory (e.g., Brady, Konkle, Alvarez, & Oliva, 2008, 2013; Cunningham, Yassa, & Egeth, 2015). Thus, in the present work we focus on these dimensions and ask whether these two aspects of objects are stored in a unitized format. Note that any number of relationship between these features are possible: They could be stored in a unitized format; as completely independent aspects of an object; in a partially bound state; or even in a hierarchical format, where remembering exemplar information is necessary for accessing object state but object state can be lost independently of exemplar information. The goal of our experiments was to examine whether these aspects of objects were unitized or stored in some manner that allows them to be stored and accessed independently of each other.

In Experiment 1A and 1B, we find that even when independently manipulated properties of a pair of objects are both successfully stored in memory, participants frequently show interference or confusion between the properties of objects from the same category. This suggests that even semantically meaningful objects in long-term memory are stored in at least a partially unbound manner, with separate features of their appearance stored and accessed separately. In Experiments 2A, 2B, and 2C, we find that changes in one property (e.g., state) cause little impairment in memory for the other property (e.g., exemplar). By contrast, for features known to be holistically represented (integral features), we find such generalization is quite difficult. This again argues strongly against a unitized, bound storage of objects. Together, our data suggest that real-world objects are not inherently stored in a bound, unitized format; but that instead at least some object features are stored independently.

Experiment 1A

As our first test of whether realistic objects are stored as holistic units in long-term memory, we used an exemplar-state memory task. We presented two exemplars from a set of object categories (e.g., Box A and Box B). We manipulated their states so that both exemplars could be in the same state (e.g., two open boxes) or in different states (Box A is open and Box B is closed). Participants subsequently had to recognize the states of each exemplar. As noted above, both object "state" and "exemplar" properties are likely quite complex and different kinds of "state" changes (i.e., different ways the pose or configuration of an object could be changed) may rely on different visual or semantic features. The current study asks whether the features that support these state versus exemplar discrimination tasks are holistic or independent—that is, for a given object, if you know the state that went with it, do you also know which exemplar it was?

If the features supporting these discriminations are distinct and independent, then ascribing correct states to exemplars should be difficult if the exemplars are shown in different states, as reporting each item's state would then require binding, potentially leading to errors. By contrast, if the features underlying exemplar and state discriminations are stored in a holistic object representation (an "integrated, holistic object" account), then participants should perform equally well when the two exemplars are shown in the same state as when they are shown in different states, as these similar within-category items would be "pattern separated" to create completely unique long-term memory traces.

Because the logic of this exemplar-state task depends only on testing state memory, we do not assess exemplar memory in the main task. Thus, we also include a control task designed to ensure that participants successfully remember the exemplars in this task. In this Exemplar task, participants study two exemplars of each category but are then asked to distinguish these exemplars from new exemplars of the same category. This allows us to distinguish between difficulty with maintaining the distinction between multiple features within a category and poor memory for exemplars.

Method

Participants. For determining the sample size, we used the statistical tool G*Power 3.0.10 (Faul, Erdfelder, Lang, & Buchner, 2007). Our sample sizes were based on those used by Brady et al. (2013) and on effect size estimates reported in the same study. In their exemplar-state task, Brady et al. (2013) used 15 participants per group. Taking into account a possibility of technical problems or poor performance in some participants, we recruited 20 participants. With this sample size, required statistical power set at .8, and a Type I error set at .05 (two-tailed t test), effect sizes with Cohen's d (for one-sample comparisons) and d_z (for paired-sample comparisons) of at least .67 could be detected, smaller than the effects reported by Brady et al. (2013). Students of the Higher School of Economics (18 female; age: M = 19.65 years, SD = 1.06) took part in the experiment for course credit or for a compensation equivalent to approximately $3. All participants reported having normal or corrected-to-normal vision and no neurological problems. Before the beginning of the experiment, they provided written informed consent.

Apparatus and stimuli. Stimulation was developed and presented through PsychoPy (Peirce, 2007) for Linux. Stimuli were presented on a standard VGA monitor (75 Hz at 1,024 × 728 resolution) on a homogeneous white field. Participants were seated approximately 57 cm from the monitor. From that distance, each item subtended approximately 10.5 degrees of visual angle.

Two image sets were used in the experiment. The first one was the set of 480 images used in Brady et al. (2013), Experiment 2 (available from https://bradylab.ucsd.edu/stimuli/StateExemplar.zip). It contained 120 unique object categories, with each category containing two exemplars (e.g., two different coffee mugs) in each of two different states (each mug empty and filled). This stimulus set was used in the exemplar-state memory task of the experiment (see below). State instances varied widely across categories and could involve image transformations across a range of scales, including local details (e.g., different positions of hands of a clock), middle-size details (e.g., open vs. closed cover of a toolkit), or the entire object (e.g., whole vs. cut cabbage). This variety of states made it unpredictable which state features are diagnostic for subsequent memory test and discouraged participants from focusing on specific features during encoding (unlike Reinitz et al.'s 1992 procedure). The second stimulus set was the object exemplar set used in the study by Konkle, Brady, Alvarez, and Oliva (2010a; available from https://bradylab.ucsd.edu/stimuli/ObjectCategories.zip). From that set, we selected 120 new categories not overlapping with the categories from the first set. From each category, we took four exemplars having no state variation. We used these images in the exemplar memory task (see below).

Procedure. The experiment consisted of exemplar-state and exemplar tasks. Each participant was exposed to both tasks, as the two tasks used nonoverlapping sets of objects. The order of the tasks was counterbalanced across participants. In addition, a practice task that was a shortened version of the exemplar task with seven categories (not used in the main experiment) preceded the main tasks. We did not have participants practice with the exemplar-state task as we did not want people to strategically and/or verbally encode exactly the relevant information about exemplar-state conjunctions; we were interested in how objects are encoded when people's task is to remember all the details of the object as best as possible.

Exemplar-state task. The task consisted of a study phase followed by a test phase. In the study phase, participants were shown 240 images at a rate of one image per 2 s. Each image was presented at the center of the screen for 1 s followed by a 1-s blank interval. Participants were instructed to memorize the images. Both the presentation rate and the instructions excluded any possibility of binding errors occurring at the perceptual level, as perceptual binding errors are known to only occur when several objects are presented briefly either under diverted attention (Treisman & Schmidt, 1982; Wolfe & Cave, 1999) or peripherally (Rosenholtz, Huang, & Ehinger, 2012). Each object category was represented by two different exemplars that appeared at a random time in the stream of images (e.g., could be separated by a random number of other images) to discourage participants from encoding the two objects from the same category relative to each other. The two exemplars from a given category could be shown in the same state (e.g., Toolboxes A and B, both open, Figure 1B) or two different states (e.g., Coffee Mug A is empty and Coffee Mug B is filled, Figure 1B). Whether an observer would be exposed to objects A and B in same or different states was varied across participants so that each category had approximately the same proportion of being shown in same or different states across the entire experiment. The variation of the states was very wide across categories (Figure 1, as a whole, provides an example of this variety), so that participants could not reliably know which particular features define a to-be-tested difference. Moreover, the participants were instructed to memorize the items in as much detail as possible. No exemplars or states were mentioned before the study phase, encouraging participants to memorize objects in their overall appearance, without special emphasis on either exemplar or state information. In the test phase, on each trial, both exemplars in both possible states were presented simultaneously. The participants had to choose one and only one correct state for each exemplar—two choices in total (each a two-alternative forced choice, 2-AFC); two different states of one exemplar could never be chosen. Even though this is effectively two different 2-AFC trials, by presenting both exemplars at once we reduce the possibility that swap errors arise simply because participants retrieve the wrong exemplar.

Exemplar task. The exemplar task was designed to ensure that participants successfully remembered information about which exemplars from a category they had seen in a similar encoding situation to the exemplar-state task, because we do not assess exemplar memory per se in the main task. In the study phase of the exemplar task, the participants were shown 120 categories, nonoverlapping with the exemplar-state task, with two exemplars shown from each category (Figure 1A). The studied list, therefore, had the same length as in exemplar-state task. In each trial of the test phase, participants were shown four exemplars of one category and had to choose the two old exemplars. They were again restricted to choosing one item from each "row" (see Figure 1A), following the same design as the Exemplar-State task (two simultaneous 2-AFC trials).

Design and analysis. The exemplar-state task consisted of two conditions: Categories where the items were presented in the same state versus categories where the items were presented in different states; thus, 60 categories were presented and tested within each condition. In the exemplar task, no states were varied, so only 60 categories out of 120 were tested to equate the number of trials and data points obtained from both tasks. From the perspective of a single participant the 60 tested categories were randomly chosen, but they were counterbalanced so that all categories were equally likely to be tested across participants.

We estimated the overall accuracy (total number of correctly chosen items) in both exemplar-state and exemplar tasks; we also compared these accuracies with chance level to test whether people remember states and/or exemplars. This allows us to measure memory for exemplars and states overall. To estimate memory for

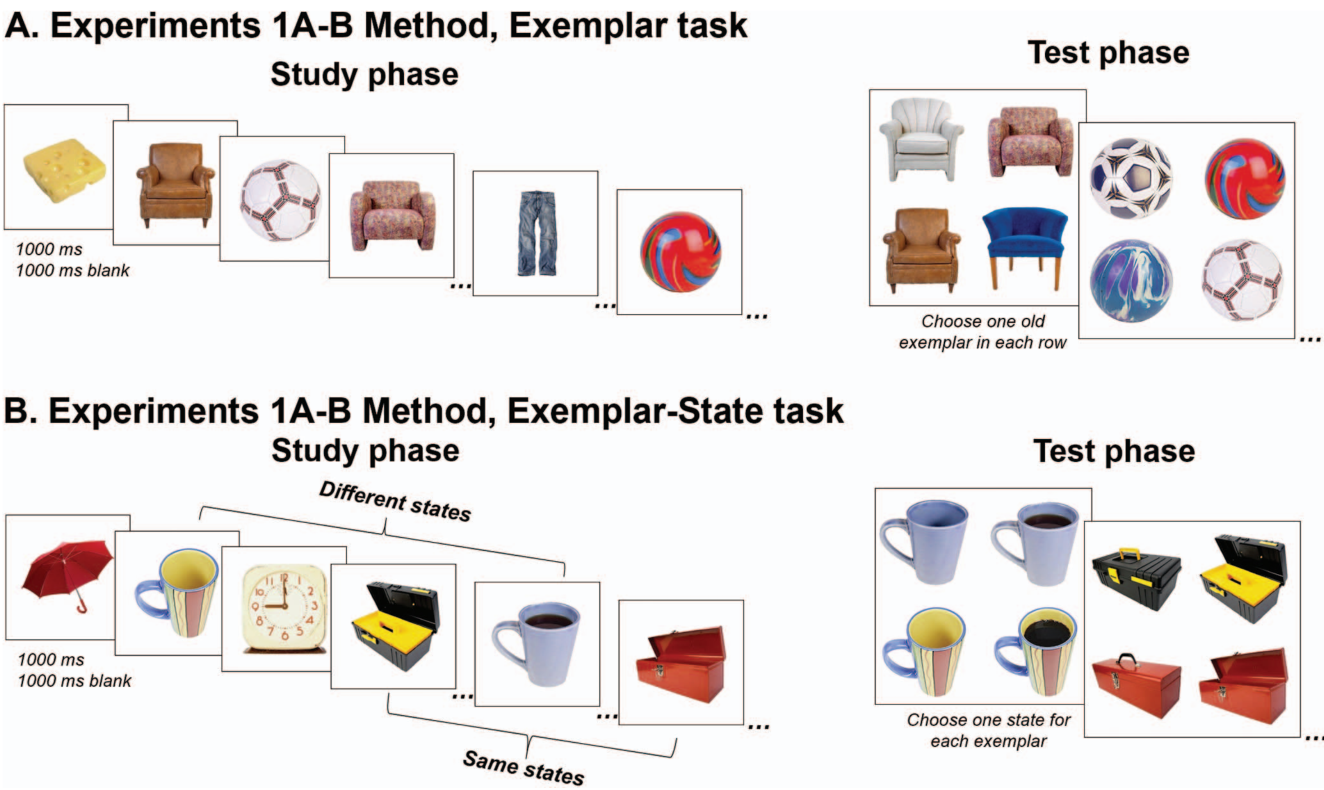

Figure 1. Study and test phases of Experiments 1A and 1B. (A) In the exemplar task, participants were instructed to remember a sequence of objects, including two object exemplars from each category. At test, they performed two simultaneous 2-AFC tasks; they had to pick the old exemplars in each row (they could choose only a single item in each row). (B) In the exemplar-state task, participants were instructed to remember a sequence of objects, including two objects from each category. The objects from the same category could be either in the same or different state (e.g., open vs. closed). At test, they again responded to two simultaneous 2-AFC tasks; they had to select correct state for each exemplar (i.e., choose a single item in each row). See the online article for the color version of this figure.

states, we measure whether participants know if the two exemplars they saw from the category had the same state or were in different states. In particular, we compare how often participants select the same state for both exemplars (in the same-state and different-state conditions).

Finally, to measure boundedness between state and exemplar memories, we estimated how often the reported states are correct for each exemplar in the exemplar-state task. If participants remember both a significant amount of information about the exemplars and about the states, then this task is a test of whether this information is bound or unbound. If it is fully bound (e.g., unitized), participants should pick the correct state for each exemplar as frequently as they remember the state information overall, whereas if it is unbound, participants should be much less accurate at this task when the states they studied each object in were different, as this condition requires successful binding.

To estimate the statistical effects of our manipulations, we performed a series of standard and Bayesian t tests. The latter is considered an alternative to classical null-hypothesis significance tests and based on evaluation of relative observed evidence for the hypothesis H_1 compared with the null hypothesis, H_0 (Bayes factor, BF_{10}; Rouder, Speckman, Sun, Morey, & Iverson, 2009). We used JASP 0.9.0.1 software (JASP Team, 2017) to run the Bayesian analysis. We interpreted Bayes factors using Jeffreys' (1961) classical scale, where $BF_{10} = 1$ shows no evidence in favor of any hypothesis, $1 < BF_{10} < 3$ is considered weak evidence in favor of H_1, $3 < BF_{10} < 10$ is considered some evidence in favor of H_1, $10 < BF_{10} < 30$ is considered strong evidence in favor of H_1, and $BF_{10} > 30$ is considered very strong evidence in favor of H_1. The inverse ratios are considered to reflect corresponding degrees of evidence in favor of H_0. In addition, we reported Cohen's d's and d_z's as estimates of effect sizes with 95% confidence intervals (CI's).

Results

Accuracy in remembering exemplars. In the Exemplar condition, when asked to remember exemplars without requiring state memory, participants performed well above chance, $M = .80$, $t(19) = 12.01$, $p < .001$, $BF_{10} = 3.80 \times 10^7$, $d = 2.68$, 95% CI [1.72, 3.63], suggesting they remembered exemplar information well when it was not required to be bound to state information.

Accuracy in remembering state. To determine how well participants remembered state information on its own, we can examine both performance in picking the correct state when both objects were shown in the same state as well as how well they discriminated between conditions where the two exemplars seen were in the same state versus in different states. When both items in a category were shown in the same state, participants were well above chance at choosing this state, $M = .74$, $t(19) = 9.05$, $p < .001$, $BF_{10} = 5.19 \times 10^5$, $d = 2.02$, 95% CI [1.24, 2.78], suggesting that participants were good at remembering states when binding was not required (Figure 2A).

In addition, participants were good at discriminating between the condition where the items were shown in the same state versus in a different state. In particular, the proportion of the time par-

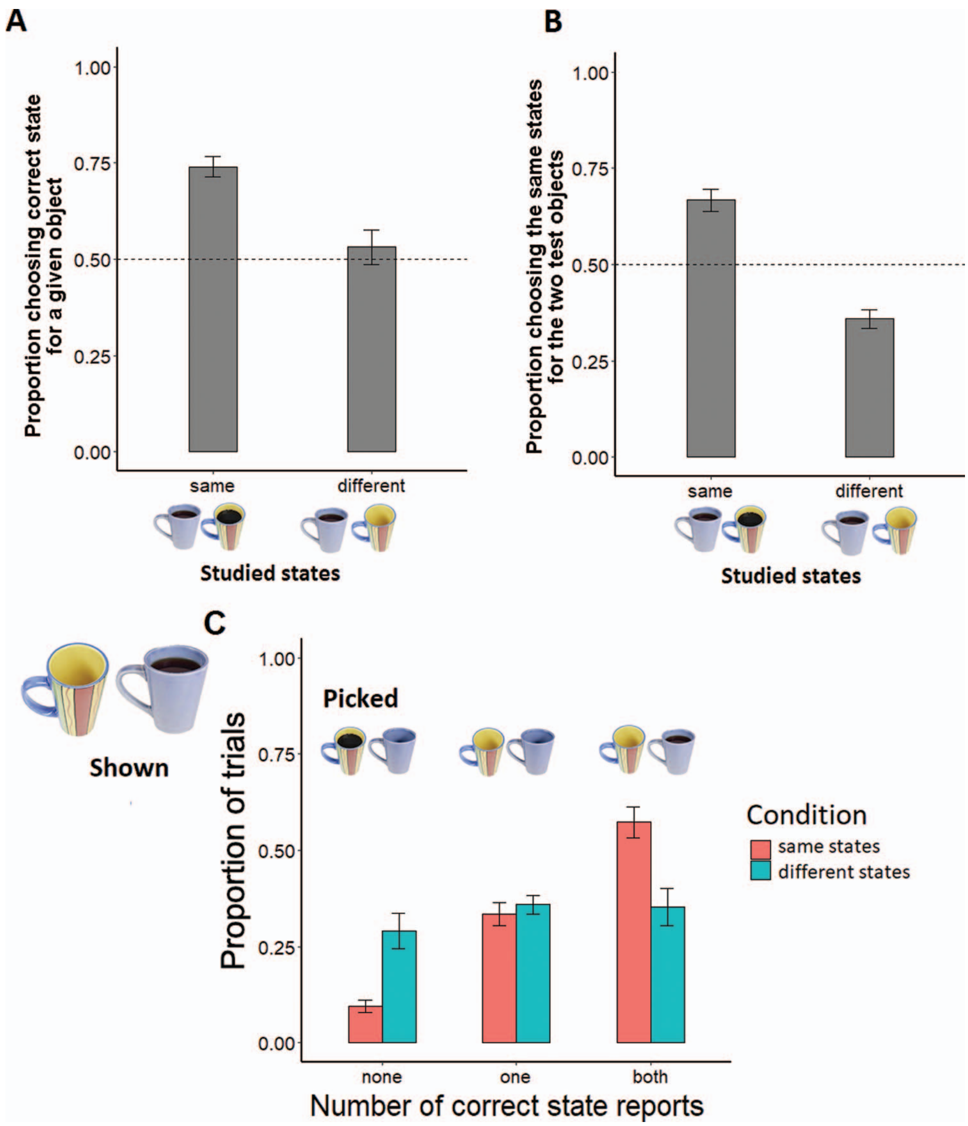

Figure 2. Results of Experiment 1A for exemplar-state task: (A) proportions choosing the correct state for a given exemplar when the two studied objects were shown in the same state (left; doesn't require binding) or different states (right; requires binding). While participants were above chance at knowing the state of objects that had been seen in the same state as each other, they were at chance when the objects differed in state; (B) proportions choosing the same states for the two test objects (regardless of whether these states are correct or incorrect). While participants were at chance (in A) in the different state condition, this was not because they did not know the states were different—participants reliably picked two different states when the states were in fact different and two same states when the states were in fact the same. Dashed lines show chance levels. (C) Breakdown of (A) into the proportion of trials where participants report both items correctly, one item correctly, or none of the items correctly as a function of the study condition. The difference between the "same state" and "different states" condition arises almost entirely from the percentage of times participants get both items correct. Error bars denote standard errors of the mean (*SEM*) in all panels. See the online article for the color version of this figure.

ticipants selected the same two states for the two exemplars was much higher for the items that actually were presented in same states as compared to presented in different states, $t(19) = 8.02$, $p < .001$, $BF_{10} = 9.59 \times 10^4$, $d_z = 1.79$, 95% CI [1.07, 2.50]. In both conditions, the proportions differed from chance level .50 and these differences were almost symmetrical (same states: $M = .67$, $t(19) = 5.61$, $p < .001$, $BF_{10} = 1{,}137$, $d = 1.25$, 95% CI [.65, 1.83]; different states: $M = .36$, $t(19) = 5.77$, $p < .001$, $BF_{10} = 1{,}553$, $d = 1.29$, 95% CI [.68, 1.87]; Figure 2B). This suggests that the participants remembered whether the exemplars within a certain category were presented in the same or in two different states, providing evidence that not only did participants remember exemplar information, but also state information.

Accuracy in exemplar-state binding. Because participants accurately remembered which exemplars they had seen and which states they had seen in a given category, we can address our main question of whether these two kinds of information were bound or unbound. In particular, we can ask how often participants correctly remembered the state of each exemplar when the items were shown in different states. Participants were significantly worse at remembering the state of each exemplar when the exemplars were shown in different states ($M = .53$) than in the same state ($M = 0.74$; comparison: $t(19) = 4.88$, $p < .001$, $BF_{10} = 271$, $d_z = 1.09$, 95% CI [.52, 1.64]; Figure 2A). In fact, memory in the different state condition did not differ from chance level (0.50; one-sample $t(19) = .68$, $p = .506$, $BF_{10} = .285$, $d = .15$, 95% CI [−.29, .59]); as compared with the same state condition, reported above, which was significantly better than chance, $t(19) = 9.05$, $p < .001$, $BF_{10} = 5.19 \times 10^5$, $d = 2.02$, 95% CI [1.24, 2.78]. This proportion included only .36 trials provided by choosing one out of two correct states, which did not differ from the proportion of choosing one correct answer in the same state condition, $M = .33$, $t(19) = .65$, $p = .526$, $BF_{10} = .280$, $d_z = .145$, 95% CI [−.298, .583] (see Figure 2C). Therefore, the difference in overall accuracy between the same state and the different state was provided by an ability to report either both or none of the states correctly.

Thus, overall, we find that people know whether the states of the objects within a category were the same or different, but are at chance in saying which state goes with which exemplar. This is strong evidence of independent storage of the features underlying state and exemplar discriminations.

Experiment 1B

To test whether the evidence for independent storage found in Experiment 1A are due to visual rather than verbal encoding, we replicated the procedure of Experiment 1A with an additional verbal interference task.

Method

Participants. Following the power analysis described in Experiment 1A, a different set of 19 psychology students of the Higher School of Economics (15 female; age: $M = 21.05$ years, $SD = 3.38$) took part in the experiment for extra course credits or for a compensation equivalent to approximately $3. All participants reported having normal or corrected-to-normal vision and no neurological problems. Before the beginning of the experiment, they provided written informed consent.

Apparatus, stimuli, and procedure. Apparatus and stimuli were the same as in Experiment 1A. The procedure also was the same with an important addition that, during the Study phases of both the exemplar task and the exemplar-state tasks, participants were instructed to repeat a syllable "ba" aloud at a rate of about 3 Hz to discourage verbal encoding of stimuli. An experimenter monitored whether the participants followed the instruction to repeat the syllable.

Design and analysis were the same as in Experiment 1A.

Results

Accuracy in remembering exemplars. In the exemplar condition, accuracy was high and differed substantially from chance, $M = .78$, $t(18) = 13.92$, $p < .001$, $BF_{10} = 1.92 \times 10^8$, $d = 3.19$, 95% CI [2.06, 4.31], suggesting they remembered exemplar information. Performance was approximately at the same level as in Experiment 1A, $t(18) = .579$, $p = .566$, $BF_{10} = .365$, $d = .185$, 95% CI [−.45, .81].

Accuracy in remembering state. As in Experiment 1A, we ran two tests to estimate how well participants remembered state information on its own. When both items in a category were shown in the same state and, hence, no binding was required, participants were well above chance at choosing this state, $M = .70$, $t(18) = 9.69$, $p < .001$, $BF_{10} = 8.82 \times 10^5$, $d = 2.23$, 95% CI [1.36, 3.06] (see Figure 3A).

Also, participants were able to discriminate between the condition where the items were shown in the same state versus in a different state. The participants chose the same states for the two exemplars more frequently when the two exemplars were actually presented in the same state than when they were actually presented in different states (same states: $M = .58$; different states: $M = .40$; comparison: $t(18) = 5.33$, $p < .001$, $BF_{10} = 567$, $d_z = 1.22$, 95% CI [.61, 1.81]; Figure 3B). This finding replicates the corresponding result from Experiment 1A.

Accuracy in exemplar-state binding. Given good memory for exemplars and states independently, we now turn to the question whether the features that support these discriminations are remembered in a bound representation. Again, the pattern was quite similar to that obtained in Experiment 1A. Participants were near chance level, $M = .47$, one-sample $t(18) = .85$, $p = .408$, $BF_{10} = .326$, $d = .19$, 95% CI [−.26, .64] (see Figure 3A) when reporting the states of exemplars that had been shown in different states. This contrasts with performance when the exemplars were presented in the same state ($M = .70$, as reported above), and the difference between these two levels of performance was significant, $t(18) = 5.43$, $p < .001$, $BF_{10} = 685$, $d_z = 1.25$, 95% CI [.63, 1.81]. Only .40 trials in the different state condition was provided by choosing one out of two correct states, which did not differ from the proportion of choosing one correct answer in the same state condition, $M = .42$, $t(18) = .33$, $p = .744$, $BF_{10} = .280$, $d_z = .076$, 95% CI [−.375, .525] (see Figure 3C). Therefore, the difference in overall accuracy between the same state and the different state conditions was provided by an ability to report either both or none of the states correctly.

Together with rather good memory for exemplars and states, this difference suggests that our participants had difficulties with correct exemplar-state binding, and that this occurs even when verbal encoding is minimized. This suggests that the lack of binding

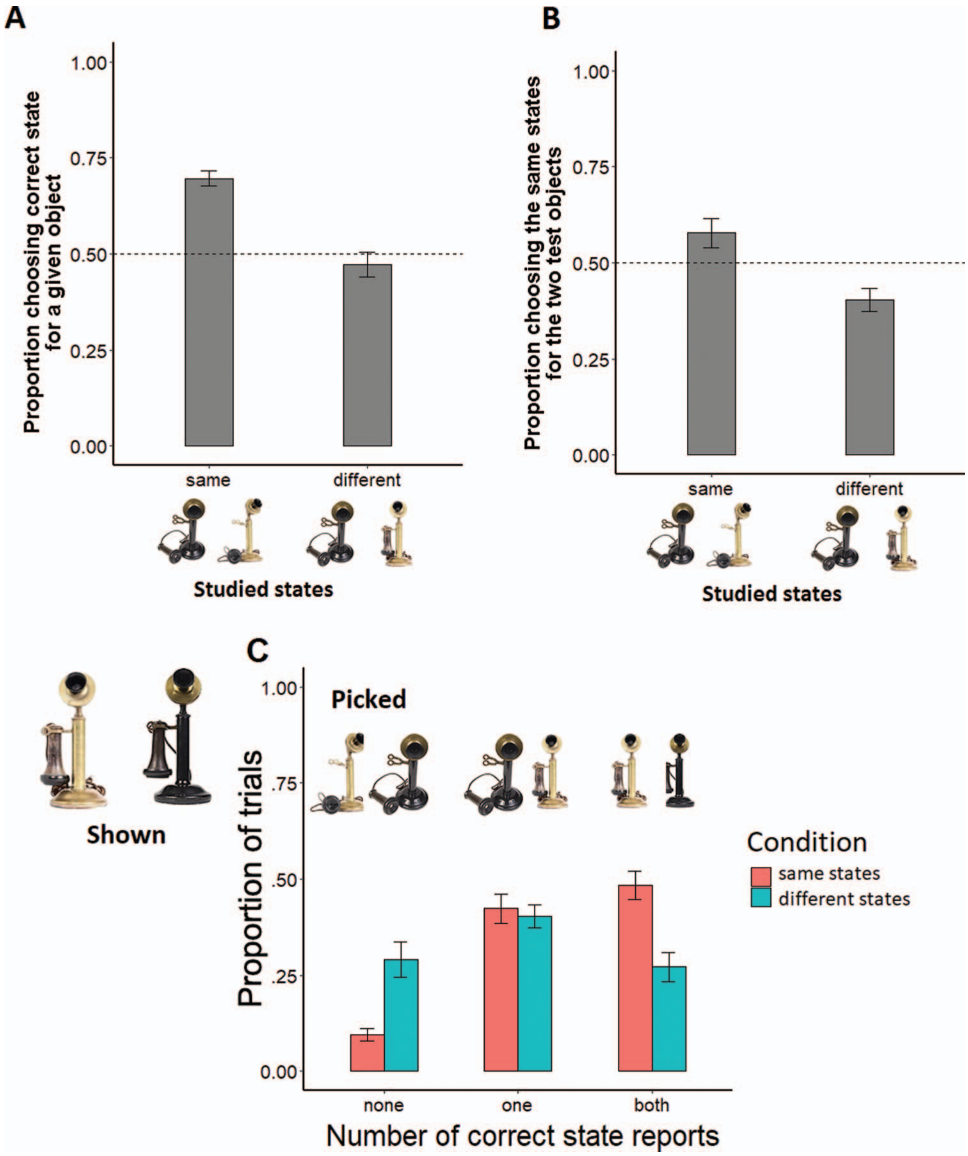

Figure 3. Results of Experiments 1B in the exemplar-state task nearly exactly replicate Experiment 1A, demonstrating verbal encoding is not the cause of unbound memories: (A) proportions choosing correct state for a given exemplar when the two studied objects were shown in the same state (left; doesn't require binding) or different states (right; requires binding); (B) proportions choosing the same states for the two test objects (regardless of whether these states are correct or incorrect). Dashed lines show chance levels. (C) The proportion of reporting both correct states, one correct state, or no correct states in a test trial as a function of the study condition. Error bars denote the *SEM*. See the online article for the color version of this figure.

between the features that support state discrimination and the features that support exemplar discrimination does not arise simply because participants are able to label some aspects but not other aspects of the stimuli.

Robustness across stimuli in Experiments 1A–B. The analyses above treat participants as random effects and average over stimuli. We have 120 distinct stimulus categories in our main analysis, and so we can also ask whether our main conclusions are robust across categories as well as across participants. As exactly the same stimuli and manipulations were used in Experiments 1A and 1B, we collapsed data across all participants from these experiments. Treating categories (e.g., abacus, book) as the random factor rather than subjects and same/different states of objects from these categories as repeated measures obtained from different subsamples of participants, we find that across categories, participants reliably choose the same state more often when the two examples of that category were in the same state than a different state, $t(119) = 15.15$, $p < .001$, $BF_{10} = 2.14 \times 10^{25}$, $d_z = 1.38$, 95% CI [1.13, 1.63] and, that when the items are in the same state, participants are more accurate in choosing the correct state for each exemplar than when the items are in different states, $t(119) = 13.93$, $p < .001$, $BF_{10} = 3.84 \times 10^{23}$, $d_z = 1.27$, 95% CI [1.13, 1.63]. This suggests the main conclusion of unbound representations robust across our stimuli as well as across our participants.

Order of blocks. To investigate whether there was an effect of the order of the exemplar versus exemplar-state task, we collapsed the data across Experiments 1A and 1B since they use effectively the same method ($N = 39$). For our main task (exemplar-state), we ran a 2-way ANOVA (task order by same/different states). We found no evidence of a main effect of task order on the correct response rate, $F(1, 37) = 1.203$, $p = .280$, $BF_{10} = .360$, $\eta^2 = .031$ or on the probability of choosing two same states, $F(1, 37) = .823$, $p = .370$, $BF_{10} = .324$, $\eta^2 = .022$. We also found no evidence for an effect of task order interacting with Same/different states on these dependent variables, $F(1, 37) = 1.78$, $p = .190$, $BF_{10} = .682$, $\eta^2 = .019$; $F(1, 37) = .460$, $p = .502$, $BF_{10} = .383$, $\eta^2 = .004$, respectively. For the exemplar task, we analyzed the recognition rate as a function of task order and also found no effect, $t(37) = 1.66$, $p = .105$, $BF_{10} = .902$, $d_z = .53$. Thus, we see no evidence of an effect of task order. This is consistent with participants trying to remembering the objects in as much detail as possible (as they were instructed) rather than specifically focusing on the same kind of details they had found relevant in previous blocks.

Discussion

Overall, in Experiment 1A and 1B, we find that people know whether the states of the objects within a category were the same or different, but are at chance in saying which state goes with which exemplar. This is strong evidence of independent storage of the features underlying state and exemplar discriminations, since unitized memory storage for real-world objects cannot allow for participants to know which features are present without knowing which objects they are part of—this can occur only if information about the state properties of the objects are stored separably from information about the exemplar properties of the objects. There are several possible accounts of independent storage that can give rise to this data pattern—including misbinding errors as well as independent forgetting of different object properties (see General Discussion section), but it is broadly inconsistent with unitized, holistic object representations. This independence is not a property of the perceptual encoding of the objects, as immediately after encoding participants are excellent even in the different state conditions (see Experiment A1 in the Appendix).

In this set of experiments we tested both items from a category at the same time—two simultaneous 2-AFC trials. We did this in order to limit the possibility that participants inadvertently accessed the wrong memory (e.g., the wrong mug). In general, gist-based false memory studies often find that people will false alarm to similar items in old-new formats, but this is generally alleviated in forced-choice situations where the correct answer is physically present (e.g., Guerin, Robbins, Gilmore, & Schacter, 2012). By explicitly contrasting the two within-category items, we hoped to prevent mistakes of this kind and ensure that participants chose the wrong state not because they mistakenly accessed the memory for the incorrect exemplar.

Thus, overall, the data from Experiment 1A and 1B provide evidence in favor of independent storage of different properties of real-world objects in long-term memory.

Experiment 2A

If the information required to discriminate exemplars and the information required to discriminate states is stored in relatively independent format, as the results of Experiment 1 suggest, this implies that when participants are instructed to retrieve only the features underlying one or the other of these discriminations, variations in the other set of features should not interfere with this retrieval. Thus, generalization to similar objects is another test of how unitized versus independent the storage of different object properties is in long-term memory. Therefore, in Experiment 2A–C, we asked participants to generalize across irrelevant changes in other object features and asked how much interference this caused with accessing memory. In Experiment 2A and 2B, we used the exemplar-state properties, consistent with Experiment 1. In Experiment 2C, we asked how much interference would occur in this generalization task for truly holistic representations by using a task with somewhat integral features (luminance and hue).

Thus, in Experiment 2A, participants recognized a studied exemplar while we varied whether the state information (a) matched the state information in memory, (b) did not match the state information in memory, or (c) the state information in memory was actively misleading. If the two sources of information are relatively independent, varying the state information should only slightly hinder the retrieval of the exemplar information from memory. The logic of this experiment is thus similar to the idea of integral versus separable features in category learning experiments (Garner & Felfoldy, 1970). Previous work provides reasons for thinking that, despite the difficulty of binding in Experiment 1, participants may not be able to access exemplar memory independent of state changes. In particular, in the domain of object memory, previous literature has shown that when participants are asked to respond purely based on categorical or conceptual information, they have difficulty ignoring exemplar information (e.g., Koutstaal, 2003; Koutstaal & Cavendish, 2006); related work shows that participants seem to be better able to retrieve memories when the exact same item is present during the test display than when only related objects are present (Guerin et al., 2012). Thus, Experiment 2A provides a strong test of the independence of the state and exemplar properties. (Experiment 2C provides context by examining memory for two features known to be somewhat holistically represented, luminance and hue).

Method

Participants. Following the power requirements described in Experiment 1, a different set of 20 psychology students of the Higher School of Economics (16 female; age: $M = 19.85$ years, $SD = .91$) took part in the experiment for extra course credits. All participants reported having normal or corrected-to-normal vision and no neurological problems. Before the beginning of the experiment, they signed an informed consent form. The sample size was matched to Experiment 1.

Apparatus and stimuli. Apparatus and stimuli were the same as in Experiment 1A. The only difference was that we used only the exemplar-state stimulus set from Brady, Konkle, Alvarez, et al. (2013). We used stimuli from Konkle et al. (2010a) only for practice trials.

Procedure. In the study phase, observers were sequentially shown 120 images, with one exemplar in one state from each of the 120 categories used in the exemplar-state memory task of Experiment 1. The speed of presentation and intertrial intervals were the same as in Experiment 1.

In the test phase, the participants performed 2-AFC recognition judgments. On each trial, an old exemplar was presented paired with a new exemplar from the same category (Figure 4A). Each exemplar could be presented in the same state as in the Study phase (old state) or in a different state (new state). The participants were instructed to choose the old exemplar regardless of the state.

Design and analysis. Our main manipulation in this experiment was the combination of states with the old and new exemplars in test phase. We tested three conditions: (a) *baseline*: old exemplar/old state versus new exemplar/old state, (b) *generalized-state*: old exemplar/new state versus new exemplar/new state, and (c) *state-misleading*: old exemplar/new state versus new exemplar/old state. Forty categories were presented to each participant in each of these conditions. We analyzed the overall accuracy in each condition.

Results

The results of Experiment 2A are summarized in Figure 4B. A one-way repeated-measures ANOVA was run to compare the total accuracy between the three conditions. It showed no significant effect, $F(2, 38) = 2.83$, $p = .071$, $\eta^2 = .130$. Our main question focused on binding, and thus the most relevant comparison is between the baseline condition and the generalized-state condition, where state is changed for both the studied and tested object, and thus cannot serve as a cue. There was no significant cost to this

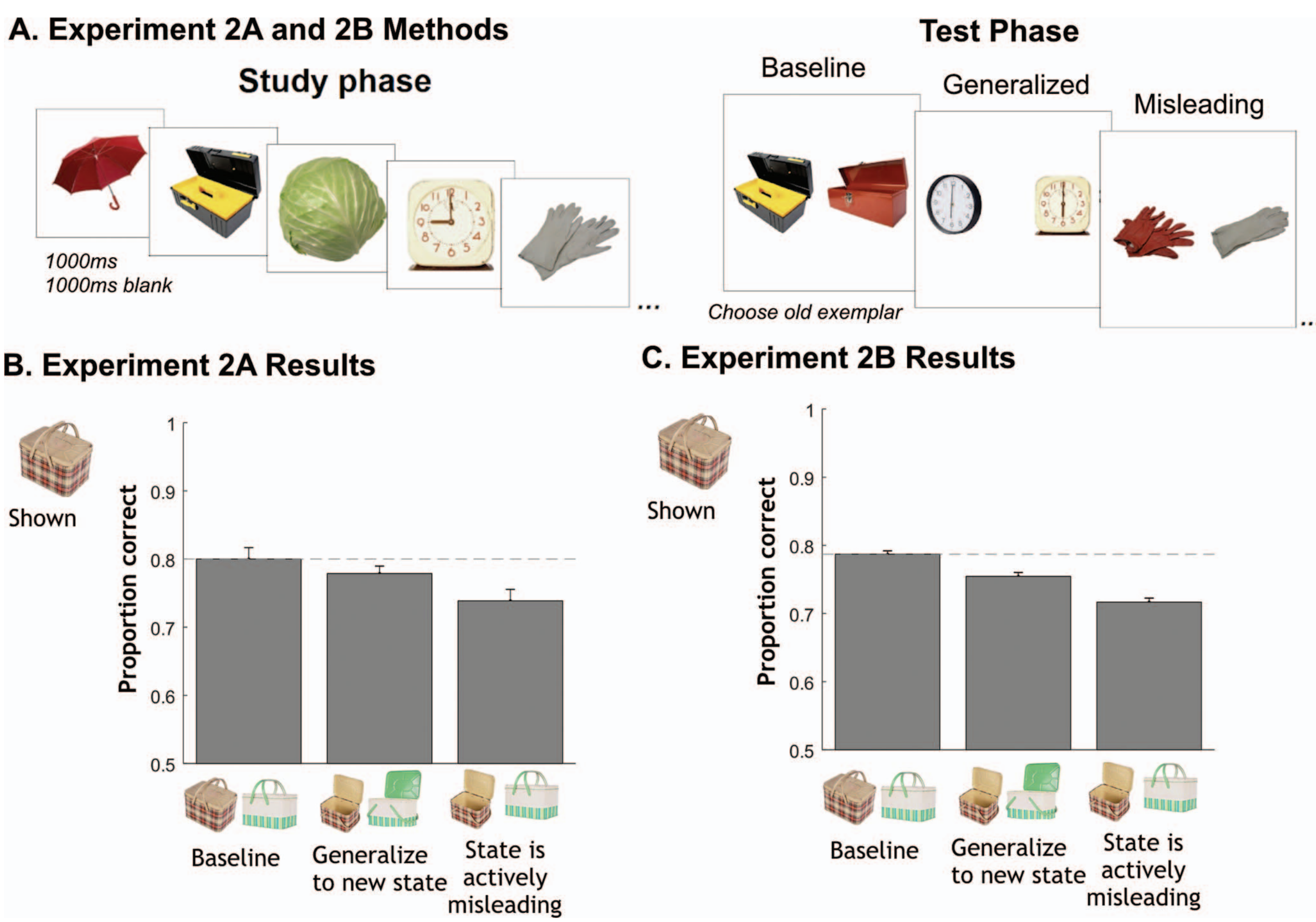

Figure 4. Methods (A) and results (B–C) of Experiments 2A–B. (A) In Experiment 2A and 2B, participants had to remember one object from each category and then pick this same exemplar during test, even if its state changed between study and test (e.g., participants were asked to remember they saw the gray gloves, even though their state/pose changed between study and test). (B) Results of Experiment 2A and 2B show that there little cost to asking participants to generalize to a new state when remembering the exemplar and performance remains high—and well above chance—even when the state is actively misleading, providing evidence for independent memory of information used in exemplar and state discriminations. Error bars denote the *SEM*. The dashed line shows performance in the baseline condition to facilitate comparisons to this condition. See the online article for the color version of this figure.

change in our sample (baseline M = .80; generalized-state M = .78; $t(19)$ = .94, p > .999, Bonferroni corrected; BF_{10} = .343, d_z = .21, 95% CI [−.23, .65]), with performance falling by only slightly less than 2% on average.

The state-misleading condition tests an even stronger claim than just unbound exemplar and state features, as participants could choose the incorrect object in this case if they simply remembered which state they had seen and not which exemplar (e.g., if there was a significant amount of independent forgetting even at such a short duration). Nevertheless, the difference between the baseline condition and the state-misleading condition was also not significant (state-misleading M = .74, $t(19)$ = 1.88, p = .225, Bonferroni corrected, d_z = .421, 95% CI [−.42, .87]), although the Bayesian analysis did not show strong evidence for the absence of any difference (BF_{10} = 1.090). The cost to this change was approximately 6% on average, a relatively small difference despite the fact that the exemplar and state information were now actively in conflict. Nevertheless, people were quite accurate at recognizing the exemplar they had previously seen, performing well above the chance-level performance you would expect from a fully unitized memory account.

Thus, we find that participants are quite good at remembering exemplar information even when state information is varied (*generalized-state*) and almost as good when state information matches the incorrect exemplar and is thus actively misleading (*state-misleading*). This provides strong evidence that the recognition of an exemplar is barely affected by changes in its state. If state and exemplar information were entirely unitized in memory, you would expect that one object which matches the state and one object which matches the exemplar (as in the state-misleading condition) would be nearly impossible to discriminate—both would match the encoded memory approximately equally well. However, participants remain well above chance at this distinction, suggesting they are able to focus on only the exemplar information with little interference from the state information.

Experiment 2B

Experiment 2B served as a direct replication of Experiment 2A with higher power.

Method

Participants. Because we wished to more precisely estimate the small effects observed in Experiment 2A, we collected data from N = 100 participants on Amazon's Mechanical Turk for additional power (an additional 15 participants were collected but excluded and replaced due to our exclusion criterion of performance below 55% averaged across all conditions; including these participants does not change any of our conclusions or the pattern of results). All participants reported having normal or corrected-to-normal vision and no neurological problems. Before the beginning of the experiment, they provided informed consent.

Procedure. The design and procedure were identical to the Experiment 2A.

Results

Experiment 2B served to more precisely estimate the effect size observed in Experiment 2A with a larger sample (N = 100 compared with N = 20). Average performance in the three conditions was: M = .79 in the baseline condition; .75 in the generalized-state condition; and .72 in the misleading-state condition (as compared with .80, .76, and .74 in Experiment 2A). Thus, we substantially replicated the pattern of data of Experiment 2A and the very high performance at recognizing exemplars even with generalized or misleading state information.

With this higher power replication, we were able to find statistically significant declines for generalizing state or misleading state information (compared to baseline, all $p \leq .001$). More importantly, we were able to estimate the effect size of this decline more precisely, confirming a decline of only approximately 4% for generalizing across states, $t(99)$ = 3.63, p = .001, Bonferroni corrected, BF_{10} = 46.53, d = .363, 95% CI [.16, .57] and of approximately 7% in performance for misleading state information, $t(99)$ = 7.63, p < .001, $BF_{10} = 5.92 \times 10^8$, d = .763, 95% CI [.54, .99].

Thus, participants are again only very slightly impaired at remembering exemplar information even when state information is varied (*generalized-state*) and still not much worse when state information matches the incorrect exemplar and is thus actively misleading (*state-misleading*). This provides evidence that the recognition of an exemplar is not strongly affected by changes in its state.

Experiment 2C

We showed in Experiments 2A–B that irrelevant manipulations of state information at test cause very little interference in the recognition of exemplars. In order to know whether this small amount of interference is consistent with relatively independent feature storage, we need to compare it to how much interference we would expect between features if they are strongly bound or unitary. Thus, in Experiment 2C, we tested how well participants can recall one feature if another feature manipulated at test is somewhat integral with the first one. One such dimension is color with its components hue and brightness that are shown to be integral (Garner & Felfoldy, 1970). Thus, in Experiment 2C, our participants studied silhouettes of real-world objects with different hues and brightnesses. In the test phase, they had to recognize the brightness of each object regardless of its hue (in analogy with the task of recognizing an exemplar regardless of its state in Experiments 2A–B). We used the same set of manipulations with the hues of a target and a foil as in Experiments 2–AB (baseline-misleading-generalized).

Method

Participants. We collected data from N = 100 participants on Amazon's Mechanical Turk (an additional 18 participants were collected but excluded and replaced due to performance below 55% in the baseline condition; including these participants does not change any of our conclusions or the pattern of results). All participants reported having normal or corrected-to-normal vision and no neurological problems. Before the beginning of the experiment, they provided informed consent.

Stimuli and procedure. The design of the study was the same as Experiment 2B, except rather than being asked to remember the exemplar of an object and generalize over state, participants were instructed to remember the luminance of an object and generalize over its hue. Stimuli consisted of 30 silhouettes taken from 30 different categories of objects (from Sutterer & Awh, 2016). These stimuli assigned random colors from a circle of radius 38 centered on A = 0, B = 0 in CIELab color space. "Bright" (high luminance) objects had colors chosen from L = 75, and "dark" objects (low luminance) had colors chosen from L = 45.

During the study phase, participants saw a sequence of 30 such silhouettes and were instructed to remember their brightness independent of their hue. Each of the 30 objects was assigned a random hue from the color wheel and was randomly assigned to be either "bright" or "dark." Objects were presented for 2,000 ms with a 1,000-ms blank interval.

After the study phase, participants were tested in a 2-AFC format on the brightness of all 30 objects they had encoded. As in Experiment 2B, these tests were split between three conditions: 10 in baseline, 10 in generalized, and 10 in misleading. On each trial, they were shown two instances of the same silhouette, one having the same luminance as the original (target brightness), another one having a different luminance (foil brightness). In the baseline condition, the hue of both items was the same as in the originally studied object. In the generalized condition, both items had a new, not studied hue (always 180 degrees away on the color wheel from the original hue, which has previously been shown to be a similar magnitude of change to the state changes used in Experiments 2A–B: Brady et al., 2013, Exp. 1). Finally, in the misleading condition, an item with the target brightness had the new hue, whereas an item with the foil brightness had the studied hue. The examples of stimuli and conditions are shown in Figure 5B.

Overall, the design and data analysis were the same as in Experiments 2A–B.

Results

Performance in the Baseline condition was as good as in Experiments 2A–B ($M = .78$ vs. $M = .79$–.80, Figure 5B). However, this performance dropped markedly down when the hue was manipulated, in contrast to the effects of Experiment 2A–B (generalized-hue: $M = .66$ in Experiment 2C vs. .75–.78 in Experiments 2A–B; misleading-hue: $M = .59$ in Experiment 2C vs. .72–.75 in Experiments 2A–B). Comparisons across the conditions of Experiment 2C showed quite a strong effect of the hue manipulations on recognition memory, $F(2, 198) = 46.61$, $p < .001$, $\eta^2 = .320$. Evidence for substantial drops in performance was corroborated by large effect sizes for the difference from baseline in the misleading and generalization conditions (generalized-hue vs. baseline: $t(99) = 7.42$, $p < .001$, Bonferroni corrected, $BF_{10} = 2.20 \times 10^8$, $d = .742$, 95% CI [.519, .962]); misleading-hue versus baseline: $t(99) = 9.41$, $p < .001$, Bonferroni corrected, $BF_{10} = 3.03 \times 10^{12}$, $d = .941$, 95% CI [.703, 1.18]. Thus, we found that our observers were overall extremely impaired by the irrelevant hue manipulations—seemingly much more than by the irrelevant state changes in Experiments 2A–B.

Comparison between Experiments 2C and 2B. To estimate how much interference state manipulations cause on exemplar memory compared to the interference predicted by strongly integral features (such as brightness and hue), we compared the losses from the state manipulation in Experiments 2B with the hue manipulations in 2C, as they used similar sample sizes and similar testing conditions. To estimate this cost, we subtracted the proportions of correct answers in the generalized and the misleading conditions from those in the baseline conditions in both experiments. The resulting differences are shown in Figure 5C. Because the two experiments had similar baseline performance, no strong assumptions of the linearity of percent correct are needed in this analysis, and the results are extremely similar if a measure of memory strength is used instead (e.g., d'). We found that both the generalization and misleading conditions caused substantially greater costs in brightness reports (Experiment 2C) than in exemplar reports (Experiment 2B). Specifically, the generalized-state manipulation caused the average loss of .03 on exemplar report accuracy, whereas the generalized-hue manipulation caused the average loss of .13 on brightness report accuracy, $t(99) = 4.85$, $p < .001$, $BF_{10} = 5{,}964$, $d = .686$, 95% CI [.400, .971]. The misleading-state manipulation caused the average loss of .07 on exemplar report accuracy, whereas the misleading-hue manipulation caused the average loss of .19 on brightness report accuracy, $t(99) = 5.85$, $p < .001$, $BF_{10} = 70.01 \times 10^3$, $d = .766$, 95% CI [.478, 1.052].

Summary. Overall, we observed much stronger interference from irrelevant changes in the highly integral dimensions of brightness and hue than we observed with exemplars and states. This is consistent with the idea that hue and brightness are "integral"—depending on the same underlying set of features, or in some way unitized—whereas state and exemplar discriminations rely on distinct features which can be independently accessed.

Another factor affecting this experiment is the extent to which states vary versus the extent to which hues vary. For example, it may be more or less difficult in general to "see past" state versus hue changes if these changes are of different magnitudes. We used maximal hue changes in the current experiment because previous work revealed that state changes in a similar set of objects were well-matched in terms of perceptual discriminability by 180 degree color changes (Brady, Konkle, Gill, Oliva, & Alvarez, 2013, Experiment 1). This suggests that both the irrelevant state changes and irrelevant hue changes were likely similar in magnitude, and thus the primary difference between Experiments 2A–B and 2C is the relationship between the features (exemplar-state and hue-brightness).

Note that although we found greater integrality in the hue-brightness experiment, our brightness memory task should, if anything, have allowed more straightforwardly independent encoding of features: because the participants knew in advance that they had to study brightness and because it was much easier to provide compact uniform verbal labels for object and brightness changes (compared with the case of unpredictably changing exemplar and state information that the observers were not warned about). Even despite this, our observers failed to report the correct brightness under hue manipulations much more often than they failed to report a correct exemplar under state manipulations. This finding provides a strong argument that the features underlying state and exemplar discriminations are stored and accessed relatively independently.

General Discussion

In Experiments 1A–B, we tested whether features of real-world objects can be successfully remembered but not as unitary, holistic

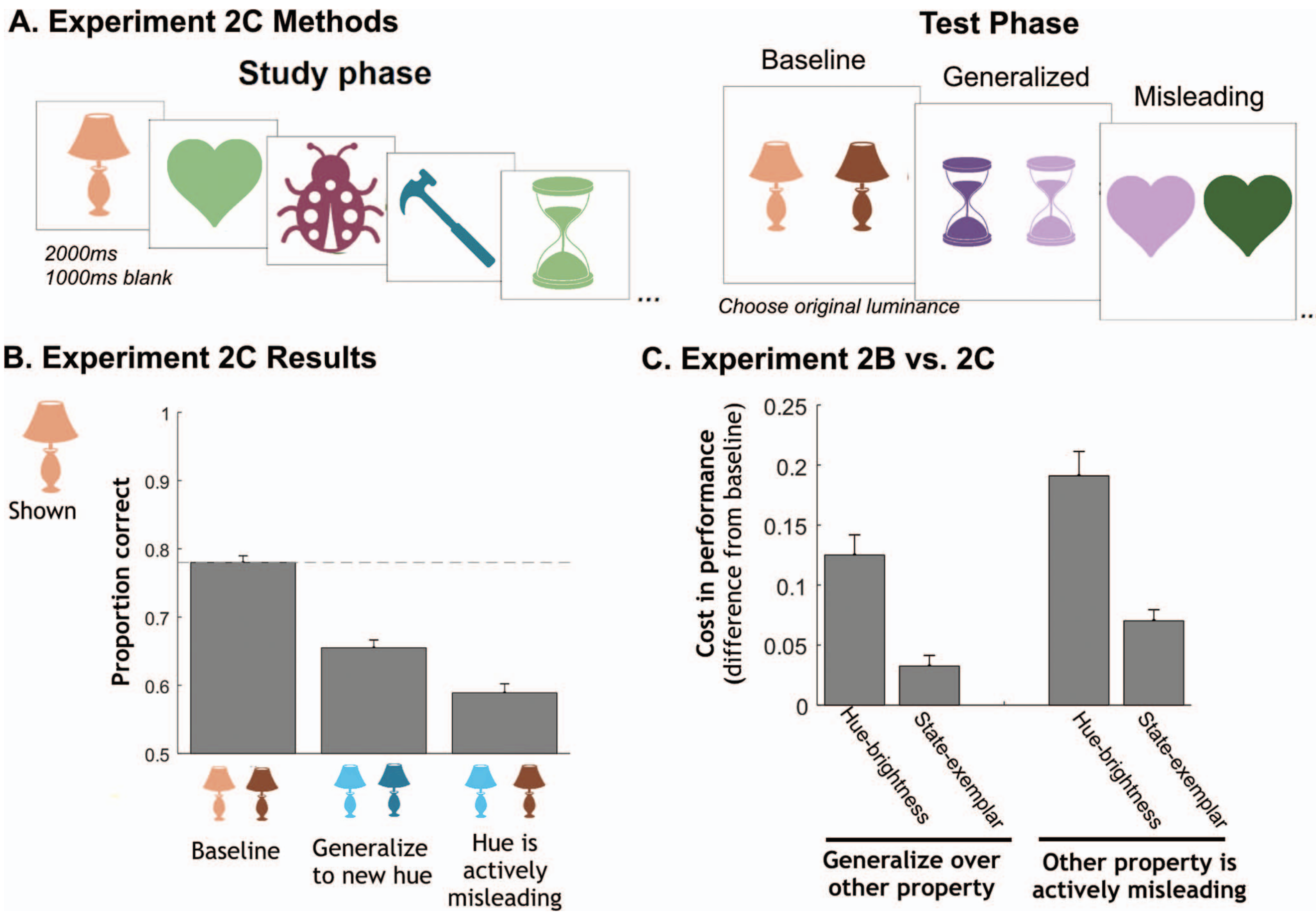

Figure 5. Methods (A) and results (B–C) of Experiments 2C. (A) Experiment 2C served as a comparison point for Experiment 2A and 2B. In this experiment, we used a set of features known to be somewhat "integral" and asked how much cost participants would incur in generalizing over one of the features. In particular, participants had to remember the luminance of each item and then pick the appropriate luminance item during test, even if its hue changed between study and test (e.g., participants were asked to remember they saw the dark hammer, even if its hue changed between study and test). (B) Results of Experiment 2C show that there is a very large cost to luminance memory when asked to generalize across hue or when hue is actively misleading. The dashed line shows performance in the baseline condition to facilitate comparisons to this condition. (C). Experiment 2B and 2C had extremely similar baseline performance. Thus, we can compare the "cost" associated with generalizing over the irrelevant dimension in these two experiments by subtracting performance in the generalization conditions from the baseline condition (greater cost = greater drop in performance relative to baseline). Comparison of Experiment 2B and 2C show that the cost in transfer across hue when remembering brightness is much larger than the cost in transfer across state when remember exemplars, consistent with the idea that state and exemplar features are relatively independent and unbound whereas hue/luminance are integral. Error bars denote the *SEM*. See the online article for the color version of this figure.

objects but instead as separate features that require binding to connect to each other. We found that participants demonstrated reasonably good memory for pairs of object exemplars and for whether the set of exemplars were presented in the same or different states. However, when participants were required to bind state to exemplar information—remember which state went with which exemplar—performance dropped to near-chance levels. Thus, people often know something about which "states" were present within a category, and which exemplars were present within a category, without knowing which exemplars the states go with, consistent with independent storage of the features underlying state and exemplar discriminations.

The flip side of this relative feature independence was demonstrated in Experiments 2A–C. While retrieving a correct exemplar-state conjunction is difficult, retrieving the exemplar alone is easy, even if its state has changed since study, and it even remains relatively easy when the familiar state is present on the foil. Indeed, the amount of interference between state and exemplar features is far less than that found with highly bound, integral features such as brightness and hue (Garner & Felfoldy, 1970). Thus, taken together, our experiments provide logically strong evidence for unboundedness based on a double dissociation between our two tasks. In Experiments 1A–B, the task required successful exemplar-state binding and it turned out to be difficult for our participants. By contrast, in Experiments 2A–B, the task rather required successful exemplar-state "unbinding" and it turned out to be easy for our participants (unlike brightness-hue "unbinding" in Experiment 2C).

The Structure of Memory Representations

In Experiment 1, participants frequently were at chance in choosing the correct state for items that had been seen in different states (e.g., where they had seen one open box and one closed box), despite reliably knowing whether two different states were present. In other words, in the "different states' condition," participants frequently "swapped" the items in their reports—getting both wrong as often as both right, but with reliable knowledge of the fact that the two items should have different states.

To understand how this experimental demonstration of "swap" errors is related to participants' underlying object representations, it is worthwhile to consider several scenarios that are schematically shown in Figure 6. Bound storage implies that an object is unitized—stored in its entirety or completely forgotten. Independent storage implies participants know information about only the state of an object or only the exemplar of the object, stored separately.

Figure 6A—the left side—focuses on what bound, unitized storage would look like for two objects of the same category, with different amounts of forgetting, that is, when both objects are stored (Figure 6A, first box), only one is stored (Figure 6A, second box), or both are forgotten (Figure 6A, third box). By necessity, all these models predict an equal likelihood for objects to be remembered or forgotten regardless of whether the two objects are in the same or different states; because the objects are unitized, there is no strong relationship of the items to each other (e.g., as though pattern separation processes had made them entirely distinct in their representation despite their within-category similarity). However, our results consistently showed that this was not the case: objects in the same states were recognized correctly more often that objects in different states, broadly inconsistent with this bound

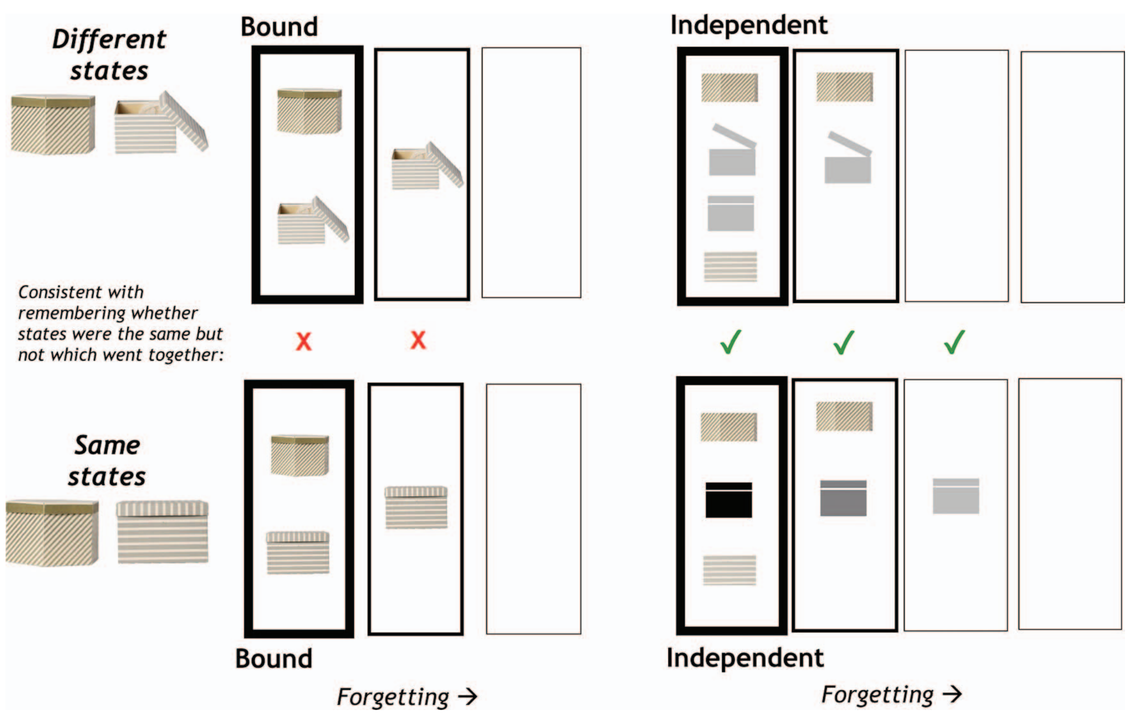

Figure 6. Schematic of memory representations and the predictions they make about Experiment 1A and 1B. (A) Bound memories are holistic item representations. When memory is strong (left), then participants are expected to know both the state and exemplar of each item. As memory gets weaker (moving right), participants lose entire objects holistically. Memory representations of this kind cannot lead to participants having above chance performance at knowing whether one or two different states were present while not knowing what those states were; once one of the items is lost, participants no longer know whether the two states were the same or different. (B) Independent memories of features that distinguish between exemplars and states as memory is lost (left to right). There are several possibilities for how independent memory can give rise to above-chance performance in remembering whether the states were the same or different but chance performance in knowing which state goes with which exemplar. If participants are aware of both states, but not the binding between them, this can cause "swap" errors, consistent with the data (left; strong memories). If participants do not remember both states, but memories for a given state are known by participants to be stronger (darker colored) when the state was the same between the objects than when it was different, this also could give rise to the data pattern we find, as participants could use the strength of the state memory to decide whether the objects had the same or different state. See the online article for the color version of this figure.

storage account. In addition, our results showed that people often knew which states were present without knowing which exemplar the states went with, again broadly inconsistent with this bound storage account.

Thus, our data strongly favor the idea of independent storage (Figure 6B, right side). However, our data in Experiment 1 do not by themselves disambiguate between accounts of misbinding and accounts of independent forgetting—although Experiment 2 speaks in favor of the misbinding account. Considering only Experiment 1, the first scenario (Figure 6B, first box) suggests that observers store both exemplars and both states but fail to remember which state went with which exemplar. This "unbound" memory scenario makes a very straightforward prediction that fits our data. If an observer has such representations, he or she would preferentially choose the same states for objects that were presented in the same states and choose different states for objects that were presented in two different states. Moreover, the repeated presentation of the same state should lead to better performance in same state trials. For different states, such a representation predicts that observers would choose either both or none of the states correctly as they do not have firm knowledge how the states and the exemplars were bound, again broadly consistent with our data. Such independent memory representations are also consistent with the independent access to state and exemplar information observed in Experiment 2.

However, accounts based on independent forgetting are also possible in Experiment 1. (Although they do not straightforwardly account for Experiment 2's generalization results, which show independence even when both features are strongly represented). In particular, if we imagine independent storage of state and exemplar information, with the traces for the "state" being stronger when both objects shared that state, this could also explain the pattern we observed in Experiment 1 (Figure 6B, second box). If the two same-category objects are presented in the same state, this could aid the consolidation of the state information, resulting in a high frequency of correct choices of both these states at test. By contrast, if two objects are presented in different states, this may result in a weaker trace for each of them which might increase the probability of forgetting one or both these states. When the states are forgotten, observers may randomly guess exemplar-state conjunctions but their choice of same versus different states could be strategic. For example, if an observer does not actually remember the states, he or she may assume this is because the objects differed in state and thus choose two different states. Unlike the previous scenario implying good memory for all features and a failure to bind them, the current scenario emphasizes the role of between-feature interactions causing consolidation or forgetting. Our data from Experiments 1A–B do not make a clear distinction between these two mechanisms.

Importantly, both of the possible mechanisms for explaining our data are incompatible with fully bound, unitized memory representations of objects. In particular, participants having a memory for the abstracted state of the items in the category independent of the particular exemplar features presupposes that the storage of state and exemplar features is independent.

Finally, one more scenario theoretically compatible with our data from the exemplar-state task alone is if participants frequently failed to remember exemplar information at all, and had only some independent memory of the state of the items in the category, again with this memory varying in strength between same-state and different-state situations (Figure 6B, third box). Again, this account implies strongly independent features. However, this scenario also does not account for the strong memory for exemplars we find in our data with the same encoding conditions, so it is largely incompatible even with the data from Experiment 1. Thus, broadly, the data from Experiment 1 are consistent with unbound state and exemplar features or independent forgetting of exemplar and state features (with differential memory strength in the "same" and "different" state conditions).

Importantly data from Experiment 2 largely speak in favor of the unbound feature account, at least in terms of access to the features: If independence only arose from the loss of some features due to differential forgetting, we would not expect to see clear and easy generalization in Experiment 2. This generalization is separate evidence in favor of independent storage, even in the absence of independent forgetting. Thus, overall, we believe our data speak strongly to the independence of different object features in memory and overall favor the view of unbound representations of these features rather than purely an independent forgetting/independent memory strength account.

Note that in the current work, we only test whether object exemplar information can be accessed independent of state information (in Experiment 2) and whether object state is always bound to object exemplar (in Experiment 1). In theory, alternative accounts of our data where there is an asymmetry between object state and object exemplar are possible—for example, hierarchical memory representations where remembering exemplar information is necessary for accessing object state but object state can be lost independently of exemplar information. In the visual working memory literature such accounts have been proposed to understand why there are costs to encoding information from multiple objects yet information within an object appears to be forgotten independently (e.g., Brady, Konkle, & Alvarez, 2011; Fougnie et al., 2010, 2013). The extent to which the nonholistic, nonunitized representations we observe may have a hierarchical structure remains an open question for future work.

State and Exemplar as Properties of Objects

The features underlying state and exemplar discrimination are not reducible to separable sensory dimensions (like color or orientation) but represent more complex pieces of information about object appearance in visual long-term memory. Previous work has used exemplar and state manipulations (Brady et al., 2008, 2013; Konkle et al., 2010a), because distinguishing between two different states or poses of an object and between two different exemplars of the same object category are two common and important tasks in the real-world, and these aspects of objects capture important regularities in how the objects we interact with actually vary in the real-world. Our results suggest that these distinctions on average rely on different aspects of the visual appearances or semantic representation of real-world objects, and these aspects of the objects are not represented together—holistically—in memory. What features underlie these discriminations? The answer is likely to vary at least in part from object to object. For example, on average state changes likely affect more about the current role or functionality of the object than do replacements of an item with another exemplar of the same category in the same state; in addition, state changes frequently result in a part of the object moving, creating more changes in global shape. In the current work, because we found that these dimensions appear to be unbound, we provide evidence that whatever the visual or semantic dimensions that are encoded about objects are, the features that underlie the ability to make "state" and "exemplar" discriminations must be at least partially independent. However, understanding the exact visual and semantic features used to represent real-world objects and to make such discriminations is a significant challenge for future work.

Importantly, our experiments contain an asymmetry with respect to the state and exemplar properties: If we did not find independence between state and exemplar properties, the data would not disambiguate between two ideas. On the one hand, this could have been evidence that objects are stored holistically in a unitized format. On the other hand, it could have been that people could simply rely on the exact same features for the state and exemplar tests (e.g., the state features are just the exact same features as the exemplar features, no binding required), whereas other feature combinations might in fact be stored independently. However, finding independence implies not only that objects are not unitized but also that the features underlying the two tests are independent. As noted above, this could be because people really store object representations semantically (e.g., in meaningful features specific to object categories, like "fullness") or simply because on average the visual features underlying state and exemplar discriminations tend to vary.

An important question regarding our results can be asked whether our pattern could be achieved via verbal encoding rather than a visual or semantic format of encoding. For example, one might suggest that two exemplars in same states could be described by a single label (e.g., two "coffee mugs") and two exemplars in different states by two labels (e.g., "a coffee mug" and "a mug of coffee"), and interference could arise from the difficulty of labeling at retrieval rather than visual unboundedness. However, we do not believe nonvisual, purely verbal encoding strongly contributes to memory performance in our experiments. First, the variety of exemplar and state changes across images was so large that participants could not know in advance which features would be relevant for a subsequent memory test. Instead, the nature of our stimuli would require them to create a long verbal description for each item to provide a considerable chance of having a proper feature description accessible at test. Given the number of items, relatively short encoding time, and the fact that the participants were not warned about relevant features in the instruction, this strategy seems unlikely and not very fruitful. In addition, we directly manipulated the availability of explicit verbal encoding via articulatory suppression (Experiment 1B) and found the same pattern of results. We therefore advocate the idea that memories for exemplars and for states, as well as a failure to bind them relied substantially on relatively independent visual or semantic representations rather than verbal labels.

Connections to Previous Work on Binding Within Objects in Visual Long-Term Memory

Our results are in line with the main conclusion of Brady et al. (2013) and Reinitz et al. (1992) that the features of real-world objects are not bound into one unitized representation in long-term memory. Our findings make this conclusion significantly stronger, particularly with respect to meaningful features of real objects. Brady et al. (2013) showed that some features of an object can be forgotten while others are still in memory, which could be explained by the idea that some features are stored better than others for particular objects. Reinitz et al. (1992) found that observers falsely recognized faces conjoining parts of studied faces more frequently than faces having new features, which is evidence that participants have at least some ability to detect novel features, but is not evidence for a lack of bound representations. Here, we show that erroneous reports about exemplar-state conjunctions frequently occur and that people are easily able to generalize across state features when considering exemplar features. This cannot be ascribed to lacking feature encoding or to differentiation between old and new features but can be explained only by independent feature storage. Thus, the current results provide direct evidence against inherently unitized representations of features of an object. They go beyond previous work in particular in showing the implications of this independence: in terms of access to those features, which appears largely independent (Experiment 2) and in terms of the ability to keep items distinct within a category, which is significantly impaired when items in a category vary (e.g., in state) due to a failure of binding each item to its relevant feature information.

Long-Term Memory Limitations

Our findings have important theoretical implications. Binding plays an important role in systems with limited capacities, such as rapid or peripheral perceptual recognition (Treisman, 1996; Wolfe, Võ, Evans, & Greene, 2011) or visual working memory (Luck & Vogel, 1997; Raffone & Wolters, 2001). In these highly limited systems, it is reasonable to suppose that participants selectively encode a limited set of separate features and have a small fraction of them bound at a time, rather than form special units tuned to each of the endless possible set of feature combinations (Tsotsos, 1988). Unlike perception or visual working memory, visual long-term memory has no such evident limits: It stores many more items than visual working memory (Brady et al., 2008; Konkle et al., 2010a; Standing, 1973). From such a massive storage capacity, we might expect an ability to hold vast numbers of items in fully integrated, bound formats. However, our results show that this is not the case, even for meaningful objects. It appears that visual long-term memory stores at least some features of real-world objects independently. This strategy can yield at least three benefits. First, it reduces the computational complexity of the storage (Tsotsos, 1988). Second, it provides a reasonable degree of memory invariance, as demonstrated in Experiment 2: An object can be recognized when some of its features naturally change. Third, even when some features are forgotten, unboundedness allows us to retrieve object information using other features. Thus, our results are in line with a model where at least some of the meaningful features making up objects are stored separately and accessed separately—as opposed to unitized into holistic object representations.

Our suggestion that binding is required even for remembering single objects also has significant repercussions for examining the role of hippocampus and other medial temporal lobe structures (e.g., Davachi, 2006). In particular, this literature often implicitly treats object-only memory as not requiring binding, in order to contrast object memory with object-context memory, which is presumed to require binding operations. Insofar as objects are themselves "bundles" of features that require binding, the distinction between object-only memory and object-context memory is likely to be more complicated than the need for binding alone. Our data also argues against an account where memory is access purely through holistic representations that have undergone hippocampal-dependent pattern separation processes that make even similar remembered items as distinct as possible (e.g., Diana et al., 2007; Norman & O'Reilly, 2003; Yassa & Stark, 2011). In particular, Experiment 1 shows that similar objects within a category interfere with each other in a predictable, feature-based way: For example, seeing a full mug and an empty mug leaves people uncertain which was full and which was empty. This is incompatible with a view where similar items are stored in fundamentally unitized, separated representations.

Our results give new insights into the processes leading to failures in visual long-term memory. It has long been known that recognition degrades if a tested item has been presented among multiple similar as compared with dissimilar items (Hunt, 2006; Konkle, Brady, Alvarez, & Oliva, 2010b; Wallace, 1965). Describing this effect, researchers often refer to a broad concept of distinctiveness (Hunt, 2006): in particular, arguing that more distinctive items are best remembered, and thus adding items that are maximally similar to a given item is most likely to impair memory for that object. Here we report the opposite finding: Observers were worse at memory when the two items to be remembered within a category were the most distinct from each other—for example, when the two items they saw in a category were different exemplars in different states. By contrast, if both items shared information—for example, were the same state—participants performed better, in spite of the fact that this meant the objects-to-be-remembered were more similar to each other. This does not conform with the standard distinctiveness framework, suggesting that the lack of binding for meaningful features of real-world objects could be an important driver of memory failures independent of distinctiveness.

Conclusion

In summary, our findings advance our view of how information about everyday objects is represented in visual long-term memory. Memories for individual objects are not the fundamental units of our memory. Rather, object memories are themselves stored in the form of independent representations that can be lost or misbound (Experiment 1) or recalled separately (Experiment 2). Our results thus provide deeper understanding of the basic mechanisms of storage in visual long-term memory.

References

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge, UK: Cambridge University Press.

Brady, T. F., Konkle, T., & Alvarez, G. A. (2011). A review of visual memory capacity: Beyond individual items and towards structured representations. *Journal of Vision, 11*(5), 4.

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences, USA, 105,* 14325–14329.

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2013). Real-world objects are not represented as bound units: Independent forgetting of different object details from visual memory. *Journal of Experimental Psychology, 142,* 792–808.

Brady, T. F., Konkle, T., Gill, J., Oliva, A., & Alvarez, G. A. (2013). Visual long-term memory has the same limit on fidelity as visual working memory. *Psychological Science, 24,* 981–990. http://dx.doi.org/10.1177/0956797612465439

Chalfonte, B. L., & Johnson, M. K. (1996). Feature memory and binding in young and older adults. *Memory & Cognition, 24,* 403–416. http://dx.doi.org/10.3758/BF03200930

Cowan, N., Blume, C. L., & Saults, J. S. (2013). Attention to attributes and objects in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 731–747. http://dx.doi.org/10.1037/a0029687

Cunningham, C. A., Yassa, M. A., & Egeth, H. E. (2015). Massive memory revisited: Limitations on storage capacity for object details in visual long-term memory. *Learning & Memory, 22,* 563–566. http://dx.doi.org/10.1101/lm.039404.115

Davachi, L. (2006). Item, context and relational episodic encoding in humans. *Current Opinion in Neurobiology, 16,* 693–700. http://dx.doi.org/10.1016/j.conb.2006.10.012

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2007). Imaging recollection and familiarity in the medial temporal lobe: A three-component model. *Trends in Cognitive Sciences, 11,* 379–386. http://dx.doi.org/10.1016/j.tics.2007.08.001

Di Lollo, V. (2012). The feature-binding problem is an ill-posed problem. *Trends in Cognitive Sciences, 16,* 317–321. http://dx.doi.org/10.1016/j.tics.2012.04.007

Erez, J., Cusack, R., Kendall, W., & Barense, M. D. (2016). Conjunctive coding of complex object features. *Cerebral Cortex, 26,* 2271–2282. http://dx.doi.org/10.1093/cercor/bhv081

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175–191. http://dx.doi.org/10.3758/BF03193146

Fougnie, D., & Alvarez, G. A. (2011). Object features fail independently in working memory: Evidence for a probabilistic feature-store model. *Journal of Vision, 11*(12), 3.

Fougnie, D., Asplund, C. L., & Marois, R. (2010). What are the units of storage in visual working memory? *Journal of Vision, 10*(12), 27.

Fougnie, D., Cormiea, S. M., & Alvarez, G. A. (2013). Object-based benefits without object-based representations. *Journal of Experimental Psychology: General, 142,* 621–626. http://dx.doi.org/10.1037/a0030300

Garner, W. R., & Felfoldy, G. L. (1970). Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology, 1,* 225–241. http://dx.doi.org/10.1016/0010-0285(70)90016-2

Gauthier, I., & Tarr, M. J. (2016). Visual object recognition: Do we (finally) know more now than we did? *Annual Review of Vision Science, 2,* 377–396. http://dx.doi.org/10.1146/annurev-vision-111815-114621

Guerin, S. A., Robbins, C. A., Gilmore, A. W., & Schacter, D. L. (2012). Retrieval failure contributes to gist-based false recognition. *Journal of Memory and Language, 66,* 68–78. http://dx.doi.org/10.1016/j.jml.2011.07.002

Hardman, K. O., & Cowan, N. (2015). Remembering complex objects in visual working memory: Do capacity limits restrict objects or features? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41,* 325–347.

Hicks, J. L., & Starns, J. J. (2015). Using multidimensional encoding and retrieval contexts to enhance our understanding of stochastic dependence in source memory. *Psychology of Learning and Motivation, 62,* 101–140. http://dx.doi.org/10.1016/bs.plm.2014.09.004

Hollingworth, A. (2006). Scene and position specificity in visual memory for objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 58–69. http://dx.doi.org/10.1037/0278-7393.32.1.58

Hunt, R. R. (2006). The concept of distinctiveness in memory research. In R. R. Hunt & J. B. Worthen (Eds.), *Distinctiveness and memory* (pp. 3–25). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195169669.003.0001

JASP Team. (2017). JASP (Version 0.8.1.1) [Computer software]. Retrieved from https://jasp-stats.org/

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford, UK: Oxford University Press.

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010a). Scene memory is more detailed than you think: The role of categories in visual long-term memory. *Psychological Science, 21,* 1551–1556. http://dx.doi.org/10.1177/0956797610385359

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010b). Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. *Journal of Experimental Psychology: General, 139,* 558–578. http://dx.doi.org/10.1037/a0019165

Koutstaal, W. (2003). Older adults encode—But do not always use—Perceptual details: Intentional versus unintentional effects of detail on memory judgments. *Psychological Science, 14,* 189–193. http://dx.doi.org/10.1111/1467-9280.01441

Koutstaal, W., & Cavendish, M. (2006). Using what we know: Consequences of intentionally retrieving gist versus item-specific information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 778–791. http://dx.doi.org/10.1037/0278-7393.32.4.778

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 19–31.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390,* 279–281. http://dx.doi.org/10.1038/36846

Norman, K. A., & O'Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. *Psychological Review, 110,* 611–646. http://dx.doi.org/10.1037/0033-295X.110.4.611

Oberauer, K., & Eichenberger, S. (2013). Visual working memory declines when more features must be remembered for each object. *Memory & Cognition, 41,* 1212–1227. http://dx.doi.org/10.3758/s13421-013-0333-6

Olson, I. R., & Jiang, Y. (2002). Is visual short-term memory object based? Rejection of the "strong-object" hypothesis. *Perception and Psychophysics, 64,* 1055–1067.

Olsen, R. K., Lee, Y., Kube, J., Rosenbaum, R. S., Grady, C. L., Moscovitch, M., & Ryan, J. D. (2015). The role of relational binding in item memory: Evidence from face recognition in a case of developmental amnesia. *The Journal of Neuroscience, 35,* 5342–5350. http://dx.doi.org/10.1523/JNEUROSCI.3987-14.2015

Peirce, J. W. (2007). PsychoPy - psychophysics software in Python. *Journal of Neuroscience Methods, 162,* 8–13.

Raffone, A., & Wolters, G. (2001). A cortical mechanism for binding in visual working memory. *Journal of Cognitive Neuroscience, 13,* 766–785. http://dx.doi.org/10.1162/08989290152541430

Ranganath, C., Yonelinas, A. P., Cohen, M. X., Dy, C. J., Tom, S. M., & D'Esposito, M. (2004). Dissociable correlates of recollection and familiarity within the medial temporal lobes. *Neuropsychologia, 42,* 2–13. http://dx.doi.org/10.1016/j.neuropsychologia.2003.07.006

Reinitz, M. T., Lammers, W. J., & Cochran, B. P. (1992). Memory-conjunction errors: Miscombination of stored stimulus features can produce illusions of memory. *Memory & Cognition, 20,* 1–11. http://dx.doi.org/10.3758/BF03208247

Rosenholtz, R., Huang, J., & Ehinger, K. A. (2012). Rethinking the role of top-down attention in vision: Effects attributable to a lossy representation in peripheral vision. *Frontiers in Psychology, 3,* 1–15. http://dx.doi.org/10.3389/fpsyg.2012.00013

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian *t* tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16,* 225–237.

Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition, 80,* 1–46. http://dx.doi.org/10.1016/S0010-0277(00)00152-9

Standing, L. (1973). Learning 10,000 pictures. *The Quarterly Journal of Experimental Psychology, 25,* 207–222. http://dx.doi.org/10.1080/14640747308400340

Sutterer, D. W., & Awh, E. (2016). Retrieval practice enhances the accessibility but not the quality of memory. *Psychonomic Bulletin & Review, 23,* 831–841.

Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology, 6,* 171–178. http://dx.doi.org/10.1016/S0959-4388(96)80070-5

Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology, 14,* 107–141. http://dx.doi.org/10.1016/0010-0285(82)90006-8

Tsotsos, J. K. (1988). A 'complexity level' analysis of immediate vision. *International Journal of Computer Vision, 2,* 303–320. http://dx.doi.org/10.1007/BF00133569

Underwood, B. J., Kapelak, S. M., & Malmi, R. A. (1976). Integration of discrete verbal units in recognition memory. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 293–300. http://dx.doi.org/10.1037/0278-7393.2.3.293

van den Honert, R. N., McCarthy, G., & Johnson, M. K. (2017). Holistic versus feature-based binding in the medial temporal lobe. *Cortex, 91,* 56–66. http://dx.doi.org/10.1016/j.cortex.2017.01.011

Wallace, W. P. (1965). Review of the historical, empirical, and theoretical status of the von Restorff phenomenon. *Psychological Bulletin, 63,* 410–424. http://dx.doi.org/10.1037/h0022001

Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General, 131,* 48–64.

Wolfe, J. M., & Cave, K. R. (1999). The psychophysical evidence for a binding problem in human vision. *Neuron, 24,* 11–17, 111–125. http://dx.doi.org/10.1016/S0896-6273(00)80818-1

Wolfe, J. M., Võ, M. L-H., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes involves selective and nonselective pathways. *Trends in Cognitive Sciences, 15,* 77–84. http://dx.doi.org/10.1016/j.tics.2010.12.001

Yassa, M. A., & Stark, C. E. (2011). Pattern separation in the hippocampus. *Trends in Neurosciences, 34,* 515–525. http://dx.doi.org/10.1016/j.tins.2011.06.006

(*Appendix follows*)

Appendix

Testing Working-Memory Encoding As a Possible Cause of Binding Failures

Experiment A1

In Experiments 1A–B, participants were required to retrieve the studied items from visual long-term memory, as they studied all 240 items before being tested. However, it is possible that the observed binding errors have nothing to do with constructive processes in long-term memory, but occur earlier—for example, participants never encode the items into long-term memory with state information bound to exemplar information at all; the binding failure occurs at perception or in working memory. In Experiment A1, we tested this possibility. Participants studied only the pairs of items from each category and were immediately tested on their memory.

Method

Participants. Following the power requirements described in Experiment 1A, a different set of 20 psychology students of the Higher School of Economics (19 female; age: M = 21.05 years, SD = 3.38) took part in the experiment for extra course credits or for a compensation equivalent to approximately $3. All participants reported having normal or corrected-to-normal vision and no neurological problems. Before the beginning of the experiment, they provided written informed consent.

Apparatus and stimuli. Apparatus and stimuli were the same as in Experiment 1A.

Procedure. Like Experiments 1A–B, Experiment A1 also consisted of exemplar-state and exemplar tasks. Each participant was exposed to both tasks. The order of the tasks was counterbalanced across participants. These two tasks were always preceded by a practice task that was a shortened version of the exemplar task with seven categories (not used in the main experiment).

As in Experiments 1A–B, the participants had to study a pair of exemplars from each category. The critical procedural difference from Experiment 1A was in how study and test were arranged. Both exemplars from the same category were presented in a row (each image was presented for 1 s with a 1-s blank interval) and then immediately tested. In particular, 1 s after the second exemplar presentation, memory for this pair of items was tested using the two simultaneous 2AFC trials (Figure A1).

Design and analysis. As in Experiment 1A from the main text, the paired exemplars in exemplar-state task could be shown in the same state or two different states (60 pairs in the same state and 60 pairs in different states). In the exemplar task, 60 pairs of exemplars were presented and tested. The method of analysis was the same as in Experiment 1A.

Results

Accuracy in remembering exemplars. In the exemplar condition, when asked to remember exemplars without requiring state memory, participants showed very high accuracy (M = .99) that differed substantially from chance, $t(19) = 223.55$, $p < .001$, $BF_{10} = 8.53 \times 10^{29}$, $d = 49.99$, 95% CI [33.43, 65.23], suggesting they remembered exemplar information very well.

Accuracy in remembering state. As in Experiment 1A from the main text, we ran two tests to estimate how well participants remembered state information on its own. First, we examined performance in picking the correct state when both objects were shown in the same state. Participants were excellent at choosing this state, $M = .96$, $t(19) = 44.72$, $p < .001$, $BF_{10} = 2.42 \times 10^{17}$, $d = 10.00$, 95% CI [6.81, 13.08], suggesting they have very good memory for states when binding was not required (Figure A1).

Second, we tested whether participants were good at discriminating between the condition where the items were shown in the same state versus in different states. We found that the proportion of the time participants selected the same two states for the two exemplars was much higher for the items that actually were presented in same states as compared to presented in different states, $t(19) = 33.55$, $p < .001$, $BF_{10} = 1.49 \times 10^{15}$, $d_z = 7.50$, 95% CI [5.09, 9.82]. In both conditions, the proportions differed from chance level .50 and these differences were almost symmetrical (same states: $M = .94$, $t(19) = 22.74$, $p < .001$, $BF_{10} = 1.62 \times 10^{12}$, $d = 5.86$, 95% CI [3.41, 6.74]; different states: $M = .07$, $t(19) = 32.43$, $p < .001$, $BF_{10} = 8.11 \times 10^{14}$, $d = 7.25$, 95% CI [4.91, 9.50]; Figure A1). This provides more evidence that the participants had very good memory for the states of the studied objects.

(*Appendix continues*)

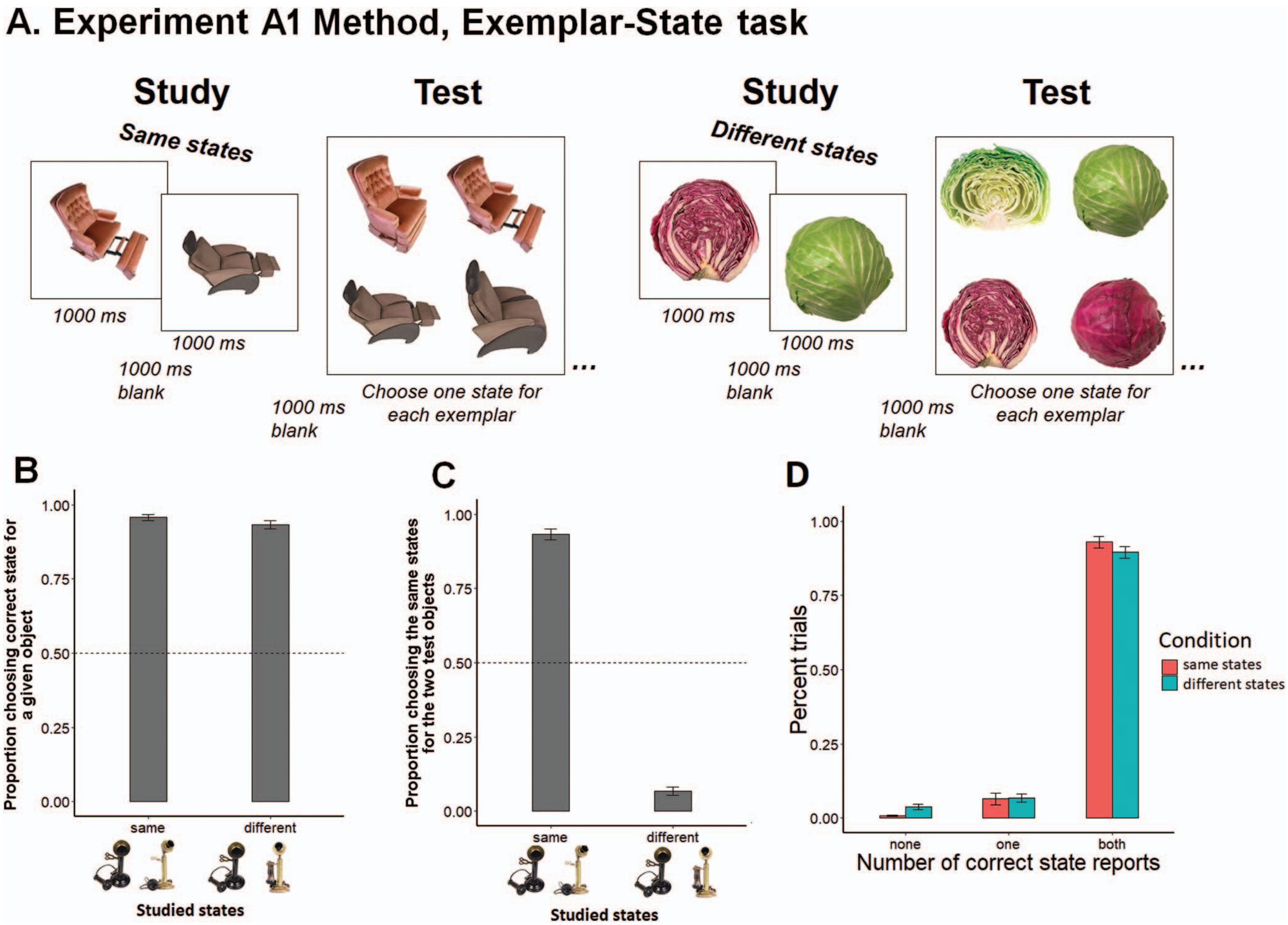

Figure A1. Methods (A) and results (B–C) of Experiment A1. (A) Experiment A1 was a working memory version of Experiments 1A–B: Two exemplars of one category were always shown in a row and followed by an immediate test. The two exemplars could be in the same or different states. (B) Proportions choosing correct state for a given exemplar when the two studied objects were shown in the same state (left; doesn't require binding) or different states (right; requires binding); (C) Experiment A1, proportions choosing the same states for the two test objects (regardless of whether these states are correct or incorrect). Error bars denote 95% CI's. Dashed lines show chance levels. See the online article for the color version of this figure.

Accuracy in exemplar-state binding. To test how well our participants remembered bound objects in an immediate working memory test, we compared their accuracy in reporting the states of the exemplars shown in different states with their accuracy in reporting the states of the exemplars shown in the same states. Memory for states of exemplars shown in different states (M = .93) was only slightly worse than memory for exemplars shown in the same state (M = .96; comparison: $t(19) = 2.66$, $p = .016$, $BF_{10} = 3.54$, $d_z = .59$, 95% CI [.11, 1.06]; Figure A1), suggesting a small number of errors can be explained by the difficulty at ascribing the remembered states to the correct exemplars even in immediate memory. Yet, the proportion correct was far above chance when participants needed to report the states of exemplars that had been presented in the different states (.50; one-sample $t(19) = 31.04$, $p < .001$, $BF_{10} = 3.75 \times 10^{14}$, $d = 6.94$, 95% CI [4.71, 9.10]). Performance was much better than the same condition in Experiments 1A–B, which was, as reported above, near chance (and consistent with the null hypothesis according to the Bayes factor analysis). We conclude that the observers do remember both exemplars, states, and the way the states went with the particular exemplars in immediate working memory tests. Although a small percentage of errors can be explained by binding difficulties even immediately after encoding, these errors appear to be much rarer than in Experiment 1A–B. This provides evidence that most of binding errors observed in Experiment 1A–B are likely to occur during retention and retrieval from visual long-term memory.

Received December 5, 2018
Revision received June 28, 2019
Accepted June 30, 2019 ■