NATIONAL RESEARCH UNIVERSITY HIGHER SCHOOL OF ECONOMICS

As a manuscript

Yuri Markov

The Structure of Visual Working Memory Units

PhD Dissertation
for the purpose of obtaining academic degree
Doctor of Philosophy in Psychology

Academic supervisor:

Doctor of Psychological Sciences

Igor Utochkin

Moscow, 2022

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ
«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»

На правах рукописи

Марков Юрий Алексеевич

Структура единиц хранения информации в зрительной рабочей памяти

Диссертация
на соискание ученой степени
кандидата психологических наук

Научный руководитель:
доктор психологических наук
Уточкин Игорь Сергеевич

Москва, 2022

The dissertation was prepared at the Laboratory for Cognitive Research, National Research University Higher School of Economics.

Three published articles were selected for the defense:

1. Markov, Y.A., Utochkin, I. S. (2022). Effects of item distinctiveness on the retrieval of objects and object-location bindings from visual working memory. Attention, Perception, & Psychophysics
2. Markov, Y.A., Utochkin, I. S., & Brady T. F. (2021). Real-world objects are not stored in holistic representations in visual working memory. Journal of Vision, 21(3): 18, 1–24. DOI: 10.1167/jov.21.3.18.
3. Markov, Y.A., Tiurina, N.A., & Utochkin, I.S. (2019). Different features are stored independently in visual working memory but mediated by object-based representations. Acta Psychologica, 197, 52-63. DOI: 10.1016/j.actpsy.2019.05.003

The results are also published in the following articles on this topic:

4. Markov, Y.M., Tiurina, N.A., Stakina, Y.M., & Utochkin, I.S. (2017). The capacity and precision of visual working memory for objects and ensembles. Psychology. Journal of HSE, 14(4), 735-756. DOI: 10.17323/1813-8918-2017-4-735-755

Table of content

1. Introduction

General research problem

At every second, we perceive and interact with the complex and rich world that contains various objects with many details. Working memory is a highly limited cognitive system (Cowan, 2001; Miller, 1956) that allows us to store and operate information about the perceived world immediately accessible for the ongoing task (Baddeley, 1986; Baddeley & Hitch, 1974). Visual working memory operates visual information and, as a subsystem of working memory, is also limited (Luck & Vogel, 1997). It is essential to understand the nature of the representations maintained by visual working memory to build a comprehensive theory of visual working memory. For the last several decades, there has been a long-lasting debate about the structural units of visual working memory. Numerous studies have provided evidence that both unitary objects (Cowan, Chen, & Rouder, 2004; Kahneman, Treisman, & Gibbs, 1992; Lee & Chun, 2001; Luck & Vogel, 1997; Luria & Vogel, 2011; Treisman, 1999; Vogel, Woodman, & Luck, 2001; Xu, 2002; Xu & Chun, 2006; Zhang & Luck, 2008) and separate features (see Brady, Konkle, & Alvarez, 2011, for review; Bays, Catalao, & Husain, 2009; Bays, Wu, & Husain, 2011; Fougnie & Alvarez, 2011; Fougnie, Cormiea, & Alvarez, 2013; Pertzov, Dong, Peich, & Husain, 2012; Shin & Ma, 2017; Wang, Cao, Theeuwes, Olivers, & Wang, 2017; Wheeler & Treisman, 2002) could be the units of visual working memory. How could different studies come to such different conclusions? What are the core units of visual working memory? Does visual working memory store whole objects representations or distinct features? Is it possible that neither objects, nor features but something more complex is a unit of visual working memory? Or does it depend on the task? How is information retrieved in various tasks from visual working memory? These are the central questions to the topic of the current work, which characterize the **problem of research**.

The **main aim** of this PhD thesis is to study the structure of visual working memory representations.

Research goals

- To analyze current research on the topic of visual working memory representations

- To empirically test feature- and object-based units of visual working memory
- To conduct base of images of real-world objects and test visual working memory for real-world objects
- To test retrieval from visual working memory under different tasks

Methodological and theoretical basis of the current work

The dissertation is based on several theoretical frameworks: feature-integration theory of visual perception and attention (Treisman, 2006); object-based visual working memory theories (Luck & Vogel, 1997; Zhang & Luck, 2008); resource-based models of visual working memory (Bays & Husain, 2008; Bays, Catalao & Husain, 2009; Bays, 2014; Bays, 2015; Schneegans & Bays, 2017); interference model of visual working memory (Oberauer & Lin, 2017); hierarchical encoding theory in visual memory (Brady, Konkle, Alvarez, 2011; Brady, Alvarez, 2011); target confusability competition model (Schurgin, Wixted, & Brady, 2020).

Methods of the research

Laboratory psychophysical experiments using methods modified for research tasks: continuous report task (Wilken & Ma, 2004; Zhang & Luck, 2008), exemplar-state task (Utochkin & Brady, 2020). We used descriptive statistics, RM ANOVA, t-tests to analyze the results. We used mixture models to process raw data (Zhang & Luck, 2008; Suchow, Brady, Fougnie, & Alvarez, 2013).

Summary of scientific novelty

- We showed that the recall of object features from working memory depends on within-, not cross-dimension load suggesting independent memory capacities for different features. Importantly, we also showed that this cross-dimensional independence is violated when different features are spatially separated and clearly belong to different objects, suggesting that object-based representations play the role of a mediator that decreases interference between the contents of visual working memory.
- We reported for the first time binding errors between representations of complex and meaningful features of real-world objects in visual working memory. These

binding errors manifested as failures to recognize which exemplar of an object from a given basic category went with which state. This suggests that even real-world objects are not stored holistically in visual working memory.

- We demonstrated that the distinctiveness of remembered objects differently affects their retrieval from visual working memory depending on a retrieval task. Specifically, the distinctiveness of memoranda does not affect simple recognition (old-new judgments), but it affects memory for object-location conjunctions, such that observers confuse where which object has been presented when the objects are similar.

Theoretical significance

The theoretical significance of the current studies could be characterized by its contribution to the discussion about the representational format of visual working memory as well as models of visual working memory. It specifically adds to the understanding of how complex, real-world objects are represented and retrieved from visual working memory.

Applied significance

Working memory is a subject of high applied interest, as working memory performance is considered a powerful predictor of subsequent academic success (Alloway & Alloway, 2008) and correlates with fluid intelligence (Fukuda, Vogel, Mayr, & Awh, 2010; Unsworth, Fukuda, Awh, & Vogel, 2015). Various working memory tests are used as a diagnostic tool for assessing various neurological disorders, e.g., Alzheimer's disease (Liang et al., 2016). Our contribution to the discussion of the representational format of visual working memory can be useful to clarify what exactly visual working memory tests measure and, thus, can improve currently available tests. Also, our results could be partially used in such practice-oriented areas as User Experience/User Interface Design in order to effectively minimize working memory load during the interaction with various virtual environments.

Reliability of the research results is ensured by the use of controlled experimental procedures in accordance with the standards of psychophysics and experimental psychology. Statistical methods of data processing are selected correctly.

The data of most studies are available online on the “Open Science Framework” platform, thus, the correctness of the conclusions could be rechecked.

Statements for the defense

- Individual objects are not represented holistically in visual working memory. Rather, their meaningfully separable feature dimensions (be they basic visual properties such as color or orientation or properties of real-world objects – exemplar or state features) can be represented relatively independently in visual working memory.
- Independent feature storage can, nevertheless, be part of the more complex hierarchical organization of visual working memory. This hierarchical organization implies that the information about independent features is accessed as a primary representational format, but the availability of whole-object information (e.g., when different features belong to the same location) can be additionally used to reduce interference from different features being remembered independently. "Feature bundles" are hierarchical and core units of visual working memory.
- The access to the representation is highly dependent on the task. Interference caused by the similarity of items could affect object-location retrieval rather than object recognition. These differences in accessibility and discriminability could be explained by the difference in target-nontarget familiarity in the two tasks.

Data collection and apparatus

We conducted ten separate experiments, with 208 observers taking part in these experiments. The observers were tested at the Cognitive Research Laboratory (HSE University, Moscow, Russia). Experiments were developed and presented via PsychoPy (Peirce et al., 2019) for Linux Ubuntu on a standard CRT monitor with a refresh frequency of 75 Hz and 1,024 × 768-pixel spatial resolution.

Approbation of the research

The results of the present work have been publicly presented in talks and posters:

- Vision Sciences Society 16th Annual Meeting (2016, St. Pete Beach, USA), *The compression of bound features in visual short-term memory*

- Theoretical and applied problems of cognitive psychology (2016, Russia), *Compression and binding in visual short-term memory*
- Vision Sciences Society 17th Annual Meeting (2017 St. Pete Beach, USA)*, An effect of categorical similarity on object-location binding in visual working memory*
- Vision Sciences Society 18th Annual Meeting (2018, St. Pete Beach, USA), *Real-world objects are not stored in bound representations in visual working memory*
- 41st European Conference on Visual Perception (2018, Trieste, Italy)*, Object distinction and object-location binding as sources of interference in visual working memory*
- Virtual Working Memory Symposium (2020, online, USA), *Different features are stored independently in visual working memory but mediated by object-based representations*
- Virtual Working Memory Symposium (2021, online, USA), *What allows an object to escape attribute amnesia?*
- 43rd European Conference on Visual Perception (2021, online)*, JURICS Stimulus base - Joint Universal Real-world Images with the Continuous States*

Six colloquium talks have been presented in the HSE Laboratory for Cognitive Research (2019), Cognitive Research Seminar HSE University (2019), Vision and Memory Laboratory at University of California San Diego (2019), Laboratory of Psychophysics École polytechnique fédérale de Lausanne (2020), Visual Attention Lab, Harvard (2021), Fougnie Lab, NYUAD (2021).

2. Features vs. objects as units of visual working memory

In their foundational study, Luck and Vogel (1997) found that the total number of presented variable features (at least 4 per object) does not affect performance in change detection, a typical visual working memory task (see also Vogel, Woodman, & Luck, 2001), whereas the number of separate objects these features belongs to limits the estimated capacity of visual working memory. Luck and Vogel concluded that objects rather than features are units of visual working memory and estimated visual working memory capacity as about 3-4 objects (see also, Cowan, 2001). These results support the "strong" object hypothesis, which states that visual working memory is restricted only by a number of objects, while features do not affect capacity independently from the objects and can only be lost when the whole object is forgotten. These findings are also in line with 'slot' models of visual working memory (Rouder et al., 2008). However, these strong claims of Luck and Vogel (1997) that the number of objects is the main limiting factor of working memory capacity were not reliably supported by later studies (Oberauer & Eichenberger, 2013; Hardman & Cowan, 2015; Wheeler & Treisman, 2002). According to those studies, increasing the number of features per object considerably reduces performance in visual working memory tasks, at least if these features are from the same sensory dimension, such as in multicolor objects (Olson & Jiang, 2002; Wheeler & Treisman, 2002; Oberauer & Eichenberger, 2013; Hardman & Cowan, 2015). The estimated capacity of visual working memory is affected by the complexity of memoranda (Alvarez & Cavanagh, 2004) and by the similarity between studied objects and those used at the retrieval stage (Awh, Barton, & Vogel, 2007). Moreover, different features of the same objects can be independently forgotten in working memory (Fougnie & Alvarez, 2011; Fougnie, Cormiea, & Alvarez, 2013) or "migrate" between representations of different objects or locations, making observers commit "swap" errors (false recall of features from different objects as belonging to the same object) or report non-targets as being presented at probed locations instead of targets (Bays, 2016; Bays, Catalao, & Husain, 2009; Bays, Wu, & Husain, 2011; Emrich & Ferber, 2012; Pertzov et al., 2012; Oberauer & Lin, 2017). On the other hand, object representations still seem to play a role in organizing information in visual working memory. For example, it was shown that feature recall could benefit from being presented within the same rather than different objects: multiple features of a single object are easier to recall than the same set of features separated across multiple objects (Fougnie et al., 2010; Fougnie et al., 2013).

As can be seen, there is still ongoing debate about the structure of visual working memory units, with new evidence for discrete slot representations (e.g., Adam, Vogel, & Awh, 2017) and against them (e.g., Oberauer, 2021). Therefore, it is important to understand how object-based benefit can occur in visual working memory along with a lot of evidence of independent feature storage.

2.1 Independent processing of features mediated by object-based representation

Article selected for the defense: Markov, Tiurina, Utochkin, 2019

But how do we process different features of the same object? While the previous studies showed that visual working memory could have separate capacities for individual features (Wang et al., 2017; Shin & Ma, 2017), it is not obvious how these features form object representations and interact. These independent feature capacities could be a reflection of purely feature-based representations of visual working memory, suggesting that we should observe feature independence for both ″bound″ (different feature dimensions belongs to the same object) and separate (different feature dimensions belong to different objects) objects. Alternatively, given the fact that feature recall benefits from the features being presented within the same object (Fougnie et al., 2010), could object representation be a mediator of independent feature storage in visual working memory?

In order to deeply investigate the relationship between memories for features from different dimensions and potential whole-object representations in visual working memory representations, we have run three experiments using a continuous report paradigm (Wilken & Ma, 2004; Zhang & Luck, 2008; Bays et al., 2009) that allowed us to estimate the capacity and precision of visual working memory representations. In each trial, we asked participants to memorize a set of items of different colors and orientations presented for 500 ms and then, after a short delay (1000 ms), report either the color (50% trials) or the orientation (another 50% trials) of a cued item (see Figure 1B). Participants had to adjust the target feature of the cued item to match the sample feature presented at that location in the original memory display. The participants were not warned about the target feature until the report; thus, they had to remember both features. Within each experiment, we manipulated feature load orthogonally for each feature dimension: the set could have either a single feature value (color/orientation) for all items or three different values. Hence, there were four possible conditions (see Figure 1A): (1) both color and orientations identical, (2) only orientation identical (colors different), (3) only color identical (orientations different), (4) all features different. Also, we included a control condition, with only one object presented on the screen (two objects in Experiment 2). Importantly, how colors and orientations were distributed among items was varied between experiments. In Experiment 1, colors and orientations were combined to form two-feature objects (isosceles randomly colored triangles turned in random directions). In Experiment 2, colors and orientations were separated into spatially distinct objects (colored circles and white oriented triangles). In Experiment 3, colors and orientations belonged to different but spatially overlapping objects (colored circles overlying white triangles).

As in a typical continuous report task, in each trial, we estimated the response error along a target circular dimension (color wheel or orientation wheel). The error was calculated as an angular difference between the correct feature value and that reported by a participant. Error distributions across trials were then analyzed separately for each condition and participant using the mixture model (Zhang & Luck, 2008; Suchow, Brady, Fougnie, & Alvarez, 2013). The mixture model decomposes the overall error distribution into two component distributions: a circular normal (von Mises) distribution around 0 (correct answer) and a uniform distribution ranging over the whole circular feature space. It is assumed that the von Mises component reflects responses based on an imperfect but existent memory representation of the probed item, whereas the uniform component reflects random guesses on items that are not in memory. Two parameters of the error distributions were estimated: the standard deviation (*SD*) of the von Mises distribution and the area below the uniform component of the mixed distribution (Zhang & Luck, 2008, but see Schurgin, Wixted, Brady, 2020). The SD parameter is an estimate of memory precision, and the area below the uniform distribution is the probability of guessing (termed P_{guess}). We used the reverse P_{guess} to estimate P_{memory}, the probability that an item is stored in visual working memory: 1- P_{guess}. Overall, we found the best performance (smallest *SD* and highest P_{memory}) in the condition with all identical features for both color and orientation reports (see Figure 1C). Importantly, the performance in this condition was comparable with that in the control condition with only one object presented (two separate objects in Experiment 2). When any of the feature dimensions took different values across objects, this significantly decreased performance for this feature dimension (increase in *SD* and decrease in P_{memory}), but it did not affect performance for another dimension keeping it on the same level as in the control condition. That is, manipulating feature load within each dimension independently affected that but not the other dimension. This pattern of independence was observed in all experiments, except for Experiment 2. In Experiment 2, we found that color load decreased performance not only in color recall but also for orientation recall: when all color objects were the same, reports about oriented objects were more precise compared to the condition when the color objects were different; also, when color objects were different, the probability of remembering orientation (P_{memory}) decreased compared with all features being identical.

In sum, we found no cross-dimensional interference in Experiments 1 and 3, where colors and orientations were represented as features of the same objects or as features of spatially overlapping objects, which could instantiate object-like units (Rensink, 2000; Trick & Pylyshyn, 1993; Wolfe & Bennett, 1997; Xu, 2002). However, we found interference between different feature dimensions in Experiment 2, where features were divided into distinct objects by clear spatial separation. These results are in line with the hypothesis that the representational units of visual working memory are not whole objects or totally separate features but rather hierarchical representations – "feature bundles" (e.g., Brady et al., 2011). In this view, both object information and feature information are available (see also, Qian, Zhang, Liu, Lei, 2019), but on different levels of hierarchy. We suggest that features are stored independently in visual working memory, but the object level of hierarchy could play the role of a mediator, decreasing cross-dimensional interference (Oberauer & Lin, 2017) and supporting the proper allocation of available resources (Bays, 2015; Wilken & Ma, 2004).

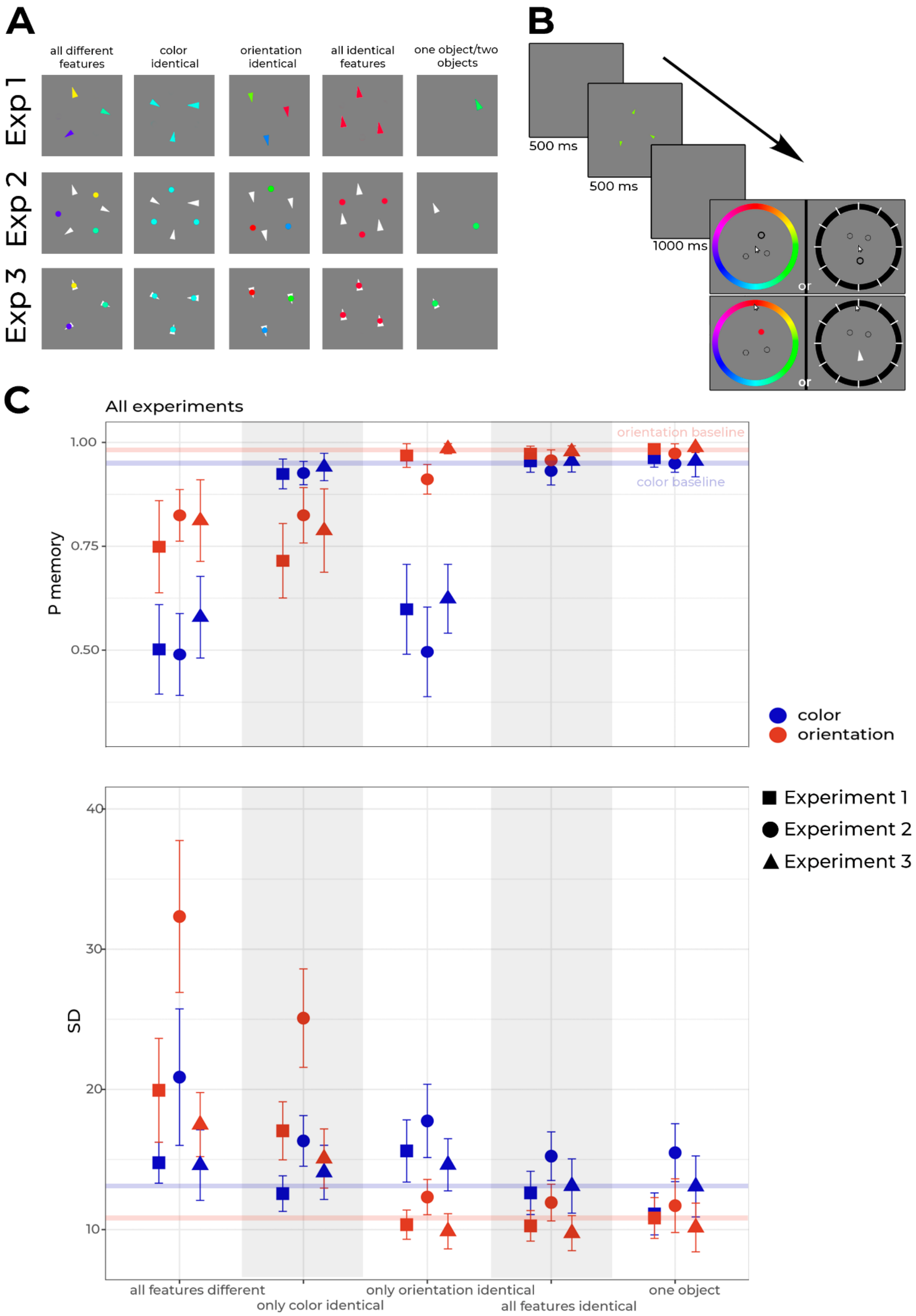


Figure 1. A: Examples of stimulus sets and conditions used in Experiments 1, 2, & 3 from Markov et al. (2019). B: The time course of a typical trial in Experiment 1. C: Results of Experiment 1-3: P_{memory} and *SD* of the mixture models as a function of Sample type and Experiment. Error bars depict 95% CIs. Adapted from Markov, Tiurina, & Utochkin 2019.

3. Representation of real-world objects in visual working memory

Most of the studies investigating the representational format of visual working memory use simplistic objects with easily manipulated independent features (geometric shapes, colors, orientations, and so on). But how are real-world objects stored in visual working memory? Real-world objects are usually complex and have a variety of features that are often hard to define and manipulate in an experiment independently. Intuitively, it seems that real-world objects are holistic, and it is possible only to remember the whole object or forget it completely. So, does working memory for real-world objects differ from working memory for simple stimuli? According to numerous studies (Asp, Störmer & Brady, in press; Brady, Störmer, & Alvarez, 2016; Brady & Störmer, 2020; Brady & Störmer, in press; Starr et al., 2020), working memory capacity for real-world objects is not so fixed as for simple stimuli. Moreover, real-world objects have some special properties that provide additional information about them: real-world size (Konkle & Oliva, 2007; Konkle & Oliva, 2012; Long, Konkle, Cohen, & Alvarez, 2016; Long, Moher, Carey, & Konkle, 2019; Wang, Janini, & Konkle, 2022), spatial position in the scene (Draschkow, Võ, 2017; Kaiser, Stein, & Peelen, 2015; O'Donnell, Clement, & Brockmole, 2018), object category (Konkle et al., 2008; Cohen et al., 2014; Markov & Utochkin 2022), familiarity and typicality that can be related to observers' specific expertise for objects from certain categories (Curby & Gauthier, 2007; Curby, Glazek, & Gauthier, 2009; Janini & Konkle, 2019; Xie & Zhang, 2017a; Xie & Zhang, 2017b; Starr et al., 2020). These can be used to enhance visual working memory via increased distinctiveness or via proper chunking of items. All in all, there is clear evidence that representations of real-world objects differ from those of simple stimuli. Here, we ask: Can independence of features similar to that found for simple stimuli (Wang et al., 2017; Shin & Ma, 2017; Markov, Tiurina & Utochkin, 2019) be observed for real-world objects? Can we remember complex real-world features but fail to bind them correctly? Or can we forget these features independently of each other? This question has been previously studied for visual long-term memory (Balaban, Assaf, Arad Meir, & Luria, 2020; Brady, Konkle, Alvarez, & Oliva, 2013; Spachtholz & Kuhbandner, 2017; Utochkin & Brady, 2020). Findings suggest that features of real-world objects (e.g., colors or specific states or poses in which objects appear) can be independently lost from visual long-term memory (Brady, Konkle, Alvarez, & Oliva, 2013). In another recent study (Utochkin & Brady, 2020), it was found that observers have a good memory for two complex features of real-world objects, state features (e.g., book open and book closed, see Figure 2A) and exemplar features (e.g., John Tolkien's "The Hobbit, or There and Back Again" book and William James' "The Principles of Psychology"). Although we refer to state and exemplar as complex features or object properties, they are not similar to the basic features, e.g., as described in the Feature Integration Theory, such as color or orientation (Treisman, 1996). The discrimination of these visual features is quite complex, and different kinds of exemplar or state changes can be provided by various changes in visual appearance and semantic relationships. However, distinguishing between different states and different exemplars of the same object category are important everyday tasks. Thus, in the present chapter, we study these properties of real-world objects while investigating the nature of these features is the topic of further research. At the same time, "swap" errors (participants report incorrect state-exemplar combinations) took place frequently, suggesting that the state information and the exemplar information are represented independently and not in a holistic manner.

Since the evidence of feature independence has been previously observed in visual long-term memory, we can ask: Is this a property of long-term memory organization that leads to the features being stored and/or forgotten independently? Or does the information about objects is consolidated into long-term memory as a set of independent features? In other words, can feature independence do with the way visual working memory represents real-world objects? We addressed this question in our following study.

3.1 Non-holistic representation of real-world objects in visual working memory

Article selected for the defense: Markov, Utochkin, Brady, 2021

In order to investigate real-world object representations in visual working memory, we adapted the paradigm from Utochkin and Brady (2020). In that study, observers had to memorize lists of objects from various basic categories. Each category was represented by two exemplars (e.g., coffee mug A and coffee mug B), each shown in the same state (e.g., both coffee mugs full) or in different states (e.g., empty coffee mug A and full coffee mug B). The observers were then asked to recall in which state each exemplar had been presented. Utochkin and Brady (2020) found that observers had no difficulties reporting exemplar-state combinations when the original states were the same, but the observers were at the chance when the original states were different (although there was evidence that the observers remembered the exemplars and the states on their own).

In the present study, we used a similar manipulation with exemplars and states in a typical visual working memory experimental setting. We used a stimulus set of object images from different basic categories (the original stimulus set from Brady et al., 2013 with an additional subset of images never used before), where each category was represented by four images, two different exemplars in two different states. For instance, a whole red apple, a whole green apple, a cut red apple, and a cut green apple (see Figure 2A). Critically, we were interested in how often observers could correctly remember both exemplars and states of presented objects but incorrectly assign states to the exemplars making a «swap» or a binding error. For instance, if the whole green apple and the cut red apple were presented and an observer reports having seen the whole red apple and the cut green apple, this is what we label a "swap". In all our experiments, we used modifications of the same method. In each trial, we presented four to be remembered real-world objects from two different categories (e.g., two different apples and two different laptops) for three seconds (see Figure 2B). After a blank one-second retention interval, two pairs of test images were presented, such that one image in each pair was always new (foil) and another was always old (target). All test images were from the same category. Therefore, the test images were four instances of objects from that category. The participants had to choose an old image from each pair, so they made a double two-alternative forced-choice in each trial.

Experiments 1 and 2 consisted of the exemplar memory task and the exemplar-state memory task. In the exemplar-state task, two exemplars from one category and two exemplars from another category were presented in a memory set. Within each category, the exemplars could be presented in the same state or in different states. At test, memory for only one of the categories was tested. Each test pair included two possible states of the same exemplar. The observers had to choose a state in which a given exemplar had been presented in the memory set. Following the logic of Utochkin and Brady (2020), if real-world objects are stored not as holistic representation in visual working memory, we expect worse performance for objects originally presented in different states because observers have to remember not only exemplars and states but also to correctly "bind" particular states with particular exemplars. In contrast, the need to bind is not a big difficulty if the exemplars are originally presented in the same state: In this case, remembering a common state for both exemplars is sufficient to perform the state-exemplar task. The exemplar task was used to obtain baseline performance for exemplar memory. In this task, the memory set also included four items, two exemplars from two categories. At test, two pairs of objects from one of the categories were shown, each pair including one old and one new exemplar. The participants had to choose exactly an old exemplar in each pair. In Experiment 1, the exemplar task and the exemplar-state task were presented in separate blocks. In Experiment 2, trials from exemplar and exemplar-state tasks were randomly mixed with discouraging observers from focusing on specific exemplar or state features.

In both experiments, the observers demonstrated good memory for exemplars, as shown by the exemplar task (86% correct). They also had a reasonably good memory for states, as they were mostly correct at choosing two same states when the objects had been presented in the same states, and they were correct at choosing two different states when the objects had been presented in different states (see Figure 2C). However, in terms of reporting exemplar-state conjunctions, performance was significantly worse in the condition where items were presented in different states compared to the same states. Interestingly, in about 15% different-state trials, observers successfully reported the states as being different but chose wrong the exemplars for these two states (committed 'swaps'). These results indicate that, even when observers have some memory for exemplars and states separately, the binding errors occurred between these two kinds of features. Additional analyses demonstrated that this result is consistent across categories, that not tested items presented in the same states improved performance, that our results could not be explained by inter-object visual similarity estimated by a pre-trained convolutional neural network (Brady & Störmer, 2020; Simonyan & Zisserman, 2014). In Experiment

3, we tested how location update at the test could influence binding errors. In two conditions, we presented items at the test at the same location as during the presentation, or locations of tested items were switched. We found that the location update did not cause more binding errors but decreased overall performance. This suggested that during location update, exemplars and states could not be bound independently to the locations and that location updating appears to act on the unitized, fully bound representations.

Overall, our results suggest that real-world objects are also prone to binding errors, like simple objects, confirming the basic non-holistic nature of object representations in visual working memory.

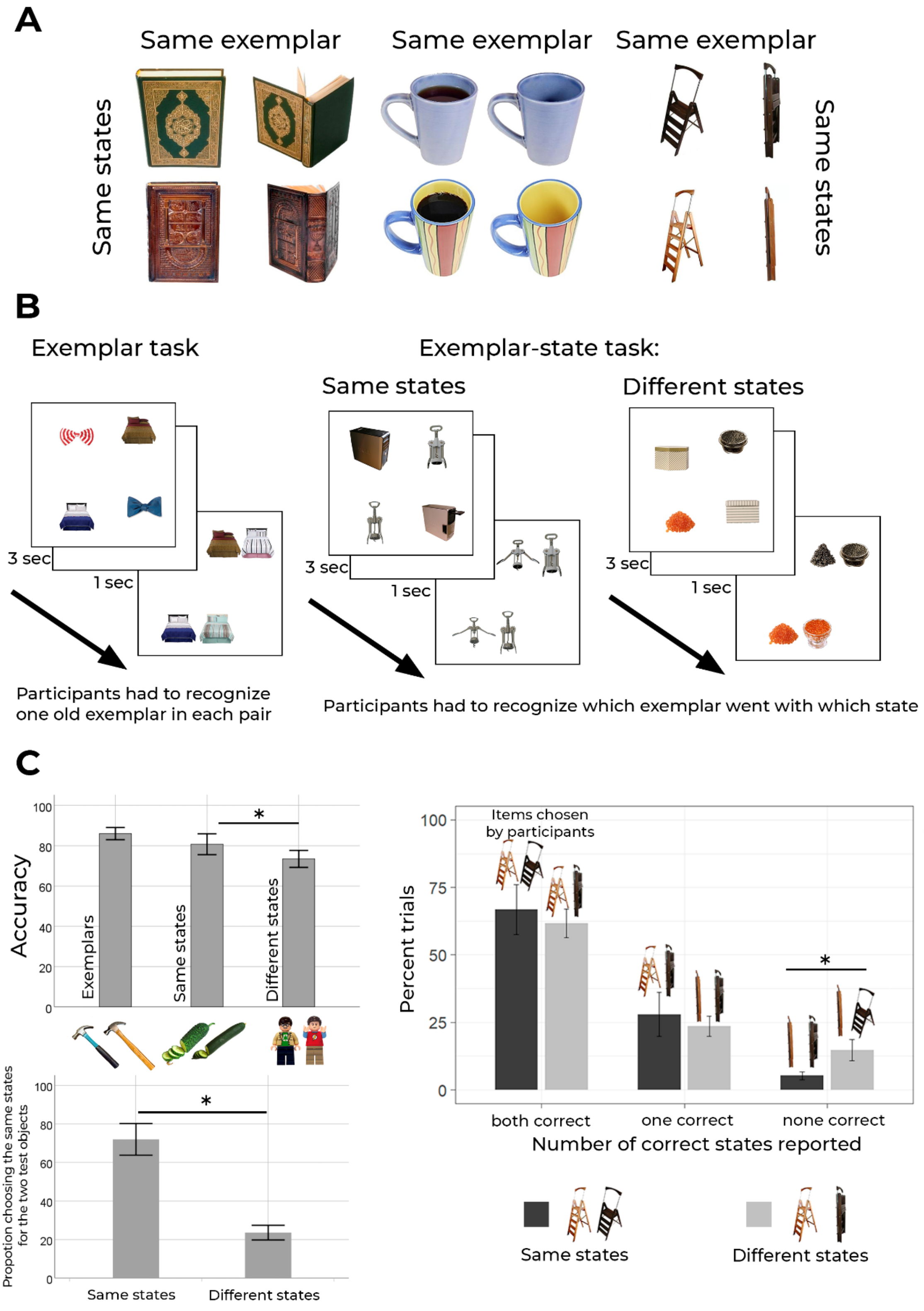


Figure 2. General methods and principal results by Markov, Utochkin, & Brady, 2021. A: Example of two different exemplars in two different states. B: The time course of a typical trial in Experiment 1 in the exemplar–state task and the exemplar task. C: Results of Experiment 1: Overall accuracy, State memory accuracy, and choosing both, one, or no correct states for exemplars. Error bars depict 95% CIs. Adapted from Markov, Utochkin, & Brady, 2021.

4. Retrieval of information from visual working memory

As can be seen from the previous chapter, interactions between different objects and their features can give us useful information about the way information in visual memory is organized. Here, we further investigate interactions between objects in visual working memory and focus on two aspects. First, real-world objects are typically remembered, not in isolation, that is, context is important. Visual working memory is frequently considered a spatially organized system (Logie, 2003; Magen & Emmanouil, 2019). Thus, we store information not only about what we saw but also about where. Spatial information plays an important role in the binding process (Swan & Wyble, 2014; Treisman, 1996) and is relevant for many tasks involving visual working memory. Remembering objects at locations (that we term object-location memory) is also prone to binding errors similar to those between different features of multiple objects (such as exemplar and state features), but they occur between objects representation and locations (Bays et al., 2009; Dent & Smyth, 2005; Hollingworth & Rasmussen, 2010; Pertzov, Dong, Peich, & Husain, 2012; Postma & De Haan, 1996; Treisman, 1996; Toh, Sisk, & Jiang, 2020; but see Pratte, 2019). For example, when a person puts a smartphone in their left jeans pocket and a wallet in their right jacket pocket, the person can subsequently recall which items they put in the pockets but can swap their locations in memory and look for the wallet in the left jeans pocket. These errors suggest that memories for objects and for scene context in which those objects have been seen are not unitized. Other than that, representations of objects can interfere with each other (feature binding errors described in the previous chapter is one example). The degree of interference as a function of inter-object relationship strongly depends on a set of factors termed distinctiveness (Hunt, 2006). Previous studies show that item distinctiveness affects performance in visual working memory tasks, but the direction of these effects can be the opposite. While some studies have suggested that high inter-item distinctiveness increases subsequent memory performance, others have suggested that high distinctiveness decreases it (Cohen et al., 2014; Jiang, Lee, Asaad, & Remington, 2016; Lin & Luck, 2009; Sims, Jacobs, & Knill, 2012). The important theoretical question that we address in the current study is how distinctiveness influences both object and object-location memory. That is, we ask how inter-item structural relationships affect object retrieval in context-free (simple object recognition) and context-dependent tasks. Does low object distinctiveness disrupt proper object-location binding, causing more swap errors? Why do swap errors for location occur? Are there significant differences between retrieval information about objects and object-location bindings?

4.1 Item distinctiveness in object and object-location memory

Article selected for the defense: Markov, Utochkin 2022

In a series of experiments, we investigated the influence of item distinctiveness on object memory, location memory, and object-location memory. We manipulated the distinctiveness of items by presenting objects that belonged to the same or different basic categories. In Experiment 1, we presented three objects located around an imaginary circumference for two seconds and asked observers to remember them and their locations (see Figure 3A). After a one-second blank retention interval, two test items were presented (always from the same category regardless of the presentation condition): one item was old (already shown in the set), and another was new. Observers were asked to choose the old item. On the next step, the observers had to localize the chosen item along a circumference so that it matched the location of this item in the original display. Feedback was given about the correctness of both recognition (whether the chosen object was old or new) and localization (how close the reported location was to the original location). We measured percent correct answers in the recognition task. For the localization task, we used a modification of the mixture model (Zhang & Luck, 2008 – see description in section 2.1) "swap model" (Bays et al., 2009). The outcome of the swap model is a set of parameters supposed to reflect various aspects of visual working memory for a given material. These parameters include the P_{memory}, (or its complement, P_{guess}), the *SD* or precision of a correctly reported memory representation (see description in section 2.1), and also the P_{swap} estimated as the area of a second von Mises distribution component with the same *SD,* reflecting the probability of reporting a really presented but not probed item. While the P_{memory} and the *SD* are characteristics of location memory (how likely and how precisely observers recall locations themselves), the P_{swap} reflects specific object-location failures (how often observers recall an object at a wrong location).

We found in the result of Experiment 1 that distinctiveness did not affect performance in the recognition task and did not affect the precision of item localization (see Figure 3B). By contrast, we found that observers made more swap errors for low-distinctive items in the localization task. Additional analysis demonstrated that this increment in non-target reports was not caused by simple forgetting of elements. We conclude that reduced item distinctiveness impaired an ability to specifically recall object-location bindings rather than abilities to recognize objects or remember locations.

In Experiment 2, we basically replicated the design of Experiment 1, but this time we tested object-location and location memory without the recognition task in order to remove a potential interference of recognition with localization. We also added a condition where we could test "pure" location memory with reduced demands on object-location binding. In that condition, we showed the same set of objects (spatially ordered pictures of hands with one, two, or three raised fingers), so that observers could successfully recall the location of any object even if they remembered the location of one object (but they still had to remember all three locations). These stimuli could be encoded by visual memory as symbols or labels and not as completely visual representations; however, we were interested in whether we could observe any location memory improvements for these simplified object-location conjunctions. We found the same results as in Experiment 1 and additionally showed that the precision of "pure" location memory was the same as in other conditions (those requiring remembering object-location conjunctions). Therefore, we conclude that task demands on remembering objects or binding them to locations did not affect the precision of localization. This corroborates our main conclusion from Experiment 1 that it is object-location memory that is specifically affected by the object distinctiveness manipulations.

In Experiments 3A and 3B, we considered two plausible explanations for the distinctiveness effects observed in Experiments 1 and 2. First, poor item distinctiveness could impair object-location memory in general and non-specifically. High demands on visual working memory to store low-distinctive items could cause a trade-off between remembering objects and object-location conjunctions. Second, object-location retrieval depends on how easily a particular target representation can be separated from non-target representations based on their inter-object differences and spatial cues. The main difference between these two accounts is what they predict about non-target reports. The first account predicts that if there are several items varying in distinctiveness from item to item (for example, some of the items are from one category and others are from another category), then non-target reports should occur randomly regardless of similarity between particular items. The second account predicts that the non-target reports specifically depend on item-to-item similarity: There should be more "swap" errors between more similar items.

In Experiment 3A, we tested object-location memory for four objects from either one, two, or four categories (Figure 3C). Instead of using a continuous localization report (as in Experiments 1 and 2), we used a four-alternative forced choice of an object by a location cue. That is, after the retention interval, the location of one of the presented objects was cued, and the observers had to choose which of the four presented objects had been presented at the cued location. In Experiment 3B, the same sets of objects were tested for recognition in a two-alternative choice task (same as in the recognition stage of Experiment 1). Similar to the previous two experiments, distinctiveness was found to affect performance in the object-location task (Experiment 3A), such that observers committed more object-location errors ("swaps") when a memory display consisted of four objects drawn from one category. At the same time, no cost to recognition performance was found (Experiment 3B), which again is in line with our previous finding (Experiment 1). We then took a closer look at the critical condition where four objects were from two different categories, that is, there always were non-targets that were more similar to the target and two non-targets that were more dissimilar with the target. Overall, observers were quite accurate at reporting correct target locations in most cases (Figure 3D). Most importantly, incorrect reports (swap errors) were distributed unevenly between non-targets as a function of their similarity to the target. We found that the non-targets from the same category as the target were chosen more often than any of the foils from the different category (Figure 3D). Therefore, we found that object-location memory is affected in a specific way which is defined by item-to-item similarity.

The results of our experiments show an interesting dissociation between the effects of object distinctiveness on simple recognition (no effect) and on object-location memory (less distinctive objects are more likely swapped). To account for this dissociation, we suggested that the crucial differences between retrieval of the object and object-location information could be explained by existing models of visual working memory as noisy representations or familiarity signals competing at multiple stages of processing (e.g., Oberauer & Lin, 2017; Schneegans & Bays, 2017; Swan & Wyble, 2014; Schurgin, Wixted, & Brady, 2020). In line with some of the previous models of attention and visual search (Duncan & Humphreys, 1989), the distinctiveness of the target and non-target plays a crucial role also in memory tasks. Since representations are all noisy, then distinctiveness affects how separable representations are relative to the noise. Here, the difference between the tasks arises from how exactly noisy familiarity judgments are made. In the 2-AFC recognition "old-new" task, the familiarity signal produced by the target is compared against that produced by a foil at the test. In this case, the distinctiveness of encoded items does not strongly affect performance in the recognition task because any single target produces a stronger signal than a foil, while comparison between different targets is not required. In the object-location task, observers had to discriminate between competing noisy representations of all items linked to various locations. In Experiment 3A, non-target items from the same category as the target were chosen more frequently than non-targets from the different category because the location cue elicited more familiarity with the former non-target. An extended signal detection model depicted in Figure 3E illustrates this theoretical idea (Macmillan & Creelman, 2005; Schurgin et al., 2020). Target-nontarget distinctiveness defines the probability of a non-target response. A location cue causes a familiarity strength for each test item drawn from a certain Gaussian distribution. The stronger the familiarity distribution is shifted to the right, the more likely a corresponding item is chosen as a target. The separation between the distributions is defined by the distinctiveness of the test items. It is easy to see that the familiarity distribution of same-category non-targets has more overlap with the target distribution than the distribution of different-category non-targets. That is, in an individual trial the same-category non-target has a greater chance to produce the strongest familiarity than any of the different-category non-targets. We, therefore, suggest that object-location errors arise due to representational competition during retrieval. The visual working memory representation itself can be considered to be a noisy signal induced by a retrieval cue (such as a location cue) which is compared against other signals provided by test alternatives. This theoretical view on visual working memory representations does not contradict our previously mentioned ideas of hierarchical bundles: These familiarity signals can arise to represent different levels of this hierarchy from separate features to objects and to groups and sets.

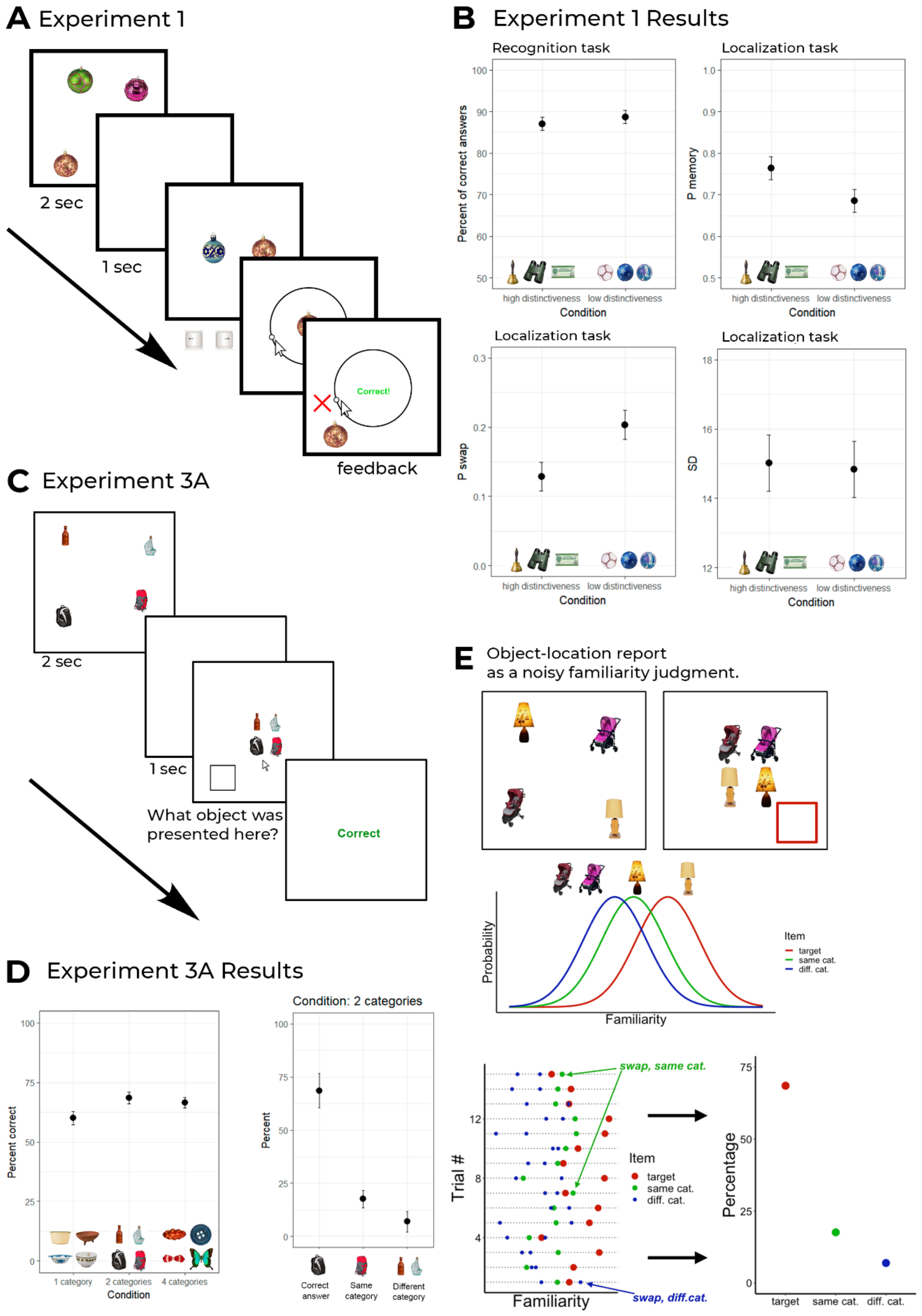


Figure 3. A: The time course of a typical trial in Experiment 1. B: Results of Experiment 1 for the recognition task (percent correct) and the localization task (P_{memory}, P_{swap}, SD). Error bars depict 95% CIs. C: The time course of a typical trial in Experiment 3A. D: Results of Experiment 3A. Percent correct for all conditions and percent of correct answers for the condition with two categories. E: Object-location report as a noisy familiarity judgment. Adapted from Markov, & Utochkin, 2022.

Conclusion

In the series of studies, we investigated the structure of visual working memory units and their interactions. We found that the units are not holistic object representations and also not completely isolated features but rather hierarchically organized feature "bundles" (Brady et al., 2011). This means that visual working memory representations in different tasks can benefit from the independence of feature representation and from their organization into whole objects. Our results suggest that these "feature bundle" structures could be applied to both simple geometrical stimuli and complex real-world objects and act as core units of visual working memory.

We suggest that information is not stored on the "shelves" in visual working memory and that retrieval of information from visual working memory depends on the task requirements on retrieval. The visual system can retrieve the information from different levels of visual working memory representations and use this information according to the current task. The visual system highly relies on competitive, noisy familiarity signals caused by retrieval cues and test alternatives available at the recall stage. In a simple recognition task (2AFC), the familiarity of the target is compared against a presented novel foil at test, and observers decide whether this or that item looks more familiar. In the tasks where the object should be remembered (bound) along with its contextual information such as location, the familiarity of the target competes with representations of other to-be-remembered items causing considerable interference between the items. Our results are in line with and elaborate the current models of visual working memory - resource-based models (e.g., Bays & Husain, 2008), interference model (Oberauer & Lin, 2017), target confusability competition model (Schurgin, Wixted, & Brady, 2020), by demonstrating similar effects for simple features and complex objects. Thus, overall, our findings suggest that visual working memory representations are flexible, hierarchical, and highly dependent on the current task.

Acknowledgments

I want to express my special thanks of gratitude to my advisor, Prof. Igor Utochkin, for providing me the opportunity to do a lot of projects and complete PhD, guiding and helping me from the start of my academic career at the Laboratory for Cognitive Research HSE. I am also thankful to the team of Cognitive Research Laboratory and Department of Psychology HSE, Prof. Vasily Klucharev, and many other teachers, colleagues, and students, and especially Prof. Maria Falikman and Dr. Elena Gorbunova for comments and suggestions on the early version of the thesis.

I want to express my sincere gratitude to Prof. Michael Herzog for the excellent opportunity to work at the Laboratory of Psychophysics EPFL, work on interesting projects together with the great team, and for navigating me in science and academia. I want to thank Dr. David Pascucci for collaboration, for many great discussions, and for a lot of help.

I want to thank Prof. Daryl Fougnie for the opportunity to be an online part of Fougnie Lab. I thank Dr. Edyta Sasin, we conducted a lot of experiments together, had great conversations, and had a lot of great ideas.

I thank Prof. Viola Störmer and Prof. Timothy Brady, and all members of Vision and Memory Lab and Perception, Attention, & Memory Laboratory for the internship in USCD, great time, and cool projects.

I also want to thank Dr. Fyodor Kondrashov, Dr. Peter Vlasov, and all team of the School of Molecular and Theoretical Biology. SMTB is the reason why I choose the academic career.

I want to thank my family. I want to thank love of my life, wife, and wonderful collaborator, Dr. Natalia Tiurina, for a lot of help and support during my research. It would not be possible without you. And thank my cat Timothy for his warmth and fluffiness.

References

Adam, K. C. S., Vogel, E. K., & Awh, E. (2017). Clear evidence for item limits in visual working memory. *Cognitive Psychology*, *97*, 79–97. https://doi.org/10.1016/j.cogpsych.2017.07.001

Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, *106*(1), 20–29. https://doi.org/10.1016/j.jecp.2009.11.003

Alvarez, G. A., & Cavanagh, P. (2004). The Capacity of Visual Short-Term Memory Is Set Both by Visual Information Load and by Number of Objects. *Psychological Science*, *15*(2), 106–111. https://doi.org/10.1111/j.0963-7214.2004.01502006.x

Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science*, *18*(7), 622–628. https://doi.org/10.1111/j.1467-9280.2007.01949.x

Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation - Advances in Research and Theory*, *8*(C), 47–89. https://doi.org/10.1016/S0079-7421(08)60452-1

Balaban, H., Assaf, D., Meir, M. A., & Luria, R. (2019). Different Features of Real-World Objects Are Represented in a Dependent Manner in Long-Term Memory. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0000716

Bays, P. M. (2014). Noise in neural populations accounts for errors in working memory. *Journal of Neuroscience*, *34*(10), 3632–3645. https://doi.org/10.1523/JNEUROSCI.3204-13.2014

Bays, P. M. (2015). Spikes not slots: Noise in neural populations limits working memory. *Trends in Cognitive Sciences*, *19*(8), 431–438. https://doi.org/10.1016/j.tics.2015.06.004

Bays, P. M. (2016). Evaluating and excluding swap errors in analogue tests of working memory. *Scientific Reports*, *6*. https://doi.org/10.1038/srep19203

Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science*, *321*(5890), 851–854. https://doi.org/10.1126/science.1158023

Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, *9*(10). https://doi.org/10.1167/9.10.7

Bays, P. M., Wu, E. Y., & Husain, M. (2011). Storage and binding of object features in visual working memory. *Neuropsychologia*, *49*(6), 1622–1631. https://doi.org/10.1016/j.neuropsychologia.2010.12.023

Brady, T. F., & Alvarez, G. A. (2011). Hierarchical encoding in visual working memory: Ensemble statistics bias memory for individual items. *Psychological Science*, *22*(3), 384–392. https://doi.org/10.1177/0956797610397956

Brady, T. F., & Störmer, V. S. (2020). Greater capacity for objects than colors in visual working memory : Comparing memory across stimulus spaces requires maximally dissimilar foils. *PsyArXiv*.

Brady, T. F., Konkle, T., & Alvarez, G. A. (2011). A review of visual memory capacity: Beyond individual items and toward structured representations. *Journal of Vision*, *11*(5), 1–34. https://doi.org/10.1167/11.5.1

Brady, T. F., Konkle, T., & Alvarez, G. A. (2011). A review of visual memory capacity: Beyond individual items and toward structured representations. *Journal of Vision*, *11*(5), 1–34. https://doi.org/10.1167/11.5.1

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: General*, *142*(3), 791–808. https://doi.org/10.1037/a0029649

Brady, T. F., Störmer, V. S., & Alvarez, G. A. (2016). Working memory is not fixed-capacity: More active storage capacity for real-world objects than for simple stimuli. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(27), 7459–7464. https://doi.org/10.1073/pnas.1520027113

Cohen, M. A., Konkle, T., Rhee, J. Y., Nakayama, K., & Alvarez, G. A. (2014). Processing multiple visual objects is limited by overlap in neural channels. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(24), 8955–8960. https://doi.org/10.1073/pnas.1317860111

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *The Behavioral and Brain Sciences*, *24*(1), 87–185. https://doi.org/10.1017/S0140525X01003922

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87–114. https://doi.org/10.1017/S0140525X01003922

Cowan, N., Chen, Z., & Rouder, J. N. (2004). Constant capacity in an immediate serial-recall task: A logical sequel to Miller (1956). *Psychological Science*, *15*(9), 634–640. https://doi.org/10.1111/j.0956-7976.2004.00732.x

Curby, K. M., & Gauthier, I. (2007). A visual short-term memory advantage for faces. *Psychonomic Bulletin and Review*, *14*(4), 620–628. https://doi.org/10.3758/BF03196811

Curby, K. M., Glazek, K., & Gauthier, I. (2009). A Visual Short-Term Memory Advantage for Objects of Expertise. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(1), 94–107. https://doi.org/10.1037/0096-1523.35.1.94

Dent, K., & Smyth, M. M. (2005). Verbal coding and the storage of form-position associations in visual-spatial short-term memory. *Acta Psychologica*, *120*(2), 113–140. https://doi.org/10.1016/j.actpsy.2005.03.004

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96*(3), 433–458. https://doi.org/10.1037/0033-295X.96.3.433

Draschkow, D., & Võ, M. L. H. (2017). Scene grammar shapes the way we interact with objects, strengthens memories, and speeds search. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-16739-x

Emrich, S. M., & Ferber, S. (2012). Competition increases binding errors in visual working memory. *Journal of Vision*, *12*(4), 1–16. https://doi.org/10.1167/12.4.12

Fougnie, D., & Alvarez, G. A. (2011). Object features fail independently in visual working memory: Evidence for a probabilistic feature-store model. *Journal of Vision*, *11*(12), 1–12. https://doi.org/10.1167/11.12.1

Fougnie, D., Asplund, C. L., & Marois, R. (2010). What are the units of storage in visual working memory? *Journal of Vision*, *10*(12), 1–11. https://doi.org/10.1167/10.12.27

Fougnie, D., Cormiea, S. M., & Alvarez, G. A. (2013). Object-based benefits without object-based representations. *Journal of Experimental Psychology: General*, *142*(3), 621–626. https://doi.org/10.1037/a0030300

Fukuda, K., Vogel, E., Mayr, U., & Awh, E. (2010). Quantity, not quality: The relationship between fluid intelligence and working memory capacity. *Psychonomic Bulletin and Review*, *17*(5), 673–679. https://doi.org/10.3758/17.5.673

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*(2), 325–347. https://doi.org/10.1037/xlm0000031

Hollingworth, A., & Rasmussen, I. P. (2010). Binding objects to locations: The relationship between object files and visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(3), 543–564. https://doi.org/10.1037/a0017836

Hunt, R. R. (2006). The Concept of Distinctiveness in Memory Research. In *Distinctiveness and memory* (p. 476). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195169669.003.0001

Isabel E. Asp, Störmer, V., & Brady, T. (2019). *Greater visual working memory capacity for visually-matched stimuli when they are recognized as meaningful*. 1–17.

Janini, D., & Konkle, T. (2019). A Pokémon-sized window into the human brain. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-019-0594-6

Jiang, Y. V., Lee, H. J., Asaad, A., & Remington, R. (2016). Similarity effects in visual working memory. *Psychonomic Bulletin & Review*, *23*(2), 476–482. https://doi.org/10.3758/s13423-015-0905-5

Kahneman, D., Treisman, a, & Gibbs, B. J. (1992). The reviewing of object files: object-specific integration of information. *Cognitive Psychology*, *24*(2), 175–219. https://doi.org/10.1016/0010-0285(92)90007-O

Kaiser, D., Stein, T., & Peelen, M. V. (2015). Real-world spatial regularities affect visual working memory for objects. *Psychonomic Bulletin & Review*, *22*(6), 1784–1790. https://doi.org/10.3758/s13423-015-0833-4

Konkle, T., & Oliva, A. (2007). Normative Representation of Objects: Evidence for an Ecological Bias in Object Perception and Memory. *Proceedings of the Twenty-Nine Annual Conference of the Cognitive Science Society*, 407–412.

Konkle, T., & Oliva, A. (2012). A Real-World Size Organization of Object Responses in Occipitotemporal Cortex. *Neuron*, *74*(6), 1114–1124. https://doi.org/10.1016/j.neuron.2012.04.036

Lee, D., & Chun, M. M. (2001). What are the units of visual short-term memory, objects or spatial locations? *Perception & Psychophysics*, *63*(2), 253–257. https://doi.org/10.3758/BF03194466

Liang, Y., Pertzov, Y., Nicholas, J. M., Henley, S. M. D., Crutch, S., Woodward, F., Leung, K., Fox, N. C., & Husain, M. (2016). Visual short-term memory binding deficit in familial Alzheimer's disease. *Cortex*, *78*, 150–164. https://doi.org/10.1016/j.cortex.2016.01.015

Lin, P.-H., & Luck, S. J. (2009). The Influence of Similarity on Visual Working Memory Representations. *Visual Cognition*, *17*(3), 356–372. https://doi.org/10.1080/13506280701766313

Logie, R. (1988). Working memory, Alan Baddeley, Oxford University Press, Oxford 1986. No. of pages: 289. Price £30.00 (Hardback), ISBN 0 19 852116 2. *Applied Cognitive Psychology*, *2*(2), 166–168. https://doi.org/10.1002/acp.2350020209

Logie, R. H. (2003). Spatial and Visual Working Memory: A Mental Workspace. *Psychology of Learning and Motivation - Advances in Research and Theory*, *42*, 37–78. https://doi.org/10.1016/S0079-7421(03)01002-8

Long, B., Konkle, T., Cohen, M. A., & Alvarez, G. A. (2016). Mid-level perceptual features distinguish objects of different real-world sizes. *Journal of Experimental Psychology: General*, *145*(1), 95–109. https://doi.org/10.1037/xge0000130

Long, B., Moher, M., Carey, S., & Konkle, T. (2019). Real-world size is automatically encoded in preschoolers' object representations. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(7), 863–876. https://doi.org/10.1037/xhp0000619

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281. https://doi.org/10.1038/36846

Luria, R., & Vogel, E. K. (2011). Shape and color conjunction stimuli are represented as bound objects in visual working memory. *Neuropsychologia*, *49*(6), 1632–1639. https://doi.org/10.1016/j.neuropsychologia.2010.11.031

Macmillan, & Creelman. (2005). Detection theory: A user's guide, 2nd ed. *Lawrence Erlbaum Associates Publishers.*

Magen, H., & Emmanouil, T. A. (2019). Spatial Organization in Self-Initiated Visual Working Memory. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02734

Markov, Y.A., Utochkin, I. S. (2022). Effects of item distinctiveness on the retrieval of objects and object-location bindings from visual working memory. *Attention, Perception, & Psychophysics*

Markov, Y.A., Utochkin, I. S., & Brady T. F. (2021). Real-world objects are not stored in holistic representations in visual working memory. *Journal of Vision, 21*(3): 18, 1–24. DOI: 10.1167/jov.21.3.18.

Markov, Y.A., Tiurina, N.A., & Utochkin, I.S. (2019). Different features are stored independently in visual working memory but mediated by object-based representations. *Acta Psychologica, 197*, 52-63. DOI: 10.1016/j.actpsy.2019.05.003

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, *101*(2), 343–352. https://doi.org/10.1037/h0043158

O'Donnell, R. E., Clement, A., & Brockmole, J. R. (2018). Semantic and functional relationships among objects increase the capacity of visual working memory. *Journal* *of Experimental Psychology: Learning Memory and Cognition*, *44*(7), 1151–1158. https://doi.org/10.1037/xlm0000508

Oberauer, K. (2021). *Little support for discrete item limits in visual working memory*.

Oberauer, K., & Eichenberger, S. (2013). Visual working memory declines when more features must be remembered for each object. *Memory & Cognition*, *41*(8), 1212–1227. https://doi.org/10.3758/s13421-013-0333-6

Oberauer, K., & Lin, H. (2017). An Interference Model of Visual Working Memory. *Psychological Review*, *124*(1), 1–39. https://doi.org/10.1037/rev0000044

Oberauer, K., & Lin, H. (2017). An Interference Model of Visual Working Memory. *Psychological Review*, *124*(1), 1–39. https://doi.org/10.1037/rev0000044

Olson, I. R., & Jiang, Y. (2002). Is visual short-term memory object based? Rejection of the "strong-object" hypothesis. *Perception and Psychophysics*, *64*(7), 1055–1067. https://doi.org/10.3758/BF03194756

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Pertzov, Y., Dong, M. Y., Peich, M. C., & Husain, M. (2012). Forgetting What Was Where: The Fragility of Object-Location Binding. *PLoS ONE*, *7*(10). https://doi.org/10.1371/journal.pone.0048214

Pertzov, Y., Dong, M. Y., Peich, M. C., & Husain, M. (2012). Forgetting What Was Where: The Fragility of Object-Location Binding. *PLoS ONE*, *7*(10). https://doi.org/10.1371/journal.pone.0048214

Postma, A., & De Haan, E. H. F. (1996). What Was Where? Memory for Object Locations. *The Quarterly Journal of Experimental Psychology Section A*, *49*(1), 178–199. https://doi.org/10.1080/713755605

Pratte, M. S. (2019). Swap errors in spatial working memory are guesses. *Psychonomic Bulletin & Review*, *26*(3), 958–966. https://doi.org/10.3758/s13423-018-1524-8

Qian, J., Zhang, K., Liu, S., & Lei, Q. (2019). The transition from feature to object: Storage unit in visual working memory depends on task difficulty. *Memory & Cognition, 47*(8), 1498–1514. https://doi.org/10.3758/s13421-019-00956-y

Rouder, J. N., Morey, R. D., Cowan, N., Zwilling, C. E., Morey, C. C., & Pratte, M. S. (2008). An assessment of fixed-capacity models of visual working memory. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(16), 5975–5979. https://doi.org/10.1073/pnas.0711295105

Rouder, J. N., Morey, R. D., Cowan, N., Zwilling, C. E., Morey, C. C., & Pratte, M. S. (2008). An assessment of fixed-capacity models of visual working memory. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(16), 5975–5979. https://doi.org/10.1073/pnas.0711295105

Schneegans, S., & Bays, P. M. (2017). Neural Architecture for Feature Binding in Visual Working Memory. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *37*(14), 3913–3925. https://doi.org/10.1523/JNEUROSCI.3493-16.2017

Schurgin, M. W., Wixted, J. T., & Brady, T. F. (2020). Psychophysical scaling reveals a unified theory of visual memory strength. *Nature Human Behaviour*, *4*(11), 1156–1172. https://doi.org/10.1038/s41562-020-00938-0

Shin, H., & Ji Ma, W. (2017). Visual short-term memory for oriented, colored objects. *Journal of Vision*, *17*(9), 1–19. https://doi.org/10.1167/17.9.12

Simonyan Karen, & Zisserman Andrew. (2015). Very deep convolutional networks for large-scale image recognition. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*. http://www.robots.ox.ac.uk/

Sims, C. R., Jacobs, R. A., & Knill, D. C. (2012). An ideal observer analysis of visual working memory. *Psychological Review*, *119*(4), 807–830. https://doi.org/10.1037/a0029856

Spachtholz, P., & Kuhbandner, C. (2017). Visual long-term memory is not unitary: Flexible storage of visual information as features or objects as a function of affect. *Cognitive, Affective and Behavioral Neuroscience*, 1–10. https://doi.org/10.3758/s13415-017-0538-4

Suchow, J. W., Brady, T. F., Fougnie, D., & Alvarez, G. A. (2013). Modeling visual working memory with the MemToolbox. *Journal of Vision, 13*(10), 9. https://doi.org/10.1167/13.10.9

Stoermer, V., & Brady, T. (2021). The role of meaning in visual working memory: Real-world objects, but not simple features, benefit from deeper processing. *Journal of Vision*, *21*(9), 2644. https://doi.org/10.1167/jov.21.9.2644

Toh, Y. N., Sisk, C. A., & Jiang, Y. V. (2020). Effects of changing object identity on location working memory. *Attention, Perception, & Psychophysics*, *82*(1), 294–311. https://doi.org/10.3758/s13414-019-01738-z

Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology*, *6*(2), 171–178. https://doi.org/10.1016/S0959-4388(96)80070-5

Treisman, A. (1999). Solutions to the binding problem: Progress through controversy and convergence. *Neuron*, *24*(1), 105–110. https://doi.org/10.1016/S0896-6273(00)80826-0

Treisman, A. (2006). How the deployment of attention determines what we see. In *Visual Cognition* (Vol. 14, Issues 4–8). https://doi.org/10.1080/13506280500195250

Trick, L. M., & Pylyshyn, Z. W. (1993). What Enumeration Studies Can Show Us About Spatial Attention: Evidence for Limited Capacity Preattentive Processing. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(2), 331–351. https://doi.org/10.1037/0096-1523.19.2.331

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2015). Working memory delay activity predicts individual differences in cognitive abilities. *Journal of Cognitive Neuroscience*, *27*(5), 853–865. https://doi.org/10.1162/jocn_a_00765

Utochkin, I. S., & Brady, T. F. (2020). Individual representations in visual working memory inherit ensemble properties. *Journal of Experimental Psychology: Human Perception and Performance*, *46*(5), 458–473. https://doi.org/10.1037/xhp0000727

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions and objects in visual working memory. *J Exp Psychol Hum Percept Perform*, *27*(1), 92–114. https://doi.org/10.1037//0096-1523.27.1.92

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions and objects in visual working memory. *J Exp Psychol Hum Percept Perform*, *27*(1), 92–114. https://doi.org/10.1037//0096-1523.27.1.92

Wang, B., Cao, X., Theeuwes, J., Olivers, C. N. L., & Wang, Z. (2017). Separate capacities for storing different features in visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(2), 226–236. https://doi.org/10.1037/xlm0000295

Wang, R., Janini, D., & Konkle, T. (2022). *Mid-level feature differences underlie early animacy and object size distinctions : Evidence from EEG decoding*. *2018*, 1–14.

Wheeler, M. E., & Treisman, A. M. (2015). Binding in Short-Term Visual Memory . *From Perception to Consciousness*, 350–368. https://doi.org/10.1093/acprof:osobl/9780199734337.003.0033

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision*, *4*(12), 11–11. https://doi.org/10.1167/4.12.11

Wolfe, J. M., & Bennett, S. C. (1997). Preattentive object files: Shapeless bundles of basic features. *Vision Research*, *37*(1), 25–43. https://doi.org/10.1016/S0042-6989(96)00111-3

Xie, W., & Zhang, W. (2017). Familiarity increases the number of remembered Pokémon in visual short-term memory. *Memory and Cognition*, *45*(4), 677–689. https://doi.org/10.3758/s13421-016-0679-7

Xie, W., & Zhang, W. (2017). Familiarity speeds up visual short-term memory consolidation. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(6), 1207–1221. https://doi.org/10.1037/xhp0000355

Xu, Y. (2002). Limitations of object-based feature encoding in visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(2), 458–468. https://doi.org/10.1037/0096-1523.28.2.458

Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, *440*(7080), 91–95. https://doi.org/10.1038/nature04262

Zhang, W., & Luck, S. J. (2008). Discrete Fixed-Resolution Representations in Visual Working Memory Weiwei. *Nature*, *453*(7192), 233–235. https://doi.org/10.1038/nature06860.Discrete

Appendices

Appendix A. Different features are stored independently in visual working memory but mediated by object-based representations

Article "Different features are stored independently in visual working memory but mediated by object-based representations"

Markov, Y.A., Tiurina, N.A., & Utochkin, I.S. (2019). Different features are stored independently in visual working memory but mediated by object-based representations. *Acta Psychologica, 197*, 52-63. DOI: 10.1016/j.actpsy.2019.05.003

Abstract. The question of whether visual working memory (VWM) stores individual features or bound objects as basic units is actively debated. Evidence exists for both feature-based and object-based storages, as well as hierarchically organized representations maintaining both types of information at different levels. One argument for feature-based storage is that features belonging to different dimensions (e.g., color and orientations) can be stored without interference suggesting independent capacities for every dimension. Here, we studied whether the lack of cross-dimensional interference reflects genuinely independent feature storages or mediated by common objects. In three experiments, participants remembered and recalled the colors and orientations of sets of objects. We independently manipulated set sizes within each feature dimension (making colors and orientations either identical or differing across objects). Critically, we assigned to-be-remembered colors and orientations either to same spatially integrated or to different spatially separated objects. We found that the precision and recall probability within each dimension was not affected by set size manipulations in a different dimension when the features belonged to integrated objects. However, manipulations with color set sizes did affect orientation memory when the features were separated. We conclude therefore that different feature dimensions can be encoded and stored independently but the advantage of the independent storages are mediated at the object-based level. This conclusion is consistent with the idea of hierarchically organized VWM.

According to the publisher's policy of Elsevier, "theses and dissertations which contain embedded published journal articles (PJAs) as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect" (https://www.elsevier.com/about/policies/sharing).

The article can be accessed at:

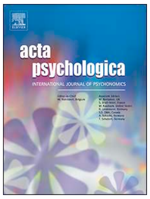
https://www.sciencedirect.com/science/article/abs/pii/S0001691818305511

Acta Psychologica 197 (2019) 52–63

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy



Different features are stored independently in visual working memory but mediated by object-based representations

Yuri A. Markov*, Natalia A. Tiurina, Igor S. Utochkin

National Research University Higher School of Economics, Russia

A R T I C L E I N F O

Keywords:
Visual working memory
Objects
Features
Hierarchical encoding

A B S T R A C T

The question of whether visual working memory (VWM) stores individual features or bound objects as basic units is actively debated. Evidence exists for both feature-based and object-based storages, as well as hierarchically organized representations maintaining both types of information at different levels. One argument for feature-based storage is that features belonging to different dimensions (e.g., color and orientations) can be stored without interference suggesting independent capacities for every dimension. Here, we studied whether the lack of cross-dimensional interference reflects genuinely independent feature storages or mediated by common objects. In three experiments, participants remembered and recalled the colors and orientations of sets of objects. We independently manipulated set sizes within each feature dimension (making colors and orientations either identical or differing across objects). Critically, we assigned to-be-remembered colors and orientations either to same spatially integrated or to different spatially separated objects. We found that the precision and recall probability within each dimension was not affected by set size manipulations in a different dimension when the features belonged to integrated objects. However, manipulations with color set sizes did affect orientation memory when the features were separated. We conclude therefore that different feature dimensions can be encoded and stored independently but the advantage of the independent storages are mediated at the object-based level. This conclusion is consistent with the idea of hierarchically organized VWM.

1. Introduction

At every moment of our perception, we interact with different objects, each having a number of various features, such as color, shape, size, etc. A limited portion of the information about these objects and their features can be used for current tasks and maintained for a short period of time in working memory (Baddeley, 1986; Baddeley & Hitch, 1974). It is consistently established that the capacity of working memory has serious limitations (e.g., Cowan, 2001; Miller, 1956). These fundamental limits are also true for the visual subsystem of working memory (VWM) that maintains and operates visual information necessary for an ongoing task (Alvarez & Cavanagh, 2004; Luck & Vogel, 1997). However, for a correct capacity estimate, it is important to determine what is represented in VWM as a basic unit of storage. There is a long-lasting debate around this question in the VWM literature: Does VWM store whole objects or separate features?

Existing studies provide evidence that both objects (Cowan, Chen, & Rouder, 2004; Kahneman, Treisman, & Gibbs, 1992; Lee & Chun, 2001; Luck & Vogel, 1997; Luria & Vogel, 2011; Treisman, 1999; Vogel, Woodman, & Luck, 2001; Xu, 2002; Xu & Chun, 2006) and features (see Brady, Konkle, & Alvarez, 2011, for review; Wang, Cao, Theeuwes, Olivers, & Wang, 2017; Wheeler & Treisman, 2002; Shin & Ma, 2017; Fougnie & Alvarez, 2011) can be the units of VWM. In their seminal study, Luck and Vogel (1997) demonstrated a strong advantage of maintaining any number of features in a limited number of spatially bound objects (at least up to four features per object). The prevailing limiting factor for capacity, as Luck and Vogel (1997) suggested, was the number of objects (~3–4) rather than the number of features. They concluded that objects are the units of VWM, showing no limitation in VWM by a number of features. However, other studies failed to support this strong version of object-based storage suggesting that the number of stored features also can be limited (see Brady et al., 2011, for review). Two major sets of evidence are used against this purely object-based account. The first set of evidence is based on findings that increasing the number of features to be remembered within an object does cause interference. For example, the increased number of features belonging to the same dimension per object significantly decreases VWM capacity for these objects (Olson & Jiang, 2002; Wheeler & Treisman, 2002; Xu, 2002). The same was found for increasing object complexity (Alvarez & Cavanagh, 2004; Hardman & Cowan, 2015;

---

* Corresponding author at: 101000, 4 Armyanskiy pereulok Building 2, Psychology Department, Higher School of Economics, Moscow, Russian Federation.
E-mail address: yuamarkov@gmail.com (Y.A. Markov).

https://doi.org/10.1016/j.actpsy.2019.05.003
Received 25 November 2018; Received in revised form 4 April 2019; Accepted 7 May 2019

Oberauer & Eichenberger, 2013). Other studies have found that remembering two features of the same objects impaired the precision of each remembered feature, whereas the capacity was seemingly intact (Fougnie, Asplund, & Marois, 2010; Fougnie & Marois, 2009). The second strong line of evidence against the purely object-based account of VWM is a number of demonstrations of relative independence between features of the same object. This independence can manifest as selective forgetting of some features rather than entire object (Fougnie & Alvarez, 2011) or as swaps between remembered features of different objects (Bays, Catalao, & Husain, 2009; Bays, Wu, & Husain, 2011; Fougnie, Cormiea, & Alvarez, 2013; Pertzov, Dong, Peich, & Husain, 2012). Whatsoever, even in the presence of these feature-based limitations, VWM still consistently benefits from object-based representations: It is easier to remember several features of one object than the same number of features distributed across several objects (Fougnie et al., 2010; Fougnie et al., 2013; Wheeler & Treisman, 2002). To account for this, theorists suggested that neither objects, nor features alone can be the units of VWM. Rather, the units are structured "feature bundles" containing both integrated object and feature representations hierarchically linked (see Brady et al., 2011, for review; Fougnie et al., 2010). Similar ideas that VWM can be constrained by both objects and features in different ways have been proposed by other authors (Olson & Jiang, 2002; Shin & Ma, 2017; Xu & Chun, 2006).

The complicated pattern of evidence for feature-based vs. object-based storage in VWM is additionally complicated by an ambiguity regarding the structure of feature memories. Specifically, it was noted that VWM performance can depend on whether remembered and tested features belong to same or different dimensions. Most experiments on features were the same dimension (Olson & Jiang, 2002; Wheeler & Treisman, 2002; Xu, 2002) which typically constitute different parts of an object (Alvarez & Cavanagh, 2004) show a significant decrement in performance with an increasing number of features per object (but see Luck & Vogel, 1997; Vogel et al., 2001 for an opposite conclusion). There is no such interference between features from the same dimension (e.g., Wheeler & Treisman, 2002). This leaves room for a theory that feature-based VWM is, in fact, a multistorage system having separate capacities for features from different dimensions. This theory was directly tested and supported in recent studies where researchers independently manipulated the memorized set size for features from two separable dimensions, color and orientation (Wang et al., 2017). They found that VWM capacity for a given feature depended on the set size in the corresponding dimension rather than joint set size in both dimensions. For example, if observers are shown six isosceles triangles, each triangle having one of two possible colors (color set size is two) and one of two possible orientations (orientation set size is also two), their ability to spot a change in either of the dimensions is rather high. If color set size becomes six (each triangle has a unique color) and orientation set size remains two, it selectively impairs change detection for color but not for orientation (and vice versa if color set size stays small and orientation set size increases). These separate storages can provide an advantage when the selective encoding of one dimension and ignoring another can be required (Shin & Ma, 2017; Woodman & Vogel, 2008).

However, it is important to note that independent set size manipulations in the experiments by Wang et al. (2017) concerned features but not objects these features belonged to. In all experiments, colors and orientations were tested in the same set of objects. If object representations facilitate feature storage in general, can they mediate the advantage of the independent feature capacities? Alternatively, these independent capacities can be purely feature-based in which case they should manifest in both unitary and separate objects.

To address this question, we have run three experiments testing VWM for color and orientation. The general approach was similar to that used by Wang et al. (2017): We orthogonally manipulated the set size within each dimension by assigning either a single or three different values and measured VWM for both dimensions. Critically, colors and orientations could be assigned to the same objects (Experiment 1), different parts of spatially integrated objects (Experiment 3), or spatially separated objects (Experiment 2). Unlike Wang et al. (2017), we used a continuous report task (Wilken & Ma, 2004; Zhang & Luck, 2008) instead of a change detection task. It is justified by a fact that the former paradigm allows parametric estimation of both capacity and fidelity of VWM (Zhang & Luck, 2008), that are both known to be sensitive to feature-based and object-based load (Fougnie et al., 2010).

2. Experiment 1

In Experiment 1, we tested VWM for colors and orientations in the same set of three objects. In different conditions of the experiment, we assigned either three different values or a single value to each object in each dimension orthogonally. This manipulation affected both within-dimension and joint set sizes in a manner similar to that in the experiments by Wang et al. (2017). Hence, the main goal of this experiment is to test whether the principal finding of independent storages for color and orientation is replicated in our paradigm.

2.1. Methods

2.1.1. Participants

Twenty students from the Higher School of Economics (17 female) participated for extra course credits. The participants ranged in age from 18 to 25 years (average age was 19.93 years) and reported having normal or corrected to normal visual acuity, no color blindness and neurological problems. Before the beginning of the experiment, they signed an informed consent form. In this and subsequent experiments, sample sizes were determined based on similar studies addressing the issue of feature storage and binding in VWM and using a continuous report task (from 10 to 16; for example, Fougnie & Alvarez, 2011; Fougnie et al., 2010; Bays et al., 2009; Pertzov et al., 2012). The planned sample size also included a few extra participants considering a possibility of technical problems or poor performance in some participants.

2.1.2. Apparatus and stimuli

Stimulation was developed and presented through PsychoPy (Pierce, 2007) for Linux Ubuntu. Stimuli were presented on a standard VGA monitor with a refresh frequency of 75 Hz and 1024 × 768-pixel spatial resolution. Stimuli were presented on a homogeneous gray field. Participants sat approximately at 47 cm from the monitor. From that distance, screen subtended approximately 42.44 × 32.5 degrees of visual angle.

Sample displays consisted of one or three colored isosceles triangles presented in randomized positions along an imagery circumference 4.35° away from a monitor center (Fig. 1). Each triangle had sides of 0.6°, 1.2°, and 1.2° in length. To set the positions of the three triangles on the imaginary circumference, we first generated a random rotational angle from 1° to 360° for a first triangle and then positioned the rest two triangles 120° and − 120° away from the first with a ± 30°-jitter. For color assignment, we used the hue wheel in the 360° HSV (hue-saturation-value) space, and for orientation assignment, we used the 360° orientational circumference. As color and orientation had the same dimensionality as spatial positions, we applied the rotational algorithm described above to set three colors and three orientations. When an experimental condition required a single color, a single orientation, or a single item to be presented, the color, orientation, or position was chosen randomly.

For memory test, outline circles were presented at the locations of sample triangles, with one thick outline indicating the location of a probed item. In trials where color was probed, the test display was surrounded by an HSV color wheel 4.35° in radius (Fig. 1). In trials where orientation was probed, the test display was surrounded by a black orientational wheel with white ticks marking 30° steps (Fig. 1).

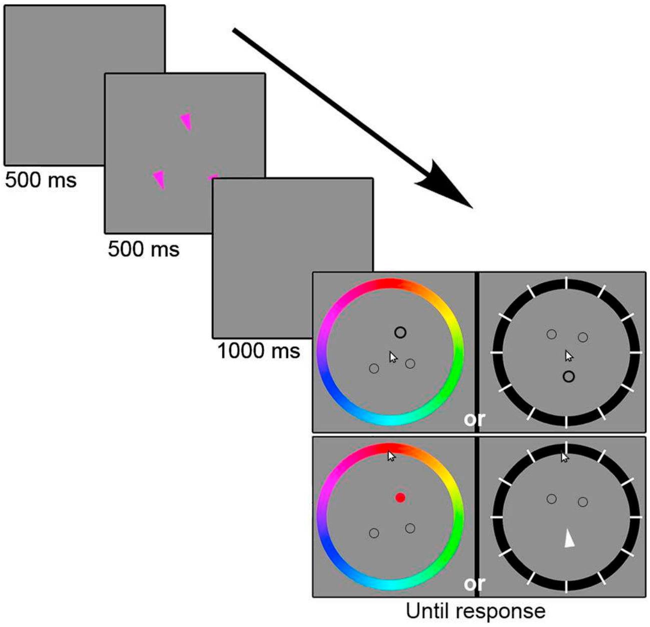


Fig. 1. The time course of a typical trial in Experiment 1.

The probed outline turned into a solid colored circle (if the color was probed) or a white-oriented triangle (if the orientation was probed) upon mouse click on a wheel (Fig. 1).

2.1.3. Procedure

Each experimental trial started with a 500-ms presentation of a sample display. Participants were instructed to memorize both the color and orientation of the triangles. The sample was followed by a 1-s delay (retention interval) that, in turn, was followed by a probe screen (Fig. 1). Clicking on a color or orientation wheel, participants had to adjust a corresponding attribute of the probe item to match the sample attribute presented at that location. At the beginning of the experiment, participants completed a practice block. The total duration of the experiment varied between 45 and 60 min.

2.1.4. Design and data analysis

Five conditions of the Sample type were tested in Experiment 1 (Fig. 2A). In four of these conditions, we orthogonally varied color and orientation set sizes in three triangles: (1) all different features (three colors and three orientations), (2) color identical (one color and three orientations), (3) orientation identical (three colors and one orientation), (4) all identical features (one color and one orientation). The latter condition can be considered a baseline estimating observers' capacity and fidelity at minimal load for each feature. Finally, condition (5) contained a single object and was used as a baseline. This baseline, in comparison with the “all identical” condition, aimed to test whether three identical feature values of three objects are indeed encoded like a single feature of one object. In a within-subject experiment, each participant was exposed to 5 (Sample type) × 2 (Probed dimension: Color vs. Orientation) × 47 repetitions = 470 trials.

For each trial, the error was calculated as an angular difference between the correct feature value and that adjusted by a participant. The distribution of errors was then analyzed using the mixture model (Zhang & Luck, 2008) implemented in MemToolbox for Matlab (Suchow, Brady, Fougnie, & Alvarez, 2013). The standard mixture model has two different parameters obtained from fitting two components of the error distribution. The first parameter is the standard deviation (*SD*) of the von Mises distributional component, that is supposed to reflect the precision of a noisy representation that is present in memory. The second parameter is the probability of random guess (P_{guess}) can be estimated as an area below the uniform component of the mixed distribution; this component is supposed to reflect randomly chosen answers when a probed item is likely to be absent in the memory (not encoded or forgotten). Reverse P_{guess} is used as an estimate for a probability that a probed element is held in VWM: $P_{memory} = 1 - P_{guess}$.

To evaluate the effect of Sample type, we applied the standard frequentist and Bayesian one-way repeated measures ANOVA to the *SD* and P_{memory} for color and orientation. The Bayes factor (BF_{10}) was calculated using JASP 0.9.0.0 (JASP Team, 2018; Wagenmakers et al., 2017) and interpreted using the standard Jeffreys's (1961) scale. The Bayesian approach estimates the odds of H_1 to H_0 (Rouder, Speckman, Sun, Morey, & Iverson, 2009).

2.2. Results and discussion

One participant was excluded from the analysis because she showed nearly 100% guess rate in all conditions. The results of Experiment 1 for P_{memory} and *SD* are summarized in Fig. 3.

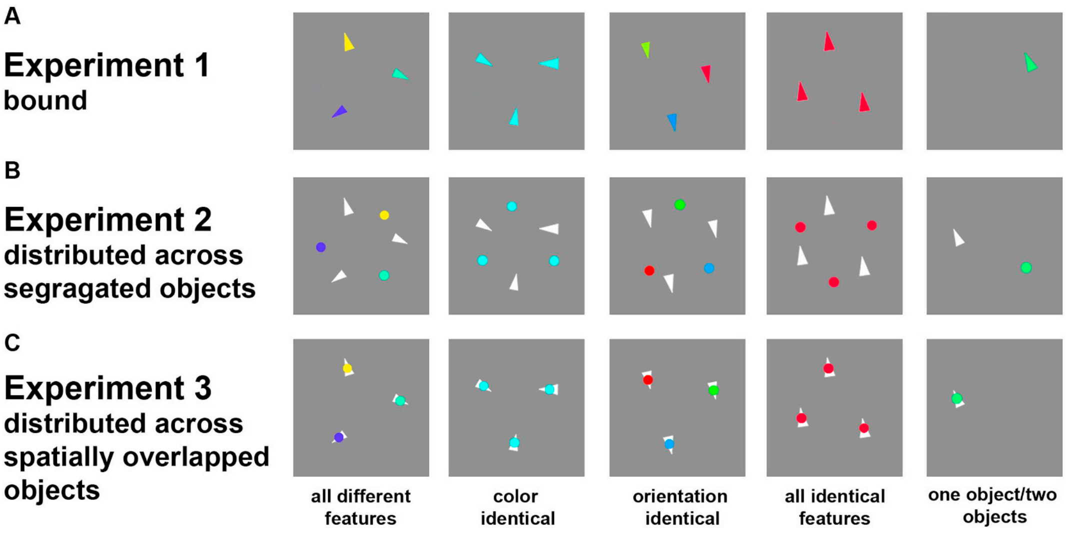


Fig. 2. Example of stimuli for three experiments for five conditions (objects with all different features, objects with different color and identical orientation, objects with different orientation and identical color, objects with all identical features, one pair of features): (A) Experiment 1 with bound features in the object, (B) Experiment 2 with features distributed across segregated objects, (C) Experiment 3 with features distributed across spatially overlapped objects.

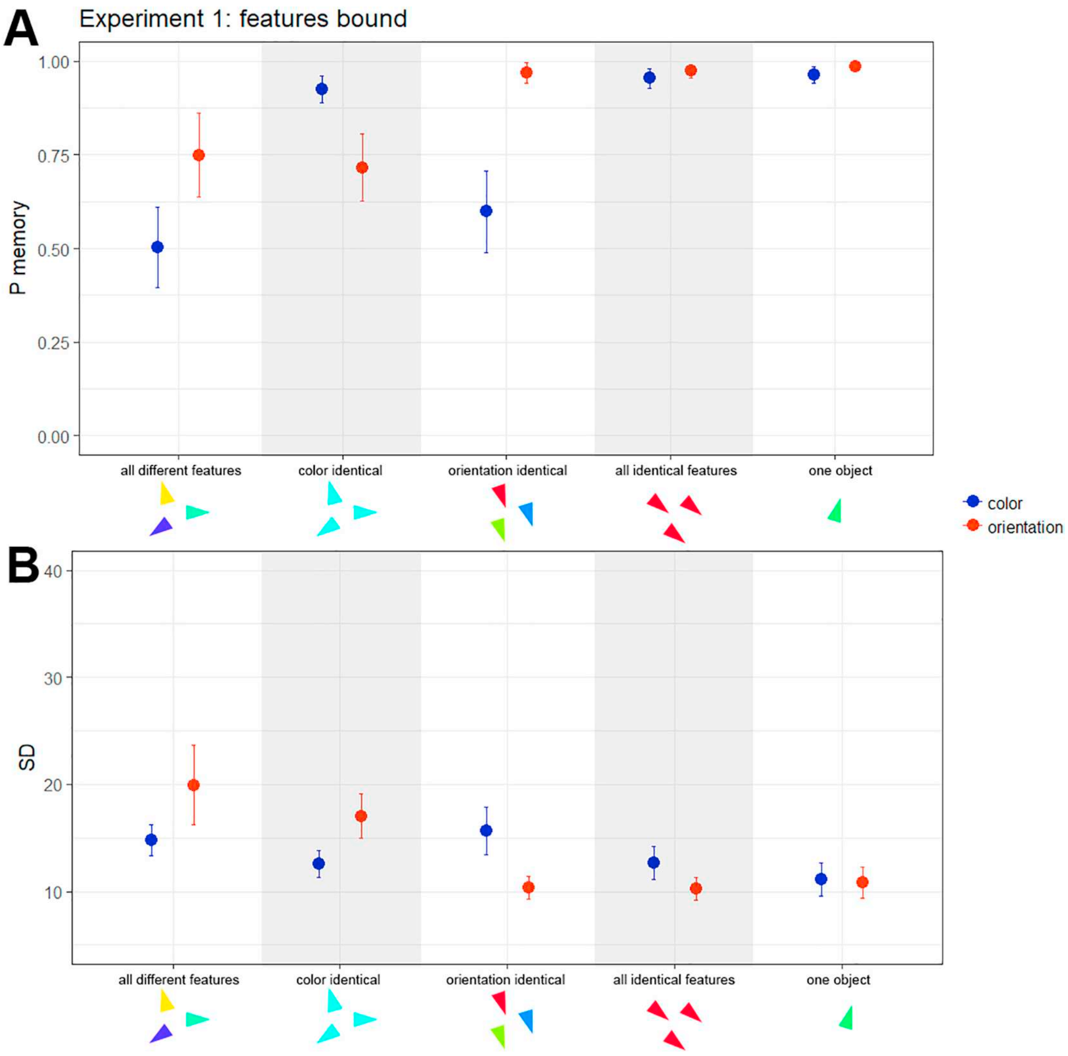


Fig. 3. Results of Experiment 1: (A) P_{memory} and (B) *SD* as a function of Sample type. Error bars depict 95% CIs.

2.3. P_{memory} for color

We found a strong effect of Sample type on P_{memory} for color (F (4, 72) = 65.92, $p < .001$, $\eta^2 = 0.786$, $BF_{10} > 10^{20}$). P_{memory} was greater in conditions where the color was identical across objects (color identical, all features identical, and one object) compared to conditions where colors differed across (all different features and orientation identical) $-7.348 \leq t(18) \leq 9.739$, $p < .001$, Bonferroni corrected $\alpha = 0.005$, $1.686 \leq$ Cohen's $d \leq 2.234$, $10^4 < BF_{10} < 10^{18}$. There were no significant differences between conditions with identical color across objects (color identical, all features identical, and one object – see Appendix A for the exact results of statistical evaluation) and also between conditions with different color across objects (all different features and orientation identical – see Appendix A). Note that, when the colors were identical, P_{memory} was always near 100% ceiling (Fig. 1A) suggesting the perfect capacity to remember one color regardless of variations in the number of orientations and the total number of physically presented objects.

2.4. P_{memory} for orientation

We found a strong effect of Sample type on P_{memory} for orientation (F (4, 72) = 28.53, $p < .001$, $\eta^2 = 0.613$, $BF_{10} > 10^{10}$). P_{memory} was greater in samples where orientation was identical across objects (orientation identical, all identical features, and one object) compared to samples where orientation differed across objects (all different features and color identical) – $4.537 \leq t(18) \leq 7.376$, $p < .001$, Bonferroni corrected $\alpha = 0.005$, $1.041 \leq$ Cohen's $d \leq 1.692$, $122 < BF_{10} < 10^9$. There were no significant differences between conditions with identical orientation across objects (orientation identical, all identical features, and one object – see Appendix A for the exact results of statistical evaluation) and also between conditions with different orientation across objects (all different features and color identical – see Appendix A). Again, P_{memory} for orientations was at ceiling in all conditions with identical orientation, suggesting perfect VWM capacity for a single orientation regardless of any variations of the number of colors or objects in total.

2.5. *SD* for color

We found a strong effect of Sample type on *SD* for color (F (4, 72) = 6.115, $p < .001$, $\eta^2 = 0.254$, $BF_{10} = 217.3$). In all different features and in identical orientations, *SD* for color was greater compared to color identical ($t(18) = 3.312$, $p = .0039$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.760$, $BF_{10} = 11.53$) and one object (t (18) = 3.312, $p = .0004$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 1.003$, $BF_{10} = 88.89$). In all other comparisons there were no significant differences (see Appendix A).

2.6. SD for orientation

We found a strong effect of Sample type on *SD* for orientation (F (4, 72) = 30.66, $p < .001$, $\eta^2 = 0.630$, $BF_{10} > 10^{10}$). *SD* was lower in samples where orientation was the same across objects (only orientation identical, all features identical and one object) compared to conditions where orientation was different through objects (all features different and only color identical): $5.327 \leq t(18) \leq 8.269$, $p < .001$, Bonferroni corrected $\alpha = 0.005$, $1.222 \leq$ Cohen's $d \leq 1.897$, $559 < BF_{10} < 10^{11}$. There were no significant differences between conditions with identical orientation across objects (orientation identical, all identical features, and one object – see Appendix A for the exact results of statistical evaluation) and also between conditions with different orientations across objects (all different features and color identical – see Appendix A).

In total, in Experiment 1 we observed a consistent pattern across both probed dimensions and both estimated VWM parameters. Specifically, we found that a greater P_{memory} (roughly corresponding to capacity in items) and a lower *SD* (corresponding to better precision) take place in those clusters of conditions where the tested features have been identical across objects or where a memorized object has been physically alone. More importantly, within these clusters, there was no effect of whether a second dimension had been represented by identical or different features. Hence, we found that both P_{memory} and *SD* for a given dimension depended only on the set size within that dimension and not on the joint set size. Additionally, we found that all identical features are encoded as efficiently as a corresponding feature in one object. Overall, the results of Experiment 1 replicate Wang et al.'s (2017) finding in favor of independent storages for features from different dimensions.

3. Experiment 2

In Experiment 2, we modified stimuli so that colors and orientations belonged to different spatially separated objects (exactly like in Fougnie et al., 2010). This would allow us to test whether dimension independence is preserved when there is no object-based advantage for storing the features together and when object-based load is increased.

3.1. Methods

3.1.1. Participants

Nineteen students from the Higher School of Economics (14 female) participated for extra course credit. They ranged in age from 18 to 22 years (average age is 18.52 years) and reported having normal or corrected to normal visual acuity, no color blindness and no neurological problems. Before the beginning of the experiment, they signed an informed consent form.

3.1.2. Apparatus and stimuli

Apparatus and stimuli were similar to Experiment 1, except that colors and orientations were distributed across spatially separated objects. This led to duplicated numbers of objects from Experiment 1 (from three to six and from one to two). Objects were located along an imaginary circumference with a radius of 4.35°. If there were six objects on a screen, each object was separated by 60° of rotation ± 15° jitter from its neighbors (Fig. 2B). When there were two objects on a screen, each object was separated by 180° of rotation from another (presented symmetrically across the center of the screen, Fig. 2B). There were two types of objects depending on which dimension was relevant for memorization. "Color" objects were the circles whose colors were set using the coloring algorithm from Experiment 1. "Orientation" objects were the isosceles triangles whose orientations were set using the orientation rotation algorithm from Experiment 1. "Color" objects alternated with "orientation" objects on the imaginary circumference forming two overlapping triangular groups (this was exactly the same method of positioning as that used by Fougnie et al., 2010, Fig. 2B). When two objects were presented, one was a "color" object and another was an "orientation" object.

3.1.3. Procedure

The procedure of Experiment 2 was the same as in Experiment 1, except for a difference in instruction. Participants were instructed to memorize only orientations of white triangles and only colors of color circles.

3.1.4. Design and data analysis

The design of Experiment 2 was the same as that of Experiment 1 in terms of Sample types, two tested dimensions, and a number of trials. The only nominal change was that the baseline "one object" condition from Experiment 1 was renamed to "two objects" for clarity (but they were equal in terms of feature set sizes). Data analysis was identical to Experiment 1.

3.2. Results and discussion

The results of Experiment 2 for P_{memory} and *SD* are summarized in Fig. 4.

3.3. P_{memory} for color

We found a strong effect of Sample type on P_{memory} for color (F (4, 72) = 69.53, $p < .001$, $\eta^2 = 0.794$, $BF_{10} > 10^{20}$). P_{memory} for was greater in samples where color was identical across "color" objects (color identical, all identical features) or belonged to a single "color" object compared to samples where colors were different across objects (all features different and orientation identical): $8.426 \leq t(18) \leq 9.129$, $p < .001$, Bonferroni corrected $\alpha = 0.005$, $1.993 \leq$ Cohen's $d \leq 2.094$, $BF_{10} > 10^6$. There were no significant differences between conditions with identical color across objects (color identical, all features identical, and single "color" object – see Appendix A for the exact results of statistical evaluation) and also between conditions with different color across objects (all different features and orientation identical – see Appendix A). This result replicates the respective pattern from Experiment 1.

3.4. P_{memory} for orientation

We found an effect of Sample type on P_{memory} for orientation (F (4, 72) = 19.03, $p < .001$, $\eta^2 = 0.514$, $BF_{10} > 10^8$). As in Experiment 1, P_{memory} was greater in samples with all identical features or in a single "orientation" object (two objects) compared to conditions where orientations differed across objects (all different features and color identical; comparison: $3.375 \leq t(18) \leq 5.838$, $p \leq .0034$, Bonferroni corrected $\alpha = 0.005$, $0.774 \leq$ Cohen's $d \leq 1.339$, $12.981 < BF_{10} < 10^4$). However, unlike Experiment 1, we found that P_{memory} for orientation suffered from the increased color set size (orientation identical condition). Specifically, P_{memory} in that condition ($M = 0.91$) was lower than in all identical features ($M = 0.96$) condition ($t(18) = 3.368$, $p = .0018$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.842$, $BF_{10} = 7.431$) but greater than in the two conditions with three different orientations (all different features ($M = 0.82$) and color identical ($M = 0.83$); comparisons: $3.375 \leq t(18) \leq 5.838$, $p \leq .0034$, Bonferroni corrected $\alpha = 0.005$, $0.774 \leq$ Cohen's $d \leq 1.339$, $12.981 < BF_{10} < 10^4$). We also found some evidence (basically, from effect size and Bayes factor estimates) that the orientation P_{memory} in the orientation identical trials ($M = 0.91$) was lower than that in the two object trials ($M = 0.97$), though this evidence was not conclusive, since the significance level did not fall below the strictly Bonferroni corrected critical ($t(18) = 3.075$, $p = .0065$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.705$, $BF_{10} = 22.689$). Overall, we conclude that orientation P_{memory} in the orientation identical condition was high, yet, we observed some evidence for slight impairment due to color

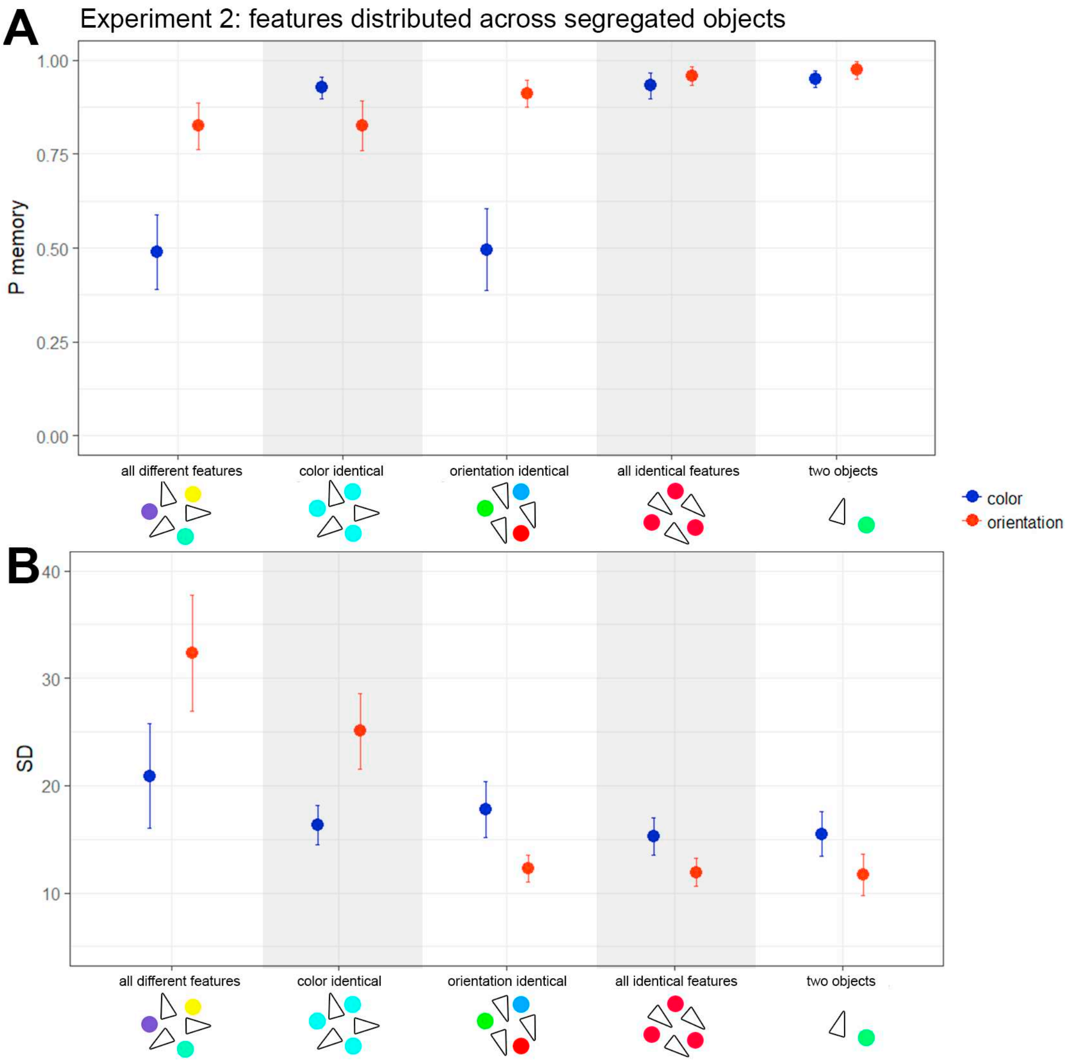


Fig. 4. Results of Experiment 2: (A) P_{memory} and (B) *SD* as a function of Sample type. Error bars depict 95% CIs.

memory load. In all other comparisons, we found no significant differences (see Appendix A).

3.5. SD for color

We did not find convincing evidence for a reliable effect of Sample type on *SD* for color (F (4, 72) = 2.865, p = .029, $\eta^2 = 0.137$, $BF_{10} = 2.735$). We conclude, therefore, that memory set size manipulation did not have a strong effect on the precision of color encoding. As we will show later, by direct comparison of results across Experiments 1–3, the lack of the effect can be associated with overall drop in color memory precision in Experiment 2.

3.6. SD for orientation

We found a strong effect of Sample type on the *SD* for orientation (F (4, 72) = 71.02, $p < .001$, $\eta^2 = 0.798$, $BF_{10} > 10^{12}$). Like in Experiment 1, *SD* was lower in the samples where orientations were identical across “orientation” objects (identical orientation and all identical features) or belonged to a single “orientation” object (two objects condition) compared to the samples where orientations were different across objects (all different features and identical color; comparisons: $8.059 \le t(18) \le 10.197$, $p < .001$, Bonferroni corrected $\alpha = 0.005$, $1.849 \le$ Cohen's $d \le 2.335$, $10^5 < BF_{10} < 10^6$). We also found that, in the identical color samples (but different orientations), *SD* was smaller than in samples with all different features (t (18) = 3.867, p = .0011, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.887$, $BF_{10} = 33.281$), that suggests that color set size interfered with the precision of VWM for orientation. In all other comparisons, we found no significant differences (see Appendix A).

Overall, in Experiment 2 we replicated the finding from Experiment 1 and from the literature (Fougnie et al., 2010; Wang et al., 2017; Zhang & Luck, 2008) that the set size increment within the same dimension makes P_{memory} for that dimension dropping and *SD* for that dimension raising. It is also quite evident that, like in Experiment 1, the VWM parameters for a given dimension depended on memory load in that dimension more than on memory load in another dimension. For example, P_{memory} for orientations dropped strongly when the orientation set size increased (all identical features vs. identical color comparison, Fig. 4A) and it dropped slightly when the color set size increased (all identical features vs. identical orientation comparison, Fig. 4A). Therefore, this result supports an idea that observers kept storing colors and orientations relatively independently, even despite a serious increment in the number of spatially separated objects.

However, unlike Experiment 1, we found some evidence for slight detrimental effects of VWM load in one dimension on feature storage in another dimension. Specifically, when orientation load was kept low but color load increased (all identical features vs. orientation identical) it impaired P_{memory} for orientations. Also, when orientation load was high and color load increased (color identical vs. all different features) the precision of orientation estimates decreased substantially (Fig. 4B). Following the interpretational logic of the mixture model for VWM (Zhang & Luck, 2008), this pattern suggests that storing more colors decreases the probability that the orientation would be remembered, whereas storing more orientations (when color set size is already big) causes the subsequent loss in the precision of each orientation memory. Another recent framework (Schurgin, Wixted, & Brady, 2018) does not make a strong qualitative difference between the *SD* and the P_{memory} parameters suggesting a single source of degradation for both these parameters, namely, the distinctiveness of familiarity signals provided by available test alternatives. In this framework, the drop of P_{memory} and *SD* of orientation memory under high color load might reflect gradual decrement in the quality of familiarity signals for orientation.

This detrimental effect of color memory load on orientation memory was not mirrored in an effect of orientation memory load on color memory. One possible explanation of this asymmetry could be that colors were more prioritized for encoding, so it did not suffer from overall feature load as much as less prioritized orientation memory. One finding can seemingly contradict to this interpretation, namely, the fact that P_{memory} for color drops much stronger when color set size increases than P_{memory} for orientation drops when orientation set size increases (Fig. 4A). However, this fact may suggest that remembering three colors is generally a more difficult task than remembering three orientations. This suggestion does not rule out the possibility that observers put a higher priority to color (note that in Experiment 1, the relative P_{memory} decrement for color was also greater despite the absence of interference between color and orientation set sizes, Fig. 3A). Color priority could be partially explained by the additional focus on colored objects, because of the match between its shape and the shape of location cues in their initial states, before observers clicked on a color or orientation wheel (Fig. 1). Although this asymmetry between color and orientation needs further research, our major result indicates that there is interference between color VWM and orientation VWM when these features are distributed between different objects.

4. Experiment 3

The interference pattern that we found in Experiment 2 for orientation memory under the increasing color memory load, can have an alternative explanation apart from the spatial separation of colors and orientations. Overall stimulus complexity was greater than in Experiment 1 that could become an extra source of noise (some items were circles and some were triangles, some were white and some had different colors). Moreover, the instruction requiring to selectively encode different features in different objects could be also more difficult than in Experiment 1. To control for these possible confounds, we have run Experiment 3. Here, we presented participants with spatially integrated objects and asked to remember the color and orientation information about each of the object, like we did in Experiment 1. However, each of the objects consisted of two overlapping parts, one corresponding to a “color” object and another corresponding to the “orientation” object from Experiment 1 (for similar manipulations, see Fougnie et al., 2010; Xu, 2002). So, each object presented in Experiment 3 had the same amount of complexity as two separate objects in Experiment 2. Also, the instruction in Experiment 3 required selective encoding of orientation information from one part of an object and of color information from another part.

4.1. Methods

4.1.1. Participants

Nineteen students from the Higher School of Economics (14 female) participated for extra course credits. They ranged in age from 18 to 22 years (average age is 19.03 years) and reported having normal or corrected to normal visual acuity, no color blindness and no neurological problems. Before the beginning of the experiment, they signed an informed consent form.

4.1.2. Apparatus, stimuli, and procedure

In general, apparatus and stimuli were the same as in two previous experiments with some differences. Each object consisted of two parts: an oriented white triangle overlaid with a color circle (see Fig. 2C for examples). Object positioning was the same as in Experiment 1. The procedure was the same as in Experiment 1 with an addition that participants were instructed to remember the color of the circular part and the orientation of the triangular part of each object.

Design and data analysis were the same as in Experiment 1

5. Results and discussion

The data from four participants were excluded from analysis because they showed nearly 100% guess rate in all conditions. The results of Experiment 3 for P_{memory} and *SD* are summarized in Fig. 5.

5.1. P_{memory} for color

We found the strong effect of Sample type on P_{memory} for color (F (4, 56) = 69.53, $p < .001$, $\eta^2 = 0.808$, $BF_{10} > 10^{20}$). P_{memory} for color was higher in all conditions where color was identical across objects (color identical, all identical features, and one object) compared to the conditions where color differed across objects (all different features and orientation identical; comparisons: $7.916 \le t(14) \le 8.898$, $p < .001$, Bonferroni corrected $\alpha = 0.005$, $2.044 \le$ Cohen's $d \le 2.298$, $BF_{10} > 10^5$). In all other comparisons, we found no significant differences (see Appendix A).

5.2. P_{memory} for orientation

We found the strong effect of Sample type on P_{memory} for orientation (F (4, 56) = 14.37, $p < .001$, $\eta^2 = 0.506$, $BF_{10} > 10^6$). P_{memory} was greater in all conditions where orientation was identical across objects (identical orientation, all identical features and one object) compared to the conditions where orientations differed across objects (all different features and identical color; comparisons: $3.450 \le t(14) \le 4.191$, $p \le .0039$, Bonferroni corrected $\alpha = 0.005$, $0.891 \le$ Cohen's $d \le 1.081$, $12.18 < BF_{10} < 42.123$). In all other comparisons, we found no significant differences (see Appendix A).

5.3. SD for color

We found no effect of Sample type on the color *SD* (F (4, 56) = 0.726, $p = .578$, $\eta^2 = 0.049$, $BF_{10} = 0.139$). There were no significant differences in SD for color between conditions.

5.4. SD for orientation

We found the strong effect of Sample type on the orientation *SD* (F (4, 56) = 40.92, $p < .001$, $\eta^2 = 0.745$, $BF_{10} > 10^{13}$). *SD* was lower in all conditions where orientation was identical across objects (identical orientations, all identical features, and one object) compared to the conditions where orientation differed across objects (all different features and identical colors; comparisons: $6.008 \le t(14) \le 10.397$, $p < .001$, Bonferroni corrected $\alpha = 0.005$, $1.551 \le$ Cohen's $d \le 2.684$, $781 < BF_{10} < 10^6$). In all other comparisons, we found no significant differences (see Appendix A).

Therefore, the results of Experiment 3 basically replicated the principal results of Experiment 1 regarding the absence of interference between color and orientation VWM parameters. We conclude that VWM can support the independent storage of features from different dimensions in spatially integrated objects.

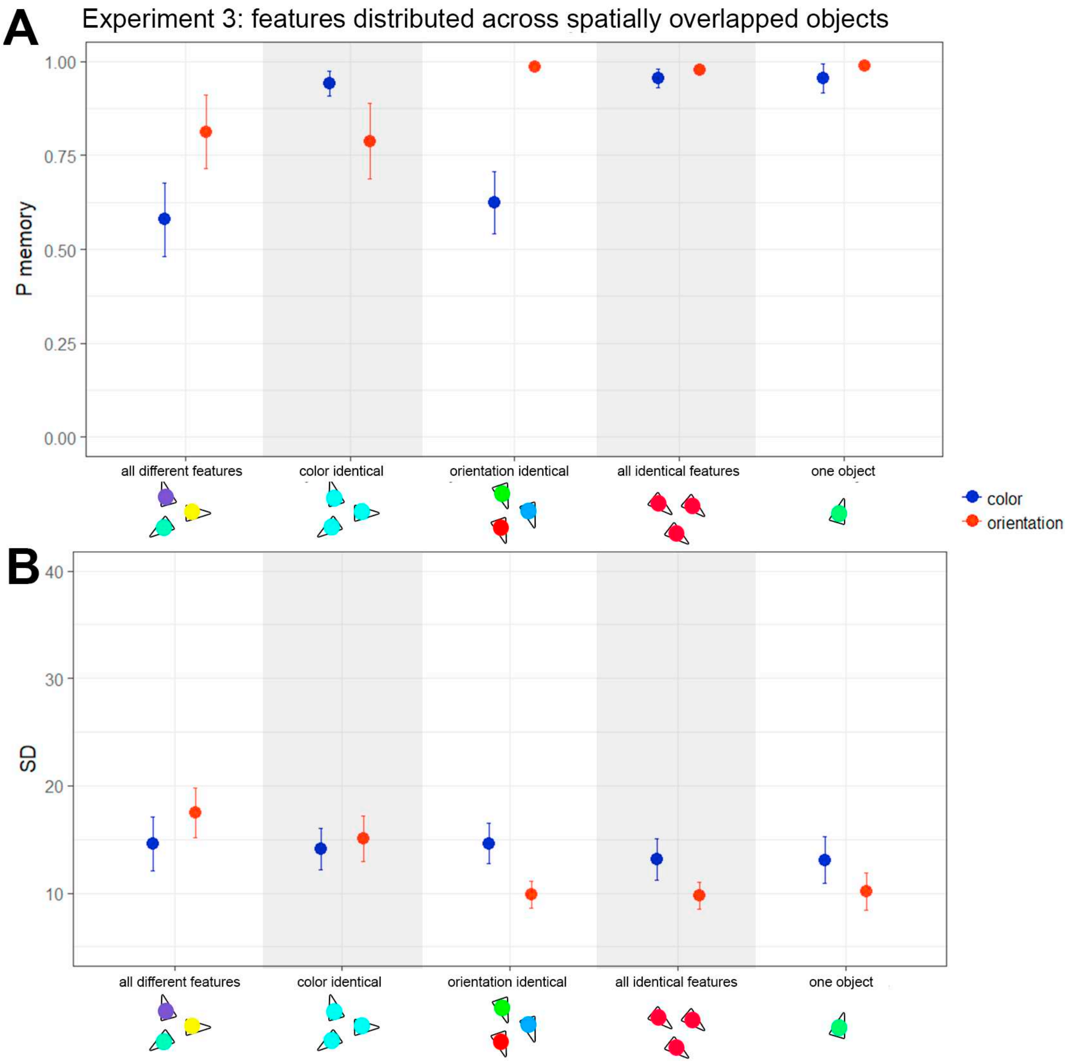


Fig. 5. Results of Experiment 3: (A) P_{memory} and (B) *SD* as a function of Sample type. Error bars depict 95% CIs.

5.5. Comparisons between experiments

To get a more comprehensive picture of the effects of feature separation vs. feature integration on VWM for both color and orientation, we directly compared the results of all three experiments. Data from 53 participants were analyzed. In Fig. 6, we plotted the results of all experiments together.

There were no significant differences between experiments in P_{memory} for both colors ($F(2, 50) = 1.434$, $p = .248$, $\eta^2 = 0.054$, $BF_{10} = 0.106$) and orientations ($F(2, 50) = 0.699$, $p = .502$, $\eta^2 = 0.027$, $BF_{10} = 0.123$). Yet, these differences were substantial in *SD* for both color ($F(2, 50) = 15.67$, $p < .001$, $\eta^2 = 0.054$, $BF_{10} = 415$) and for orientation ($F(2, 50) = 14.51$, $p < .001$, $\eta^2 = 0.367$, $BF_{10} = 246$). These differences were provided by Experiment 2 (Fig. 5B) where *SD*'s were overall greater than in Experiment 1 (color SD: $t = 5.242$, $p < .001$, Bonferroni corrected $\alpha = 0.017$, Cohen's $d = 0.720$, $BF_{10} = 20{,}176$; orientation SD: $t = 4.224$, $p < .001$, Bonferroni corrected $\alpha = 0.017$, Cohen's $d = 0.580$, $BF_{10} = 218$) and Experiment 3 (color SD: $t = 4.196$, $p < .001$, Bonferroni corrected $\alpha = 0.017$, Cohen's $d = 0.576$, $BF_{10} = 213$; orientation SD: $t = 4.936$, $p < .001$, Bonferroni corrected $\alpha = 0.017$, Cohen's $d = 0.678$, $BF_{10} = 3293$). Together these results demonstrate that both color and orientation were encoded and stored with a substantial loss in precision when they belonged to different rather than same objects. This finding is in line with the previous evidence for object-based advantage for storing features in VWM (Fougnie et al., 2010; Fougnie et al., 2013; Wheeler & Treisman, 2002).

We found evidence for a small effect of Sample type × Experiment on P_{memory} for orientations ($F(8, 38) = 2.640$, $p = .009$, $\eta^2 = 0.046$, $BF_{10} = 4.457$). It is provided by a smaller P_{memory} found for trials with identical orientations (and different colors) in Experiment 2 compared to Experiments 1 ($t(36) = 2.633$, $p = .012$, Bonferroni corrected $\alpha = 0.017$, Cohen's $d = 0.854$, $BF_{10} = 4.234$) and 3 ($t(32) = 3.721$, $p < .001$, Bonferroni corrected $\alpha = 0.017$, Cohen's $d = 1.285$, $BF_{10} = 38.705$). This result suggests that the probability of not having an orientation in VWM is slightly impaired by increasing color load but only when colors and orientations belong to spatially separated objects. To remind, this very decrement was earlier supported by the within-experiment comparison between the identical orientation and the all identical displays in Experiment 2. We also found evidence for an effect of Sample type × Experiment on the *SD* for orientations ($F(8, 200) = 12.58$, $p < .001$, $\eta^2 = 0.127$, $BF_{10} > 10^{11}$). The effect was mostly provided by disproportional absolute increment of the *SD* as a function of the sample type in Experiment 2 compared to Experiments 1 and 3 (Fig. 6B). When observers had to remember a single orientation in Experiments 1 or 3 (samples: one object, all identical features, or identical orientation) the orientation SD were ~10°, and they reached only ~12° in Experiment 2 showing a bare or no statistical evidence of growth ($0.77 \le t \le 2.87$, $0.007 \le p \le .45$, Bonferroni corrected $\alpha = 0.017$, $0.249 \le$ Cohen's $d \le 0.991$, $BF_{10} < 7$). When observers

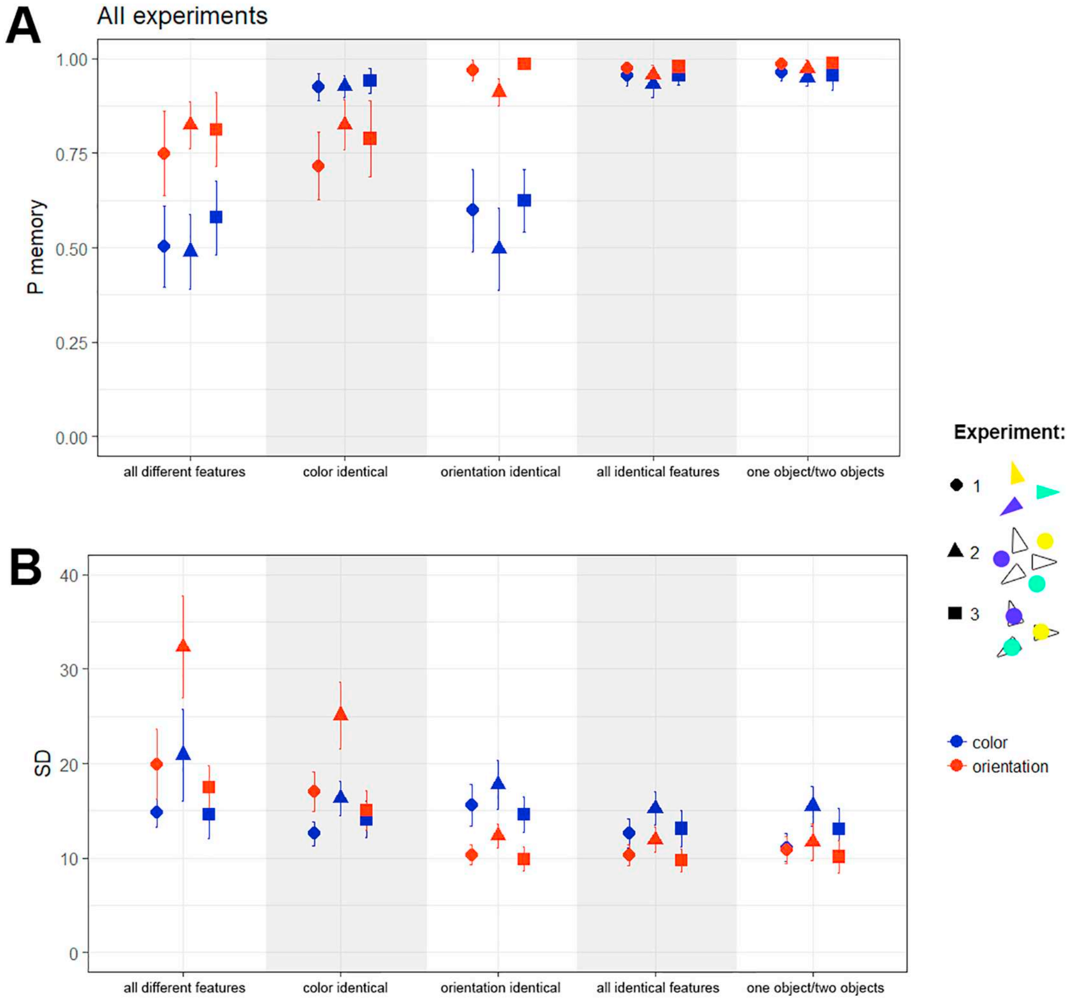


Fig. 6. Results of Experiments 1–3: (A) P_{memory} and (B) SD as a function of Sample type and Experiment. Error bars depict 95% CIs.

had to remember three orientations, the growth in orientation SD was much substantial in Experiment 2 compared to Experiments 1 and 3: from $\sim$15 to 17° to $\sim$25° in identical colors (t's ≥ 4.1, $p < .001$, Bonferroni corrected $\alpha = 0.017$, Cohen's $d \geq 1.3$, $\mathrm{BF}_{10} \geq 91$), and from $\sim$18–19° to $\sim$32° in all different features (t's ≥ 3.9, $p < .001$, Bonferroni corrected $\alpha = 0.017$, Cohen's $d \geq 1.3$, $\mathrm{BF}_{10} \geq 78$). In other words, memory precision for a single orientation almost did not suffer from assigning this orientation to a set of objects separate from a "color" set. Together with the insensitivity of the orientation SD to manipulating the color size (see the comparisons between all identical and identical orientation displays in Experiment 2), this keeps supporting the independent storage of colors and orientations regardless of object integration. However, orientation precision suffers much more from object separation when three orientations are to be remembered. This suggests that, under high load, the task to remember both colors and orientations becomes more detrimental, at least for the orientation memory, if the colors and orientations are spatially separated. Together with the demonstration that higher color load stronger impairs the orientation *SD* in separated objects (see the comparisons between identical color and all different features displays in Experiment 2), this suggests that spatial separation of features limits their independent storage.

6. General discussion

Our principal research question was about the relationship between feature-based and object-based unit organization in VWM. In particular, we tested whether the finding that features from two different dimensions, color and orientation, can be stored without substantial interference (Wang et al., 2017; Wheeler & Treisman, 2002) is related to object-based coordination between these features. In other words, we tested whether the absence of interference is due to the fact that each particular color goes with a certain orientation within a unitary object (Duncan, 1984; Luck & Vogel, 1997). In our experiments, we implemented the same approach as Wang et al. (2017) used in their work to test independence or interdependence of VWM resources for color and orientation. This approach is based on the orthogonal manipulation with set sizes in each dimension. Our critical addition to this manipulation was spatial separation vs. spatial integration of features from different dimensions in a paradigm very much resembling that used by Fougnie et al. (2010). It is supposed that spatial separation would cause features to be perceived and encoded as belonging to different objects, whereas spatial integration would cause the features to be encoded as belonging to the same objects. One could question object unity in Experiment 3 where two geometrical shapes were overlaid, but in fact spatial overlap seems to be a strong factor that aids the formation of object-like units (Rensink, 2000; Trick & Pylyshyn, 1993; Wolfe & Bennett, 1997; Xu, 2002). Overlaying could indeed create single object representation (Xu, 2002), but connections between different dimensions represented by different parts of the object are weaker than connections between different features represented by one object (Fougnie et al., 2010; Xu, 2002). We suggest that overlapped objects are still able to create a two-part representation of a more complex object, similar to beachball-like or Saturn-like objects in Xu's (2002) study, so that the originally separate features of the more elementary objects can become the features of a single complex objects.

Using the continuous report paradigm, we replicated the basic finding made by Wang et al. (2017) in the change detection paradigm. When colors and orientations belonged to the same set of objects (Experiments 1 and 3), we found no evidence of cross-dimensional interference. Both capacity (P_{memory}) and precision (*SD*) for colors stayed intact when the number of orientations increased, and vice versa. Together with intra-dimensional interference remarkably growing with a set size, this supports the conclusion about the independent capacities for features from different dimensions (Shin & Ma, 2017; Wang et al., 2017; Wheeler & Treisman, 2002). Moreover, the pattern was mirrored quite consistently by another parameter, *SD* indicating the precision of a VWM trace. Together, these findings corroborate the robustness of the main conclusion made by Wang et al. (2017). It is also important to note that this pattern is still observable in Experiment 2 (Fig. 6) when colors and orientations were distributed between spatially separated objects. This supports the idea that observers could selectively extract different relevant features from different objects and “put” these features to relatively separate storages. This selective encoding could be aided by the way the objects were organized in Experiment 2 (see also Fougnie et al., 2010). The round shape of “color” objects made them “orientation-free”, and the white color of “orientation” objects made them “color-free”. Moreover, shape differences between the “color” and the “orientation” objects could provide stronger grouping between all objects of the same type and weaker grouping between objects of different types, thus encouraging independent feature encoding. Apart from our main topic, it is an interesting question for future research whether observers are capable of selective encoding of different features from different objects when the objects are variable in both dimensions. Whatsoever, the basically consistent pattern across all three experiments suggests that observers are more efficient in dealing with an increased VWM load across feature dimensions than within these dimensions, both within spatially integrated and spatially separated objects. This conclusion supports multiple feature-based storages in VWM.

Having said that, we also found that the prevailing pattern of the independent color and orientation storages was modulated by the spatial separation of these features in Experiment 2. Here, we found some signs of cross-dimensional interference, although they manifested only in the orientation domain. Moreover, we found that the precision of orientation reports is substantially more prone to the detrimental effect of the increased orientation set size when colors are to be remembered from different rather than same objects. This can indicate that object separation limits the totally independent storage of features from different dimensions.

This pattern of results leads us to a conclusion that may seem paradoxical. On one hand, we demonstrated that features from different dimensions can be stored independently from each other. On the other hand, this independence is better supported by their belongingness to shared objects. In general, this supports the idea both separate features and integrated feature “bundles” can be hierarchically stored by VWM (Brady et al., 2011; Fougnie et al., 2010, 2013) in such a way that the “bundles” facilitate the encoding and retrieval of features. Interpreting their results from the paradigm similar to our present paradigm, Wang et al. (2017) also speculated about the possibility of the hierarchically organized memories about features and objects. Our experimental manipulations with feature separation and integration provided empirical support for this suggestion.

How can the object-based advantage mediate the feature independence? One possibility is that, when features are separated between different objects, observers have to spread their attention and VWM resources across a greater number of locations and, thus, each feature representation is noisier than when two features are integrated into one location. We did find evidence that all features, in general, were represented with the greater noise in Experiment 2 with feature separation (see also Fougnie et al., 2010). Viewing the noise as an important source of interference in VWM (Bays, 2015; Wilken & Ma, 2004), we could explain the cross-dimensional interference in spatially separated features by overall noisier representations. In this view, adding a new spatially separable item involves increased firing in a neural population with receptive fields corresponding to the location of this item during the entire retention interval (Buschman, Siegel, Roy, & Miller, 2011; Sprague, Ester, & Serences, 2014). The firing rate within each population is normalized (divided) by the activity in the rest of firing populations that respond to other encoded objects, thus attenuating responses in each population and reducing the signal-to-overall noise ratio of each encoded item (Bays, 2014, 2015). However, this explanation can be insufficient. Most importantly, it does not account for interference specificity towards a feature dimension. Therefore, structural links between individual feature representations can be important for understanding the difference between integrated and separated features. Our experiments were not designed to explore particular structures. Future theoretical analysis and following experiments would be necessary for that field to advance our understanding of VWM beyond the dichotomous “feature-based vs. object-based” scale.

Acknowledgements

The study was conducted within the Basic Research Program of the National Research University Higher School of Economics in 2018 (TZ-57) and 2019 (TZ-54).

Contributions

Y.A.M. designed the experiments, prepared the experimental scripts, collected and analyzed data, and wrote the manuscript. N.A.T. collected and analyzed data and wrote the manuscript. I.S.U. conceptualized the basic ideas, designed the experiments and wrote the manuscript.

Appendix A

In Appendix A, additional statistical results are provided showing the exact estimates of effects mentioned as non-significant in Result sections of Experiments 1–3.

Experiment 1

P_{memory} for color

There were no significant differences between conditions with identical color across objects (color identical vs all identical: $t(18) = 2.819$, $p = .0114$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.647$, $BF_{10} = 0.801$; color identical vs one object: $t(18) = 2.281$, $p = .0349$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.523$, $BF_{10} = 4.091$; all identical vs one object: $t(18) = 0.565$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.13$, $BF_{10} = 0.257$) and also between conditions with different color across objects (all different features vs orientation identical: $t(18) = 2.938$, $p = .0088$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.674$, $BF_{10} = 2.265$).

P_{memory} for orientation

There were no significant differences between conditions with identical orientation across objects (orientation identical vs all identical features: $t(18) = 0.473$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.109$, $BF_{10} = 2.287$; orientation identical vs one object: $t(18) = 2.011$, $p = .0595$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.461$, $BF_{10} = 16.141$; all identical vs one object: $t(18) = 2.428$, $p = .0259$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.557$, $BF_{10} = 4.222$) and also between conditions with different orientation across objects (all different features vs color identical: $t(18) = 0.964$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.221$, $BF_{10} = 0.280$).

SD for color

There were no significant differences between following conditions (all different features vs orientation identical: $t(18) = 0.632$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.145$, $BF_{10} = 0.186$; all different features vs all identical: $t(18) = 2.653$, $p = .0162$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.609$, $BF_{10} = 10.277$; color identical vs orientation identical: $t(18) = 3.063$, $p = .0067$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.703$, $BF_{10} = 1.32$; color identical vs all identical: $t(18) = 0.063$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.015$, $BF_{10} = 0.271$; color identical vs one object: $t(18) = 1.482$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.34$, $BF_{10} = 0.888$; orientation identical vs all identical: $t(18) = 2.096$, $p = .0505$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.481$, $BF_{10} = 18.888$; orientation identical vs one object: $t(18) = 3.094$, $p = .0063$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.71$, $BF_{10} = 10.975$; all identical vs one object: $t(18) = 1.988$, $p = .0622$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.456$, $BF_{10} = 0.202$).

SD for orientation

There were no significant differences between conditions with identical orientation across objects (orientation identical vs all identical features: $t(18) = 0.228$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.052$, $BF_{10} = 0.221$; orientation identical vs one object: $t(18) = 0.997$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.229$, $BF_{10} = 0.15$; all identical vs one object: $t(18) = 1.102$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.253$, $BF_{10} = 0.195$) and also between conditions with different orientation across objects (all different features vs color identical: $t(18) = 1.817$, $p = .0859$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.417$, $BF_{10} = 430.193$).

Experiment 2

P_{memory} for color

There were no significant differences between conditions with identical color across objects (color identical vs all identical: $t(18) = 0.287$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.066$, $BF_{10} = 0.246$; color identical vs one object: $t(18) = 1.666$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's d = 0.382, $BF_{10} = 0.762$; all identical vs two objects: $t(18) = 1.104$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.253$, $BF_{10} = 0.404$) and also between conditions with different color across objects (all different features vs orientation identical: $t(18) = 0.168$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.039$, $BF_{10} = 0.241$).

P_{memory} for orientation

We found no significant differences between following conditions (all different features vs color identical: $t(18) = 0.007$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.002$, $BF_{10} = 0.237$; all identical vs two objects: $t(18) = 1.474$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.338$, $BF_{10} = 0.6$).

SD for orientation

We found no significant differences between following conditions (orientation identical vs all identical: $t(18) = 0.861$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.198$, $BF_{10} = 0.33$; orientation identical vs two objects: $t(18) = 1.068$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.245$, $BF_{10} = 0.391$; all identical vs two objects: $t(18) = 0.427$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.098$, $BF_{10} = 0.258$).

Experiment 3

P_{memory} for color

There were no significant differences between conditions with identical color across objects (color identical vs all identical: $t(14) = 0.878$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.227$, $BF_{10} = 0.366$; color identical vs one object: $t(14) = 0.71$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.183$, $BF_{10} = 0.327$; all identical vs one object: $t(14) = 0.022$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.006$, $BF_{10} = 0.262$) and also between conditions with different color across objects (all different features vs orientation identical: $t(14) = 1.446$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.373$, $BF_{10} = 0.624$).

P_{memory} for orientation

There were no significant differences between conditions with identical orientation across objects (orientation identical vs all identical: $t(14) = 1.447$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.374$, $BF_{10} = 0.624$; orientation identical vs one object: $t(14) = 0.334$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.086$, $BF_{10} = 0.276$; all identical vs one object: $t(14) = 1.469$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.379$, $BF_{10} = 0.64$) and also between conditions with different orientation across objects (all different features vs color identical: $t(14) = 1.284$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.331$, $BF_{10} = 0.524$).

SD for orientation

There were no significant differences between conditions with identical orientation across objects (orientation identical vs all identical: $t(14) = 0.379$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.098$, $BF_{10} = 0.28$; orientation identical vs one object: $t(14) = 0.495$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.128$, $BF_{10} = 0.292$; all identical vs one object: $t(14) = 0.639$, $p = .1$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.165$, $BF_{10} = 0.314$) and also between conditions with different orientation across objects (all different features vs color identical: $t(14) = 2.014$, $p = .0636$, Bonferroni corrected $\alpha = 0.005$, Cohen's $d = 0.52$, $BF_{10} = 1.278$).

References

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science: A Journal of the American Psychological Society/APS, 15*(2), 106–111. https://doi.org/10.1167/2.7.273.

Baddeley, A. D. (1986). *Working memory.* Oxford, UK: Clarendon Press.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.). *The psychology of learning and motivation: Advances in research and theory* (pp. 47–89). New York: Academic.

Bays, P. M. (2014). Noise in neural populations accounts for errors in working memory. *Journal of Neuroscience, 34*(10), 3632–3645. https://doi.org/10.1523/jneurosci.3204-13.2014.

Bays, P. M. (2015). Spikes not slots: Noise in neural populations limits working memory. *Trends in Cognitive Sciences, 19*(8), 431–438. https://doi.org/10.1016/j.tics.2015.06.004.

Bays, P. M., Catalao, R. F., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision, 9*(10), 1–11. https://doi.org/10.1167/9.10.7.

Bays, P. M., Wu, E. Y., & Husain, M. (2011). Storage and binding of object features in visual working memory. *Neuropsychologia, 49*(6), 1622–1631. https://doi.org/10.1016/j.neuropsychologia.2010.12.023.

Brady, T. F., Konkle, T., & Alvarez, G. A. (2011). A review of visual memory capacity: Beyond individual items and toward structured representations. *Journal of Vision, 11*(5), 4. https://doi.org/10.1167/11.5.4.

Buschman, T. J., Siegel, M., Roy, J. E., & Miller, E. K. (2011). Neural substrates of cognitive capacity limitations. *Proceedings of the National Academy of Sciences of USA, 108*(27), 11252–11255. https://doi.org/10.1073/pnas.1104666108.

Cowan, N. (2001). The magical number 4 in short term memory. A reconsideration of storage capacity. *Behavioral and Brain Sciences, 24*(4), 87–186. https://doi.org/10.1017/S0140525X01003922.

Cowan, N., Chen, Z., & Rouder, J. N. (2004). Constant capacity in an immediate serial-recall task: A logical sequel to Miller (1956). *Psychological Science, 15*(9), 634–640. https://doi.org/10.1111/j.0956-7976.2004.00732.x.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113*(4), 501–517. https://doi.org/10.1037/0096-3445.113.4.501.

Fougnie, D., & Alvarez, G. A. (2011). Object features fail independently in visual working memory: evidence for a probabilistic feature-store model. *Journal of Vision, 11*(12), https://doi.org/10.1167/11.12.3 3. doi:10.1167/11.12.3.

Fougnie, D., Asplund, C. L., & Marois, R. (2010). What are the units of storage in visual working memory? *Journal of Vision, 10*(12), 27. https://doi.org/10.1167/10.12.27.

Fougnie, D., Cormiea, S. M., & Alvarez, G. A (2013). Object-based benefits without object-based representations. *Journal of Experimental Psychology: General, 142*(3), 621–626. https://doi.org/10.1037/a0030300.

Fougnie, D., & Marois, R. (2009). Attentive tracking disrupts feature binding in visual working memory. *Visual Cognition, 17*(1–2), 48–66. https://doi.org/10.1080/13506280802281337.

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*(2), 325–347. https://doi.org/10.1037/xlm0000031.

JASP Team (2018). *JASP (Version 0.9.0.0).* [Computer software].

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

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24*(2), 175–219. https://doi.org/10.1016/0010-0285(92)90007-O.

Lee, D., & Chun, M. M. (2001). What are the units of visual short-term memory, objects or spatial locations? *Perception & Psychophysics, 63*(2), 253–257. https://doi.org/10.3758/BF03194466.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*(6657), 279–281. https://doi.org/10.1038/36846.

Luria, R., & Vogel, E. K. (2011). Shape and color conjunction stimuli are represented as bound objects in visual working memory. *Neuropsychologia, 49*(6), 1632–1639. https://doi.org/10.1016/j.neuropsychologia.2010.11.031.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 101*(2), 343–352. https://doi.org/10.1037/h0043158.

Oberauer, K., & Eichenberger, S. (2013). Visual working memory declines when more features must be remembered for each object. *Memory and Cognition, 41*(8), 1212–1227. https://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 & Psychophysics, 64*(7), 1055–1067. https://doi.org/10.3758/BF03194756.

Pertzov, Y., Dong, M. Y., Peich, M. C., & Husain, M. (2012). Forgetting what was where: The fragility of object-location binding. *PLoS One, 7*(10), https://doi.org/10.1371/journal.pone.0048214.

Pierce, J. W. (2007). PsychoPy – Psychophysics software in python. *Journal of Neuroscience Methods, 162*, 8–13.

Rensink, R. A. (2000). The dynamic representation of scenes. *Visual Cognition, 7*(1–3), 17–42. https://doi.org/10.1080/135062800394667.

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*(2), 225–237. https://doi.org/10.3758/PBR.16.2.225.

Schurgin, M. W., Wixted, J. T., & Brady, T. F. (2018). Psychological Scaling Reveals a Single Parameter Framework For Visual Working Memory. *BioRxiv,* 325472. https://doi.org/10.1101/325472.

Shin, H., & Ma, W. J. (2017). Visual short-term memory for oriented, colored objects. *Journal of Vision, 17*(9), 1–19. https://doi.org/10.1167/17.9.12.

Sprague, T. C., Ester, E. F., & Serences, J. T. (2014). Reconstructions of information in visual spatial working memory degrade with memory load. *Current Biology, 24*(18), 2174–2180. https://doi.org/10.1016/j.cub.2014.07.066.

Suchow, J. W., Brady, T. F., Fougnie, D., & Alvarez, G. A. (2013). Modeling visual working memory with the MemToolbox. *Journal of Vision, 13*(10), 9. https://doi.org/10.1167/13.10.9.

Treisman, A. (1999). Solutions to the binding problem: Progress through controversy and convergence. *Neuron, 24*(1), 105–110. https://doi.org/10.1016/S0896-6273(00)80826-0.

Trick, L. M., & Pylyshyn, Z. W. (1993). What enumeration studies can show us about spatial attention: Evidence for limited capacity preattentive processing. *Journal of Experimental Psychology: Human Perception and Performance, 19*(2), 331–351. https://doi.org/10.1037/0096-1523.19.2.331.

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions and objects in visual working memory. *Journal of Experimantal Psychology: Human Perception and Performance, 27*(1), 92–114. https://doi.org/10.1037//0096-1523.27.1.92.

Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., ... Morey, R. D. (2017). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review,* 1–19. https://doi.org/10.3758/s13423-017-1323-7.

Wang, B., Cao, X., Theeuwes, J., Olivers, C. N. L., & Wang, Z. (2017). Separate capacities for storing different features in visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*(2), 226–236. https://doi.org/10.1037/xlm0000295.

Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology. General, 131*(1), 48–64. https://doi.org/10.1037//0096-3445.131.1.48.

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision, 4*(12), 11. https://doi.org/10.1167/4.12.11.

Wolfe, J. M., & Bennett, S. C. (1997). Preattentive object files: Shapeless bundles of basic features. *Vision Research, 37*(1), 25–43. https://doi.org/10.1016/S0042-6989(96)00111-3.

Woodman, G. F., & Vogel, E. K. (2008). Selective storage and maintenance of an object's features in visual working memory. *Psychonomic Bulletin and Review, 15*, 223–229. https://doi.org/10.3758/PBR.15.1.223.

Xu, Y. (2002). Encoding color and shape from different parts of an object in visual short-term memory. *Perception & Psychophysics, 64*(8), 1260–1280. https://doi.org/10.3758/BF03194770.

Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature, 440*(7080), 91–95. https://doi.org/10.1038/nature04262.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature, 453*(7192), 233–235. https://doi.org/10.1038/nature06860.

Appendix B. Real-world objects are not stored in holistic representations in visual working memory

Article "Real-world objects are not stored in holistic representations in visual working memory"

Markov, Y.A., Utochkin, I. S., & Brady T. F. (2021). Real-world objects are not stored in holistic representations in visual working memory. *Journal of Vision, 21*(3): 18, 1–24. DOI: 10.1167/jov.21.3.18.

Abstract. When storing multiple objects in visual working memory, observers sometimes misattribute perceived features to incorrect locations or objects. These misattributions are called binding errors (or swaps) and have been previously demonstrated mostly in simple objects whose features are easy to encode independently and arbitrarily chosen, like colors and orientations. Here, we tested whether similar swaps can occur with real-world objects, where the connection between features is meaningful rather than arbitrary. In Experiments 1 and 2, observers were simultaneously shown four items from two object categories. Within a category, the two exemplars could be presented in either the same or different states (e.g., open/closed; full/empty). After a delay, both exemplars from one of the categories were probed, and participants had to recognize which exemplar went with which state. We found good memory for state information and exemplar information on their own, but a significant memory decrement for exemplar–state combinations, suggesting that binding was difficult for observers and swap errors occurred even for meaningful real-world objects. In Experiment 3, we used the same task, but in one-half of the trials, the locations of the exemplars were swapped at test. We found that there are more errors in general when the locations of exemplars were swapped. We concluded that the internal features of real-world objects are not perfectly bound in working memory, and location updates impair object and feature representations. Overall, we provide evidence that even real-world objects are not stored in an entirely unitized format in working memory.

Real-world objects are not stored in holistic representations in visual working memory

Yuri A. Markov HSE University, Moscow, Russia

Igor S. Utochkin HSE University, Moscow, Russia

Timothy F. Brady Psychology Department, University of California, San Diego, La Jolla, CA, USA

When storing multiple objects in visual working memory, observers sometimes misattribute perceived features to incorrect locations or objects. These misattributions are called binding errors (or swaps) and have been previously demonstrated mostly in simple objects whose features are easy to encode independently and arbitrarily chosen, like colors and orientations. Here, we tested whether similar swaps can occur with real-world objects, where the connection between features is meaningful rather than arbitrary. In Experiments 1 and 2, observers were simultaneously shown four items from two object categories. Within a category, the two exemplars could be presented in either the same or different states (e.g., open/closed; full/empty). After a delay, both exemplars from one of the categories were probed, and participants had to recognize which exemplar went with which state. We found good memory for state information and exemplar information on their own, but a significant memory decrement for exemplar–state combinations, suggesting that binding was difficult for observers and swap errors occurred even for meaningful real-world objects. In Experiment 3, we used the same task, but in one-half of the trials, the locations of the exemplars were swapped at test. We found that there are more errors in general when the locations of exemplars were swapped. We concluded that the internal features of real-world objects are not perfectly bound in working memory, and location updates impair object and feature representations. Overall, we provide evidence that even real-world objects are not stored in an entirely unitized format in working memory.

Introduction

Working memory is a limited capacity system (Cowan, 2001; Miller, 1956) used to actively maintain and work with the information necessary for our current goals and tasks (Baddeley, 1986; Baddeley & Hitch, 1974). However, it is unclear what the nature is of the representations stored in visual working memory. Most early work suggested that information about entire objects is represented in intrinsically holistic, totally bound units, with all features stored or forgotten together (Cowan, Chen, & Rouder, 2004; Kahneman, Treisman, & Gibbs, 1992; Lee & Chun, 2001; Luck & Vogel, 1997; Luria & Vogel, 2011; Treisman, 1999; Vogel, Woodman, & Luck, 2001; Xu, 2002; Xu & Chun, 2006); however, since those studies were undertaken, many other studies have provided evidence that there is—either also or instead—relatively independent feature storage (see Brady, Konkle, & Alvarez, 2011, for review; ; Bays, Catalao, & Husain, 2009; Bays, Wu, & Husain, 2011; Fougnie & Alvarez, 2011; Fougnie, Cormiea, & Alvarez, 2013; Markov, Tiurina, & Utochkin, 2019; Pertzov, Dong, Peich, & Husain, 2012; Shin & Ma, 2017; Wang, Cao, Theeuwes, Olivers, & Wang, 2017; Wheeler & Treisman, 2002).

In the foundational study about this issue, Luck and Vogel (1997) claimed that only objects, not features, limit the capacity of visual working memory, because they found no decrement in performance with additional features per object, even within the same dimension. Luck and Vogel (1997) suggested, therefore, that unlike in perception, where illusory conjunctions occur and features seem to be unbound to some extent (Treisman, 2006), unitized objects are the "units" of visual working memory. This finding was in line with the "strong" object hypothesis, which claims that visual working memory is limited only by a number of objects and that features play no role in working memory limits and are only forgotten when entire objects are forgotten. However, further research provided evidence against this "strong" object view: multiple features from the same dimension cannot be stored without cost, even if they are on the same objects (Olson & Jiang,

Citation: Markov, Y. A., Utochkin, I. S., & Brady, T. F. (2021). Real-world objects are not stored in holistic representations in visual working memory. *Journal of Vision*, *21*(3):18, 1–24, https://doi.org/10.1167/jov.21.3.18.

https://doi.org/10.1167/jov.21.3.18 Received July 20, 2020; published March 17, 2021 ISSN 1534-7362

2002; Wheeler & Treisman, 2002); detecting changes that require binding is harder than detecting changes that do not (van Lamsweerde, et al. 2015); additional resources are needed for keeping bound features in memory (Fougnie, Asplund, & Marois, 2010; Fougnie & Marois, 2009); features can independently fade from memory (Fougnie & Alvarez, 2011; Fougnie, Cormiea, & Alvarez, 2013); and memories for different sensory dimensions rely on independent storage capacities (Markov et al., 2019; Wang, Cao, Theeuwes, Olivers, & Wang, 2017). Thus, although the early evidence was mixed on this issue, it is now clear that for simple stimuli like colors and orientations, items in working memory are not inherently represented in a solely object-based, unitized manner (e.g., Cowan et al., 2013; Park et al., 2017). Although participants can maintain the binding between features (at least partially by using location as a cue), this is not because the objects are themselves stored in a single holistic representation that by necessity is encoded and forgotten in an all-or-none manner, as suggested by early work; instead, features rely on distinct capacities, accumulate independent noise, and can be lost independently.

One result of this is that, similar to perceptual illusory conjunctions, binding errors also occur in visual working memory studies (Bays, 2016; Bays, Catalao, & Husain, 2009; Bays, Wu, & Husain, 2011; Dent & Smyth, 2005; Emrich & Ferber, 2012; Pertzov et al., 2012), and this occurs even though these studies use displays where such errors are unlikely to have perceptual origins (e.g., for a set of four objects the percentage of binding errors is approximately 10%; see Bays et al., 2009; Emrich & Ferber, 2012). This finding suggests that memory representations are also prone to the binding problem, at least in circumstances where location noise is considerable (Oberauer & Lin, 2017). Thus, the evidence suggests that both the binding of features within an object and the binding to location of objects are often imperfect in visual working memory.

Are real-world objects likely to be stored holistically in visual working memory?

Existing studies have almost exclusively tested simple objects with features that are easy to manipulate in experiments (geometrical shapes with various colors, orientations, etc.). There is very little research examining the storage of real-world objects in visual working memory. In comparison with the simple objects that are usually used in standard visual working memory tasks, real-world objects have many more features, both visual and semantic. and the connections between these features are meaningful rather than arbitrary. Although simple features (e.g., color and orientation) each have somewhat or complete independence in their underlying storage capacity (Markov et al., 2019; Wang et al., 2017) and are in some cases even represented by separate neurons or structures (Conway, 2009; Paik & Ringach, 2011), the complex features of real-world objects are not nearly as separable from each other as basic visual features like color and orientation, at least in terms of their visual properties: a cabinet being "open", for example (rather than "closed"), results in changes to spatial frequency, color, orientation, shape, and many other visual aspects of the object. Thus, one possibility is that real-world objects are stored in memory in a way that is effectively unitized—that, rather than distinct features being encoded and lost separately, and requiring effort or resources to bind, the objects are stored and remembered in a wholly all-or-none manner.

There are several bodies of work from outside of visual working memory that are consistent with this possibility and that can be interpreted as predicting that object memories should be holistic and all-or-none. For example, a large body of work looking at ventral visual processing shows that the individual low-level features that make something an object (e.g., a mug)—the curves and colors and spatial frequencies—are untangled during visual processing into a more general mug representation as processing proceeds to higher level visual areas (DiCarlo, Zoccolan, & Rust, 2012). Thus, unlike low-level features with arbitrary bindings, there are preexisting mid- and high-level representations of many aspects of real-world objects that could be used for working memory storage, perhaps suggesting that, if memory relies on such high-level representations, memory should consist of a relatively unitized object representation (e.g., a mug representation), rather than separable memories for separate properties of objects. Indeed, some studies on cortical representation of objects argues in that the medial temporal lobe as well as more ventral visual regions, 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 hypothesis 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 designed to prevent confusions of similar items.

In addition to the question of binding, there are also other pieces of evidence consistent with the idea that participants store real-world objects differently than simpler objects, which could result in qualitatively distinct representations. For example, real-world objects, compared with simple stimuli, allow access to significant additional information (e.g., the real size of the objects: Konkle & Oliva, 2012; Long, Konkle, Cohen, & Alvarez, 2016; Long, Moher, Carey, & Konkle, 2019; expected nearby objects and their spatial position: Kaiser, Stein, & Peelen, 2015); O'Donnell, Clement, & Blockmole, 2018, and people can have specific expertise with certain object categories (Curby & Gauthier, 2007; Curby, Glazek, & Gauthier, 2009; Janini & Konkle, 2019; Xie & Zhang, 2017), all of which may be used to enhance working memory. In fact, several studies have shown that the capacity of visual working memory for real-world objects differs from that of simple stimuli, in particular being less fixed and more dependent on the particular stimuli used and how much meaningful information about them can be processed (Asp, Störmer & Brady, 2019; Brady, Störmer, & Alvarez, 2016). For example, Brady et al. (2016) showed a boost in performance for real-world objects that was attributable to more active storage in visual working memory, consistent with a theory where additional high-level information about such objects, perhaps in the ventral stream, is maintained in working memory in addition to low-level information. Some recent studies (Li, Xiong, Theeuwes, & Wang, 2020; Quirk, Adam, & Vogel, 2020) instead found no difference between storing simple features and real-world objects in visual working memory, but these results were likely due to a lack of control for similarity between targets and foils in the color versus real-world object tasks (Brady & Störmer, 2020; Brady & Störmer, in press). With better control for target–foil similarity (Brady & Störmer, 2020), real-world objects result in significantly better performance compared with simple features (Brady & Störmer, 2020; Brady & Störmer, in press).

Altogether, then, there is significant evidence that real-world objects differ from simple stimuli in working memory and there are reasons to believe that real-world objects may be stored in a more holistic manner because they depend on more high-level representations that have been argued to be based on unitized object representations. Are real-world objects, then, stored in a unitized, all-or-none format in visual working memory? Or can different features of such objects be lost independently, or misbound?

This subject has been largely unaddressed to date. In fact, there are only a few studies that have used real-world objects to investigate binding in visual working memory, and this work has been done mainly in the context of object–location binding (e.g., Lew & Vul, 2015; Pertzov et al., 2012). In such tasks, researchers test memory for item identities and for their locations. In such object–location tests, memory failures can come from forgetting objects, forgetting locations, or forgetting which object was in which location. Because it is widely believed that objects and locations are stored relatively independently, even for realistic objects (e.g., ventral and dorsal pathways; Haxby et al., 1991; Mishkin & Ungerleider, 1982), it is perhaps not surprising that such independence has been found in such studies. However, whether the object's internal features are themselves stored in a holistic, bound representation is not considered within the scope of this existing object–location binding research.

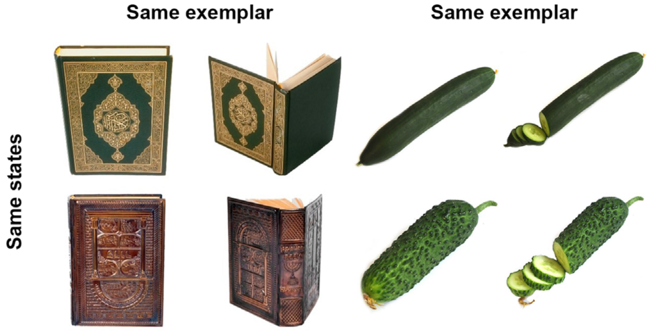


Figure 1. Example of two different exemplars in two different states.

The question of whether real-world objects are stored in bound units has, however, been previously addressed in the studies of visual long-term memory (Balaban, Assaf, Arad Meir, & Luria, 2020; Brady, Konkle, Alvarez, & Oliva, 2013; Spachtholz & Kuhbandner, 2017; Utochkin & Brady, 2020). These studies have largely found evidence supporting a "weak" object view; what Olson and Jiang (2002) used to designate representations that are somewhat bound—that is, people know which features went with which others to some extent—but that are not intrinsically all-or-none representations, and certainly not totally holistic representations (as in Luck & Vogel, 1997).

For example, Brady et al. (2013) showed that information about an object's color is forgotten faster than information about the state the object was in (whether the cucumber was whole or sliced; whether the book was open or closed; Figure 1). Furthermore, they showed forgetting is independent for information about object state and information about the specific object exemplar (whether I saw this or that cucumber; or this book or that one; see Figure 1). In other words, Brady et al. showed that the features of real-world objects can be forgotten independently. This finding is consistent with the demonstrations of independent forgetting of simple features in working memory (Fougnie & Alvarez, 2011; Fougnie, Cormiea, & Alvarez, 2013), and suggestive that the representation of real-world objects, at least in long-term memory, is higher level; that is, that information about state and exemplar are encoded distinctly. In their recent work, Utochkin and Brady (2020) tested other predictions that derive from the idea of independent storage of real-world object features in long-term memory. In particular, they demonstrated that people often commit "swap" errors, when they show good memory for both which exemplars and which object states they have seen, but frequently choose an incorrect combination of the exemplar and the state. In particular, Utochkin and Brady (2020) asked participants to remember different exemplars from the same category presented in the same state (e.g., two different coffee mugs, both empty) or in different states (e.g., two different coffee mugs, one empty and one full). They found that participants had good memories for the states and exemplars alone, but when the two exemplars had been seen in different states, participants were at chance in correctly matching which state went with which exemplar, often reporting swapped states for the two exemplars. Thus, Utochkin and Brady (2020) concluded that state and exemplar information are represented independently, rather than in an all-or-none, holistic representation in long-term memory—rejecting the "strong" object claim and casting doubt on even "weak" object-based representations in favor of largely independent storage of high-level object properties like exemplar and state.

Although this work argues for extremely independent representations, other work provides evidence in favor of at least a "weak" object-based view, that is, representations that are dependent between features, even if not all-or-none or holistic. For example, Balaban et al. (2020) argued that there is a dependence between different features of objects in long-term memory, and that at minimum state information cannot be stored without storing the exemplar information, so that exemplars and states are stored in a hierarchical structure or some other form of dependence between them is present. However, the hierarchical structure is not consistent with strong, holistic all-or-none object-based views and leaves the nature of the dependence unclear.

This evidence of at least partially independent long-term memory storage of different features of real-world objects leads to important questions about the representation of these objects in working memory. Are high-level properties of these objects (like object state vs. object exemplar) more holistic and integrated when stored in working memory, which may rely more on sustained perceptual activation in high-level visual areas than does long-term memory? Many accounts of working memory suggest that working memory is more focused on the storage of perceptual features per se, rather than semantic features, suggesting that items are stored in a more perceptual format in working memory than long-term memory (e.g., Baddeley, 1966). As applied to visual working memory, this account would suggest that even real-world objects might be stored solely as perceptual features. This account would be broadly consistent with models of visual working memory that argue memory storage occurs sustained activity in sensory visual cortex; that is, even real-world objects are stored in terms of their colors and shapes and orientations, not in terms of semantically meaningful object features (e.g., Harrison & Tong, 2009; Serences, Ester, Vogel, & Awh, 2009; for review, see Serences, 2016). If this is the case—that visual working memory is strictly perceptual—then nearly all high-level changes to real-world objects (like a change of the state or exemplar) would be expected to appear strongly bound in this memory system, because almost any change in either dimension would change the color, shape, orientation, and so on. By contrast, 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 that we perform every day in the context of real-world objects, if objects are stored in more meaningful representations, rather than purely as perceptual features, they could appear quite independent. Indeed, this would be consistent with some evidence that the inferior temporal cortex, where objects are likely to be represented in richer ways (e.g., via parts) is involved in visual working memory (Fiebach, Rissman, & D'Esposito, 2006; Li, Miller, & Desimone, 1993; Miller, Erickson, & Desimone, 1996; Ranganath, DeGutis, & D'Esposito, 2004), and with suggestions that visual working memory involves more abstracted representations in parietal cortex rather than low-level sensory representations (e.g., Xu, 2017). Thus, to some extent the question of how bound they are in working memory is a question about the nature of visual working memory storage for real-world objects more generally.

In addition, if we consider working memory representation as a precursor for consolidation into long-term memory, then it could be that the independent long-term memory storage of such high-level features like object state and object exemplar could plausibly be accounted for by limited binding in working memory. Utochkin and Brady (2020) briefly addressed this issue using a short-term memory version of their exemplar–state memory task. However, their task was extremely straightforward, with only two objects to remember, and thus did not load working memory capacity in the ways that are known to cause more misbinding reports for simple stimuli and basic features (Bays et al., 2009, 2011; Emrich & Ferber, 2012; Pertzov et al., 2012).

Thus, in the current study, we tested the boundedness of real-world object representations in visual working memory using a relatively high-load task requiring participants to remember four items presented simultaneously, akin to standard visual working memory studies with simple stimuli (e.g., Luck & Vogel, 1997). As the possibly independent features of real-world objects, we rely on the previous long-term memory 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) can be forgotten or represented independent 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), as in Utochkin and Brady (2020). Note that although we refer to state and exemplar as object properties, they are not like the simple features of color and orientation often used in working memory tasks. Instead, as noted elsewhere in this article, the visual features used to discriminate them 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 both common and important tasks that need to be regularly performed with real-world objects. Thus, in the present work we focus on these dimensions and ask whether these two aspects of objects are stored in a unitized format in visual working memory. Thus, in all experiments, our participants performed different modifications of the exemplar–state task. This task required participants to report the state (open/closed, full/empty, etc.) of each of two exemplars from the same category that were presented in the working memory display. The idea of this double recognition test was to dissociate exemplar memory and state memory alone from memory for exemplar–state conjunctions (e.g., binding). In Experiments 1 and 2, we asked whether observers are more accurate with exemplar and state memory alone than their conjunction. In Experiment 3, we tested how this relates to spatial location, that is, whether observers might separately update the location of two features when the object moves spatial locations (with only a single feature staying attached to an old location).

Experiment 1

In Experiment 1, we asked participants to remember four objects in visual working memory: two exemplars from two object categories. The two exemplars from each category could be shown in either the same state or different states. At test, both items of a single category were probed and observers had to recognize the states in which each of the exemplars of that category had been presented. If the information used to discriminate exemplars and the information used to discriminate states are represented in memory in a fully unitized, all-or-none format, we should not observe any differences between the performance of remembering objects in same or different states, because each object would be self-contained, with the features perfectly bound. However, if representations of real-world objects are not unitized and holistic, we anticipate that swap errors could occur between the features of objects (e.g., a mug that was full gets reported as empty because the other mug was empty). Thus, if there is independence in the representation of the features of the objects we predict worse performance for objects presented in different states compared with objects where the two items from the same category are in the same state.

The experiment included two tasks. In the exemplar–state task, we evaluated both whether participants knew which exemplar was in which state (state–exemplar conjunctions), and also whether participants knew whether the two objects had been in the same state or different states than each other (an index of state memory independent of binding). We also measured memory for exemplars alone in a separate exemplar task. Here, participants had to remember two exemplars from two categories (four items in total), but rather than them differing in states and participants needing to recognize the states, instead the test pitted these two previously seen exemplars against two new exemplars from the same category. This task helps to dissociate poor memory for exemplars alone from genuine swap errors.

Method

Participants

Twenty psychology students from the Higher School of Economics, 19 female, age, 18-22 years, $M = 18.8$, took part in the experiment for course credit. All participants reported having normal color vision, normal or corrected-to-normal visual acuity, and no neurological problems. Before the beginning of the experiment, they signed an informed consent form. The sample size was estimated using G*Power 3.1.9.2 (Faul, Erdfelder, Lang, & Buchner, 2009). Our sample size was based on previously reported samples in a similar study of exemplar–state memory (Brady et al., 2013; Utochkin & Brady, 2020)—–15 to 20 participants in one group. The planned sample size also included a few extra participants taking into account the possibility of technical problems or poor performance in some participants. With this sample size, we are able to detect η^2 equal to 0.08 (for repeated measures analysis of variance [ANOVA]) and Cohen's d's equal to 0.7 (two-tailed t test) with an α of 0.05 and power $(1–\beta)$ of 0.8. This is smaller than the effect size reported by previous studies investigating binding errors in visual working memory and visual long-term memory (Bays et al., 2009; Emrich & Ferber, 2012; Pertzov et al., 2012; Utochkin & Brady, 2020), which ranges from 1.1 to 1.9 (Cohen's dz).

Apparatus and stimuli

The experiment was developed and presented via PsychoPy (Peirce et al., 2019) for Linux Ubuntu. Stimuli were presented on a standard CRT monitor with a refresh frequency of 75 Hz and 1,024 × 768-pixel spatial resolution. Stimuli were presented on a homogeneous white field. Participants sat approximately 47 cm from the monitor. From that distance, the screen subtended approximately 42.4° × 32.5° of visual angle.

Three image sets were used in the experiment. For the exemplar task, we used the image set from the study by Konkle, Brady, Alvarez, and Oliva (2010) including more than 360 categories with 2 to 16 exemplars per category. We selected 120 object categories for the exemplar task. For the exemplar–state task, as the items that tested, we used the image set from the study by Brady et al. (2013; also used by Utochkin & Brady, 2020). It contained 120 unique object categories and each category contained two exemplars (e.g., two different books, Figure 1) and each exemplar was represented by two different states (e.g., open books and closed books, Figure 1). For the items that served as distractors on each trial (e.g.., that were not tested), we created a new image set consisting of 60 categories not overlapping with the categories from Brady et al. (2013) and the categories used in the exemplar task. This image set had the same exemplar–state organization as that of Brady et al. (2013), yet not all categories always had the full set of exemplar and state instances. It was sufficient to have at least one exemplar in one state and another exemplar in a different state, so they could be used as learned but not tested items in different states shown together with two subsequently tested items. For studied but not tested items shown in the same states, we used 60 additional categories from Konkle et al. (2010) not overlapping with those 120 used for the exemplar task (two exemplars were drawn from each category).

Sample and test displays: Each to be remembered set (sample) contained four items, each presented at approximately 6.22° × 6.22° of visual angle. The centers of the images laid on an invisible circle with radius 10.3°. The only parameter defining the position of each object was the rotational angle on the imaginary circle. These angles were chosen randomly for each object in each trial with the only restriction that the minimum distance between the centers of any two objects was 30° of rotation. This was done to avoid overlap or clustering between the objects. Two items in the sample were always drawn from one object category and the other two objects were drawn from a different category (Figure 2). At test, two locations corresponding with the centers of two originally presented items—always of the same category—were marked by dots. At 3.3° to the right and to the left of each dot, two images were presented. One of the images at each dot was the exactly the item presented at that location in a sample and another item was always a foil item not presented in the sample. Therefore, one and only one correct answer was present at each probed location and participants had to make two simultaneous two-alternative forced choice judgments. By presenting both tests simultaneously, we decrease the possibility that swap errors arise simply because participants retrieve the wrong exemplar, by making it completely explicit that there are two different relevant items that must be independently remembered.

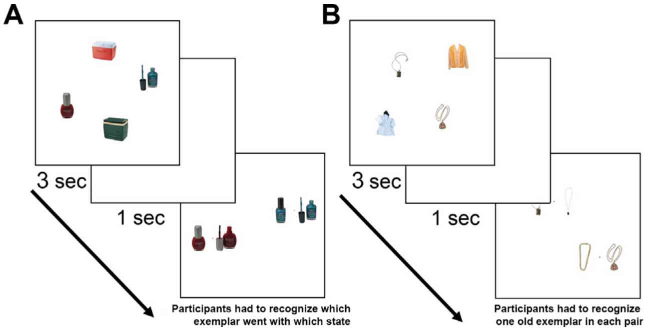


Figure 2. The time course of typical trial in Experiment 1 in (A) the exemplar–state task and (B) the exemplar task. In both tasks, participants had to remember all four initial objects as accurately as possible. Then, after a delay, they had to perform two simultaneous two-alternative forced choice memory tests, picking the correct item at each spatial location. In the exemplar–state task, this meant picking the correct state for each exemplar. In the exemplar task, this meant picking the correct exemplar in each spatial location.

Procedure

The experiment contained two tasks, the exemplar–state and exemplar tasks. The order of the tasks was counterbalanced across participants. During testing, participants were instructed to repeat a syllable “ba” aloud at a rate of about 3 Hz to diminish verbal encoding of stimuli. An experimenter was present in the testing room during the entire experiment and monitored whether the participants followed the instruction to pronounce the syllable.

Exemplar–state task: In each trial, two exemplars from two object categories were presented for 3 seconds (e.g., nail polish A opened, nail polish B closed, cooler A closed, and cooler B closed, Figure 2A). This sample display was followed by a 1-second retention interval (blank screen) and then the test screen was presented. Each of the two probed locations contained two possible states of the same exemplar, exactly the exemplar originally presented at that location in the sample (nail polish A opened vs. nail polish A closed at one location; nail polish B opened vs. nail polish B closed at another location, Figure 2A). Participants had to choose the correct state for each exemplar (double two-alternative forced choice). Exemplars of both the subsequently tested and subsequently not-tested category could be presented in either the same states or in different states. There were 60 trials with exemplars from a tested category presented in the same states and 60 with exemplars from the tested category in different states. Whether the nontested exemplars were presented in the same or different states was manipulated orthogonally to the state manipulation of the tested items. Because the participants did not know in advance which exemplars would be tested and because any categorical pair of objects had an equal chance to appear in the same or different states, our participants needed to remember all four items. For each tested category, exemplars were presented in the same states to one-half of the participants and in different states to another one-half of the participants.

Exemplar task: The procedure of the exemplar task was similar to that of the exemplar–state task in terms of display structure and time course. The main difference was that two items from the same category were always presented in the same state, and the foil items at test were new exemplars from the same category rather than different states of old exemplars. Therefore, on each trial, observers had to pick a single exemplar at each probed location (Figure 2B). There were 60 trials in the exemplar task.

Data analysis

We estimated the overall accuracy (the total number of correctly chosen items) in both the exemplar–state and the exemplar tasks. Report accuracy in the exemplar task was used as a measure of memory for exemplars. Report accuracy in the two conditions of the exemplar–state task was used to estimate the memory for exemplar–state conjunctions. Finally, to estimate state memory, we compared how often the participants reported both items as being in the same states when the studied items had been presented in the same states compared with the trials when the studied items had been presented in different states. This logic was similar to that used by Utochkin and Brady (2020).

We also estimated how often the reported states matched their exemplars in the exemplar–state task. There were three possible outcomes: both correct, one correct, and none correct. If real-world objects are stored in a fully holistic form, we should not observe any difference between these three outcomes as a function of the condition in the exemplar–state task. However, if the features underlying exemplar and state discrimination are stored in some way that is nonholistic, and thus somewhat independent, we should observe an increase in the number of no correct answers for the different states condition, with a concurrent decrease in the number of both correct answers.

The standard frequentist and Bayesian t tests were performed. The Bayesian t-test is a direct way to estimate evidence for H_1 against H_0 (Rouder, Speckman, Sun, Morey, & Iverson, 2009). The Bayes factor (BF10) was calculated using JASP 0.9.0.0 (JASP Team, 2018; Wagenmakers et al., 2017) and interpreted using the standard Jeffreys, 1961. Theory of probability (3rd ed.), Oxford University Press, Oxford. The Cauchy distribution with a width of 0.707 was used as a prior distribution of effect sizes under H_0. A Bonferroni correction was made for multiple comparisons in calculating the statistical significance level.

Results

Overall accuracy

A one-way repeated-measure ANOVA was run to compare the total accuracy between the exemplar task and the two conditions of the exemplar–state tasks (same states vs. different states). We found a substantial effect, $F\,(2{,}18) = 28.99$, $p < 0.001$, $BF_{10} > 10^5$, $\eta^2 = 0.604$, Figure 3A. Comparisons between the exemplar task and two conditions of the exemplar–state tasks found differences between all three. The accuracy of exemplar recognition, $M = 0.86$, was higher than the accuracy of state recognition when the objects were presented in same states, $t(19) = 2.924$, $p = 0.009$, Bonferroni corrected $\alpha = 0.017$, $BF_{10} = 5.749$, Cohen's d = 0.654, and when they were presented in different states, $t(19) = 7.702$, $p < 0.001$, Bonferroni corrected $\alpha = 0.017$, $BF_{10} > 10^5$, Cohen's d = 1.722. Most important, the accuracy of reporting correct exemplar–state conjunctions, $M = 0.81$, was greater in trials when two exemplars were presented in the same states compared with trials when they were presented in two different states, $M = 0.74$; $t\,(19) = 4.772$, $p < 0.001$, Bonferroni corrected $\alpha = 0.017$, $BF_{10} = 216$, Cohen's d = 1.067, Figure 3A.

State memory

The percentage of time participants choose two responses in the same states is our index of state memory. We found that participants did so more often when the objects had actually been presented in the same states compared with when they had been presented in different states, same states: $M = 0.72$, different states: $M = 0.23$; comparison: $t\,(19) = 10.7$, $p < 0.001$, $BF_{10} > 10^6$, Cohen's d = 2.393. This finding suggests that our participants had good memory for which states were presented, regardless of their ability to report which exemplars these states belonged to.

Accuracy of conjunction memory within paired choices

Given the good memory for whether two exemplars had been presented in the same or in different states, we can ask how often these memories were correctly bound to the exemplars. To estimate that, we analyzed the proportions of three possible outcomes of the

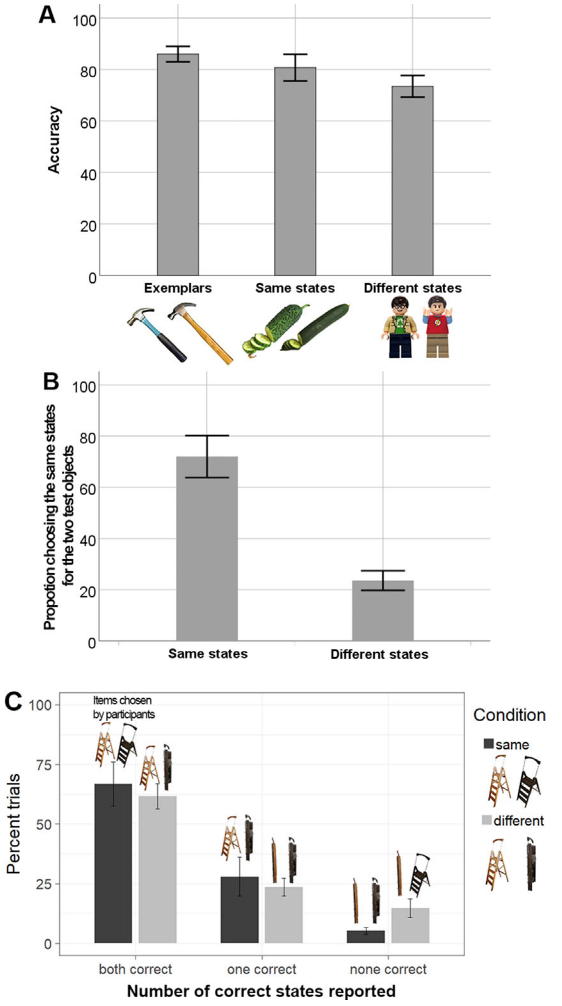


Figure 3. Results for Experiment 1 for overall accuracy (A), state memory accuracy (B) and choosing both, one, or no correct states for exemplars (C). Images on top of the bars in (C) are examples of participants' answers in each of the response outcomes. Note that the axis labels show two the different exemplars of the category (e.g., two different Lego people), either in the same state as each other, or different states than each other. Error bars depict 95% CIs. *LEGO Group. This is an independent site not authorized or sponsored by the LEGO Group*.

paired choices that our observers made on each trial: both states correct, one state correct, or zero correct. Swap errors occur if participants often report both items incorrectly (e.g., knows the items had different states but not which state went with which exemplar). Thus, an excess of 0 correct trials is evidence of swap errors (e.g., knowledge of the states separate from the exemplars they go with).

If the overall accuracy is simply impaired by the objects being in different states (perhaps from greater difficulty encoding them, for example), with no change in binding difficulty per se, participants overall performance should predict their performance for picking both items wrong. For example, with an 81% overall correct response rate, as in the same state condition, if the two objects are responded to independently we expect the proportion of trials with two correct answers to be $(0.81)^2$ (e.g., an independent chance to get each object correct); with item one correct and item two incorrect to be $(0.81)(1 - 0.81)$; with item one incorrect and item two correct to be $(1 - 0.81)(0.81)$; and with zero correct answers to be $(1 - 0.81)^2$. The small decrease in percent correct to 74% in the different state condition can also be analyzed in this way, to see if relatively speaking, participants make more double errors (swaps) than would be expected given how often they get none and one correct. These errors should change predictably if the only difference in conditions is an overall accuracy effect from more difficult encoding rather than independent knowledge of states and exemplars (resulting in a change from $[1 - 0.81]^2 = 3.6\%$ zero correct trials, to $[1 - 0.74]^2 = 6.7\%$ zero correct trials). By contrast, if different states result in binding difficulties, then when the objects are presented in different states, participants should show an abundance of zero correct trials, which represent swap errors (where they knew the states, but not how they went together).

Overall, we found that there was no significant difference between the proportions of choosing both correct answers between the two conditions of exemplar–state task, same states: $M = 0.67$, different states: $M = 0.62$; comparison: $t(19) = 1.609$, $p = 0.124$, $BF_{10} = 0.698$, Cohen's d =.360, and also between the proportions of choosing only one correct answer, same states: $M = 0.28$, different states: $M = 0.24$, comparison: $t(19) = 1.073$, $p = 0.297$, $BF_{10} = 0.386$, Cohen's d = 0.240, because both displayed small decreases that did not reach significance. However, there were significantly more none correct trials for objects shown in different states compared with same states, same states: $M = 0.05$, different states: $M = 0.15$; comparison: $t\ (19) = 5.345$, $p < 0.001$, $BF_{10} = 676.7$, Cohen's d =1.195. This result is indicative of swap errors.

We can also compare this level of no correct answers with the independent responses baseline. In the same state condition, there was no significant excess of zero correct answers relative to the individual performance level, same states none correct vs. predicted from individual performance level: $t(19)=0.468$, $p = 0.645$, $BF_{10} = 0.256$, Cohen's d = 0.11. However, for the different state condition there was a significant excess of such trials, different states none correct vs. predicted from individual performance level: $t\ (19) = 7.14$, $p < 0.001$, $BF_{10} > 10^5$, Cohen's d = 1.59, consistent with swaps. This is because zero correct choices in the different state condition mean that a participant reported the states as being different (which is correct in terms of states alone) but ascribed them to the wrong exemplars (which is incorrect in terms of conjunctions or swap errors). In comparison with the failure to report any correct conjunctions for same state objects (which is more consistent with the absence of state memory), the "swap responses, according to our analysis, are observed in a considerable number of trials.

Overall, our results showed that our observers were quite good at recognizing exemplars and at discriminating whether the objects were presented in the same or different states. However, in a significant number of trials, they showed difficulties with reporting correct exemplar–state conjunctions. As a specific sign of a binding failure, this difficulty manifested as an increased fraction of trials within the different state condition where observers successfully reported the states as being different but chose wrong the exemplars for these two states. Note that the failure to report any conjunction correctly is rare in the same state trials where people do not actually need to remember exact conjunctions to perform the exemplar–state task and an ability to report the conjunction depends on memory only for the state itself. Therefore, we conclude that the difference between memory performance in the two conditions of the exemplar–state task is a result of binding failures, which is consistent with nonholistic, at least partially independent storage of exemplar and state features of real-world objects.

Experiment 2

The exemplar and the exemplar–state tasks were separated into two different blocks in Experiment 1. This practice could artificially encourage our observers to particularly focus on exemplar or state features during encoding, which could result in an inflated rate of swap responses and overestimate the independence of exemplar and state memories. Therefore, in Experiment 2, we randomly mixed trials from the exemplar and exemplar–state tasks to discourage our participants from selective encoding of the corresponding features.

Method

Participants

Twenty-five psychology students from the Higher School of Economics, 21 female; age, 18 to 33 years; $M = 19.7$, took part in the experiment for course credit. All participants reported having normal color vision, normal or corrected-to-normal visual acuity, and no neurological problems. The apparatus, stimuli, and procedure were the same as in Experiment 1. The main difference from Experiment 1 is that trials from the exemplar task were randomly mixed with trials from exemplar–state task.

Results

One participant showed less than 50% accuracy in all conditions and was excluded from the analysis.

Overall accuracy

We found evidence for a strong effect of the task and same/different-state manipulation on recognition accuracy, $F(2,21) = 39.23$, $p < 0.001$, $BF_{10} > 10^8$, $\eta^2 = 0.630$ (Figure 4). Participants were more accurate in the exemplar task, $M = 0.86$, than in the exemplar–state task, both when the states were same, $M = 0.80$; comparison: $t(23) = 4.247$, $p < 0.001$, Bonferroni corrected $\alpha = 0.017$, $BF_{10} = 100.2$, Cohen's d = 0.867, and when the states were different, $M = 0.73$; comparison: $t(23) = 9.024$, $p < 0.001$, Bonferroni corrected $\alpha = 0.017$, $BF_{10} > 10^6$, Cohen's d = 1.842. Most important, within the exemplar–state task, our participants were more accurate when tested exemplars were shown in the same state, $M = 0.80$, than when they were in different states, $M = 0.73$; comparison: $t(23) = 4.512$, $p < 0.001$, Bonferroni corrected $\alpha = 0.017$, $BF_{10} = 180.5$, Cohen's d = 0.921.

State memory

Like in Experiment 1, our participants were quite good at discriminating whether the two tested exemplars had been presented in the same or different states. For objects presented in the same states, they chose two response options of the same states much more frequently, $M = 0.69$, than when the objects had been presented in different states, $M = 0.21$; comparison: $t(23) = 13.93$, $p < 0.001$, $BF_{10} > 10^8$, Cohen's d = 2.843.

Accuracy of conjunction memory within paired choices

Our participants chose both correct exemplar–state conjunctions with approximately equal frequencies in the two conditions of the exemplar–state task, same states: $M = 0.64$, different states: $M = 0.62$; comparison: $t(23) = 0.747$, $p = 0.463$, $BF_{10} = 0.276$, Cohen's d =.152. The participants chose only one correct conjunction for items shown in same states more frequently than for items shown in different states, same states: $M = 0.31$, different states: $M = 0.21$, comparison: $t(23) = 2.649$, $p = 0.014$, $BF_{10} = 3.581$, Cohen's d = 0.541. Critically, they were substantially more likely to choose no correct conjunctions for items shown in different states than for items shown in same states, same states: $M = 0.05$, different states: $M = 0.17$, comparison: $t\ (23) = 7.846$, $p < 0.001$,

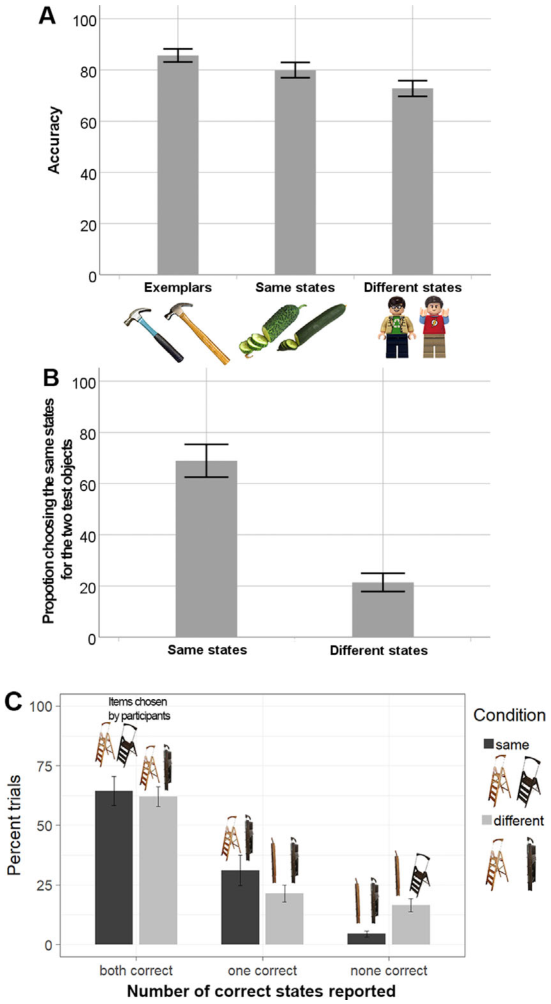


Figure 4. Results for Experiment 2 for overall accuracy (A), state memory accuracy (B), and choosing both, one, or no correct states for exemplars (C). Images on top of the bars in (C) are examples of participants' answers in each of the response outcomes. Note that the axis labels show two the different exemplars of the category (e.g., two different Lego people), either in the same state as each other, or different states than each other. Error bars depict 95% CIs. *LEGO Group. This is an independent site not authorized or sponsored by the LEGO Group*.

$BF_{10} > 10^6$, Cohen's d = 1.6. Again following the logic of what is expected from independent responses based on participants' individual percent correct, we find an excess of none correct trials in the different state condition only, comparison for same states: $t(23) = 0.008$, $p = 0.994$, $BF_{10} = 0.215$, Cohen's d = 0.002; comparison for different states: $t(23) = 10.676$, $p < 0.001$, $BF_{10} > 10^7$, Cohen's d = 2.18. This result replicates the finding from Experiment 1, showing that, in a substantial number of trials with a correct report about tested items being in different states, participants committed swap errors failing to correctly report which exemplars these states had gone with, but being—correctly—aware that there were two different states present.

The results of Experiment 2 closely replicated the pattern of results from Experiment 1. In both experiments, there were swap errors, where participants failed at reporting the state–exemplar conjunctions but had accurate memory for whether the objects' states had been the same or different. Because the mixed design of Experiment 2 discouraged selective encoding of exemplar features or state features, it seems that swap errors are not strongly dependent on an encoding strategy. These results are compatible with the idea of there being significant independence in the features that underly exemplar and state discrimination in visual working memory.

Robustness across categories in Experiments 1 and 2

To test whether the differences between the two conditions of the exemplar–state task are caused by our central manipulation of same versus different state, not by internal characteristics of individual images potentially affecting the memorability of objects, we analyzed the proportions of correct answers and the likelihood of choosing the same states across all tested images. In other words, we treated categories as a random effect rather than participants to ensure robustness not only across participants, but also across individual images. In particular, we can estimate how many observers chose the same states and correct exemplar–state conjunctions for every given category as a function of whether items in this category were presented in the same or different states (rather than how many observers did so, regardless of category). As the results of Experiments 1 and 2 showed highly similar patterns, we merged responses from all observers taking part in these two experiments. Overall, each category was seen by 22 participants in the same state and by 22 participants in different states.

We found that the overall accuracy across categories was lower when objects from these categories were shown in different states than when they were shown in same states, $t(119) = 6.180$, $p < 0.001$, Cohen's d = 0.564, $BF_{10} > 10^6$. The probability of choosing the same states for objects shown in different states was lower than the probability of choosing same states for objects shown in same states, $t(119) = 26.52$, $p < 0.001$, Cohen's d = 2.421, $BF_{10} > 10^{40}$. Therefore, our pattern of results shows the robustness of the pattern described in Experiments 1 and 2 not only across observers but also across stimuli.

The effect of not tested states in Experiments 1 and 2

In each trial of our exemplar–state task, we tested only one category (e.g., two of the four objects), and

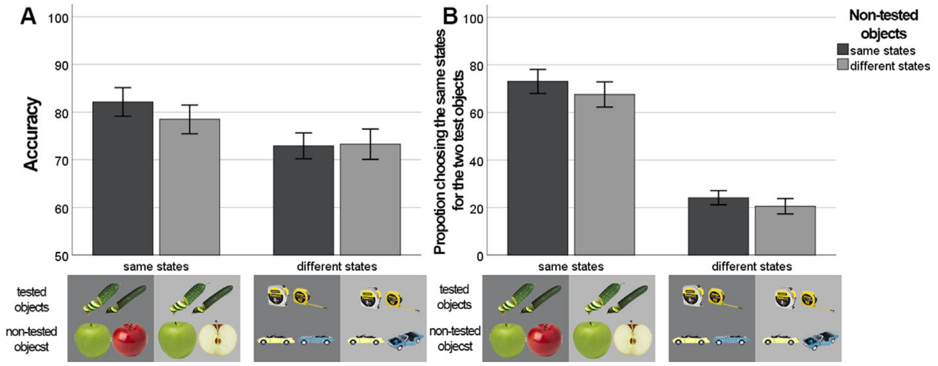


Figure 5. Results for the influence of states of tested and not tested objects on overall accuracy (A), state memory accuracy (B). Error bars depict 95% CIs.

the main manipulation was whether these objects were in the same state as each other or different states. However, items that have not been tested could also appear in either the same or different states. Because the same versus different status of the states of the not tested items were purposefully made orthogonal to the states of the tested items, it is possible to estimate the contribution of the former to the accuracy of reporting the latter. Given that performance estimates for such an analysis were built on 30 trials per combination of tested and not tested states (rather than 60 trials for our main analysis), we merged the data from Experiments 1 and 2 to compensate for some possible loss in the precision of individual estimates owing to the reduced number of trials.

Overall accuracy: A repeated-measure two-way ANOVA was run to estimate the effect of state manipulations in tested and not-tested items. We found a significant effect of the states of tested items on accuracy, $F(1,43) = 43.305$, $p < 0.001$, $BF_{10} > 10^7$, $\eta^2 = 0.502$ (Figure 5), reflecting the trend reported separately for each experiment: observers were less accurate when tested items were presented in different states. More importantly, we found evidence for the effect of tested items states × not tested items states on accuracy, $F(1,43) = 5.429$, $p = 0.025$, $BF_{10} = 1.02$, $\eta^2 = 0.112$. This effect arose because the accuracy of reporting tested conjunctions presented in same states was lower when the not-tested objects were presented in different states compared with not-tested objects presented in same states, $t(43) = 2.991$, $p = 0.005$, Cohen's d =.451, $BF_{10} = 7.731$ (Figure 5A). When the tested objects were presented in different states, the states of not tested objects had no effect, $t(43) = 0.206$, $p = 0.838$, Cohen's d =.031, $BF_{10} = 0.167$. In other words, there was an overall performance improvement particularly in the condition where both categories were more homogenous, each having only one state present. This finding is broadly consistent with the idea that state and exemplar properties are not automatically represented in a single holistic representation, but that binding the features used to discriminate state and exemplar features is difficult.

State memory: Participants reported the tested items were in the same states less when the tested objects were in different states compared with same states, $F(1,43) = 305.784$, $p < 0.001$, $BF_{10} > 10^{54}$, $\eta^2 = 0.877$, reflecting the result shown separately in Experiments 1 and 2. Importantly, the percentage reporting both tested items were in the same state was lower when not-tested objects were in different states compared with when the not tested objects were in the same states, $F(1,43) = 22.424$, $p < 0.001$, $BF_{10} = 0.282$, $\eta^2 = 0.343$ (Figure 5B). Therefore, we found that the states of not-tested but also memorized objects had a slight effect on reported states of tested objects in the same direction as the effect of tested objects states. Again, the finding that observers' answers were sensitive to state manipulations (although less than within the tested categories) supports the idea of somewhat independent storage of the features underlying state and exemplar discrimination in real-world objects.

Similarity between exemplar and state pairs

Looking at the performance on the exemplar task compared with the exemplar–state task, one could argue that there are differences in the exemplar and state manipulations themselves that account for this effect. For example, it intuitively seems that two different states of the same object might be more visually similar than two different exemplars in the same state, and that this could affect the two-alternative forced choice task performance (e.g., Brady & Störmer, 2020). In this case, the exemplar–state task would be harder than the exemplar task based on the images alone, rather than because of binding difficulties. Although previous work has found these two kinds of test tend to be similar in difficulty (Brady et al., 2008), to test this intuition more systematically, we quantitatively estimated

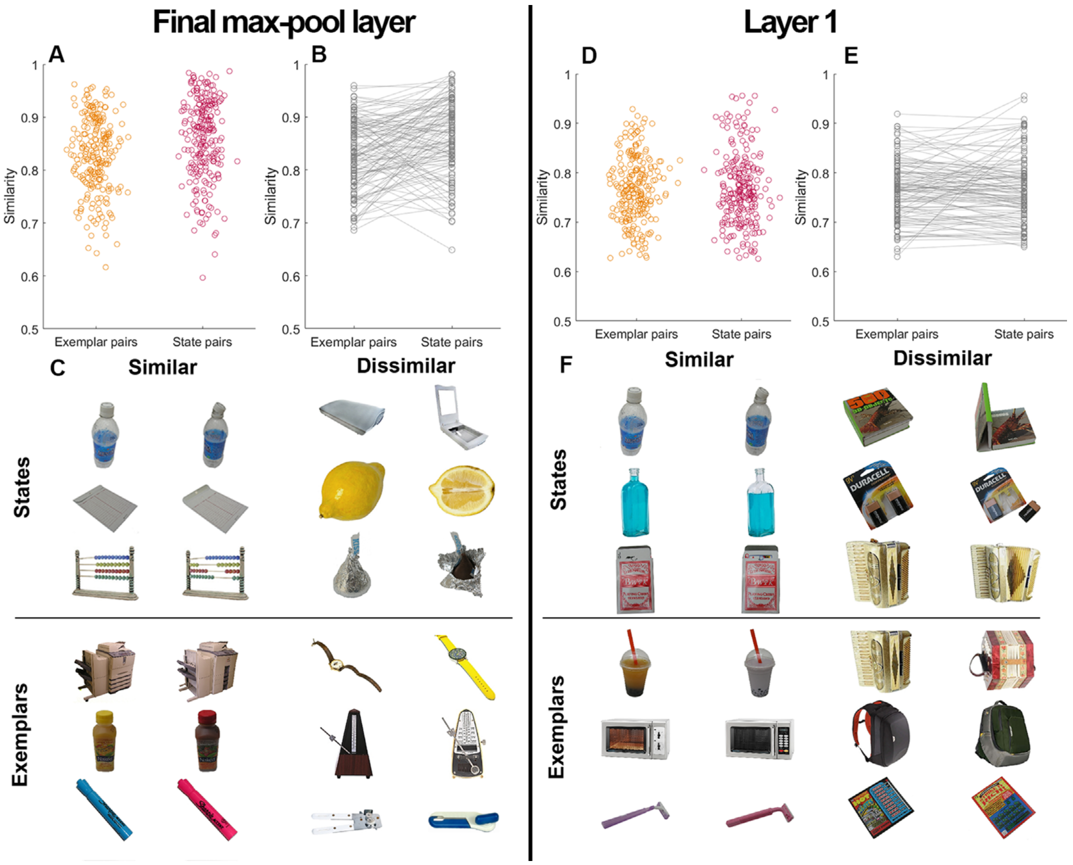


Figure 6. Similarity of exemplar and state pairs (0 = completely dissimilar objects; 1 = completely similar objects). Final max-pool layer: Estimation of similarity for paired exemplars and paired states (A) and average values for each category show significant heterogeneity by category (B). The examples of most similar and dissimilar pairs of exemplars and states according to the max pool layer. (C). Layer 1: Estimation of similarity for paired exemplars and paired states (D) and average values for each category show significant heterogeneity by category (E). The examples of most similar and dissimilar pairs of exemplars and states according to layer 1 (F).

similarity using the VGG16 pretrained convolutional neural network (Brady & Störmer, 2020; Simonyan & Zisserman, 2014). We were particularly focused on the top max-pool layer, which allowed us to retrieve high-level features that are more invariant to low-level transforms, because previous work demonstrates that this provides a useful proxy for object similarity (Brady & Störmer, 2020) with extremely similar object stimuli. However, we also estimated similarity values based only on layer 1, which are more related to very low-level image similarity and should have little invariance. We used our target stimuli (Brady et al., 2013) and estimated similarity between paired exemplars in the same states (exemplar pairs) and between the two states of the same exemplar (state pairs). Similarity for the 240 exemplar pairs and 240 state pairs (Figure 6A) and the average similarity for each category (Figure 6B) were estimated.

We found quite small (approximately 3%) but significant differences in the similarity based on final max-pool layer estimates between exemplar pairs, $M = 0.83$, and state pairs, $M = 0.86$; comparison: $t\ (119) = 4.272$, $p < 0.001$, $BF_{10} = 411.6$, Cohen's $d = 0.39$ (Figure 6B), with state pairs being slightly more similar on average. However, we found no significant differences in the similarity based on layer 1, exemplar pairs: $M = 0.77$; state pairs: $M = 0.77$; comparison: $t\ (119) = 0.314$, $p = 0.754$, $\mathrm{BF}_{10} = 0.1$, Cohen's $d = 0.029$) (Figure 6E). We also found significant, although weak, correlations between the similarity estimates based on the final max-pool layer and the similarity estimates based on layer 1, exemplar pairs: $r\ (118) = 0.39\ p < 0.001$; state pairs: $r\ (118) = 0.44\ p < 0.001$. To estimate the effect of similarity on task performance—and validate the similarity metric—we calculated a correlation between the average performance on the task and the similarity estimates for exemplars and state pairs based on both final max-pool layer and layer 1. We found that layer 1 did not significant predict performance in the task ($p > 0.10$), but that the max pool layer did, with a significant negative correlation between similarity values for state pairs based on the final max-pool layer and accuracy, r (118) = 0.31 $p < 0.001$. Therefore, state similarity—as measured by the more invariant max pool layer—indeed affected overall performance and the difficulty of the two-alternative forced choice task (e.g., Brady & Störmer, 2020).

Thus, to test whether these differences in similarity affected the comparisons we made between the exemplar task and the exemplar–state task, we selected the categories with most similar exemplars and most dissimilar states, according to the max pool layer (we chose categories where exemplar similarity was higher than state similarity, 40 categories overall with exemplar similarity of 0.87 and state similarity of 0.81) and redid the analysis from Experiments 1 and 2. In this sample, the states should actually be "easier" to discriminate than the exemplars. Thus, this provides a strong test of whether similarity alone accounts for the difference between our conditions.

Overall accuracy: Observers were less accurate when exemplars were in different states, $M = 0.75$, compared with same states, $M = 0.84$; comparison: $t(43) = 4.305$, $p < 0.001$, Cohen's $d = 0.649$, $BF_{10} = 249$.

State memory accuracy: Observers reported that exemplars were in the same states more frequently when exemplars were presented in the same states, $M = 0.75$, compared with different states, $M = 0.20$; comparison: $t(43) = 17.352$, $p < 0.001$, Cohen's $d = 2.616$, $\mathrm{BF}_{10} > 1017$.

Accuracy of conjunction memory within paired choices: There were no differences between conditions in choosing both correct, same states: $M = 0.72$, different states: $M = 0.66$, comparison: $t(43) = 1.985$, $p = 0.054$, Cohen's $d = 0.299$, $\mathrm{BF}_{10} = 0.975$, and one correct, same states: $M = 0.25$, different states: $M = 0.19$, comparison: $t(43) = 1.811$, $p = 0.077$ Cohen's $d = 0.273$, $\mathrm{BF}_{10} = 0.729$. The percentage of none correct answers was significantly higher in the different states condition compared with the same states condition, same states: $M = 0.04$, different states: $M = 0.15$, comparison: $t(43) = 6.046$, $p < 0.001$, Cohen's $d = .911$, $\mathrm{BF}_{10} = 49{,}614$.

This additional analysis demonstrates the same pattern as our main analysis, even for objects with highly dissimilar states. This finding suggests that binding errors still occur when discrimination at test is made easier. Most important, the patterns are preserved even when the state pairs are more dissimilar than the exemplar pairs. Therefore, inferior performance in the exemplar–state task cannot be explained by interobject similarity. Instead, we conclude that it has to do with the need to report exemplar–state conjunctions.

Discussion of Experiments 1 and 2

In Experiments 1 and 2 we found that our participants were less accurate at reporting the conjunctions of exemplars and states when these exemplars had been shown in different states compared with exemplars shown in the same state. This decrement in conjunction recognition was combined with good recognition memory for exemplars and good ability to discriminate—without the need for binding—whether states were the same or different. Importantly, the accuracy decrement that we observed for objects presented in the different states was mostly provided by trials where observers correctly reported the states being different, but swapped these states between exemplars. Moreover, we found some benefits for performance when the nontested objects were in the same states rather than different states, suggesting the task was easier when the nontested objects did not require binding states to exemplars. Therefore, we conclude that the features underlying exemplar and state discrimination are represented in some sense independently and that remembering their conjunctions causes additional difficulty compared with remembering these features per se. A fully unitized, all-or-none, totally bound representation account fails to account for our results, because it predicts that the working memory traces should be indifferent to whether two different exemplars are presented in same or different states: in any case, two separate records are created with equal likelihood to be stored or forgotten.

In fact, binding was also not necessary in the same states condition to recall both exemplar–state conjunctions. It was sufficient to remember a common state for a category instead, which is also consistent with the idea of independence, rather than each object being a holistic, unitized object encoded separately from each other. In Experiment 1, observers could discard exemplar information in the state–exemplar condition and remember only the common state, as the exemplar memory was not tested in the same block as state memory. However, in Experiment 2 with its mixed design observers did not know in advance whether their exemplar or exemplar–state memory would be tested. The nearly identical results of Experiments 1 and 2 suggests, therefore, that observers were encoding relevant details sufficient to discriminate both state and exemplar comparisons in both experiments. However, encoding the relevant features for both (I remember both these mugs and that they both were full) does not seem to entail that these features were fully unitized and bound (if I remember these mugs, I also remember that they were full, and vice versa). The results from the different state condition demonstrate that while participants appear to encode all the relevant features in both Experiments 1 and 2, these features are nevertheless somewhat independent (e.g., I sometimes remember the mugs and I remember one full and one empty, but I do not remember which one was full and which was empty). In the General Discussion, we consider possible exemplar and state representations leading to the observed pattern of results in more detail.

Experiment 3

In Experiment 3, we tested another interesting prediction following from the idea of at least partially independently stored properties of real-world objects. In particular, we looked at how people update previously studied information, one of the crucial functions of working memory (Ecker, Lewandowsky, Oberauer, & Chee, 2010; Nyberg & Eriksson, 2016). If a task requires participants to remember an item and then update it, taking into account a subsequent change to the item (update) can often cause confusions between the initial and updated representations, resembling binding errors (Gorgoraptis, Catalao, Bays, & Husain, 2011). An example of such errors in visual working memory can be a failure to completely update location changes during retention. Hollingworth and Rasmussen (2010) showed that binding to new locations after motion is nevertheless impacted by a remaining binding to the original locations. Other work on the spatial congruency bias also suggests that a location of an object is automatically attended and that the identity of an object is bound to this location even after updates (Bapat, Shafer-Skelton, Kupitz, & Golomb, 2017; Golomb, Kupitz, & Thiemann, 2014; Shafer-Skelton, Kupitz, & Golomb, 2017). With respect to our main research interest, this point raises an important question about real-world object representation: When an object changes location, will observers update or fail to update the entire set of object properties to a new location? Or is it possible that separate properties can separately fail to be updated? For example, imagine I am shown a full coffee mug A in a location X and an empty coffee mug B at a location Y. If my memory for the mug A is then tested at the location Y (originally belonging to the mug B), will I fail to update both the mug A and its "fullness" (as expected if updating is based on unitized memories) or I can update the mug A but remember the emptiness encoded from that location (which should cause a swap report as we defined it in Experiment 1)? We addressed this question in Experiment 3. We tested whether observers commit more swaps between exemplars and states of real-world objects when updating of locations is required. In particular, we tested whether observers more often choose the wrong state for a studied exemplar if at test it takes the location of a different exemplar shown in a different state. If features can be independently bound to a certain location, we predict that we will find no difference between original and updated locations for exemplars in the same states, but will find them for exemplars in the different states. This prediction follows the similar logic for Experiments 1 and 2. In other words, if two exemplars in the same state change their locations at test it should cause no confusions, regardless of whether the state is updated or not, because of the commonality between the two states at the two locations. By contrast, if exemplars in different states change their locations at test, this could cause more binding errors if updating is independent for state information (e.g., I remember a full mug at this location but do not remember which mug it was, so I choose a full mug here). By contrast, if updating works on unitized, fully bound representations then we expect that exemplar swap at test should produce an effect on exemplar–state reports both when the states are same and when they are different.

Method

Participants

Twenty-five psychology students from the Higher School of Economics, 22 female; age, 18 to 22 years, $M = 19.24$, took part in the experiment for extra course credits. All participants reported having normal color vision, normal or corrected to normal visual acuity, and no neurological problems.

The apparatus, stimuli, and spatial layout were similar to Experiments 1 and 2. Three image sets were used in this experiment: 120 categories from an image set (Brady et al., 2013) and 40 categories from an image set created by ourselves were used as tested categories in the exemplar–state task. Eighty-two categories from our image set and 78 categories from (Konkle et al., 2010) were used as nontested categories in the exemplar–state task.

Procedure

In this experiment, we used only the exemplar–state task. As in Experiments 1 and 2, observers had to remember two pairs of exemplars from two object categories and each pair of exemplars could be presented in the same or different states (Figure 7). The critical difference from Experiments 1 and 2 was at test. In one-half of trials, the presented exemplars from the tested category were tested at their original locations. If exemplar A was shown at location X and exemplar B was shown at location Y, then at test the old and new states of exemplar A were also shown at location and the old and new states of exemplar B were also shown at location Y (Figure 7), as in the previous experiments. In another one-half of the trials, the two tested exemplars swapped their locations. If exemplar A was studied at location X, its new and old states were tested at location Y, and if exemplar B was studied at location Y its states were tested at location X (Figure 7). Participants were warned that objects could swap their locations at test

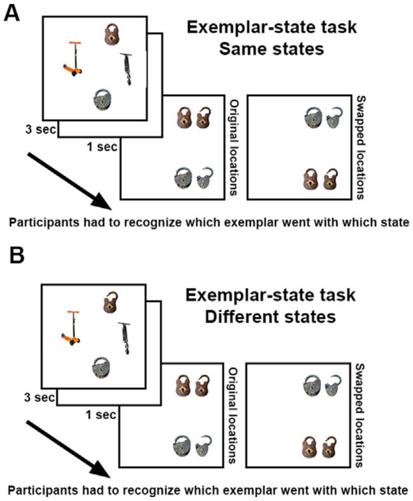


Figure 7. Example trials of Experiment 3 where two exemplars (locks) are (A) studied in the same states and tested in their original locations, or (B) studied in different states and tested in swapped locations.

and were instructed to recognize in which state each exemplar in the category was presented, regardless of test location.

In 80 trials, the tested exemplars were studied in different states and in 80 trials the tested exemplars were studied in the same states. In 80 trials, objects were tested in their original locations and in the other 80 trials in swapped locations. Categories were counterbalanced across conditions between participants using a Latin square.

Design and analysis

We had a 2 (objects in same and different states) × 2 (original or swapped test locations) within-subject design. We estimated the overall accuracy (total number of correctly chosen items) and state memory for the two location conditions (original vs. swapped) and for same versus different states. We also estimated the accuracy of conjunction memory within paired choices similarly to Experiments 1 and 2. For Bayesian $\underline{t}$ tests the same priors as in Experiments 1 and 2 were used.

Results and discussion

Overall accuracy

A two-way repeated-measures ANOVA was run to estimate the effects of studied exemplars being in the same versus different states and of test location. We found an overall effect of test location, $F\,(1{,}24) = 30.230$, $p < 0.001$, $BF_{10} = 12.678$, $\eta^2 = 0.557$. Observers were overall less accurate reporting exemplar–state conjunctions at swapped locations (Figure 8). We also found a strong effect of the studied states, $F\,(1{,}24) = 117.380$, $p < 0.001$, $BF_{10} > 10^{10}$, $\eta^2 = 0.830$. Observers were less accurate when objects were presented in different states compared with same states (Figure 7), which is in line with the corresponding findings from Experiments 1 and 2. However, we found no evidence for the interaction effect between studied states and test location, $F\,(1{,}24) = 1.896$, $p = 0.181$, $BF_{10} = 0.517$, $\eta^2 = 0.073$. That is, participants did not seem to have additional difficulty recognizing the different state items at updated locations compared with the same states at updated locations, broadly consistent with whole-object updating rather than separate updating for each property.

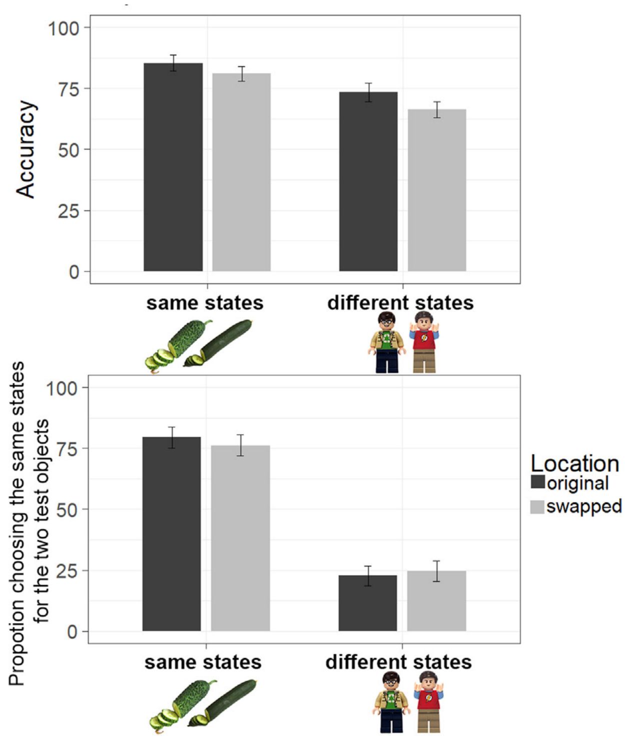


Figure 8. Results for Experiment 3 for overall accuracy and state memory for objects in same and different states. Error bars depict 95% CIs. *LEGO Group. This is an independent site not authorized or sponsored by the LEGO Group.*

State memory

There was a strong effect of the state of studied objects, $F\,(1{,}24) = 288.742$, $p < 0.001$, $BF_{10} > 10^{10}$, $\eta^2 = 0.923$. Participants more often chose two same states when objects were presented in the same states and this effect was flipped for objects presented in two different states. We found no evidence of the effect of test location, $F\,(1{,}24) = 0.563$, $p = 0.46$, $BF_{10} = 0.208$, $\eta^2 = 0.023$, but did find a weak test location × studied states interaction, $F\,(1{,}24) = 6.197$, $p = 0.02$, $BF_{10} = 0.128$, $\eta^2 = 0.205$, consistent with less accurate memory

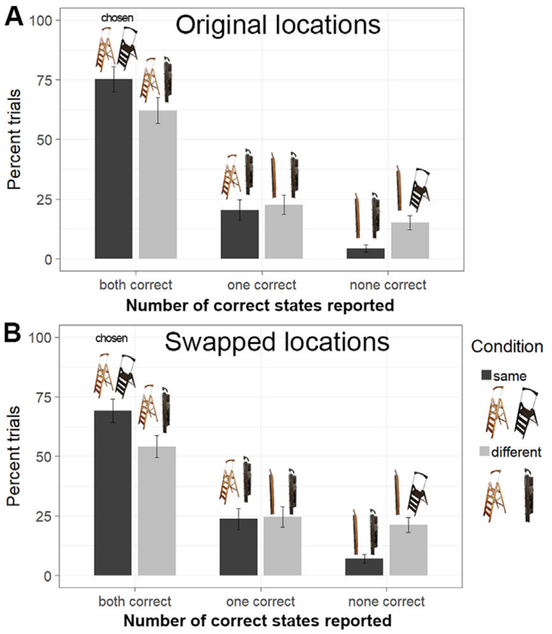


Figure 9. Results for Experiment 3 for choosing both, one or none correct states for exemplars tested in (A) original locations and (B) swapped locations. Error bars depict 95% CIs.

for states when the items flipped location (e.g., both conditions being closer to 0.50).

Accuracy of conjunction memory within paired choices

We analyzed the frequencies of paired report outcomes (both, one, or none correct) separately for original (Figure 9A) and swapped locations (Figure 9B). Overall, the patterns were strongly similar across these two conditions. Participants more frequently chose both of the correct exemplar–state conjunctions when the objects has been shown in the same state compared with different states, original locations: $M = 0.75$ for same states, $M = 0.62$ for different states; comparison: $t(24) = 6.381$, $p < 0.001$, $BF_{10} = 13{,}605$, Cohen's d = 1.276; swapped locations: $M = 0.69$ for same states, $M = 0.54$ for different states; comparison: $t(24) = 5.918$, $p < 0.001$, $BF_{10} = 4{,}811$, Cohen's d = 1.184. This result was well-mirrored by the pattern of choosing no conjunctions correct: Participants were more likely to choose no correct conjunctions for items shown in different states than for items shown in same states, original locations: $M = 0.04$ for same states, $M = 0.15$ for different states; comparison: $t\,(24) = 7.712$, $p < 0.001$, $BF_{10} > 10^6$, Cohen's d = 1.542; swapped locations: $M = 0.07$ for same states, $M = 0.21$ for different states; comparison: $t\,(24) = 8.067$, $p < 0.001$, $BF_{10} > 10^6$, Cohen's d = 1.613. There were no differences between the proportions of choosing only one correct conjunction between the two state combinations, original locations: $M = 0.20$ for same states, $M = 0.23$ for different states; comparison: $t(24) = 0.912$, $p = 0.371$, $BF_{10} = 0.307$, Cohen's d =.182; swapped locations: $M = 0.24$ for same states, $M = 0.25$ for different states; comparison: $t(24) = 0.358$, $p = 0.723$, $BF_{10} = 0.224$, Cohen's d = 0.072. This result replicates the corresponding pattern from Experiments 1 and 2 showing that observers tend to swap states between exemplars when these exemplars are presented in different states, with no additional reliable effect of switching location.

Thus, in Experiment 3 we strongly replicated the basic pattern from Experiments 1 and 2 that observers show better memory for states (whether these states are same or different) than for exemplar–state conjunctions (which state goes with which exemplar), especially when conjunction discrimination is critical for doing the task as in the case of exemplars presented in different states. Therefore, Experiment 3 confirms the robustness of this basic pattern indicating the relative independence of internal object features (features underlying exemplar and state discrimination). In addition to this point, we found some cost of location swaps that can be interpreted in terms of updating failures. This result replicates the previously reported tendency of an object to be bound to its original location after motion or during the update (Bapat et al., 2017; Golomb et al., 2014; Hollingworth & Rasmussen, 2010; Shafer-Skelton et al., 2017). The cost of location swaps was not very strong (difference of approximately 6% between original and swapped locations), which is also consistent with the previous demonstrations (e.g., approximately 4% between original and updated locations in Hollingworth & Rasmussen, 2010).

Having succeeded with inducing updating failures, we can turn to the main question of Experiment 3. According to our prediction, independent updating of different features could be inferred only if we found no impairment for objects presented in the same states and found this impairment in the different states condition. Because we found no difference in the amount of updating failures for exemplars shown in same versus different states, we conclude that state features are not bound to locations independently from exemplar features. Instead, it seems that location updating has something to do with the whole object representation. This finding is consistent with the previous demonstrations and the object file theory (Bapat et al., 2017; Golomb et al., 2014; Hollingworth & Rasmussen, 2010; Shafer-Skelton et al., 2017; Kahneman, Treisman, & Gibbs, 1992), and especially consistent with recent work from Dowd and Golomb (2019) showing that updating items does not break binding in the case of simple objects.

Thus, on the one hand, we see that the features underlying state discriminations behave relatively independently from those underlying exemplar discriminations, as revealed by the same versus different states manipulation. On the other hand, location updating appears to act on the entire set of object features.

General discussion

Experiments 1 and 2 showed that real-world objects are not necessarily stored as completely unitized, fully bound units in visual working memory, as swap errors occur between features of different objects of the same category. These results are in line with theories based on simple features which argue that visual working memory is not based on fully bound representations (e.g., with the weak object hypothesis, Olson & Jiang, 2002). In particular, they are consistent with theories that suggest that memory is object based only in that instantiating a new object representation is costly, but that within an object, memory for features is somewhat independent(see Brady, Konkle, & Alvarez, 2011, for review; Fougnie, Asplund, & Marois, 2010; Markov, Tiurina, & Utochkin, 2019; Shin & Ma, 2017; Wang, Cao, Theeuwes, Olivers, & Wang, 2017; Wheeler & Treisman, 2002). Unlike previous work, we showed that feature independence can be found not only for simple stimuli with basic features naturally assumed to be separable (such as color, orientation, or shape), but also for real-world objects whose properties are more complex and meaningfully connected. Therefore, our study suggests that at least partially independent feature storage is a common property of visual working memory representations. It is possible that, under certain conditions, objects could be encoded more holistically (e.g., when visual working memory load is lower), but our results demonstrate that objects are not, by necessity, encoded in a holistic, all-or-none manner.

Because we used considerably more ecologically valid stimuli in these experiments than in previous work, it is important to compare our results with the results reported in the previous literature with simple stimuli. In a study using simple features such as color and orientation (Bays et al., 2009), the percentage of swap errors for set size 4, with similar encoding duration and similar items localization on the screen to the current study, was around 0.11, which is similar to the percentage of swap errors in our experiments (around 0.10). So, the complexity of the features and meaningfulness of remembered stimuli did not make a very large difference in the frequency of binding errors, suggesting that the features of real-world objects are stored with a similar degree of independence to colors and orientations. Note that this study did not aim to discriminate between the visual and semantic features which underlie the state and exemplar discriminations, and therefore which are somewhat independently represented. This complicated question is open for future investigation.

However, these results do provide some evidence for the idea that objects are stored in a more abstract way than just basic visual features. That is, if participants solely stored shape, color, spatial frequency, orientation, and other basic visual features, then—because state and exemplar comparisons inevitably both rely on a combination of these features—memory for the state and exemplar properties would end up looking "bound." This is because, even if color or orientation or another of the basic visual features was selectively lost, this would impact both state and exemplar comparisons, so the two would be expected to change together for the most part. Because swap errors are about as common for these properties as with simple features like color and orientation, this provides some evidence that people are representing higher-level, perhaps more semantic features of the objects (which can be misbound or independently forgotten, as in color/orientation).

Although our results show that the information about exemplar and state features of real-world objects is not stored in working memory in a totally holistic, all-or-none manner, it does not mean that these features are stored on completely independent "shelves" somewhere in memory. One promising way of thinking about this is that connection between exemplars and states can be a hierarchically linked structure (Balaban et al., 2020; Brady et al., 2011; Fougnie et al., 2013), still leaving a possibility that these links can be incorrect or lost leading to the observed misreports of state–exemplar conjunctions. Compared with long-term memory (Utochkin & Brady, 2020), these misreports are rarer in working memory, suggesting that the feature representations can be linked more strongly in the latter case (Pertzov et al., 2012).

It is important to note that, in our experiments, we used a relatively extended presentation time (3 seconds), which might raise a question whether observers could go beyond the "pure" encoding capacity of visual working memory putting some part of information into long-term memory. Indeed, there is no gold standard in the literature regarding the critical presentation time for pure working memory, with durations ranging from hundreds milliseconds (Alvarez & Cavanagh, 2004; Bays et al., 2009; Luck & Vogel, 1997; Olson & Jiang, 2002) to several seconds (Bays et al., 2009; Bays et al., 2011; Brady et al., 2013; Brady et al., 2016; Fougnie & Alvarez, 2011; Pertzov et al., 2012), and with no strong evidence that encoding time fundamentally changes the relative contribution of different memory systems. However, we do not completely rule out a long-term component in our experiments, especially given the nature of our real-world stimuli, because they are inherently linked to existing knowledge and could allow for the use of newly encoded long-term representations, although because both the studied items and lures are real-world objects, to discriminate them, participants would need display-specific, newly formed long-term memories, no more than a few seconds old; without encoding the particular items in a particular display, they would be a chance in the test, even using long-term memory. As we discussed elsewhere in this article, both visual and semantic features can underlie the exemplar–state discrimination and some of this information could be supported and enhanced by existing knowledge (Curby, Glazek, & Gauthier, 2009; Brady et al., 2016; Schurgin et al., 2018). However, even with the potential support from long-term memory, we still observe around 10% binding errors, which is comparable with the data from other studies using meaningless stimuli and shorter encoding times (e.g., Bays et al., 2009; Emrich & Ferber, 2012).

It is important to note that accuracy in the exemplar task was always higher than the accuracy in the exemplar–state task, even when two exemplars were studied in same states (Experiments 1 and 2), when memory for conjunctions is in fact not required and memory for states is sufficient. One explanation of such results is that, when both exemplars are presented in the same states, it is hard for observers to know which state features they should remember (e.g., if they saw two empty glasses of water it may not be easy to anticipate that the tested state will be fullness [empty glass and full glass] or brokenness [empty glass and broken glass]). In a condition with different states, observers are more likely to realize which states will be tested and which features they should encode. Thus, it is possible that there are even more binding errors, but we could not detect them because the performance in the same states condition is impaired by a failure to anticipate and encode the proper states. However, it also could be thought that the exemplar task was easier than the exemplar–state task, because of the discrimination required at test—that is, that states are simply harder to discriminate than exemplars. It may seem that recognizing an old exemplar against a new exemplar is, by necessity, always easier than recognizing an old state of an exemplar against a new state of the same exemplar. But in fact, two states of the same exemplar often could be more perceptually different (e.g., a sliced and nonsliced instances of the same apple differ by color, texture, shape, presence or absence of pips, etc.) than two different exemplars in the same state (e.g., two different exemplars of apples that could differ only by color), and most previous work has found that observers are about equally good at recognizing old exemplars versus new exemplars and old states of same exemplars versus their new states (Brady et al., 2008).

To test this idea more directly, we performed an additional analysis of similarity using a deep convolutional neural network that has been previously shown to be a useful model of similarity of real-world objects, successfully predicting which foils are more or less difficult to discriminate (Brady & Störmer, 2020). This additional analysis of similarity confirmed that, in fact, there was quite a lot of heterogeneity in the difficulty of both state and exemplar comparisons. Furthermore, this analysis showed that, even when paired states were more dissimilar than paired exemplars, this did not affect the general pattern we observed, suggesting the patterns arise from binding difficulties rather than differences in the stimuli. Therefore, it is not very likely that overall object similarity can explain the superiority of exemplar memory over state memory. From our additional analysis of exemplar–state report accuracy as a function of nontested item states, we can conclude that a fraction of misreported same-state conjunctions (tested objects) can be accounted for by trials where nontested exemplars were shown in different states, perhaps taking up more encoding time/working memory resources.

What kind of representations could underlie the swap errors observed in our experiments? We suggest a few possible scenarios compatible with our data, all consistent with some form of nonholistic, at least partially independent storage: 1) participants might have strong feature memories (when both exemplar and state information are present) and a failure to bind them. The binding failures, in turn, can take the form of false bindings (remembering full mug *A* and/or empty mug *B*, whereas mug *A* was in fact empty and mug *B* was in fact full) or not remembered bindings (remembering seeing both mugs, as well as fullness and emptiness but not remembering which mug was presented full and which one empty). 2) Another possibility is that some of the features could be (independently) forgotten but observers strategically guess these features—that is, participants might not know, at test, whether one of the mugs was full or empty, but remember seeing that mug. For example, the superior performance in the same state condition of the exemplar–state task could be explained by better memory for repeated states (perhaps owing to chunking these states) and worse memory for different states. The trials where observers correctly reported different states, but reported none of the conjunctions correctly could be the result of the strategic guesses (if I do not remember states, I can randomly choose two different ones). This scenario is possible both with and without good exemplar memory. Although our data do not allow us to distinguish between these scenarios and future research is necessary for it, our principal conclusion is that any of these scenarios require state and exemplar memories to be stored or lost at least partially independently. Any of the scenarios is inconsistent with strongly holistic, bound representations (the strong object hypothesis), as such representations should be indifferent to whether objects are shown in same or different states.

Our results add to the picture of how objects are stored and forgotten across memory systems. Previous work has suggested that long-term memory is likely to store quite a lot of independently represented features of real-world objects (Brady et. al., 2013; Utochkin & Brady, 2020). In particular, using the exemplar and exemplar–state tasks for long-term memory, Utochkin and Brady (2020) showed that observers were at chance when reporting state–exemplar conjunctions of objects presented in different states, although they chose two different states for such objects well above chance and their exemplar memory also was good. Importantly, the difference in recognition accuracies between same-state and different-state pairs of studied exemplars was dramatic (0.74 vs. 0.53, respectively). Our current results do not show such a strongly disruptive effect of different-state objects. Neither exemplar–state report accuracy was near chance in that condition, nor was its difference from same-state trials that large. We, therefore, can conclude that visual working memory can provide more boundedness of representations than visual long-term memory—perhaps in part because of representations of color, shape, orientation, and so on, that by necessity provide useful information about both state and exemplar discriminations. At the same time, the fact that some binding failures can occur after a 1-second retention period in working memory is suggestive that part of the failed conjunction representations in long-term memory may arise when these features are consolidated from working memory. It is also consistent with previous demonstrations that object–location bindings are most susceptible to forgetting in working memory (Pertzov et al., 2012). The links between binding errors in visual working memory and visual long-term memory is an interesting subject for future research.

In Experiment 3, we tested an additional hypothesis following from the idea of independence—namely, that manipulating object locations at encoding and retrieval would produce specific updating failures when an observer reports a feature remembered at that location when another feature at this location is changed. Although this experiment allowed a strong replication of the independence pattern in terms of the same–different state manipulations, and location swaps caused additional failures, we found no interaction between the location manipulation and state manipulation in Experiment 3. From these results we concluded that location updating appears more whole object based (consistent with what Dowd & Golomb, 2019 found for simple features).

Several explanations can be considered to account for the finding of independence of features combined with the demonstration of the whole object location updating. First, the object–location binding problem is a separate problem from feature binding (Treisman, 1996), in that objects and locations are processed to some extent via two separate pathways, ventral and dorsal (Haxby, et al., 1991; Mishkin & Ungerleider, 1982; Wilson, O'Scalaidhe & Goldman-Rakic, 1993). Thus, object–location binding is also a separate process from storing objects and locations (Postma & De Haan, 1996; Postma, Kessels, & van Asselen, 2008), so it is possible that location swaps did not influence binding of the features underlying state and exemplar discrimination. According to our results, it is possible that object–location binding could happen after feature binding, which is consistent with object file theories (Hollingworth & Rasmussen, 2010; Kahneman, Treisman, & Gibbs, 1992) and with the general invariance of location tracking (e.g., in multiple object tracking) to feature information (Flombaum, Scholl, & Santos, 2009; Pylyshyn, 2000) Another explanation is discussed by Utochkin and Brady (2020). The interaction between exemplar and state information could be more complicated than the interaction between parallel representations of low-level features (such as color and orientation). Instead, the somewhat independent storage of features supporting exemplar discrimination and state discrimination could result not from a fully parallel organization but from a hierarchical one (Brady, Konkle, & Alrvarez, 2011; Utochkin & Brady, 2020), with exemplar information on a higher level, while access to state information is possible only when exemplar information is not lost (if I do not remember this mug, I also do not remember whether it was full or empty), but not vice versa (I do not remember whether the mug was full or empty, but I remember it was this high yellow mug). Note that this hierarchy does not contradict the idea that these features are not holistic or all-or-none but partially independent: State information still can be forgotten independently from its exemplar or "migrate" to another exemplar in such a model, and objects are not forgotten in an all-or-none way in such a model. This form of hierarchy can potentially explain the results Experiment 3: location swaps could impair exemplar recognition at a new location (because it was always the exemplar whose location has been manipulated at test in Experiment 3), which entails the loss of access to state information. This structure is similar to the one proposed by familiarity/recollection dichotomies in working memory, where item memory is necessary to access context information or other episodic details (e.g., Mickes, Wais, Wixted, 2009). The hierarchical organization of real-world object storage in working memory is an intriguing possibility that needs further investigation.

Overall, in this work we not only show that binding errors occur for real-world objects, but also investigated how working memory updates information about such objects, thus providing new information about how real-world objects are both maintained and updated, the two most critical features of working memory

(Baddeley, 1986; Baddeley & Hitch, 1974; Ecker et al., 2010; Nyberg & Eriksson, 2016). We showed that the features underlying two different discrimination tests about real-world objects are somewhat independently represented inside the object representation, as opposed to entirely holistic and all-or-none, but that location updates appear to work at the level of whole object representations rather than impairing links between internal features.

Keywords: visual working memory, feature binding, real-world objects

Acknowledgments

The study was implemented within the framework of the Basic Research Program at the National Research University Higher School of Economics in 2019 and was partially supported by NSF BCS-1829434 to TFB. The authors thank Natalia Tiurina for help in making stimuli.

Experimental data can be publicly accessed via Open Science Framework at https://osf.io/p9gzh/ and the object stimuli are available from T.F.B.'s website (https://bradylab.ucsd.edu/stimuli.html), Y.A.M.'s website (https://www.ymarkov.com/demosstimuli) and via Open Science Framework at https://osf.io/p9gzh/).

Commercial relationships: none.
Corresponding author: Yuri A. Markov.
Email: yuamarkov@gmail.com.
Address: Psychology Department, Higher School of Economics, 101000, Armyansky per., 4, room 419, Moscow, Russian Federation.

References

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science: A Journal of the American Psychological Society/APS, 15*(2), 106–111, https://doi.org/10.1167/2.7.273.

Asp, I. E., Störmer, V. S., & Brady, T. (2019). Greater visual working memory capacity for visually-matched stimuli when they are recognized as meaningful. *Journal of Cognitive Neuroscience, 11*, 1–17.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (ed). *The psychology of learning and motivation: advances in research and theory,* pp. 47–89. New York: Academic Press.

Baddeley, A. D. (1986). *Working memory.* Oxford, UK: Clarendon Press.

Baddeley, A. D. (1966). The influence of acoustic and semantic similarity on long-term memory for word sequences. *Quarterly Journal of Experimental Psychology, 18*(4), 302–309.

Balaban, H., Assaf, D., Arad Meir, M., & Luria, R. (2020). Different features of real-world objects are represented in a dependent manner in long-term memory. *Journal of Experimental Psychology: General, 149*(7), 1275–1293, https://doi.org/10.1037/xge0000716.

Bapat, A. N., Shafer-Skelton, A., Kupitz, C. N., & Golomb, J. D. (2017). Binding object features to locations: Does the "spatial congruency bias" update with object movement? *Attention, Perception, & Psychophysics, 79*(6), 1682–1694, https://doi.org/10.3758/s13414-017-1350-5.

Bays, P. M. (2016). Evaluating and excluding swap errors in analogue tests of working memory. *Scientific Reports, 6*(2015), 1–14, https://doi.org/10.1038/srep19203.

Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision, 9*(10), 7–7, https://doi.org/10.1167/9.10.7.

Bays, P. M., Wu, E. Y., & Husain, M. (2011). Storage and binding of object features in visual working memory. *Neuropsychologia, 49*(6), 1622–1631, https://doi.org/10.1016/j.neuropsychologia.2010.12.023.

Brady, T. F., Konkle, T., & Alvarez, G. A. (2011). A review of visual memory capacity: Beyond individual items and toward structured representations. *Journal of Vision, 11*(5), 4–4, https://doi.org/10.1167/11.5.4.

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 of the United States of America, 105*(38), 14325–14329, https://doi.org/10.1073/pnas.0803390105.

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. General, 142*(3), 791–808, https://doi.org/10.1037/a0029649.

Brady, T. F., & Störmer, V. S. (2020). Greater capacity for objects than colors in visual working memory: Comparing memory across stimulus spaces requires maximally dissimilar foils. PsyArXiv, https://doi.org/10.31234/osf.io/25t76.

Brady, T. F., & Störmer, V. S. (in press). The role of meaning in visual working memory: Real-world objects, but not simple features, benefit from deeper processing. *Journal of Experimental Psychology: Learning, Memory and Cognition*, doi:10.31234/osf.io/kzvdg.

Brady, T. F., Störmer, V. S., & Alvarez, G. A. (2016). Working memory is not fixed-capacity: More active storage capacity for real-world objects than for simple stimuli. *Proceedings of the National Academy of Sciences of the United States of America, 113*(27), 7459–7464, https://doi.org/10.1073/pnas.1520027113.

Conway, B. R. (2009). Color vision, cones, and color-coding in the cortex. *Neuroscientist, 15*(3), 274–290, https://doi.org/10.1177/1073858408331369.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*(1), 87–185, https://doi.org/10.1017/S0140525=01003922.

Cowan, N., Chen, Z., & Rouder, J. N. (2004). Constant capacity in an immediate serial-recall task: A logical sequel to Miller (1956). *Psychological Science, 15*(9), 634–640, https://doi.org/10.1111/j.0956-7976.2004.00732.x.

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*(3), 731–747, https://doi.org/10.1037/a0029687.

Curby, K. M., & Gauthier, I. (2007). A visual short-term memory advantage for faces. *Psychonomic Bulletin & Review, 14*(4), 620–628, https://doi.org/10.3758/BF03196811.

Curby, K. M., Glazek, K., & Gauthier, I. (2009). A visual short-term memory advantage for objects of expertise. *Journal of Experimental Psychology: Human Perception and Performance, 35*(1), 94–107, https://doi.org/10.1037/0096-1523.35.1.94.

Dent, K., & Smyth, M. M. (2005). Verbal coding and the storage of form-position associations in visual–spatial short-term memory. *Acta Psychologica, 120*(2), 113–140, https://doi.org/10.1016/J.ACTPSY.2005.03.004.

DiCarlo, J. J., Zoccolan, D., & Rust, N. C. (2012). How does the brain solve visual object recognition? *Neuron, 73*(3), 415–434, https://doi.org/10.1016/j.neuron.2012.01.010.

Dowd, E. W., & Golomb, J. D. (2019). Object-feature binding survives dynamic shifts of spatial attention. *Psychological Science, 30*(3), 343–361, https://doi.org/10.1177/0956797618818481.

Ecker, U. K. H., Lewandowsky, S., Oberauer, K., & Chee, A. E. H. (2010). The components of working memory updating: An experimental decomposition and individual differences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(1), 170–189, https://doi.org/10.1037/a0017891.

Emrich, S. M., & Ferber, S. (2012). Competition increases binding errors in visual working memory. *Journal of Vision, 12*(4), 12–12, https://doi.org/10.1167/12.4.12.

Erez, J., Cusack, R., Kendall, W., & Barense, M. D. (2016). Conjunctive coding of complex object features. *Cerebral Cortex, 26*(5), 2271–2282, https://doi.org/10.1093/cercor/bhv081.

Faul, F., Erdfelder, E., Bucher, A., & Lang, A.-G. (2009). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavioral Research Methods, 39*, 175–191, Retrieved from http://www.psycho.uni-duesseldorf.de/abteilungen/aap/gpower3/download-and-register/Dokumente/GPower3-BRM-Paper.pdf.

Fiebach, C. J., Rissman, J., & D'Esposito, M. (2006). Modulation of inferotemporal cortex activation during verbal working memory maintenance. *Neuron, 51*(2), 251–261, https://doi.org/10.1016/j.neuron.2006.06.007.

Flombaum, J. I., Scholl, B. J., & Santos, L. R. (2009). Spatiotemporal priority as a fundamental principle of object persistence. In: B. M. Hood, & L. R. Santos, editors. *The origins of object knowledge*. Oxford, UK: Oxford University Press, https://doi.org/10.1093/acprof:oso/9780199216895.003.0006.

Fougnie, D., & Alvarez, G. A. (2011). Object features fail independently in visual working memory: Evidence for a probabilistic feature-store model. *Journal of Vision, 11*(12), 3, https://doi.org/10.1167/11.12.3.

Fougnie, D., Asplund, C. L., & Marois, R. (2010). What are the units of storage in visual working memory? *Journal of Vision, 10*(12), 27, https://doi.org/10.1167/10.12.27.

Fougnie, D., Cormiea, S. M., & Alvarez, G. A. (2013). Object-based benefits without object-based representations. *Journal of Experimental Psychology: General, 142*(3), 621–626, https://doi.org/10.1037/a0030300.

Fougnie, D., & Marois, R. (2009). Attentive tracking disrupts feature binding in visual working memory. *Visual Cognition, 17*(1–2), 48–66, https://doi.org/10.1080/13506280802281337.

Golomb, J. D., Kupitz, C. N., & Thiemann, C. T. (2014). The influence of object location on identity: A "spatial congruency bias". *Journal of Experimental Psychology: General, 143*(6), 2262–2278, https://doi.org/10.1037/xge0000017.

Gorgoraptis, N., Catalao, R. F. G., Bays, P. M., & Husain, M. (2011). Dynamic Updating of

Working Memory Resources for Visual Objects. *Journal of Neuroscience, 31*(23), 8502–8511, https://doi.org/10.1523/JNEUROSCI.0208-11.2011.

Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature, 458*(7238), 632–635, https://doi.org/10.1038/nature07832.

Haxby, J. V., Grady, C. L., Horwitz, B., Ungerleider, L. G., Mishkin, M., & Carson, R. E., …Rapoport, S. I. (1991). Dissociation of object and spatial visual processing pathways in human extrastriate cortex. *Proceedings of the National Academy of Sciences of the United States of America, 88*(5), 1621–1625, https://doi.org/10.1073/pnas.88.5.1621.

Hollingworth, A., & Rasmussen, I. P. (2010). Binding objects to locations: The relationship between object files and visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 36*(3), 543–564, https://doi.org/10.1037/a0017836.

Janini, D., & Konkle, T. (2019). A Pokémon-sized window into the human brain. *Nature Human Behaviour, 3*, 552–553, https://doi.org/10.1038/s41562-019-0594-6.

JASP Team. (2018). *JASP (Version 0.8.5.1) [Computer software]*. Amsterdam, the Netherlands: JASP.

Kahneman, D., Treisman, A, & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24*(2), 175–219, https://doi.org/10.1016/0010-0285(92)90007-O.

Kaiser, D., Stein, T., & Peelen, M. V. (2015). Real-world spatial regularities affect visual working memory for objects. *Psychonomic Bulletin & Review, 22*(6), 1784–1790, https://doi.org/10.3758/s13423-015-0833-4.

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. *Journal of Experimental Psychology: General, 139*(3), 558–578, https://doi.org/10.1037/a0019165.

Konkle, T., & Oliva, A. (2012). A Real-world size organization of object responses in occipitotemporal cortex. *Neuron, 74*(6), 1114–1124, https://doi.org/10.1016/j.neuron.2012.04.036.

Lee, D., & Chun, M. M. (2001). What are the units of visual short-term memory, objects or spatial locations? *Perception & Psychophysics, 63*(2), 253–257, https://doi.org/10.3758/BF03194466.

Lew, T. F., & Vul, E. (2015). Ensemble clustering in visual working memory biases location memories and reduces the Weber noise of relative positions. *Journal of Vision, 15*(4), 10, https://doi.org/10.1167/15.4.10.

Li, L., Miller, E. K., & Desimone, R. (1993). The representation of stimulus familiarity in anterior inferior temporal cortex. *Journal of Neurophysiology, 69*(6), 1918–1929, https://doi.org/10.1152/jn.1993.69.6.1918.

Li, X., Xiong, Z., Theeuwes, J., & Wang, B. (2020). Visual memory benefits from prolonged encoding time regardless of stimulus type. *Journal of Experimental psychology. Learning, Memory, and Cognition, 46*(10), 1998–2005.

Long, B., Konkle, T., Cohen, M. A., & Alvarez, G. A. (2016). Mid-level perceptual features distinguish objects of different real-world sizes. *Journal of Experimental Psychology: General, 145*(1), 95–109, https://doi.org/10.1037/xge0000130.

Long, B., Moher, M., Carey, S., & Konkle, T. (2019). Real-world size is automatically encoded in preschoolers' object representations. *Journal of Experimental Psychology: Human Perception and Performance, 45*(7), 863–876, https://doi.org/10.1037/xhp0000619.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*(6657), 279–281, https://doi.org/10.1038/36846.

Luria, R., & Vogel, E. K. (2011). Shape and color conjunction stimuli are represented as bound objects in visual working memory. *Neuropsychologia, 49*(6), 1632–1639, https://doi.org/10.1016/j.neuropsychologia.2010.11.031.

Markov, Y. A., Tiurina, N. A., & Utochkin, I. S. (2019). Different features are stored independently in visual working memory but mediated by object-based representations. *Acta Psychologica, 197*, 52–63, https://doi.org/10.1016/j.actpsy.2019.05.003.

Mickes, L., Wais, P. E., & Wixted, J. T. (2009). Recollection is a continuous process: Implications for dual-process theories of recognition memory: Research Article. *Psychological Science, 20*(4), 509–515, https://doi.org/10.1111/j.1467-9280.2009.02324.x.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 101*(2), 343–352, https://doi.org/10.1037/h0043158.

Miller, E. K., Erickson, C. A., & Desimone, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *Journal of Neuroscience, 16*(16), 5154–5167, https://doi.org/10.1523/JNEUROSCI.16-16-05154.1996.

Mishkin, M., & Ungerleider, L. G. (1982). Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. *Behavioural Brain Research, 6*(1), 57–77, https://doi.org/10.1016/0166-4328(82)90081-X.

Nyberg, L., & Eriksson, J. (2016). Working memory: Maintenance, updating, and the realization of intentions. *Cold Spring Harbor Perspectives in Biology, 8*(2), a021816, https://doi.org/10.1101/cshperspect.a021816.

O'Donnell, R. E., Clement, A., & Brockmole, J. R. (2018). Semantic and functional relationships among objects increase the capacity of visual working memory. *Journal of Experimental Psychology: Learning Memory and Cognition, 44*(7), 1151–1158, https://doi.org/10.1037/xlm0000508.

Oberauer, K., & Lin, H. (2017). An interference model of visual working memory. *Psychological Review, 124*(1), 1–39, https://doi.org/10.1037/rev0000044.

Olson, I. R., & Jiang, Y. (2002). Is visual short-term memory object based? Rejection of the "strong-object" hypothesis. *Perception and Psychophysics, 64*(7), 1055–1067, https://doi.org/10.3758/BF03194756.

Paik, S. B., & Ringach, D. L. (2011). Retinal origin of orientation maps in visual cortex. *Nature Neuroscience, 14*(7), 919–925, https://doi.org/10.1038/nn.2824.

Park, Y. E., Sy, J. L., Hong, S. W., & Tong, F. (2017). Reprioritization of features of multidimensional objects stored in visual working memory. *Psychological Science, 28*(12), 1773–1785.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., & Sogo, H., …Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods, 51*(1), 195–203, https://doi.org/10.3758/s13428-018-01193-y.

Pertzov, Y., Dong, M. Y., Peich, M. C., & Husain, M. (2012). Forgetting what was where: The fragility of object-location binding. *PLoS One, 7*(10), e48214, https://doi.org/10.1371/journal.pone.0048214.

Postma, A., & De Haan, E. H. F. (1996). What Was Where? Memory for Object Locations. *Quarterly Journal of Experimental Psychology Section A, 49*(1), 178–199, https://doi.org/10.1080/713755605.

Postma, A., Kessels, R. P. C., & van Asselen, M. (2008). How the brain remembers and forgets where things are: The neurocognition of object-location memory. *Neuroscience and Biobehavioral Reviews, 32*(8), 1339–1345, https://doi.org/10.1016/j.neubiorev.2008.05.001.

Pylyshyn, Z. W. (2000). Situating vision in the world. *Trends in Cognitive Sciences, 4*(5), 197–207, https://doi.org/10.1016/S1364-6613(00)01477-7.

Quirk, C., Adam, K. C. S., & Vogel, E. K. (2020). No evidence for an object working memory capacity benefit with extended viewing time. *Eneuro, 7*(5), ENEURO.0150–20.2020, https://doi.org/10.1523/ENEURO.0150-20.2020.

Ranganath, C., DeGutis, J., & D'Esposito, M. (2004). Category-specific modulation of inferior temporal activity during working memory encoding and maintenance. *Cognitive Brain Research, 20*(1), 37–45, https://doi.org/10.1016/j.cogbrainres.2003.11.017.

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*(2), 225–237, https://doi.org/10.3758/PBR.16.2.225.

Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science, 20*(2), 207–214, https://doi.org/10.1111/j.1467-9280.2009.02276.x.

Serences, J. (2016). Neural mechanisms of information storage in visual short-term memory. *Vision Research, 128*, 53–67, https://doi.org/10.1016/j.visres.2016.09.010..

Shafer-Skelton, A., Kupitz, C. N., & Golomb, J. D. (2017). Object-location binding across a saccade: A retinotopic spatial congruency bias. *Attention, Perception, & Psychophysics, 79*(3), 765–781, https://doi.org/10.3758/s13414-016-1263-8.

Shin, H., & Ma, W. J. (2017). Visual short-term memory for oriented, colored objects. *Journal of Vision, 17*(9), 1–19, https://doi.org/10.1167/17.9.12.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv,1409.1556

Schurgin, M. W., Cunningham, C. A., Egeth, H. E., & Brady, T. F. (2018). Visual long-term memory can replace active maintenance in visual working memory. *BioRxiv*, 381848, https://doi.org/10.1101/381848.

Spachtholz, P., & Kuhbandner, C. (2017). Visual long-term memory is not unitary: Flexible storage of visual information as features or objects as a function of affect. *Cognitive, Affective and Behavioral Neuroscience, 17*, 1141–1150, https://doi.org/10.3758/s13415-017-0538-4.

Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology, 6*(2), 171–178, https://doi.org/10.1016/S0959-4388(96)80070-5.

Treisman, A. (1999). Solutions to the binding problem: Progress through controversy and convergence. *Neuron, 24*(1), 105–110, https://doi.org/10.1016/S0896-6273(00)80826-0.

Treisman, A. (2006). How the deployment of attention determines what we see. *Visual Cognition, 14*, 411–443, https://doi.org/10.1080/13506280500195250.

Utochkin, I. S., & Brady, T. F. (2020). Independent storage of different features of real-world objects in long-term memory. *Journal of Experimental Psychology: General, 149*(3), 530–549, https://doi.org/10.1037/xge0000664.

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, https://doi.org/10.1016/j.cortex.2017.01.011.

van Lamsweerde, A. E., Beck, M. R., & Elliott, E. M. (2015). Retrieval from long-term memory reduces working memory representations for visual features and their bindings. *Memory & Cognition, 43*, 237–246, doi:10.3758/s13421-014-0468-0.

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions and objects in visual working memory. *Journal of Experimental Psychology. Human Perception and Performance, 27*(1), 92–114, https://doi.org/10.1037//0096-1523.27.1.92.

Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., & Verhagen, J., . . .Morey, R. D. (2017). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review, 25*, 58–76, http://doi.org/10.3758/s13423-017-1323-7.

Wang, B., Cao, X., Theeuwes, J., Olivers, C. N. L., & Wang, Z. (2017). Separate capacities for storing different features in visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*(2), 226–236, https://doi.org/10.1037/xlm0000295.

Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology. General, 131*(1), 48–64, https://doi.org/10.1037//0096-3445.131.1.48.

Wilson, F. A. W., Scalaidhe, S. P., & Goldman-Rakic, P. S. (1993). Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science, 260*(5116), 1955–1958, https://doi.org/10.1126/science.8316836.

Xie, W., & Zhang, W. (2017). Familiarity increases the number of remembered Pokémon in visual short-term memory. *Memory & Cognition, 45*(4), 677–689, https://doi.org/10.3758/s13421-016-0679-7.

Xu, Y. (2002). Encoding color and shape from different parts of an object in visual short-term memory. *Perception & Psychophysics, 64*(8), 1260–1280, https://doi.org/10.3758/BF03194770.

Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature, 440*(7080), 91–95, https://doi.org/10.1038/nature04262.

Xu, Y. (2017). Reevaluating the sensory account of visual working memory storage. *Trends in Cognitive Sciences, 21*(10), 794–815.

Appendix C. Effects of item distinctiveness on the retrieval of objects and object-location bindings from visual working memory

Article "Effects of item distinctiveness on the retrieval of objects and object-location bindings from visual working memory"

Markov, Y.M. & Utochkin, I.S. (2022). Effects of item distinctiveness on the retrieval of objects and object-location bindings from visual working memory. *Attention, Perception, & Psychophysics*

Abstract. Visual working memory (VWM) is prone to interference from stored items competing for its limited capacity. Distinctiveness or similarity of the items is acknowledged to affect this competition, such that poor item distinctiveness causes a failure to discriminate between items sharing common features. In three experiments, we studied how the distinctiveness of studied real-world objects (i.e., whether the objects belong to the same or different basic categories) affects the retrieval of objects themselves (simple recognition) and object-location conjunctions (information about which object was where in a display, cued recall). In Experiments 1 and 2, we found that distinctiveness did not affect memories for objects or for locations, but low-distinctive objects were more frequently reported at "swapped" locations that originally contained other objects, showing object-location memory swaps. In Experiment 3 we found that observers swapped the location of a tested object with another object from the same category more frequently than with any of the objects from another category. This suggests that more similar studied objects cause more retrieval competition in object-location judgments than in simple recognition. Additionally, we discuss a possible role of categorical labeling of locations that can support object-location retrieval when the studied objects are highly distinct.

https://doi.org/10.3758/s13414-022-02451-0

Attention, Perception, & Psychophysics
https://doi.org/10.3758/s13414-022-02451-0

Effects of item distinctiveness on the retrieval of objects and object-location bindings from visual working memory

Yuri A. Markov[1,2] • Igor S. Utochkin[1]

Accepted: 25 January 2022

Abstract

Visual working memory (VWM) is prone to interference from stored items competing for its limited capacity. Distinctiveness or similarity of the items is acknowledged to affect this competition, such that poor item distinctiveness causes a failure to discriminate between items sharing common features. In three experiments, we studied how the distinctiveness of studied real-world objects (i.e., whether the objects belong to the same or different basic categories) affects the retrieval of objects themselves (simple recognition) and object-location conjunctions (information about which object was where in a display, cued recall). In Experiments 1 and 2, we found that distinctiveness did not affect memories for objects or for locations, but low-distinctive objects were more frequently reported at "swapped" locations that originally contained other objects, showing object-location memory swaps. In Experiments 3 we found that observers swapped the location of a tested object with another object from the same category more frequently than with any of the objects from another category. This suggests that more similar studied objects cause more retrieval competition in object-location judgments than in simple recognition. Additionally, we discuss a possible role of categorical labeling of locations that can support object-location retrieval when the studied objects are highly distinct.

Keywords visual working memory · memory distinctiveness · recognition memory · object-location memory · swap errors

Working memory is often referred to as a system that actively maintains and operates the information necessary for current goals and tasks (Baddeley, 1986; Baddeley & Hitch, 1974). One of the most important attributes of working memory is its limited capacity (Cowan, 2001; Miller, 1956). This seems to be true for visual working memory (VWM) as well, a subsystem holding visual information in the working state (Alvarez & Cavanagh, 2004; Luck & Vogel, 1997), though for real-world objects the capacity is not fixed (Asp et al., 2021; Brady et al., 2016; Brady & Störmer, 2021).

Apart from the limited capacity and the limited time of storage, interactions between specific contents of VWM can also influence an ability to retrieve these contents. One of the strong determinants of these interactions sources of interference is *distinctiveness* (Hunt, 2006) referring to the ability of an item to produce a reliable memory relative to other items or the context. It is usually described and measured as the likelihood of successful recall or the recognition of an item as a function of its similarity or dissimilarity with other memory items. In the current study, we will address effects of item distinctiveness/similarity on two different kinds of VWM contents, namely, memories for objects themselves and object-location memories, that is, memories for both the objects and where they have been originally presented (which refers to an ability to store and retrieve richer information about objects and context under which these objects are encoded).

Distinctiveness in memory for objects

Early research in this area focused on how the distinctiveness/similarity affects retrieval from long-term memory for semantically meaningful material. Indicative examples include the von Restorff isolation effect, a greater recall probability for semantically odd items studied among multiple semantically

---

Author Note The study was conducted within the Basic Research Program of the National Research University Higher School of Economics in 2020 (RS-60)

✉ Yuri A. Markov
yuamarkov@gmail.com

[1] HSE University, Moscow, Russia
[2] Laboratory of Psychophysics, Brain Mind Institute, EPFL, Lausanne, Switzerland

related items (Hunt, 1995), or the Deese-Roediger-McDermott (DRM) effect, a greater probability to false alarm a never presented foil if it is semantically related to a number of presented items from the same category (Deese, 1959; Roediger & McDermott, 1995). The role of picture distinctiveness was also shown in visual long-term memory (Hall et al., 2021; Konkle et al., 2010; Standing, 1973). Similarly, early studies in the field of working memory were focused on the distinctiveness effects on verbal memories (Conrad, 1964; Baddeley, 1966a; Baddeley, 1966b) and later research extended this focus to visual working memory for objects (Avons & Mason, 1999; Cohen et al., 2014; Jalbert et al., 2008). Retrieval decrements in VWM caused by the low distinctiveness of remembered objects, similar to those in long-term memory (Konkle et al., 2010), were documented in various tasks (Avons & Mason, 1999; Jalbert et al., 2008; Cohen et al., 2014; Jiang, Remington, Asaad, Lee, & Mikkalson, 2016b; Yang & Mo, 2017, etc). Another landmark study (Awh, Barton, & Vogel, 2007) also pointed to a critical role of categorical similarity between encoded sample items and test items in capacity estimates of VWM. The effects of item distinctiveness and similarity on VWM for these items can be linked to the degree of separation between neural representations within the higher levels of the visual cortex, such as the occipito-temporal cortex (Cohen et al., 2014). However, the existing data regarding object distinctiveness in VWM are controversial. There are studies reporting that low item distinctiveness can increase rather than decrease performance in VWM tasks (Jiang, Lee, et al., 2016a; Lin & Luck, 2009; Sims et al., 2012). Jiang, Lee, et al. (2016a) suggested that the low distinctiveness advantage is usually observed for stimuli with feature variation along basic continuous feature dimensions (RGB for color, 360° for orientation, facial morphs, etc.), while the high-distinctiveness advantage could be observed for complex stimuli whose differences are categorical and cannot be presented on a continuum (for example, faces and scenes). Avital-Cohen and Gronau (2021) suggested that the high distinctiveness advantage can be explained by an attentional bias that affects the allocation of VWM resources to certain categories, which is, perhaps, is the reason why this advantage was not observed for all types of stimuli in Jiang et al.'s (2016b) study. Thus, the low distinctiveness advantage in VWM rather works for stimuli whose differences are defined perceptually, whereas the high distinctiveness advantage occurs in memory for categories and could be referred to as conceptual distinctiveness. The idea of conceptual and perceptual distinctiveness/similarity independently contributing to object memory is supported by studies using images of real-world objects. For example, Konkle et al. (2010) have demonstrated that conceptual distinctiveness (how different remembered objects are in kind) rather than perceptual distinctiveness (how different the objects are in color, shape, etc.) mediates subsequent object recognition in a LTM test. Furthermore, Brady and Störmer (2020, 2021) demonstrated that the meaningfulness of to be remembered stimuli, that is, the ability of these stimuli to be recognized as everyday objects, human faces, etc., boosts VWM capacity and that this boost cannot be explained solely by perceptual differences between meaningful and meaningless stimuli (Asp et al., 2021). Based on these findings, we suggest that conceptual distinctiveness, along with perceptual similarity (Hovhannisyan et al., 2021; Hu et al., 2020; Liu et al., 2020; Naspi et al., 2021; Otsuka et al., 2013; Xie & Zaghloul, 2021), can also be an important determinant of VWM for real-world objects.

Object-location memory

Apart from storing the information about objects themselves, VWM is often considered as a strongly spatially referenced system (Logie, 2003; Magen & Emmanouil, 2019). This is the case for some of the paradigms used for VWM research (for review, see Brady et al., 2011; Suchow et al., 2014). For example, in a change detection task, observers have to report whether there is a difference between two consecutively presented sets of items. The whole essence of this task is the idea that if an observer has a critical item in VWM, they will be clearly aware of its change in a certain location (Luck & Vogel, 1997; Pashler, 1988; Phillips, 1974). In a continuous report task, observers also memorize a set of spatially distributed items, and then the location of one of them is directly used as a cue to recall a particular item presented exactly in that location (e.g., Wilken & Ma, 2004; Zhang & Luck, 2008). Remembering certain objects at certain locations (what was where) in the VWM has received much attention and was recently formalized by theories of binding in working memory (Oberauer & Lin, 2017; Schneegans & Bays, 2017; Swan & Wyble, 2014).

Along with errors that occur at recognition tests for object memory, object-location is associated with a specific kind of error termed swaps. They refer to a failure to correctly retrieve both an object and the location of that object, or one given another (Dent & Smyth, 2005; Hollingworth & Rasmussen, 2010; Pertzov et al., 2012; Postma & De Haan, 1996; Toh et al., 2020; but see Pratte, 2019). For example, a report can be identified as a swap when an observer is asked to report an item presented at a cued location but reports instead an item presented at another location (Bays et al., 2009). The swaps can be observed in change detection (Donkin et al., 2015) and continuous report (Bays et al., 2009) tasks. They are perhaps the most frequent and introspectively obvious binding failures (Treisman, 1996) that can be found in the everyday memory experience. For example, when a person places a smartphone and a wallet in their left and right pockets, respectively, the person can later recollect which items are in the pockets but can swap their locations in memory and look for the wallet in the left pocket. As we recently showed, object-location swaps are quite common in VWM tasks with everyday real-world objects (Markov et al., 2021). Due to the relevance of object-location memory for everyday performance, it is extensively studied from the individual differences (Cohen-Dallal et al., 2021; Liang et al., 2016; Lu et al., 2020; Pavisic et al., 2018; Pertzov et al., 2015; Pertzov et al., 2013; Zokaei et al., 2017) and clinical perspectives (Pavisic et al., 2020). From the fundamental perspective, some theories view location-based binding as a critically important mechanism to maintain and recall bindings between internal features (e.g., color and orientation) that make a whole object (Schneegans & Bays, 2017).

Object identity and object location appear to be appropriate representational "units" for object-location binding in VWM. The object-location binding problem has a solid neuronal background. The information about *which* objects we see and the information about *where* we see these objects are processed by separate pathways of the visual system (Haxby et al., 1991; Mishkin & Ungerleider, 1982). A number of behavioral (Lee & Chun, 2001; Li et al., 2015; Wood, 2011) and neurophysiological (Darling et al., 2006; Smith et al., 1995) studies showed that memory for objects and memory for locations have separate capacities. Moreover, there is also evidence that object-location memory can be a separate process from storing only objects or only locations (Postma & De Haan, 1996; Postma, Kessels, & van Asselen, 2008). Pertzov et al. (2012) suggested that object-location swaps play an important role in delay-related forgetting in VWM. They showed that while objects themselves are still in VWM, their binding to correct locations can suffer from a delay between encoding and retrieval.

How does distinctiveness/similarity come to play in object-location-memory and swaps? Assuming both objects and locations are encoded and stored imprecisely, there can be some overlap between these representations. These overlaps can cause erroneous recall of an object given a cued location or vice versa. For example, even if an object and its location are successfully bound (which means the association between representations of that object and that location is stronger than between that object and any other location, Oberauer & Lin, 2017), an object from a different location can be reported because of overlap between the representation of the cued location and that of another location. The degree of overlap is a matter of distinctiveness. For example, Bays et al. (2009) found that the swap rate in a standard color continuous report task increases with memory set size. They suggested that the imprecision (representational indistinctiveness) of both color and locations grows with set size, and this causes less efficient of a location cue to retrieve a correct color. Similarly, Oberauer and Lin (2017) suggested that presenting a recall cue (e.g., cueing a certain location) probabilistically increases activation in the to be reported feature space (e.g., color), and the amount of this activation for each feature is modulated by the similarity between the cued location and that associated with a given color. Having said that, object-location swaps are not always considered to reflect binding itself. Alternatively, swaps can be thought of as a result of the way memory gets access to an object representation based on location (or vice versa). This leads to more frequent swaps between items presented at closer locations or having more similar features (Emrich & Ferber, 2012; Oberauer & Lin, 2017).

Our study

Our study aims to investigate how object distinctiveness affects both object recognition and object-location memory when observers have to remember a set of real-world objects and their spatial positions. Although some of the previous studies have reported observing some negative correlations between item distinctiveness and the probability of swap errors (e.g., Dent & Smyth, 2005; Postma & De Haan, 1996; Oberauer & Lin, 2017), they did not address the nature of their correlation as a major issue. Does poor distinctiveness cause less likely object retrieval and the inaccessibility of its location as a consequence? Or is it possible that location memories become noisier under the increased demands for object distinction, which makes the representations at these locations more penetrable to neighbor representations (Emrich & Ferber, 2012)? Or does low object distinctiveness somehow impair object-location memory itself, although object memories and location memories are intact? To answer these questions, all three components should be comparatively tested under low and high object distinctiveness. Moreover, specific "swap" errors should be thoroughly analyzed in terms of how likely they are caused by a failure associated with object recognition or object-location memory. Also importantly, previous studies that tested the role of item distinctiveness or similarity in various object-location tasks were based on either simple objects and continuous perceptual features (e.g., colors, Oberauer & Lin, 2017) or complex but unfamiliar objects whose differences are defined mostly visually (e.g., Dent & Smyth, 2005). However, it is less clear how object-location memory is affected by the distinctiveness of real-world objects that have a strong semantic component in them, given that object recall can behave quite differently for objects whose differences are defined by visual and conceptual features. We implemented this program in three experiments. They were designed to carefully test how the requirements to distinguish between real-world objects and remember them at certain locations affect memory for object-location conjunctions. In Experiment 1, we directly tested the influence of distinctiveness (that we operationalized as all studied objects being from the same or from different basic categories) on object recognition memory and object-location memory. In Experiment 2, we compared object-location memory (that is, when the task requires to actively remember object identities and which locations these identities belonged to) against "pure" location memory (when demands on object memory and object-location memory are diminished by the high consistency of object identities and their spatial order, but location themselves change unpredictably). To anticipate, Experiments 1 and 2 showed that observers commit more object-location swap errors when their memory set is low-distinctive, although their recognition memory for objects was not affected by item distinctiveness. In Experiment 3, we tested two plausible accounts of the distinctiveness effect on swaps. One account suggests that low distinctiveness impairs object-location memory in general, leading to more likely forgetting of object-location information in an all-or-none manner, which should increase the number of swap errors regardless of their similarity to the target. Another account suggests that observers can maintain reliable object-location information when the studied objects have low distinctiveness, but swaps occur more often because of a stronger competition between the target and nontargets when they are called by a retrieval cue.

Experiment 1

Experiment 1 aimed to investigate how object distinctiveness interacts with object-location memory. We presented sets of real-world objects to our participants, asking them to remember all of the objects and where each object was located. We subsequently tested object recognition and how precisely the object location was adjusted. This gave us estimates of all constituents of object-location memory: object memory, location memory, and object-location conjunction memory. We manipulated object distinctiveness via conceptual (categorical) homo/heterogeneity of object sets which is relevant for complex, real-world objects. In accordance with the previous findings (Konkle et al., 2010), objects belonging to the same basic category have low conceptual distinctiveness and are expected to be harder to recognize against a novel foil, whereas objects from all different categories have high conceptual distinctiveness and are expected to be easier to recognize. Therefore, we ask whether the need to remember low-distinctive objects from the same category also affects an ability to remember objects at correct locations.

Method

Participants

Twenty-two psychology students from the Higher School of Economics (19 female; age *M* = 19.5 years) took part in the experiment for extra course credits. All participants reported having a normal color vision, normal or corrected to normal visual acuity, and no neurological problems. Before the beginning of the experiment, they signed an informed consent form. In this and subsequent experiments, sample sizes were determined based on similar studies addressing the issue of feature storage and binding in VWM and using a continuous report task (from 10 to 16; for example, Fougnie & Alvarez 2011; Fougnie et al., 2010; Bays et al., 2009; Pertzov et al., 2012). The planned sample size also included a few extra participants taking into account the possibility of technical problems or poor performance in some participants.

Apparatus and stimuli

Stimulation was developed and presented through PsychoPy (Peirce, 2007; Peirce et al., 2019) for Linux Ubuntu. Stimuli were presented on a standard VGA monitor with a refresh frequency of 75 Hz and 1024×768-pixel spatial resolution. Stimuli were presented on a homogeneous white field. Participants sat approximately 47 cm from the monitor. From that distance, the screen subtended approximately 42.4 × 32.5 degrees of visual angle.

Objects We used the real-world object database created by Konkle and colleagues (Konkle, Brady, Alvarez, & Oliva, 2010). 200 unique categories and 16 unique exemplars from each category were chosen from the database. Two example categories ("apple" and "toy soldier") with four exemplars are shown in Fig. 1. The objects were scaled to subtend approximately 4.34° of visual angle.

Spatial layout Each sample screen contained three objects. The centers of the objects lay on an imaginary circle with a radius of 12.8°. The only parameter defining the position of each object was the rotational angle on the imaginary circle. These angles were chosen randomly for each object in each trial. The only restriction was that the minimum distance between the centers of any two objects was 60° of rotation. This was done to avoid overlap or clustering between the objects.

Object distinctiveness The three sample objects could have low or high distinctiveness. In low-distinctiveness samples, all three items were different exemplars drawn from the same object category. In high-distinctiveness samples, the items were drawn from three different categories. 100 unique categories were used to make 100 sample trials in the low-distinctiveness condition, so that each category was presented only once per experiment. Another 100 unique categories were used to make 100 samples in the high-distinctiveness trials, so that each category could appear three times per experiment, but each time using a new exemplar. Overall, this warranted that each particular exemplar could appear only

Fig. 1 Examples of two categories (“apple” and “toy soldier”) and four exemplars from each category (from the database originally developed by Konkle et al., 2010)

once in the entire experiment as a sample. To test recognition memory for object identity, one of the sample (“old”) objects was presented paired with a “new” object. Both old and new objects were always from the same category, regardless of sample distinctiveness. This allowed keeping the test difficulty fixed while manipulating the difficulty of encoding and storage in working memory (Awh et al., 2007). Target sample items for subsequent testing were randomly assigned across participants.

Our claim that sample sets consisting of objects from the same category represent low distinctiveness and sample sets consisting of objects from different categories represent high distinctiveness is not based only on a-priori assumptions but also empirically grounded. In their original study of massive memory, Konkle et al. (2010), whose stimuli we use in this study, showed that remembering several exemplars from the same category impairs subsequent recognition memory for any such exemplar compared to remembering unique exemplars.

Procedure

Figure 2 shows the organization of a trial in Experiment 1. At the beginning of each trial, three sample objects were presented for 2 seconds. Then, after a 1-second blank interval, participants were shown two objects, one “old” and one “new”. Participants had to choose an object that they had thought to be “old”, that is, presented in the sample set (recognition task). The new and the old test objects were presented randomly either to the left or to the right of the center of the screen. Participants answered pressing a left or a right arrow key indicating the location of the old object. The chosen object then remained on the screen and participants had to set its remembered location (localization task). To do that, the participants had to drag the object along a positional ring with a mouse (Fig. 2). The initial position of the probed item was at the center of the screen until a first mouse click, which moved the object to an imaginary sample location circle. When the location was set, the participants pressed “SPACE” to confirm the response and terminate the trial. After response confirmation, feedback informed the participants how close their response had been to the true object location and their accuracy regarding object recognition. The location set by the participant was shown by a red cross centered at the set angle; the

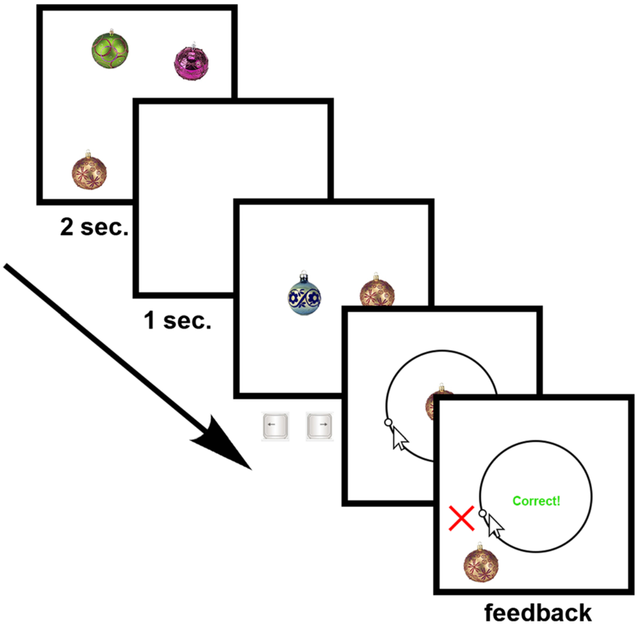


Fig. 2 The time course of a trial in Experiment 1

true location was shown by the probed object presented at that location (Fig. 2).

100 low-distinctiveness and 100 high-distinctiveness trials were presented in total in random order. Each particular object (exemplar from a category) was presented only once during the experiment.

During the entire experiment, participants were instructed to repeat a syllable "ba" aloud at a rate of about 3 Hz to prevent verbal encoding of stimuli. The experiment was preceded by ten practice trials in order to familiarize participants with the task. The total time of the experiment was between 30 and 45 minutes.

Data analysis

For the object recognition task, the percentage of correct answers was calculated. For the localization task, localization errors were calculated in each trial. The error was estimated as the angular difference between a participant's response and the true location of a probed object. The distribution of errors was then analyzed using the mixture model (Zhang & Luck, 2008) with a modification suggested by Bays and colleagues (Bays et al., 2009), the "swap" model. For modeling, we used MemToolbox for MATLAB (Suchow, Brady, Fougnie, & Alvarez, 2013). The model has three parameters derived from the three decomposed components of the error distribution. The first parameter is the standard deviation (*SD*) of the von Mises distribution built around the mean of 0, which is supposed to reflect responses made about the items whose locations are present in memory with some noise. In representational terms, *SD* estimates the precision of memory trace for the location of a probed item. The second parameter is the probability of random guess (P_{guess}) estimated as an area below the uniform component of the mixed distribution that reflects the random picking of locations in the absence of memory for the location of a probed item. The third parameter is the probability of a "swap" (P_{swap}), estimated as the area of a second von Mises distribution component. This second von Mises distribution is assumed to have a mean that equals the location of a distractor item (each of the sample objects that are memorized but not probed) and the same *SD* as the first von Mises distribution. P_{swap} accumulates the responses originating from a misreport of a distractor location instead of a probe location, which is the object-location binding error. Knowing P_{guess} and P_{swap}, it is easy to calculate the probability of correctly bound objects and locations held in VWM (P_{memory}) using the following formula: $P_{memory} = 1 - (P_{guess} + P_{swap})$.

Since the quality of parameter estimation in a mixture model is sensitive to the amount of data, we included trials with both correctly and incorrectly recognized objects in this analysis. It is obvious that misrecognized items could differently contribute to the overall probabilities of the localization outcomes. For example, these misrecognized items could inflate the swap rate, as a failure to recognize an object might lead to a random choice among three locations remembered as being occupied by some object. Therefore, further analysis within the mixture model approach was designed to test how often various types of object-location outcomes occurred in trials where a tested object was recognized correctly. This required each trial to be classified in terms of two outcomes, object recognition (correct vs. incorrect) and object localization (correct location vs. swapped location vs. guessed location). To identify the latter type of the outcome, we transformed continuous localization errors into discrete labels indicating that a given pointed location is correct, swapped, or randomly guessed. The rule for labeling was taken from Fougnie & Alvarez (2011) and is based on the parameters of the mixture model. If a localization error fell within ±3 SD around a true target location, it was labeled as a likely correct localization response. If that error fell within ±3 SD around one of the distractor locations, it was labeled as a swap response. The rest of the errors falling beyond ±3 SD around both target and distractor locations were labeled as likely random guesses of a location.

To statistically estimate the effect of distinctiveness on object recognition and object localization, we applied the standard frequentist and Bayesian t-tests to the percentage of correct object recognition, as well as *SD*, P_{memory}, P_{swap}, and P_{guess} of object localization. The Bayesian t-test is a direct way to estimate evidence for H_1 against H_0 (Rouder, Speckman, Sun, Morey, & Iverson, 2009). The Bayes factor (BF_{10}), is the odds between the relative likelihoods of H_1 and H_0 under the observed data, was calculated using JASP 0.8.2 (JASP Team, 2017; Wagenmakers et al., 2017). The Cauchy distribution with a default width of .707 was used as a prior distribution of effect sizes under H_0 (JASP Team, 2017; Wagenmakers et al., 2017).

Results and discussion

In general, participants showed reasonably good memory both for object identities (percent correct recognition: $M = 87\%$ for high-distinctiveness trials, $M = 88.6\%$ for low-distinctiveness trials) and object locations (P_{memory}: $M = .764$ for high-distinctiveness trials, $M = .685$ for low-distinctiveness trials). We found no evidence for the effect of distinctiveness on object recognition ($t(21) = 1.550$, $p = .136$, $BF_{10} = .629$, Cohen's $d_z = .330$, Fig. 3a). For localization, we found that low distinctiveness decreased P_{memory} ($M = .685$) compared to high distinctiveness ($M = .764$; comparison: $t(21) = 4.148$, $p < .001$, $BF_{10} = 71.738$, $d_z = .884$, Fig. 3b). By contrast, low distinctiveness increased P_{swap} ($M = .203$) compared to high distinctiveness ($M = .129$; comparison: $t(21) = 5.257$, $p < .001$, $BF_{10} > 10^3$, $d_z = 1.121$, Fig. 3c). Finally, there was no evidence of any effect of distinctiveness on P_{guess} ($M = .108$ for

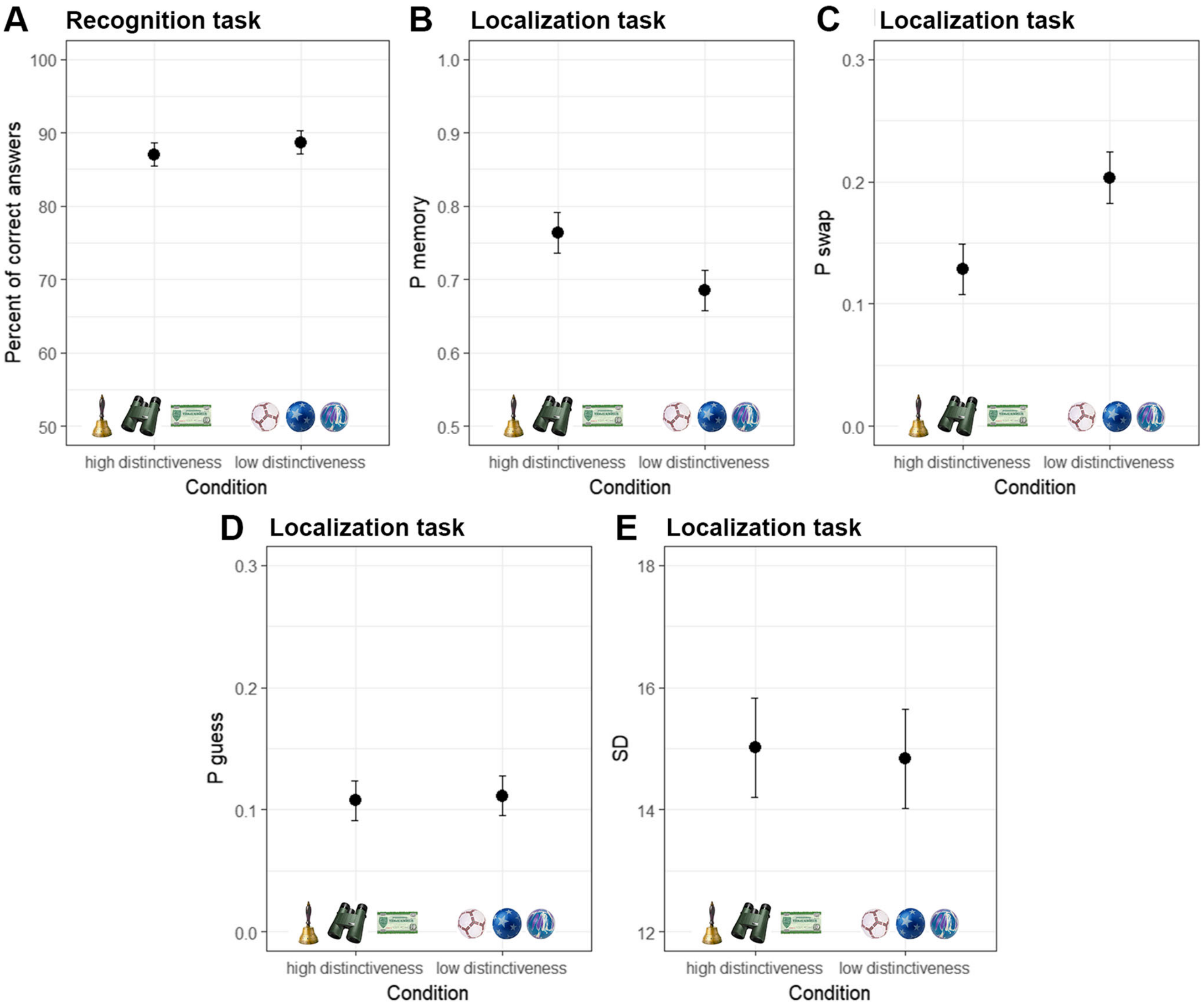


Fig. 3 The results of Experiment 1: The effect of distinctiveness on (**a**) percentage of correct answers on the object recognition task, (**b**) P_{memory}, (**c**) P_{swap}, (**d**) P_{guess} and (**e**) *SD* in the binding task. Error bars depict 95% CIs

high-distinctiveness trials, M = .112 for low distinctiveness trials; comparison: $t(21) = .348$, $p = .731$, $BF_{10} = .236$, d_z = .074, Fig. 3d) and on the *SD* (M = 15 for high-distinctiveness trials, M = 14.8 for low distinctiveness trials; comparison: $t(21) = .327$, $p = .747$, $BF_{10} = .234$, $d_z = .070$, Fig. 3e).

Object-location memory of correctly recognized items One participant was excluded from this analysis (their SD of localization error was too large to classify trials). Our analysis showed that the probability of swap errors in trials with correct recognition responses was greater in low-distinctive sets (M = .19) than in highly-distinctive sets (M = .13; comparison: $t(20) = 4.634$, $p < .001$, $BF_{10} = 182$, $d_z = 1.011$). The proportions of location guesses (P_{guess}) did not differ between the conditions (low-distinctive sets: M = .023, highly-distinctive sets: M = .029; comparison: $t(20) = 1.101$, $p = .284$, $BF_{10} = .388$, d_z = 0.240). The proportion of correctly localized objects was greater in the highly-distinctive sets (M = .71) than in the low-distinctive sets (M = .68; comparison: $t(20) = 2.117$, p =.047, $BF_{10} = 1.434$, $d_z = 0.462$). These differences are quite similar to those observed using the overall mixture model for all trials. Therefore, we conclude that the distinctiveness effect on object-location memory is preserved for correctly reported items and cannot be ascribed merely to object forgetting.

The results of Experiment 1 showed that the distinctiveness of sample objects did not affect object recognition. At the same time, we observed a specific effect on localization. Object distinctiveness did not change memory for locations themselves in terms of precision (*SD*) and the probability of remembering or forgetting the location (P_{guess}). Rather, it affected the probability of ascribing an object to its true location or to a location occupied by another object (swap error). There were fewer correctly localized items and more swap errors when distinctiveness was low.

Experiment 2

In Experiment 1, we combined tests for object recognition and object-location memory. There are two potential limitations of the testing method we used. First, the order of tests was fixed: recognition always came first, and object-location report always came second (as it was logically not possible to ask observers to localize a previously seen object before asking which of the objects the observers had seen). Therefore, the first order of the object recognition task could interfere with the subsequent object-location test. Second, whereas Experiment 1 allowed us to measure object memory as a necessary component of object-location memory, it did not allow us to measure another necessary component in isolation, spatial memory. Although object distinctiveness did not affect the *SD* or P_{guess}, that is, an ability to remember which locations were occupied by any objects, we cannot say that the need to remember the objects and their locations did not affect memories for locations (e.g., by making them less precise). In Experiment 2, we tested object-location memory without the preceding object recognition task. We also compared object-location memory for low-distinctive and high-distinctive objects (as in Experiment 1) with location memory in a task where object and object-location binding are not particularly relevant. In this version of the task, we used a consistent set of objects and located them in a consistent spatial order. Therefore, observers did not actually need to remember object identities and their relative locations ("bindings"). Absolute locations were the only features that unpredictably changed between trials and had to be encoded into working memory.

Method

Participants

Twenty-one psychology students of the Higher School of Economics (19 female; age: 18-21, $M = 20.07$) took part in the experiment for extra course credits. All participants reported having a normal color vision, normal or corrected to normal visual acuity, and no neurological problems. The sample size was determined based on the same rules as in Experiment 1. Before the beginning of the experiment, they signed an informed consent form.

Apparatus and stimuli

Apparatus and stimuli were the same as in Experiment 1. For the new task requiring localization without a strong need for remembering objects and object-location bindings, we made three pictures of hands depicting numbers from one to three. We used this kind of stimuli instead of showing regular Arabic or Roman numbers because they looked more like real-world objects and, thus, were more similar to Konkle et al.'s (2010) objects used in Experiment 1. On each sample display, all three "hand numbers" were located the same way as in the object-location version of the task. Additionally, their order was fixed: the "numbers" followed clockwise from "one". We assumed that this sort of display reduced the demands on memory for objects and object-location binding. The reduced demands on object VWM were provided by the fact that the "hand numbers" repeated consistently across the experiment and could be easily learned during a practice session. Moreover, we used the straightforward association between the number of raised fingers and the well-trained numerical representation in long-term memory. The reduced demands on object-location binding were provided by the consistent order of the "hand numbers". It allowed the recovery of the location of any given hand from memory for only one object-location combination. Therefore, by having reduced the uncertainty about object identities and about which locations they belong to, we kept the uncertainty regarding the locations themselves.

Procedure

The time course of a trial was the same as in Experiment 1, but no object recognition test was used after the retention interval. Instead, observers were shown by a single object from the sample immediately after the retention interval, and they had to localize that object with a mouse click. The object-location block of the experiment had the same design as in Experiment 1 (low-distinctive samples vs. highly-distinctive samples, 100 trials per condition). The location memory block (hereinafter - "hand localization" task) consisted of 200 trials. The number of trials was equated with the object-location task to control for serial position effects. However, only 100 random trials were drawn for the subsequent analysis, which was equal to the number of trials per condition in the object-location blocks. The trials were organized the same way as in the object-location task: Observers were presented with three "hand numbers" and had to recall the original location of a single probed hand (Fig. 4). The serial order of the object-location and the hand localization blocks was counterbalanced across participants.

Data analysis

For both the object-location and the hand localization tasks, we analyzed localization errors using the swap model (Bays et al., 2009), as described in Experiment 1. For each dependent variable (SD, P_{memory}, P_{guess}, P_{swap}), we ran the following planned comparisons. First, we compared localization performance between highly-distinctive and low-distinctive trials, which was a direct replication of the analysis for the same task in Experiment 1. Second, we compared object-location memory under each of the distinctiveness conditions with that obtained from the hand localization task. A Holm correction was

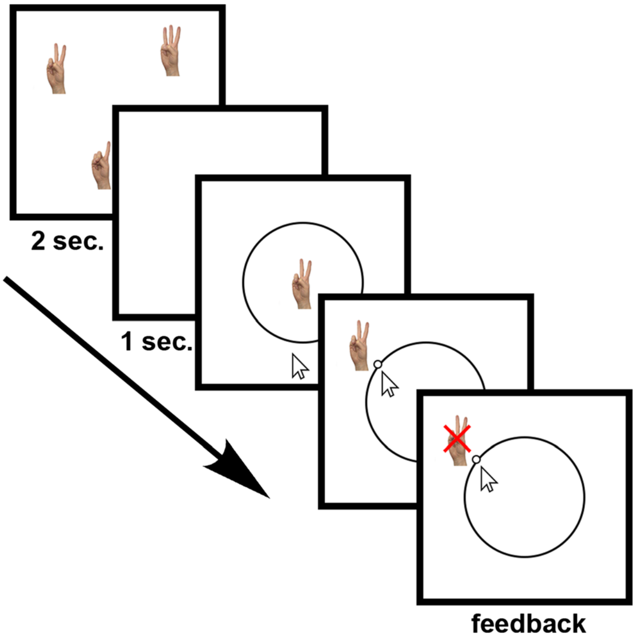


Fig. 4 The time course of a trial in the hand localization task

made for multiple comparisons in calculating the statistical significance level. For Bayesian *t*-tests, the same prior, as in Experiment 1, was used.

Results and discussion

The participants were above chance at remembering the locations of the hands (P_{memory}: $M = .92$) and of object-location conjunctions (P_{memory}: $M = .95$ for high-distinctiveness trials, $M = .89$ for low distinctiveness trials).

The results obtained in the object-location task of Experiment 1 were well replicated in Experiment 2. Low object distinctiveness decreased P_{memory} compared to high distinctiveness ($t(20) = 7.317$, $p <.001$, Bonferroni-Holm corrected $\alpha = .017$, $BF_{10} > 10^3$, $d_z = 1.597$, Fig. 5a). Symmetrically, low distinctiveness increased P_{swap} ($M = .07$) compared to high distinctiveness ($M = .03$; comparison: $t(20) = 6.164$, $p <.001$, Bonferroni-Holm corrected $\alpha = .017$, $BF_{10} > 10^3$, $d_z = 1.345$, Fig. 5b). Distinctiveness had no effect on P_{guess} (for low-distinctive objects $M = .04$; for highly distinctive objects $M = .02$; comparison: $t(20) = 2.030$, $p = .056$, Bonferroni-Holm corrected $\alpha = .017$, $BF_{10} = 1.253$, $d_z = .443$, Fig. 5c) and *SD* ($t(20) = .583$, $p = .566$, Bonferroni-Holm corrected $\alpha = .05$, $BF_{10} = .265$, $d_z = .127$, Fig. 5d).

The comparison between the object-location and the hand localization tasks showed no differences for P_{memory} (low distinctive objects vs. "hand numbers": $t(20) = .944$, $p = .36$, Bonferroni-Holm corrected $\alpha = .025$, $BF_{10} = .338$, $d_z = .206$; highly distinctive objects vs. "hand numbers": $t(20) = .677$, $p = .51$, Bonferroni-Holm corrected $\alpha = .05$, $BF_{10} = .28$, $d_z = .148$; Fig. 5a), P_{guess} (low distinctive objects vs. "hand numbers": $t(20) = .302$, $p = .77$, Bonferroni-Holm corrected $\alpha = .05$, $BF_{10} = .237$, $d_z = .066$; highly distinctive objects vs. "hand numbers": $t(20) = .899$, $p = .38$, Bonferroni-Holm corrected $\alpha = .025$, $BF_{10} = .326$, $d_z = .196$) and *SD* (low-distinctive objects vs. "hand numbers": $t(20) = .730$, $p = .38$, Bonferroni-Holm corrected $\alpha = .017$, $BF_{10} = .318$, $d_z = .159$; highly distinctive objects vs. "hand numbers": $t(20) = .865$, $p = .47$, Bonferroni-Holm corrected $\alpha = .025$, $BF_{10} = .289$, $d_z = .189$; Fig. 5c).

The proportion of swap errors (P_{swap}) in the hand localization task was very low ($M = .03$). This finding demonstrates that our hand localization task was easy in terms of remembering object-location bindings and probably can serve a proper tool for measuring spatial memory alone. However, proportions of swaps for locating low distinctive objects were greater than this baseline (low-distinctive objects vs. "hand numbers": $t(20) = 3.316$, $p = .003$, Bonferroni-Holm corrected $\alpha = .025$, $BF_{10} = 12.421$, $d_z = .724$; Fig. 5b), but not highly distinctive objects (high-distinctiveness objects vs. "hand numbers": $t(20) = .333$, $p = .743$, Bonferroni-Holm corrected $\alpha = .05$, $BF_{10} = .239$, $d_z = .073$; Fig. 5b).

Our results show that memory for locations is basically not affected by the need to also remember which object goes with which location. Critically for this conclusion, the precision (*SD*) and overall capacity (reverse P_{guess}) parameters did not change across the tasks and conditions. Object distinctiveness also did not affect the precision and capacity of location representations in memory. These findings suggest that additional demands on memory associated with harder object distinction do not impair location memory per se. As the results from both Experiments 1 and 2 show, observers demonstrate more swap errors (P_{swap}) for low distinctive objects. These conclusions replicate the conclusions from Experiment 1. In addition to Experiment 1, we ruled out a possibility that the effect of distinctiveness on the probability of the swap errors could be caused by the preceding object recognition test. We addressed the nature of the distinctiveness effect on object-location swaps more rigorously in Experiment 3.

Experiment 3A

To address the differential effects of object distinctiveness on object recognition and object-location memory observed in Experiment 1 and Experiment 2, we considered two plausible explanations. The first potential explanation is a non-specific impairment of memory caused by more difficult object distinction. To recall different object location-conjunctions, one should memorize objects themselves, locations, and, finally, which object goes with which location. Since low-distinctive objects are harder to discriminate (Konkle et al., 2010), the maintenance of a consistently high recognition rate may require more effort. In other words, this can cause a trade-off between remembering objects and object-location

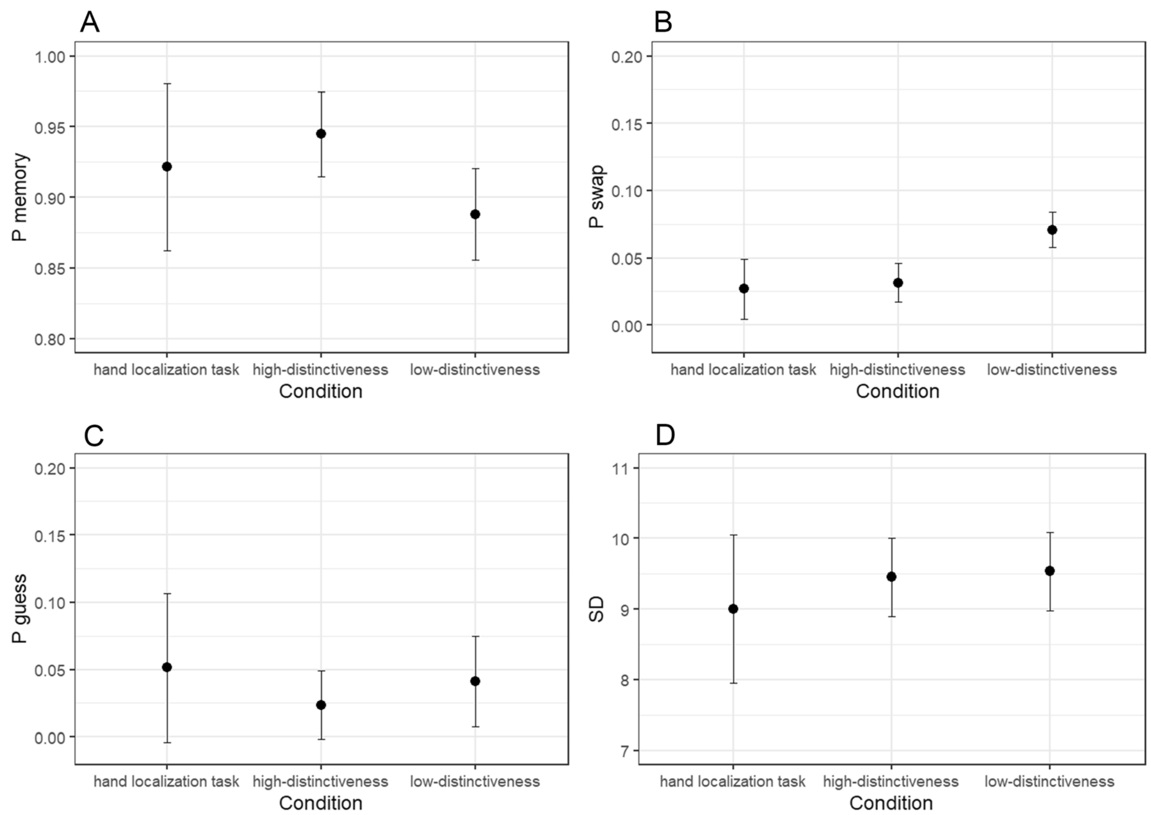


Fig. 5 The results of Experiment 2: The effect of distinctiveness on memory parameters in the binding task and parameters of hand localization task: (**a**) P_{swap}, (**b**) P_{memory}, (**c**) $P_{guess,}$ and (**d**) *SD*. Error bars depict 95% CIs

conjunctions biased in favor of the former (perhaps, because it is hardly possible to store conjunction if object memory fails). This kind of trade-off can yield a greater rate of swaps in the low-distinctive condition when observers were trying to better remember subtle differences between various exemplars more often fail to remember where each of the exemplars had been located. The second potential explanation of the advantageous effect of high distinctiveness on object-location memory is that the retrieval cue acts more effectively to tease apart targets from distractors when the items are more distinct. That is when a target object is probed, the observer checks their object-location associations for how familiar they seem given the target. If nontarget locations are associated with objects highly distinct from the probed target, then they provide a weaker familiarity signal (Oberauer & Lin, 2017; Schneegans & Bays, 2017; Swan & Wyble, 2014; Schurgin et al., 2020) and, thus, a swap is less likely. The critical difference between the two accounts is the representational "fate" of objects that are swapped at the report. The first account suggests that object-location recall in an all-or-none fashion and that remembering low-distinctive objects increases the probability of such a failure non-specifically. That is, when the observers do not remember where exactly an object belonged, they randomly guess, choosing one of the locations (which they remember per se) as a "home" for the object. The second account suggests that when observers fail to recall correct object-location conjunction, they still have some noisy representation that can be separated from nontarget conjunctions, and the degree of separation is a matter of distinctiveness. In this second case, we predict that the observers will more likely swap between more similar objects (same category) than between more distinct objects (different categories).

In Experiment 3A, we tested these two accounts against each other. We presented observers with sets of four objects at four different locations and then asked them to identify which of the four objects had been presented at one particular location. No object recognition and no continuous report were required in this task. As in Experiments 1 and 2, all objects could be highly distinctive (drawn from four different categories) or low-distinctive (drawn from one category). Critically, we added a third condition with two categories and two exemplars in each category. Here, the distribution of responses between test alternatives, in particular between incorrect alternatives (non-targets), was a sensitive measure to distinguish between the two accounts of the distinctiveness effect from Experiment 1. If distinctiveness affects an ability to store object-location information in an all-or-none fashion (either remember the exact conjunction, or pick it randomly), then we can expect that all of the nontargets would be chosen with the same probability. But if distinctiveness affects object-location memory via the relative familiarity of nontargets given a retrieval cue, then we can expect that observers would choose a nontarget from the same category as the target more frequently than the foils from the different category.

Method

Participants

Nineteen psychology students of the Higher School of Economics (19 female; *M* = 20.07) took part in the experiment for extra course credits. All participants reported having normal color vision, normal or corrected to normal visual acuity, and no neurological problems. The sample size was determined based on the same rules as in Experiment 1. Before the beginning of the experiment, they signed an informed consent form.

Apparatus and stimuli

Apparatus and stimuli were basically the same as in Experiment 1. An important difference was that sample sets consisted of four rather than three objects. Four objects were placed on an imaginary circle (radius of 12.8°) at the positions of 45°, 90°, 135° and 270° with a random jitter within ±2.17° along a radius for each object. The size of each object was 4.34°. The four objects presented in a sample set could be drawn from one category (all four objects were different exemplars of this category), from two categories (two exemplars from one category and two exemplars from another category), or from four categories (all exemplars belonged to different categories).

Procedure

Four objects were presented for 2 seconds. After a 1-second blank interval, all four sample objects were presented in a 2X2 square matrix around the center of the screen. At the same time with the matrix of four objects, a square box indicating one of the original locations of the sample set appeared. Participants had to click on an object that had been presented at the cued location in the sample. Feedback appeared after each trial as to whether the object had been correctly ascribed to the cued location (Fig. 6).

Data analysis

We analyzed the percentage of correct answers for the three conditions: one, two, or four categories. Importantly, in the condition with two categories, we also analyzed the percentages of choices broken down by all alternatives: correct object, incorrect object belonging to the same category as the correct one, and two incorrect objects belonging to a different category. This analysis was critical to test whether observers prefer a correct category over incorrect one, even when they fail to determine the exact identity of an object presented at a cued location. A Bonferroni correction was made for multiple comparisons in calculating the statistical significance level. For Bayesian *t*-tests, the same prior, as in Experiment 1, was used.

Results

*Accuracy.*We found a significant effect of the number of categories on the overall accuracy ($F(2,36) = 13.97$, $p < .001$, $\eta^2 = .437$, $BF_{10} = 3.125$). The percentage of correct answers was overall lower when participants had to remember all items from one category compared to two or four categories (one category, $M = 60.05\%$, vs. two categories, $M = 68.5\%$: $t(18) = 5.056$, $p < .001$, Bonferroni corrected $\alpha = .017$, $BF_{10} = 124.077$, Cohen's d = 1.16; one category, $M = 60.05\%$ vs. four categories, $M = 66.5\%$: $t(18) = 3.863$, $p < .001$, Bonferroni corrected $\alpha = .017$, $BF_{10} = 29.434$, Cohen's d = .886; two categories, M = 68.5% vs. four categories, M = 66.5%: t(18) = 1.193, p = .241, Bonferroni corrected $\alpha = .017$, BF10 = 0.548, Cohen's d = .274; see Fig. 7a). This result replicates one of the principal findings from Experiment 1, where participants also committed more localization errors for objects belonging to the same category.

Within-category vs. across-category localization errors Our analysis of responses broken down by outcomes in the two-categories condition has found a strong effect ($F(2,36) = 136.444$, $p < .001$, $\eta^2 = .883$, $BF_{10} > 10^{22}$). The strength of the effect was predominantly provided by the high prevalence of correct answers ($M = 68.5$ vs. foil from the same category, $M = 17.6$: $t(18) = 12.784$, $p < .001$, Bonferroni corrected $\alpha = .0167$, $BF_{10} > 10^6$, Cohen's d = 2.933; correct answer vs. foil from the different category, $M = 6.95$: $t(18) = 15.452$, $p < .001$, Bonferroni corrected $\alpha = .0167$, $BF_{10} > 10^7$, Cohen's d = 3.545). More interestingly, the foils were also chosen with different probabilities: the foil from the same category as target was chosen more often than any of the foils from the different category (foil from the same category vs. foil from the different category: $t(18) = 2.67$, $p = .011$, Bonferroni corrected $\alpha = .017$, $BF_{10} = 3746$, Cohen's d = .612 see Fig. 7b).

Discussion

The results of Experiment 3A replicated one of the major findings from Experiments 1 and 2 using a different version of object-location memory test. We found that an ability to correctly localize a remembered object was impaired when all objects had low distinctiveness (all items belonged to the same object category). The novel finding from Experiment 3A was that incorrect localizations were distributed non-uniformly across the rest of the items. The nontarget from the same category as the target was chosen about twice as frequently as each of the two foils from the different category. That is, even when participants failed to remember which particular object had been presented at a certain location and, thus, committed a swap, they still could rely on some memory about a category presented at that location. This can account for the greater swap rate in the displays consisting of objects from the same category compared to the displays consisting of objects from different categories. It can be easier to store multiple categories related to certain locations than multiple within-category variations provided by different exemplars of the same category.

Experiment 3B

In Experiment 1, we tested both object recognition memory and object localization for sample sets of three objects. In Experiment 3A, we tested only localization accuracy for sample sets of four objects, but we did not test recognition. Experiment 3B aimed to test object recognition alone, without any test of object-location memory, in the same setting as was used in Experiment 3A.

Method

Participants

Eighteen psychology students from the Higher School of Economics (17 female; age: 18-39, $M = 20.88$) took part in the experiment for extra course credits. All participants reported having normal color vision, normal or corrected to normal visual acuity, and no neurological problems. The sample size was determined based on the same rules as in Experiment 1. Before the beginning of the experiment, they signed an informed consent form.

Apparatus, stimuli, spatial layout and procedure

The sample stimuli were similar to Experiment 3A in terms of the number of objects, their layout and distinctiveness. In Experiment 3B spatial memory was not tested, observers were asked to remember only objects. On the test display following the one-second retention interval, two objects were presented to the right and to the left from the center. As in the recognition task of Experiment 1, one of the objects was “old” and the other object was “new”. Both the old and the new objects were drawn from the same category. The spatial positions of these two objects were randomized across trials. The observers had to click on the object they considered to be old.

Data analysis

The percentage of correct recognition was calculated. In other aspects, the data analysis was similar to Experiment 1.

Results and discussion

We found no significant differences between the three conditions (4 categories: $M = 75.3\%$; 2 categories: $M = 76.2\%$; 1 category: $M = 76.2\%$; comparison: $F(2,34) = .251$, $p = .779$, $\eta^2 = .015$, $BF_{10} = .169$). This result replicates the finding of equal recognition accuracy in all distinctiveness conditions (Experiment 1). It also suggests that the effect of distinctiveness on localization accuracy found in Experiment 3A is not caused by object recognition decrement. Together with Experiment 3A, Experiment 3B also suggests that distinctiveness affected object-location reports, but not object recognition.

One might argue that the difference between the effects of distinctiveness on object recognition (Experiment 3B) and object-location memory (Experiment 3A) could be due to the difference in the ways distinctiveness was manipulated at tests. Indeed, at the recognition test, a foil item was always an item from the same category as a target, which was necessary to keep test difficulty fixed (Awh et al., 2007). In contrast, object-location memory was tested using only items from a sample set as test options (otherwise, the object-location task would have been turned into another version of the recognition task, when observers would have just chosen familiar objects instead of trying to recall where these objects belonged to). As a consequence, the options in the object-location test inherited the distinctiveness from the sample set. However, there is an important argument against the idea that the differential effects of distinctiveness were due to the difference between the distinctiveness manipulations used in the two tasks. The basic pattern replicates the pattern observed in Experiment 1, which we consider to be condition-invariant for both object and object-location memories. To remind, in Experiment 1 observers were tested for object memory using the same 2-AFC task as in Experiment 3B. As for object-location memory, the observers had to locate a single item remaining on the screen after they had performed the 2-AFC; so, there were no other test options that could provide contextual differences in terms of distinctiveness.

General Discussion

Our main goal was to test the effects of distinctiveness on object recognition and object-location memory. The question of principal interest was how the distinctiveness of real-world objects stored in VWM affects the ability to recognize the objects, remember locations and report object-location conjunctions, and whether these effects are similar. To this end, we tested recognition memory for high-distinctive vs. low-distinctive objects in Experiments 1 and 3B. Our results from both experiments showed equally good performances regardless of the distinctiveness. It suggests that our participants had reasonably good visual memory for different objects, even when they belonged to the same category and therefore were more similar and potentially more mutually interfering (Cohen et al., 2014; Konkle et al., 2010). To remind, the previous research with the same stimulus set has found a detrimental, though the not dramatic effect of within-category similarity on recognition in massive long-term memory (Konkle et al., 2010). Experiments 1 and 2 showed that spatial memory per se was good in all distinctiveness conditions, as shown by mixture modeling of localization errors (Bays et al., 2009; Zhang & Luck, 2008). From consistently very low P_{guess}, we conclude that observers had no substantial problem with storing all three locations (Experiment 1), which is in line with the previous estimates of spatial VWM (Postma & De Haan, 1996) and VWM in general (Alvarez & Cavanagh, 2004; Cowan, 2001; Luck & Vogel, 1997). More importantly, we found no evidence that spatial memory suffered from the need to store categorically similar objects compared to distinct objects, as the P_{guess} did not depend on object distinctiveness. This suggests that the requirement to store less distinct objects in VWM did not cause more location forgetting. Similarly, we found practically no effect of object distinctiveness on the precision of memory for locations. Experiment 2 additionally showed that the representations of locations were not strongly affected by the need to remember object identities and their relative positions ("bindings"). This pattern is basically in line with a claim that object memory and location memory have separate capacities and appear to be independent, as has been shown in a number of previous studies (Lee & Chun, 2001; Li et al., 2015; Wood, 2011). However, although there was no effect of object distinctiveness on object recognition and location memory, object-location memory was affected by distinctiveness (Experiments 1 and 2), as we found more swap errors when the objects were low-distinctive.

To account for the differential effects of object distinctiveness on the different aspects of tested memories, we can turn to existing models of recognition and binding in VWM, that rely on the idea of noisy representations or noisy familiarity judgments in continuous feature spaces (e.g., Oberauer & Lin, 2017; Schneegans & Bays, 2017; Swan & Wyble, 2014; Schurgin, Wixted, & Brady, 2020). From this perspective, the differential effects of distinctiveness on object recognition and object-location swaps can reflect some important differences in the accessibility and discriminability of object or location representations depending on the task. Since simple object recognition in a 2-AFC requires only a familiarity judgment (which of the alternatives looks more familiar) it should naturally depend on target-foil distinctiveness (Awh et al., 2007; Schurgin et al., 2020) that we objectively kept fixed across conditions, as we always used foils from the same category as a target. That is, the familiarity of the target compared to the foil was about the same. We should note, however, that when all studied items belong to the same category (low distinctiveness) the subjective target-foil distinctiveness still can decrease in theory (for example, by increasing the familiarity of the whole category including the foil), causing more false alarms to foils (as in the DRM effect). We did not observe this in our VWM recognition task. This finding is different from the existent (though not dramatic) distinctiveness effect on LTM for the same stimulus set (Konkle et al., 2010). This interesting difference between distinctiveness effects on two memory systems can be a subject of further research.

While object recognition is not affected by the distinctiveness due to the fixed familiarity ratio between the target and the foil, object-location memory is tested using a cued recall task where one of the studied "items" (object as in Experiments 1 and 2 or location as in Experiment 3A) is presented as a cue and another one is to be reported. Here, both the cues and the to-be-reported items are familiar. Hence, the critical discrimination here is not between more familiar and less familiar features but between equally familiar features that have to be correctly linked with the cued feature. Here, target distinctiveness plays a greater role. The probability that a certain item (target or nontarget) is recalled will depend on the similarity between the cued and uncued features and/or similarity between the items these cues address. This idea is most clearly illustrated by the two-category condition from Experiment 3A (Fig. 7b). When observers memorize four items from categories A and B in four locations, and then one of the items from category A is tested for its location memory, the observers will more frequently misreport the location of another exemplar from category A because this location is associated with an item that is more similar to the cue. Figure 8 depicts this as probabilistic familiarity judgments using a signal detection model (Macmillan & Creelman, 2005; Schurgin et al., 2020). Here, a location cue (Fig. 8A) makes each item in a 4-AFC array produce a familiarity signal randomly drawn from a normal distribution whose mean is defined by the reliability of object-location binding. If this binding is reliable, then the target item distribution is the rightmost (having the strongest familiarity on average, Fig. 8b). Target-nontarget similarity defines how much nontarget distributions are shifted to the left relative to the target (that we can term $\Delta d`$ which is the difference between the means of the target and the nontarget distributions measured in the units of standard deviation). Obviously, this shift is greater if the nontarget is highly distinctive from the target. On each trial, the observer chooses an item that produced the highest familiarity signal. The overlap between the distributions predicts that there will be trials when one or several distractors produce stronger signals than the target, in which case an observer will commit a swap error (Fig. 8c). The proportion of swaps depends on the distance between the distributions – therefore, it will be greater for more similar items. Our model fits showed that the percentages of swaps we observed in this condition of Experiment 3A (17.6% for the same category and 6.95% for the different category) are accomplished under Δd` = 1.08 and 1.67, respectively (Fig. 8d). The same logic can be applied to other conditions, that is when all objects are from the same or different categories.

Existing quantitative models of binding in VWM also predict that feature similarity and distinctiveness should affect performance in an object-location task (Oberauer & Lin, 2017; Schneegans & Bays, 2017; Swan & Wyble, 2014). However, as these models are mostly built to account for recall of simple continuous features (such as continuous color reports based on item location), they should be applied to our data with meaningful objects with caution. Indeed, our data confirm some of the predictions from these models but are at odds with others. Specifically, the models predict that cue similarity increases competition between associated items, which should result in a greater proportion of swaps. This is exactly what was observed in Experiments 1 and 2 when the objects were used as cues and locations were to be reported. On the other hand, the models predict that similarity between to-be-reported features should decrease interference between them (Oberauer & Lin, 2017; Swan & Wyble, 2014). This seems to be not the case for our data, especially in Experiment 3A, when more similar to-be-reported objects yielded more swaps. To remind, the discrepancy between the directions of distinctiveness effects on simple visual features and semantically meaningful real-world objects has been observed in the previous literature (e.g., Jiang et al., 2016b). A possible explanation for this discrepancy is continuous vs. discrete nature of the objects (Jiang et al., 2016b, consider this to be one of the crucial factors mediating the direction of distinctiveness effects). In our Experiment 3, targets and nontargets were discrete items, so reporting one instead of another was unambiguously interpreted as swaps. In contrast, the binding models (Oberauer & Lin, 2017; Swan & Wyble, 2014) make their predictions for continuous features for which error distributions can be built, and their precision (SD) can be estimated with the mixture model. If targets and nontargets are similar and both influence current responses, observers usually show sharper error distributions compared to dissimilar targets and nontargets. This occurs because a similar nontarget will not pull the response away from the correct one as strongly as a distinct target. Moreover, when the target and nontarget are similar, it is harder to decompose a corresponding error distribution into correct answers and swaps (Bays et al., 2009). Therefore, the discrepancy between our data and

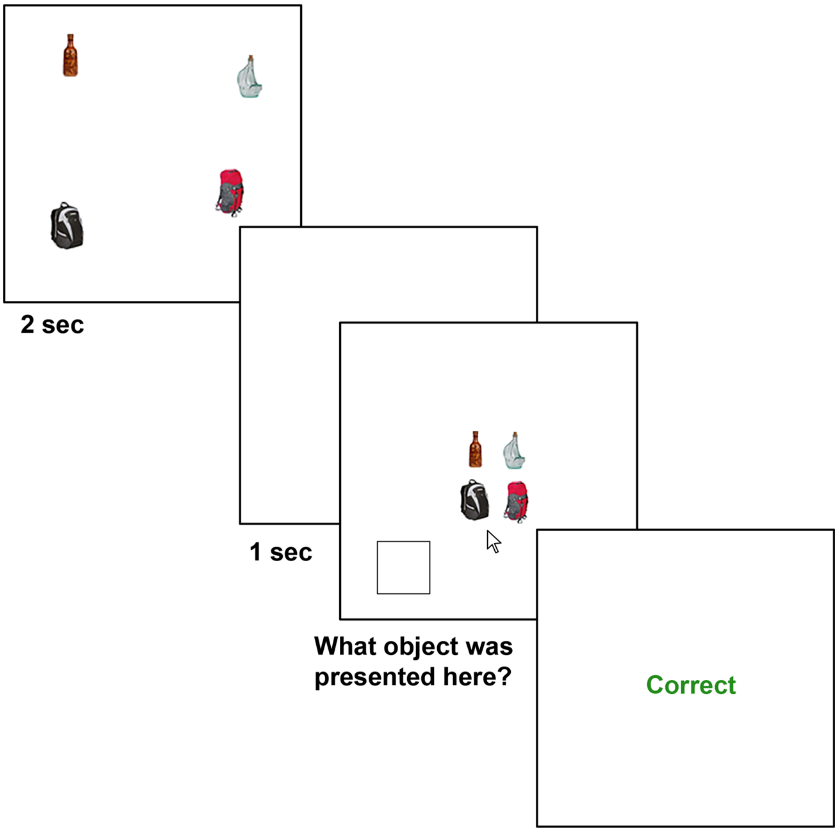


Fig. 6 The time course of a trial in Experiment 3A

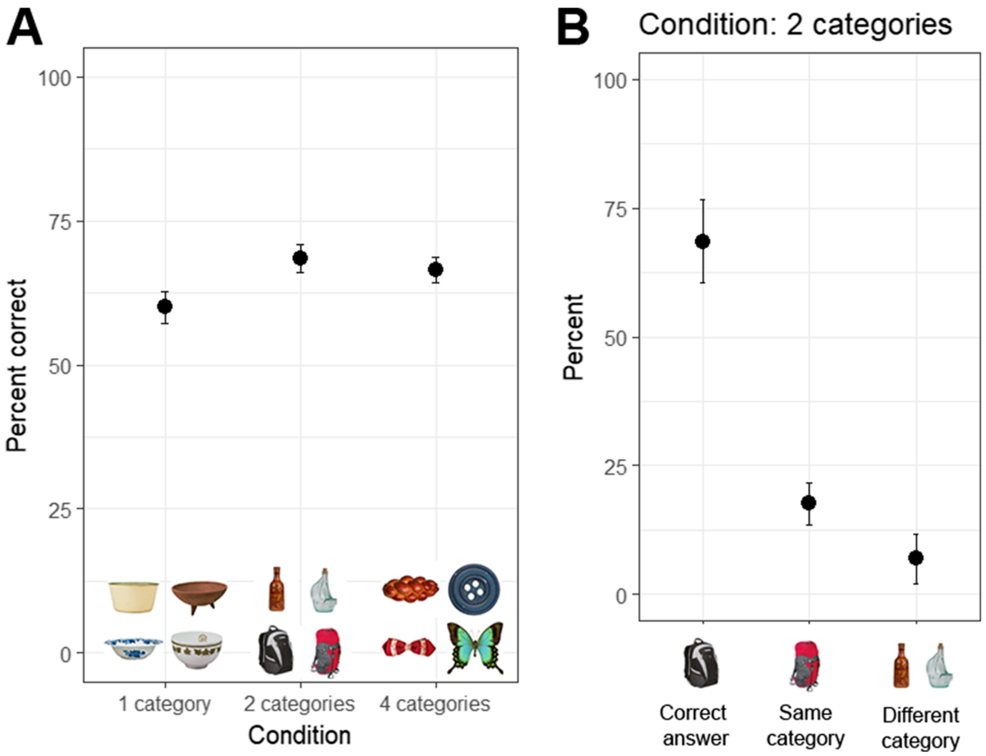


Fig. 7 The results of Experiment 3A: percentage of correct object localizations for three conditions (**a**); percent of correct answers for the condition with 2 categories (**b**)

predictions of the existing models can reflect a difference between approaches to data analysis and interpretation, whereas the true direction of the effect can be the same. Future research can focus on further figuring out what other mechanisms can cause the difference between the simple features and real-world objects in terms of distinctiveness effects. This also poses a request to the existing quantitative binding models for a unified account of VWM for simple features and complex objects.

Up to this point, we discussed the effects of object distinctiveness on object and object-location memory in terms of a single recognition mechanism that takes into account only quantitative differences in familiarity produced by targets and foils. However, since our stimuli were meaningful real-world objects, there is a possibility that object-location memories could be affected by the use of specific encoding and/or retrieval strategies working on the conceptual level. We suggest that when objects from different categories are presented, observers can rely on coarse category-location knowledge even when they fail to rely on the precise object-location knowledge. For example, looking at Fig. 6, observers can remember which two out of four locations contained backpacks even if they fail to remember which particular backpack was in which of these two "backpack locations". As a result, when one of these "backpack locations" is probed, the observers would choose a random backpack (sometimes it will be a correct answer, and sometimes it will be a swap) more often than a random bottle. This will cause within-category swaps more often than between-category swaps. Overall, this coarse knowledge of category-location associations can be useful in all cases when different categories are shown in different places, which provides an advantage to object-location reports in all high-distinctiveness conditions of our experiments. Noteworthy, this advantage of different categories does not necessarily mean that observers verbally label locations with category names, especially given the articulatory suppression task we employed in our experiments to prevent observers from verbal labeling. This can be more abstract, conceptual labeling less dependent on encoding modality. This is an intriguing possibility that can be tested in future research (for example, comparing effects of different categories with and without articulatory suppression, as in Dent & Smyth, 2005; Postma & De Haan, 1996).

In sum, in our experiments, we observed that simple object recognition in VWM does not suffer from the low distinctiveness of studied objects (although previous research has shown the opposite for LTM – e.g., Konkle et al., 2010). However, distinctiveness did have an effect on object-location memory, so more swaps occurred when the objects were more similar. These differential effects can be indicative of important differences between two ways of access to the contents of VWM. In simple recognition, the observer decides which items look more familiar (old) and which items look less familiar (new), whereas in an object location-task, the observer chooses which of equally familiar representations better matches a retrieval cue (which is the essence of binding). The object-location retrieval, therefore, involves more competition between representations which, as we assume, is mediated by distinctiveness (Oberauer & Lin, 2017; Schneegans & Bays, 2017; Schurgin et al., 2020; Swan & Wyble, 2014). In addition, the nature of distinctiveness in objects we tested keeps a possibility of using categorical labeling of locations to maintain coarse object-location information even when finer information about a particular object at a particular location fails.

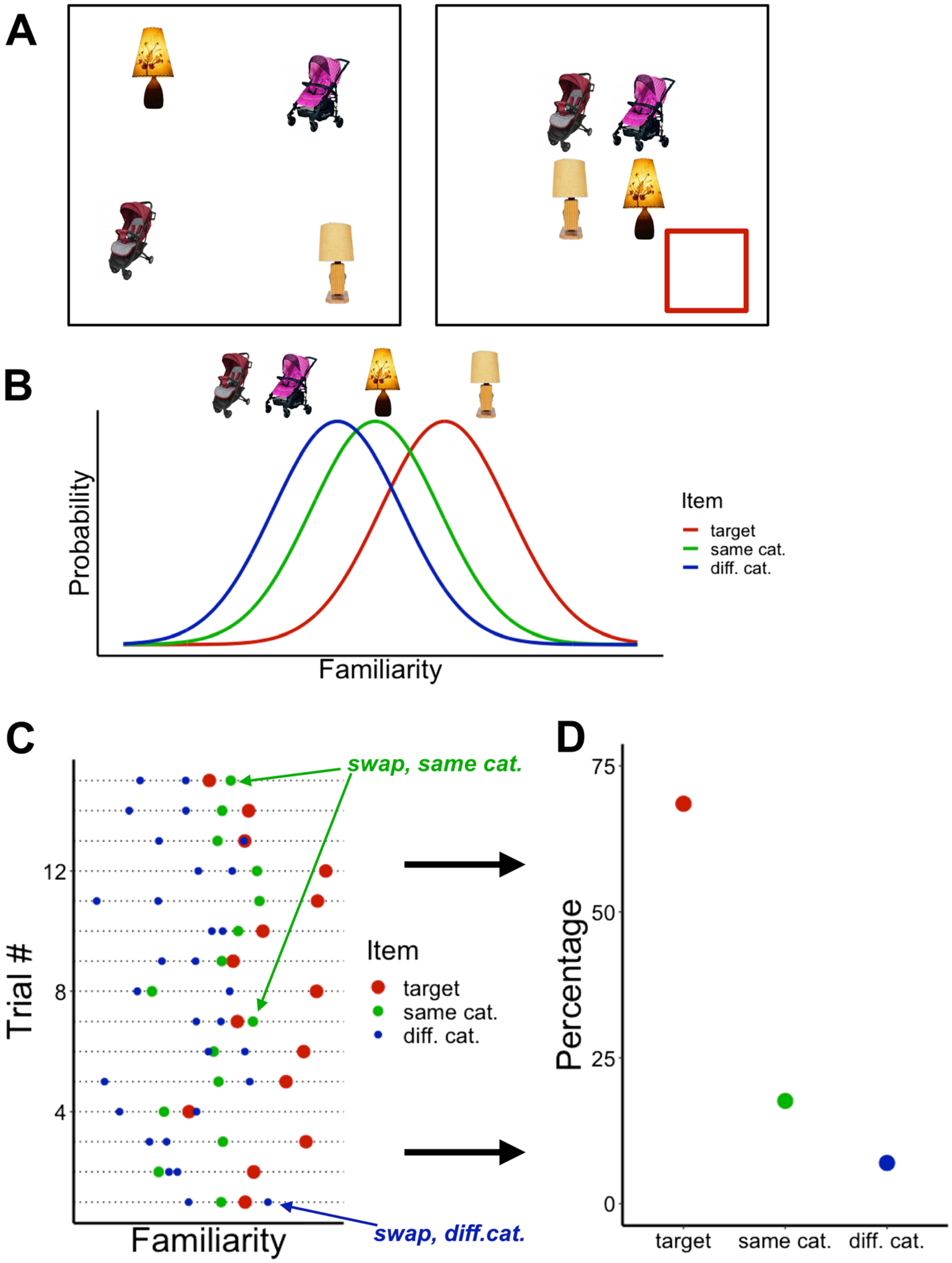


Fig. 8 Object-location report as a noisy familiarity judgment. **a** An example two-category trial from Experiment 3A. **b** A signal-detection model of object retrieval given the location cue from (A). Each distribution corresponds to overall internal representations of the four tested objects along the familiarity axis. Relative shifts between the distributions ($\Delta d'$) are a function of similarity between an item and a target. **c** In each trial, each object produces a random familiarity value from a corresponding distribution. The item producing the maximum value is chosen as an answer. If the target produces the maximum familiarity, then the response counts as correct, otherwise the response counts as swap. **d** Best-fit predictions of different response outcomes from the two-category condition of Experiment 3A.

Further research is required to figure out the potential role of conceptual encoding of categories in object-location memory.

Acknowledgements The study was implemented within the framework of the Basic Research Program at the National Research University Higher School of Economics in 2020 (RS-60). The authors thank two anonymous reviewers for their helpful comments on earlier drafts of the article.

Author contributions Y.A.M. designed the experiments, prepared the experimental scripts, collected and analyzed data, and wrote the manuscript. I.S.U. conceptualized the basic ideas, designed the experiments, and wrote the manuscript.

References

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science: A Journal of the American Psychological Society / APS*, *15*(2), 106–111. https://doi.org/10.1167/2.7.273

Asp, I. E., Störmer, V. S., & Brady, T. F. (2021). Greater Visual Working Memory Capacity for Visually Matched Stimuli When They Are Perceived as Meaningful. *Journal of Cognitive Neuroscience, 33*(5), 902–918. https://doi.org/10.1162/jocn_a_01693

Avital-Cohen, R., & Gronau, N. (2021). The asymmetric mixed-category advantage in visual working memory: An attentional, not perceptual (face-specific) account. *Journal of Experimental Psychology. Human Perception and Performance, 47*(6), 852–868. https://doi.org/10.1037/xhp0000921

Avons, S. E., & Mason, A. (1999). Effects of visual similarity on serial report and item recognition. *The Quarterly Journal of Experimental Psychology*, (1), 217–240.

Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science, 18*(7), 622–628. https://doi.org/10.1111/j.1467-9280.2007.01949.x

Baddeley, A. D. (1966a) Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quarterly Journal of Experimental Psychology*, 18, 362-365

Baddeley, A. D. (1966b) The influence of acoustic and semantic similarity on long-term memory for word sequences. *Quarterly Journal of Experimental Psychology*, 18, 302-309.

Baddeley, A. D. (1986). Working memory. Clarendon Press.

Baddeley, A. D. & Hitch, G. J. (1974). Working memory. In G. A. Bower (ed). The Psychology of Learning and Motivation: Advances in Research and Theory*,* pp. 47–89. Academic.

Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, *9*(10), 7–7. https://doi.org/10.1167/9.10.7

Brady, T. F., & Störmer, V. S. (2020). Comparing memory capacity across stimuli requires maximally dissimilar foils: Using deep convolutional neural networks to understand visual working memory capacity for real-world objects. PsyArXiv https://psyarxiv.com/25t76/

Brady, T. F., & Störmer, V. S. (2021). The role of meaning in visual working memory: Real-world objects, but not simple features, benefit from deeper processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. https://doi.org/10.1037/xlm0001014

Brady, T. F., Konkle, T., & Alvarez, G. A. (2011). A review of visual memory capacity: Beyond individual items and toward structured representations. *Journal of Vision*, *11*(5), 4. https://doi.org/10.1167/11.5.4

Brady, T. F., Störmer, V. S., & Alvarez, G. A. (2016). Working memory is not fixed-capacity: More active storage capacity for real-world objects than for simple stimuli. *Proceedings of the National Academy of Sciences*, *113*(27), 7459–7464. https://doi.org/10.1073/pnas.1520027113

Cohen, M. A., Konkle, T., Rhee, J. Y., Nakayama, K., & Alvarez, G. A. (2014). Processing multiple visual objects is limited by overlap in neural channels. *Proceedings of the National Academy of Sciences*, *111*(24), 8955–8960. https://doi.org/10.1073/pnas.1317860111

Cohen-Dallal, H., Soroker, N., & Pertzov, Y. (2021). Working Memory in Unilateral Spatial Neglect: Evidence for Impaired Binding of Object Identity and Object Location. *Journal of Cognitive Neuroscience, 33*(1), 46–62. https://doi.org/10.1162/jocn_a_01631

Conrad, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology, 55*(1), 75–84.

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *The Behavioral and Brain Sciences*, *24*(1), 87–185. https://doi.org/10.1017/S0140525X01003922

Darling, S., Della Sala, S., Logie, R. H., & Cantagallo, A. (2006). Neuropsychological evidence for separating components of visuo–spatial working memory. *Journal of Neurology*, *253*(2), 176–180. https://doi.org/10.1007/s00415-005-0944-3

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*(1), 17–22. https://doi.org/10.1037/h0046671

Dent, K., & Smyth, M. M. (2005). Verbal coding and the storage of form-position associations in visual–spatial short-term memory. *Acta Psychologica, 120*(2), 113–140. https://doi.org/10.1016/J.ACTPSY.2005.03.004

Donkin, C., Tran, S. C., & Le Pelley, M. (2015). Location-based errors in change detection: A challenge for the slots model of visual working memory. *Memory & Cognition, 43*(3), 421–431. https://doi.org/10.3758/s13421-014-0487-x

Emrich, S. M., & Ferber, S. (2012). Competition increases binding errors in visual working memory. *Journal of Vision, 12*(4), 12–12. https://doi.org/10.1167/12.4.12

Fougnie, D., & Alvarez, G. A. (2011). Object features fail independently in visual working memory: Evidence for a probabilistic feature-store model. *Journal of Vision*, *11*(12), 3. https://doi.org/10.1167/11.12.3

Fougnie, D., Asplund, C. L., & Marois, R. (2010). What are the units of storage in visual working memory? *Journal of Vision*, *10*(12), 27. https://doi.org/10.1167/10.12.27

Hall, E. H., Bainbridge, W. A., Baker, C. I. (2021). Highly similar and competing visual scenes lead to diminished object but not spatial detail in memory drawings. Memory. https://doi.org/10.1080/09658211.2021.2010761

Haxby, J. V., Grady, C. L., Horwitz, B., Ungerleider, L. G., Mishkin, M., Carson, R. E., … Rapoport, S. I. (1991). Dissociation of object and spatial visual processing pathways in human extrastriate cortex. *Proceedings of the National Academy of Sciences*, *88*(5), 1621–1625. https://doi.org/10.1073/pnas.88.5.1621

Hollingworth, A., & Rasmussen, I. P. (2010). Binding objects to locations: The relationship between object files and visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 36*(3), 543–564. https://doi.org/10.1037/a0017836

Hovhannisyan, M., Clarke, A., Geib, B. R., Cicchinelli, R., Monge, Z., Worth, T., … Davis, S. W. (2021). The visual and semantic features that predict object memory: Concept property norms for 1,000 object images. *Memory and Cognition*, *49*(4), 712–731. https://doi.org/10.3758/s13421-020-01130-5

Hu, S., Liu, D., Song, F., Wang, Y., & Zhao, J. (2020). The influence of object similarity on real object-based attention: The disassociation of perceptual and semantic similarity. *Acta Psychologica*, *205*,103046. https://doi.org/10.1016/j.actpsy.2020.103046

Hunt, R. Reed. (1995). The subtlety of distinctiveness: What von Restorff really did. *Psychonomic Bulletin & Review, 2*(1), 105–112. https://doi.org/10.3758/BF03214414

Hunt, R. R. (2006). The concept of distinctiveness in memory research. In Distinctiveness and memory (p. 476). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195169669.003.0001

Jalbert, A., Saint-Aubin, J., & Tremblay, S. (2008). Visual similarity in short-term recall for where and when. *Quarterly Journal of Experimental Psychology (2006)*, *61*(3), 353–360. https://doi.org/10.1080/17470210701634537

JASP Team (2017). JASP (Version 0.8.2) [Computer software]. https://jasp-stats.org/faq/how-do-i-cite-jasp/

Jiang, Y. V., Lee, H. J., Asaad, A., & Remington, R. (2016a). Similarity effects in visual working memory. *Psychonomic Bulletin & Review*, *23*(2), 476–482. https://doi.org/10.3758/s13423-015-0905-5

Jiang, Y. V., Remington, R. W., Asaad, A., Lee, H. J., & Mikkalson, T. C. (2016b). Remembering faces and scenes: The mixed-category advantage in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 42*(9), 1399–1411. https://doi.org/10.1037/xhp0000228

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. *Journal of Experimental Psychology: General*, *139*(3), 558–578. https://doi.org/10.1037/a0019165

Lee, D., & Chun, M. M. (2001). What are the units of visual short-term memory, objects or spatial locations? *Perception & Psychophysics*, *63*(2), 253–7. https://doi.org/10.3758/BF03194466

Li, J., Zhou, Y., Shui, R., & Shen, M. (2015). Visual working memory for dynamic objects: Impaired binding between object feature and location. *Visual Cognition*, *23*(3), 357–378. https://doi.org/10.1080/13506285.2014.1001010

Liang, Y., Pertzov, Y., Nicholas, J. M., Henley, S. M. D., Crutch, S., Woodward, F., … Husain, M. (2016). Visual short-term memory binding deficit in familial Alzheimer's disease. *Cortex, 78*, 150–164. https://doi.org/10.1016/j.cortex.2016.01.015

Lin, P.-H., & Luck, S. J. (2009). The influence of similarity on visual working memory representations. *Visual Cognition*, *17*(3), 356–372. https://doi.org/10.1080/13506280701766313

Liu, J., Zhang, H., Yu, T., Ni, D., Ren, L., Yang, Q., … Xue, G. (2020). Stable maintenance of multiple representational formats in human visual short-term memory. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(51), 32329–32339. https://doi.org/10.1073/pnas.2006752117

Logie, R. H. (2003). Spatial and Visual Working Memory: A Mental Workspace. *Psychology of Learning and Motivation - Advances in Research and Theory, 42*, 37–78. https://doi.org/10.1016/S0079-7421(03)01002-8

Lu, K., Nicholas, J. M., Pertzov, Y., Grogan, J., Husain, M., Pavisic, I. M., … Schott, J. M. (2020). APOE-ε4 carriers have superior recall on the 'What was where?' visual short-term memory binding test at age 70, despite a detrimental effect of β-amyloid. *Alzheimer's & Dementia, 16*(S6),e041090. https://doi.org/10.1002/alz.041090

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–81. https://doi.org/10.1038/36846

Macmillan N. A., & Creelman C. D. (2005). Detection theory: A user's guide, 2nd ed. Lawrence Erlbaum Associates Publishers.

Magen, H., & Emmanouil, T. A. (2019). Spatial Organization in Self-Initiated Visual Working Memory. *Frontiers in Psychology, 10*, 2734. https://doi.org/10.3389/fpsyg.2019.02734

Markov, Y. A., Utochkin I. S., & Brady T. F. (2021) Real-world objects are not stored in holistic representations in visual working memory. Journal of Vision, 21(3):18, 1-24.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, *101*(2), 343–352. https://doi.org/10.1037/h0043158

Mishkin, M., & Ungerleider, L. G. (1982). Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. *Behavioural Brain Research*, *6*(1), 57–77. https://doi.org/10.1016/0166-4328(82)90081-X

Naspi, L., Hoffman, P., Devereux, B., Thejll-Madsen, T., Doumas, L., & Morcom, A. (2021). Multiple dimensions of semantic and perceptual similarity contribute to mnemonic discrimination for pictures. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 47*(12), 1903–1923. https://doi.org/10.1037/xlm0001032

Oberauer, K., & Lin, H. (2017). An Interference Model of Visual Working Memory. *Psychological Review, 124*(1), 1–39. https://doi.org/10.1037/rev0000044

Otsuka, S., Nishiyama, M., Nakahara, F., & Kawaguchi, J. (2013). Visual statistical learning based on the perceptual and semantic information of objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(1), 196–207. https://doi.org/10.1037/a0028645

Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics, 44*(4), 369–378. https://doi.org/10.3758/BF03210419

Pavisic, I. M., Pertzov, Y., O'Connor, A., Nicholas, J. M., Macpherson, K., Collins, J., … Crutch, S. J. (2018). P1-524: VISUAL SHORT-TERM BINDING DEFICIT IN FAMILIAL ALZHEIMER'S DISEASE: A LONGITUDINAL STUDY. *Alzheimer's & Dementia, 14*(7S_Part_9), P532–P533. https://doi.org/10.1016/j.jalz.2018.06.535

Pavisic, I. M., Suarez-Gonzalez, A., & Pertzov, Y. (2020). Translating Visual Short-Term Memory Binding Tasks to Clinical Practice: From Theory to Practice. *Frontiers in Neurology, 11*,458. https://doi.org/10.3389/fneur.2020.00458

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

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., Lindeløv, J. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods 51(1): 195-203.*https://doi.org/10.3758/s13428-018-01193-y

Pertzov, Y., Dong, M. Y., Peich, M. C., & Husain, M. (2012). Forgetting what was where: the fragility of object-location binding. *PLoS ONE*, *7*(10),e48214. https://doi.org/10.1371/journal.pone.0048214

Pertzov, Y., Miller, T. D., Gorgoraptis, N., Caine, D., Schott, J. M., Butler, C., & Husain, M. (2013). Binding deficits in memory following medial temporal lobe damage in patients with voltage-gated potassium channel complex antibody-associated limbic encephalitis. *Brain, 136*(8), 2474–2485. https://doi.org/10.1093/brain/awt129

Pertzov, Y., Heider, M., Liang, Y., & Husain, M. (2015). Effects of healthy ageing on precision and binding of object location in visual short term memory. *Psychology and Aging, 30*(1), 26–35. https://doi.org/10.1037/a0038396

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics, 16*(2), 283–290. https://doi.org/10.3758/BF03203943

Postma, A., & De Haan, E. H. F. (1996). What Was Where? Memory for Object Locations. *The Quarterly Journal of Experimental Psychology Section A, 49*(1), 178–199. https://doi.org/10.1080/713755605

Postma, A., Kessels, R. P. C., & van Asselen, M. (2008). How the brain remembers and forgets where things are: The neurocognition of object-location memory. *Neuroscience and Biobehavioral Reviews*, *32*(8), 1339–1345. https://doi.org/10.1016/j.neubiorev.2008.05.001

Pratte, M. S. (2019). Swap errors in spatial working memory are guesses. *Psychonomic Bulletin & Review, 26*(3), 958–966. https://doi.org/10.3758/s13423-018-1524-8

Roediger, H. L., & McDermott. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(4), 803–814.

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(2), 225-237. doi:https://doi.org/10.3758/PBR.16.2.225.

Schneegans, S., & Bays, P. M. (2017). Neural Architecture for Feature Binding in Visual Working Memory. *The Journal of Neuroscience, 37*(14), 3913–3925. https://doi.org/10.1523/JNEUROSCI.3493-16.2017

Schurgin, M. W., Wixted, J. T., & Brady, T. F. (2020). Psychophysical scaling reveals a unified theory of visual memory strength. *Nature Human Behaviour, 4*(11), 1156–1172. https://doi.org/10.1038/s41562-020-00938-0

Sims, C. R., Jacobs, R. A., & Knill, D. C. (2012). An ideal observer analysis of visual working memory. *Psychological Review*, *119*(4), 807–30. https://doi.org/10.1037/a0029856

Smith, E. E., Jonides, J., Koeppe, R. A., Awh, E., Schumacher, E. H., & Minoshima, S. (1995). Spatial versus Object Working Memory: PET Investigations. *Journal of Cognitive Neuroscience, 7*(3), 337–356. https://doi.org/10.1162/jocn.1995.7.3.337

Standing, L. (1973). Learning 10,000 pictures. Quarterly Journal of Experimental Psychology, 25(2), 207–222. https://doi.org/10.1080/14640747308400340

Suchow, J. W., Brady, T. F., Fougnie, D., & Alvarez, G. A. (2013). Modeling visual working memory with the MemToolbox. *Journal of Vision*, *13*(10), 9. https://doi.org/10.1167/13.10.9

Suchow, J. W., Fougnie, D., Brady, T. F., & Alvarez, G. A. (2014). Terms of the debate on the format and structure of visual memory. *Attention, Perception, & Psychophysics, 76*(7), 2071–2079. https://doi.org/10.3758/s13414-014-0690-7

Swan, G., & Wyble, B. (2014). The binding pool: A model of shared neural resources for distinct items in visual working memory. *Attention, Perception, & Psychophysics, 76*(7), 2136–2157. https://doi.org/10.3758/s13414-014-0633-3

Toh, Y. N., Sisk, C. A., & Jiang, Y. V. (2020). Effects of changing object identity on location working memory. *Attention, Perception, & Psychophysics, 82*(1), 294–311. https://doi.org/10.3758/s13414-019-01738-z

Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology, 6*(2), 171–178. https://doi.org/10.1016/S0959-4388(96)80070-5

Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Selker, R., Gronau, Q. F., Dropmann, D., Boutin, B., Meerhoff, F., Knight, P., Raj, A., van Kesteren, E.-J., van Doorn, J., Šmíra, M., Epskamp, S., Etz, A., Matzke, D., de Jong, T., van den Bergh, D., Sarafoglou, A., Steingroever, H., Derks, K., Rouder, J. N., Morey, R. D. (2017). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review*, 1–19. https://doi.org/10.3758/s13423-017-1323-7

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision*, *4*(12), 11–11. https://doi.org/10.1167/4.12.11

Wood, J. N. (2011). When do spatial and visual working memory interact? *Attention, Perception, & Psychophysics*, *73*(2), 420–439. https://doi.org/10.3758/s13414-010-0048-8

Xie, W., & Zaghloul, K. A. (2021). Visual and Semantic Contributions to Visual Short-Term Memory. *Trends in Cognitive Sciences*, *25*(4), 270–271. https://doi.org/10.1016/j.tics.2021.01.010

Yang, L., & Mo, L. (2017). The Effects of Similarity on High-Level Visual Working Memory Processing. *Advances in Cognitive Psychology, 13*(4), 296–305. https://doi.org/10.5709/acp-0229-8

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233-U13. https://doi.org/10.1038/Nature06860

Zokaei, N., Giehl, K., Sillence, A., Neville, M. J., Karpe, F., Nobre, A. C., & Husain, M. (2017). Sex and APOE: A memory advantage in male APOE ε4 carriers in midlife. *Cortex, 88*, 98–105. https://doi.org/10.1016/j.cortex.2016.12.016