Journal of Experimental Psychology: Learning, Memory, and Cognition

The Role of Meaning in Visual Working Memory: Real-World Objects, But Not Simple Features, Benefit From Deeper Processing

Timothy F. Brady and Viola S. Störmer

Online First Publication, March 25, 2021. http://dx.doi.org/10.1037/xlm0001014

CITATION

Brady, T. F., & Störmer, V. S. (2021, March 25). 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*. Advance online publication. http://dx.doi.org/10.1037/xlm0001014

AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology:
Learning, Memory, and Cognition

ISSN: 0278-7393

https://doi.org/10.1037/xlm0001014

The Role of Meaning in Visual Working Memory: Real-World Objects, But Not Simple Features, Benefit From Deeper Processing

Timothy F. Brady[1] and Viola S. Störmer[2]
[1] Department of Psychology, University of California, San Diego
[2] Department of Psychological and Brain Sciences, Dartmouth College

Visual working memory is a capacity-limited cognitive system used to actively store and manipulate visual information. Visual working memory capacity is not fixed, but varies by stimulus type: Stimuli that are more meaningful are better remembered. In the current work, we investigate what conditions lead to the strongest benefits for meaningful stimuli. We propose that in some situations participants may try to encode the entire display holistically (i.e., in a quick "snapshot"). This may lead them to treat objects as simply meaningless, colored "blobs", rather than individually and in a high-level way, which could reduce benefits of meaningful stimuli. In a series of experiments, we directly test whether real-world objects, colors, perceptually matched less-meaningful objects, and fully scrambled objects benefit from deeper processing. We systematically vary the presentation format of stimuli at encoding to be either simultaneous—encouraging a parallel, "take-a-quick-snapshot" strategy—or present the stimuli sequentially, promoting a serial, each-item-at-once strategy. We find large advantages for meaningful objects in all conditions, but find that real-world objects—and to a lesser degree lightly scrambled, still meaningful versions of the objects—benefit from the sequential encoding and thus deeper, focused-on-individual-items processing, while colors do not. Our results suggest single-feature objects may be an outlier in their affordance of parallel, quick processing, and that in more realistic memory situations, visual working memory likely relies upon representations resulting from in-depth processing of objects (e.g., in higher-level visual areas) rather than solely being represented in terms of their low-level features.

Keywords: visual working memory, visual long-term memory, memory encoding, depth of processing

Visual working memory is a capacity-limited cognitive system used to actively store and manipulate visual information (Baddeley, 2012; Cowan, 2001). While theories generally agree that its capacity is limited, they differ in terms of what these limits are and how they arise. Prominent models of working memory have promoted the view of a "fixed" limit of working memory, arguing that a particular number of objects can be stored at once, regardless of what these objects are (e.g., Awh et al., 2007; Luck & Vogel, 1997), or that a fixed amount of resources can be distributed among the to-be-remembered stimuli (e.g., Bays et al., 2009) within each of a small number of feature dimensions (i.e., color, orientation). Support for these strong fixed-capacity models comes from numerous studies examining visual working memory limits using simple stimuli like colored squares, oriented lines, or novel shapes (e.g., Zhang & Luck, 2008), all stimuli about which participants have minimal background knowledge or expectations. These simple, meaningless stimuli are often assumed to best assess the core capacity of working memory because they have no semantic associations and are repeated from trial to trial, which minimizes participants' ability to use other memory systems, like episodic visual long-term memory, to support memory performance (Cowan, 2001; Lin & Luck, 2012).

Using such simple stimuli, past studies have often used short encoding times (generally < 500 ms), assuming that working memory fills up quickly, and have argued that the stark limits on performance are truly limits of working memory—that is, the limits on performance do not arise from limited encoding times or limits in perceptual processing (e.g., Alvarez & Cavanagh, 2004; Bays et al., 2011; Luck & Vogel, 1997; Tsubomi et al., 2013; Vogel et al., 2006). For example, Luck and Vogel (1997) argued that the same capacity limits appeared regardless of encoding time, which they said meant such limits arose from "limitations in storage capacity rather than limitations in perceiving or encoding the stimuli" (p. 279). Similarly, Alvarez and Cavanagh (2004) claimed that their results were from a storage limit, not an encoding limit, because "all of the information that can be stored is acquired in less than 500 ms" (p. 109).

In stark contrast to these findings of fixed performance regardless of encoding time in simple stimuli, we recently found working memory performance to be higher for real-world objects than for simple stimuli, particularly when participants were given a longer

---

Timothy F. Brady https://orcid.org/0000-0001-5924-5211

This work was supported by an NSF grant (BCS-1829434) to Timothy F. Brady and Viola S. Störmer. We thank Kevin Sayed for help with data collection.

Correspondence concerning this article should be addressed to Timothy F. Brady, Department of Psychology, University of California, San Diego, 9500 Gilman Drive #0109, McGill Hall 5322, La Jolla, CA 92093, United States. Email: timbrady@ucsd.edu

time to encode these items (Brady et al., 2016). In that study, participants were asked to remember either objects or colors over a short delay, and discriminate one of the remembered stimuli in a two-alternative forced choice against a maximally distinct foil object or color. Specifically, at long encoding times (1s and 2s) participants better remembered real-world objects than colors. What drives these differences in capacity? One possible explanation is that working memory operates equally well on both stimulus types, but real-world objects can additionally benefit from the high-capacity episodic long-term memory system or a form of more accessible long-term memories (Cowan, 1988; Quirk et al., 2020). It is sometimes speculated that such additional systems could particularly play a role with long encoding times (e.g., Lin & Luck, 2012), although there is little direct evidence to suggest this. Notably, claims of long-term memory involvement are not suggesting any use of "long"-term storage: The only way to correctly respond to the test probe in such studies is to pick the item that was seen on that particular trial at that particular location, as both the studied items and foil items are equally familiar real-world objects. Thus, information must be used about that specific trial. Instead, such objections are based on the suggestion that people can perform such binding of an object to a location from 1 second ago not only using an online, active maintenance system, but also, in some situations, by using a fundamentally different, offline system, and that the usage of such a system applies only to some stimuli in some situations and it cannot be distinguished behaviorally whether this additional offline system was used.

To directly test this idea, and examine whether the performance benefits for real-world objects at long encoding times were due to the recruitment of nonactive memory systems, such as "long"-term memory, in our previous study we examined the contralateral delay activity (CDA)—a neural marker of how much information is stored actively in working memory (Vogel & Machizawa, 2004). We found, as in many previous CDA studies, that the CDA amplitude tracked behavioral performance increases. This provides evidence that real-world objects were stored actively in visual working memory—just like colors or other basic features. Specifically, we found that in line with behavioral performance, objects showed greater CDA at longer encoding times than shorter encoding times, consistent with additional information being actively held in mind, and that at long encoding the CDA was reliably greater for remembering five objects than for remembering five colors (but not different when the amount remembered was the same for each stimulus set; e.g., with three of each presented). We hypothesized that real-world objects particularly benefited from longer encoding time in our study because this enabled a deeper processing of these stimuli, which may facilitate accessing existing knowledge of these stimuli, which can be used to help hold them "online," something that would not be useful for simple stimuli such as colors (Brady et al., 2016).

In particular, people may maintain information in visual working memory not solely in terms of colors and shapes and other "basic" visual dimensions in low-level visual cortex (e.g., Serences, 2016), but also maintain active representations of the stimuli in higher-level visual regions (e.g., fusiform face area [FFA] for face stimuli, Druzgal & D'Esposito, 2001; somatosensory regions for hand images, Galvez-Pol et al., 2018), resulting in stronger memories for items that can be meaningfully represented in higher-level brain regions. Consistent with this, Stojanoski et al. (2019) have shown greater ventral stream involvement in visual working memory tasks for meaningful rather than perceptually matched nonmeaningful stimuli. In addition, we have shown that perceptually matched images that are perceived as a face are not only better remembered in a working memory task than those not perceived as a face, but also elicit a larger CDA (Asp et al., in press), once again showing that "online" storage in working memory nearly always tracks behavioral performance in these tasks, rather than participants relying on a mix of memory systems only for some stimuli at some encoding times but not others. Consistent with this model of greater engagement of higher-level regions with meaningful stimuli, Salmela et al. (2019) have shown that storing faces in memory results in the storage of both low- and high-level information about them, whereas simple orientation stimuli are stored in a solely low-level way. Furthermore, a significant literature has shown, using behavior alone, that familiarity and knowledge improve performance in short-term memory tasks even with perceptually well-matched or even identical stimuli (e.g., Alvarez & Cavanagh, 2004; Brady et al., 2009; Curby et al., 2009; Jackson & Raymond, 2008; Ngiam et al., 2019; O'Donnell et al., 2018; Sahar et al., 2020; Starr et al., 2020). For example, familiar faces appear to be easier to remember than unfamiliar faces (Jackson & Raymond, 2008), and familiar letters, rather than letters from unfamiliar alphabets, are more easily remembered (Ngiam et al., 2019), conceivably due to the ability to recruit high-level features when processing such stimuli.

Several recent studies have, however, challenged these previous findings of benefits of knowledge and familiarity for working memory. In particular, two recent studies have challenged the claim of a selective benefit at long encoding times for real-world objects, instead finding benefits from long encoding time for both real-world objects and simple colors (Li et al., 2020; Quirk, et al., 2020). These results contest not only the "higher capacity for real-world objects" account, but pose a serious problem for fixed storage capacity models in general, as they strongly contrast with the standard claim that limits in performance in these paradigms arise primarily from storage limits, not encoding limits (e.g., Alvarez & Cavanagh, 2004; Luck & Vogel, 1997; Tsubomi et al., 2013; Vogel et al., 2006). If performance generally increases with additional encoding time for all stimuli, this would cast doubt on almost all research claiming to measure limits in working memory, instead suggesting that so-called capacity limits of working memory may actually arise primarily from capacity limits at encoding or during perception (e.g., Stojanoski et al., 2019). Thus, whether and how encoding time influences working memory performance for different kinds of stimuli is central to understanding not just how working memory capacity is affected by stimulus type, but the nature of working memory capacity limits more broadly.

In sum, there are several discrepant results as to how encoding time and meaning affect working memory performance, with the majority of studies using meaningless and often simple stimuli claiming that encoding time past a few hundred milliseconds does not affect visual working memory performance (e.g., Alvarez & Cavanagh, 2004; Bays et al., 2011; Luck & Vogel, 1997; Tsubomi et al., 2013), some suggesting that longer encoding times increase working memory performance for both simple stimuli and real-world objects (Li et al., 2020; Quirk et al., 2020), and others reporting selective benefits of long encoding times for real-world objects only (Brady et al., 2016). These mixed results depict the lack of clarity on a seemingly basic issue in the working memory literature. They raise the question of why results differ so greatly among studies, and demand a deeper understanding of the processes involved during working memory encoding, as a function not just of encoding time but of how people process the stimuli.

It is unlikely there is only one reason for the discrepant results with respect to real objects in particular. For example, in other recent work, we have shown that a major reason not all studies have found object benefits at long encoding times (e.g., Li et al., 2020; Quirk et al., 2020) is that they chose foils unfairly, in a way that disadvantages objects relative to colors (Brady & Störmer, 2020). Once this is accounted for, significant benefits for memory performance with real-world objects emerge reliably compared to both colors (Brady & Störmer, 2020) and perceptually matched meaningless stimuli (e.g., Brady & Störmer, 2020; Sahar et al., 2020; Stojanoski et al., 2019; Veldsman et al., 2017). In addition, verbal re-encoding is always a potential concern with long encoding times, and studies have differed in how they have prevented this (e.g., Brady et al., 2016; Li et al., 2020; Quirk et al., 2020).

In the current work, we investigate additional reasons why object benefits may appear somewhat heterogeneous: We show that longer encoding time is simply one possible way to allow for deeper processing of each of the to-be-remembered stimuli, and argue that any form of deeper processing is particularly beneficial for more meaningful stimuli. We propose that in some situations (or in some instruction conditions), participants may be prone to try to encode the entire display holistically (i.e., in a quick "snapshot") rather than process the items individually. If this encourages participants to treat objects simply as meaningless colored "blobs," rather than process them at a high level, connecting them to prior knowledge, this would clearly reduce the ability to find benefits in memory capacity for meaningful stimuli. Thus, in a series of experiments we directly test whether real-world objects, colors, perceptually matched less-meaningful objects, and fully scrambled objects benefit from a deeper processing, while manipulating the degree of such deeper processing. To do so, we systematically vary the presentation format of stimuli at encoding to be either simultaneous—encouraging a parallel, "take-a-quick-snapshot" strategy (similar to how we imagine participants do the task at short simultaneous encoding)—or present the stimuli sequentially, promoting a serial, each-item-at-once strategy.

We find that real-world objects result in higher memory performance than simple stimuli across all experiments (nine total replications of the object benefit in four experiments), in support of the account that working memory capacity is higher for meaningful and real-world objects relative to meaningless simple colors and meaningless perceptually matched stimuli. We also find that real-world objects—and to a lesser degree lightly scrambled versions of the objects—benefit from the sequential encoding and thus deeper, focused-on-individual-items processing, while colors and fully scrambled objects do not. Thus, our results demonstrate that different encoding situations during working memory tasks—previously only indirectly manipulated by changing encoding times—play an important role in constraining working memory capacity, and suggest that different encoding situations can have differential effects for different stimulus sets. Most broadly, our results indicate that single-feature objects may be outliers, not representative of real-world memory situations: Due to their unique role in feature-based attention, such stimuli are unlike any realistic stimuli in their affordance of parallel, holistic encoding. Thus, in more realistic situations, memory likely nearly always benefits from in-depth processing of objects (e.g., in higher-level visual areas) rather than processing them solely in terms of their low-level features (in a quick "snapshot").

Experiment 1: How Does the Memory Benefit for Objects Relative to Colors Interact With Encoding Strategy?

Single colors and other simple features can be processed quickly and in parallel (Treisman & Gelade, 1980; White et al., 2017), and people tend to make use of ensemble encoding, grouping, and other strategies when encoding such simple features into memory (e.g., Brady & Alvarez, 2015a). By contrast, object recognition is more serial (e.g., Rousselet et al., 2004), or at least severely bottlenecked, with objects benefiting from being individually selected to deeply process them and subject to severe limits from visual crowding (Whitney & Levi, 2011). Thus, we hypothesized that typical visual working memory studies may be the least advantageous situation for eliciting benefits for encoding meaningful objects: Such studies use simultaneous presentations of many stimuli at once, with all being equally task relevant. In the extreme, if people attempt to encode the entire display at once, effectively treating the stimuli as "colored, textured blobs" without processing their meaning at all, it could even be possible to eliminate object benefits. In contrast, presenting a set of colored circles all at once, which would encourage participants to "zoom out" and take a "snapshot" of the entire scene, might encourage the use of global feature-based attention (e.g., Treisman & Gelade, 1980; White et al., 2017), leading to ensemble processing and chunking of the features (e.g., Brady & Alvarez, 2015a; Nassar et al., 2018) that would be particularly beneficial for color memory and other low-level features that can easily be processed in parallel. Similarly, some work has shown that simultaneous exposure to similar stimuli allows participants to better encode, even after a delay, the differences and relations between these stimuli (e.g., Mundy et al., 2007, 2009), consistent with the idea that in brief simultaneous exposures of many similar stimuli, as in the case of color working memory, participants may focus on the relations between items as much as the items themselves (e.g., Chunharas et al., 2019; Ding et al., 2017).

Thus, we hypothesized that although there are object benefits even in studies with simultaneous presentations of a large number of items (i.e., Brady & Störmer, 2020), such conditions may nevertheless be among the most favorable conditions for simple stimuli and least favorable for realistic meaningful stimuli. In real-world scenarios where participants use visual working memory to perform tasks (e.g., holding in mind the target of an eye movement, Hollingworth et al., 2008; or the target of an action, Ballard et al., 1995; Hayhoe et al., 2003), they are relatively unlikely to try to equally encode many stimuli at once, and instead focus additional resources on more task-relevant stimuli (e.g., Salahub et al., 2019) and encode stimuli sequentially (e.g., Ballard et al., 1995).

In the current experiment, we sought to directly test the hypothesis that real-world objects, but not colored circles, benefit from a serial, item-based encoding. To do so, we compared memory performance across two encoding scenarios for objects and colors: We presented items either simultaneously, encouraging a parallel processing strategy, or sequentially, encouraging a serial processing strategy, while keeping the amount of time each item could be processed constant. We reasoned that this was a direct manipulation of different encoding strategies, but was likely similar to the difference that is tapped indirectly when encoding times are manipulated, with longer encoding times generally facilitating serial, item-based processing, and short encoding times—or even the possibility of short encoding times—generally encouraging simultaneous, parallel processing of the items.

Method

The Study design, hypothesis, analysis plan, and exclusion criteria were preregistered: https://aspredicted.org/blind.php?x=gh9md2. Materials and data are available at https://osf.io/va2te/.

Participants

Fifty participants were used in the final data analysis. Data from four participants were excluded and replaced per preregistered exclusion criteria. All participants were run in the laboratory and gave written informed consent prior to starting the experiment as approved by the Institutional Review Board at the University of California, San Diego.

Power

At long encoding times as in the current study, previous work (Brady et al., 2016) found the difference between object performance and color performance (in terms of d′) had an effect size of dz = .74. We had twice as many trials per condition, which would be expected to increase this effect size significantly, but at the same time, effects tend to be smaller in replications, and other work has disputed this effect (e.g., Quirk et al., 2020). Thus, we used this effect size as-is to calculate our power. By this calculation, the current study had 99.9% power to find this effect. We were also interested in the potential interaction, which would have twice as much variability in its estimate, halving the effect size. We thus had approximately 73% power to detect such an interaction.

Stimuli and Procedure

We contrasted memory for objects and colors in a working memory task modeled after Brady et al. (2016). The memory task used a 2-AFC memory probe (“Which of these two items did you see?”) to allow us to control foil similarity and avoid the need to model response criterion differences.

For colors, we used a standard color circle (Schurgin et al., 2020; Suchow et al., 2013) of radius 49 in the CIE L*a*b space (centered at L = 54, a = 21.5, b = 11.5). Both shown items and test foils were required to be a minimum of 15° apart on the color wheel to reduce chunking and grouping across items. Target colors and test foils were chosen to be maximally dissimilar (180° apart on the color wheel).

For object stimuli, we used the Brady et al. (2008) object image database, as in Brady et al. (2016). We aimed to exactly replicate Brady et al. (2016), by using their particular set of objects and foils (categorically distinct and hand-pruned to be visually distinct; see publicly available materials).

Participants remembered six colors or six real-world objects on each trial. Stimuli were presented either simultaneously for 1200 ms, or one at a time for 200 ms each (followed by a 200-ms interstimulus interval [ISI]). In sequential conditions, the items always appeared in the same order (clockwise starting at 9 p.m.), making them strongly temporally and spatially predictable. The long encoding time, combined with fixed spatial positions with placeholders present during the delay, helped ensure there was little to no location noise that can could misbinding.

After a delay (800 ms), a location probe was shown to indicate which location was being probed, and two test stimuli appeared in the center of the screen. Participants performed a 2-AFC, indicating which of the two stimuli appeared at the probed location during encoding (see Figure 1). This required information about exactly which object was at a particular location on this particular trial.

We used a within-subject design such that each participant encoded colors simultaneously, colors sequentially, objects simultaneously, and objects sequentially. Conditions were blocked (four blocks overall), and their order counterbalanced across participants subject to the constraint that participants either did both simultaneous conditions first (order counterbalanced) or both sequential conditions first. We counterbalanced in this way as we hoped to encourage participants to apply the same encoding strategy for both stimuli sets in each of the sequential and simultaneous conditions. Participants completed 60 trials of each condition.

To further ensure that any object benefit was not caused by verbal encoding, during the entire experiment, participants performed a concurrent articulatory suppression task designed to prevent them from verbally encoding any of the items. In particular, they were required to say “the” out loud continuously for the entire duration of the experiment, and this was monitored by an experimenter throughout the experiment.

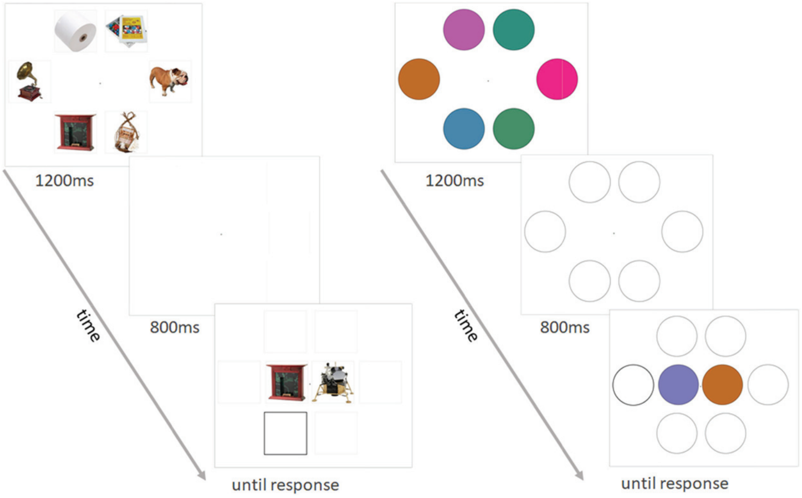
Analysis

Working memory performance was quantified using d′ for a 2-AFC task, $[zH-zFA]/\sqrt{2}$ where P is percent correct and Φ is the Gaussian cumulative distribution, $zH = \Phi(P)$ and $zFA = \Phi(1-P)$. Per the preregistration, data were excluded if the d′ averaged across all conditions was below .5, or if greater than 10% of individual trials were excluded. Individual trials were excluded if a) a response occurred less than 150 ms after the response screen appeared, or b) the response occurred more than 5 s after the response screen appeared.

Results

We found a main effect of stimulus type, indicating that objects were remembered better than colors overall ($F(1, 49) = 88.275$, $p < .0001$; $\eta_p^2 = .64$). There was no main effect of encoding type (simultaneous vs. sequential), $F(1, 49) = .572$, $p = .45$, $\eta_p^2 = .015$, but there was an interaction between encoding type and stimulus type ($F(1, 49) = 12.114$, $p = .001$, $\eta_p^2 = .20$), such that objects were remembered better during sequential encoding and colors were better remembered when encoded simultaneously (see Figure 2). Follow-up pairwise comparisons confirmed this: When objects were encoded sequentially, memory performance was higher than when encoded simultaneously $t(49) = 2.19$, $p = .033$, dz = .31, but when colors were encoded sequentially, memory performance was

Figure 1
Experimental Methods



Note. In all experiments, participants saw either six real-world objects or six colors and had to remember them over a brief delay, followed by a 2-AFC memory test. In simultaneous encoding conditions (shown), they saw all the objects at once. In sequential encoding conditions, objects appeared at the same spatial locations and in a spatially and temporally predictable sequence, but for 200 ms each, and they were probed in the same way. See the online article for the color version of this figure.

lower than when colors were encoded simultaneously ($t(49) = -2.65, p = .011$, dz = .37).

This pattern of results is consistent with our hypothesis that simple features—such as colors—can be processed efficiently in parallel, likely benefiting from ensemble and chunking processes when shown at the same time, while real-world objects benefit from a deeper item-based processing that is facilitated by sequential encoding, where each item can be focused on one at a time. Notably, the crossover interaction we found suggests a qualitative difference in the best way for people to encode meaningful stimuli versus simple features. Given existing evidence suggestive of different mechanisms available only for simple features (e.g., feature-based attention), this raises the strong possibility that the models researchers have developed to explain memory capacity in simple features (particularly at high set sizes, where parallel encoding is necessary) may not apply at all to more meaningful objects or more realistic situations where visual working memory is used.

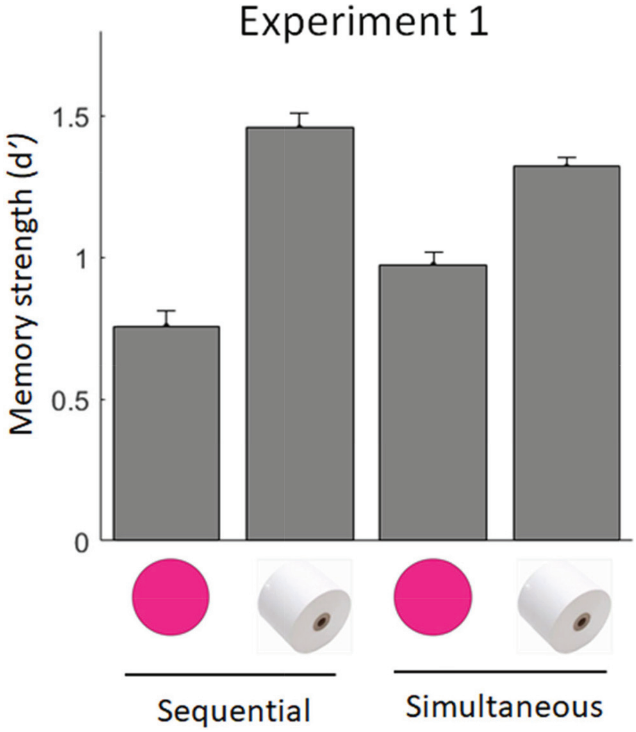
Experiment 2: The Role of Spatial Location: Do Objects in Particular Benefit From Deeper Processing Afforded by Sequential Presentations, and Is This Impacted by Spatial Locations Being Present?

To replicate our results from Experiment 1 of sequential encoding boosting memory performance for real-world objects, and to eliminate the possibility of any practice or strategy effects that could arise in a within-subject design, Experiment 2 was a between-subjects version of that experiment. This helps eliminate the concern that participants may have adapted to one encoding strategy based on what they were exposed to first and continued to use it even when the presentation format changed. For example, if a participant began with the sequential presentation that encourages encoding each item separately, they may continue focally attending to each item during simultaneous presentation trials (or vice versa). In this experiment, we also increased the encoding time in the simultaneous condition to be greater than in the sequential conditions (1,200 vs. 2,000 ms), to reduce the possibility that items are encoded somewhat longer during the sequential condition because participants may continue to process them during the time between objects (200 ms ISI)—which could, at least in theory, differentially impact objects versus colors. Finally, Experiment 2 also added a new condition in which we presented all items sequentially at the center of the display, instead of presenting them at six different locations. This allowed us to test the role of spatial information during encoding and provided another test for the robustness of the object benefit. Manipulating spatial location also provides an indirect window into whether participants might be relying on more "long-term storage" for objects. This is because when items are presented at the same location sequentially, research has shown enhanced proactive interference across trials (Makovski, 2016). Such enhanced proactive interference could in theory be caused by the usage of more "long"-term memory information, rather than purely working memory, in such conditions. Thus, if the object benefit was much larger when spatial locations were not available, this could point to the possibility that it arises from more passive long-term storage rather than active storage in working memory.

Method

Materials and data are available at https://osf.io/va2te/.

Figure 2
Results of Experiment 1



Note. Memory performance for objects was overall higher than for colors, but there was a qualitative difference between the two stimuli in terms of encoding strategy: Objects show a larger benefit from being encoded sequentially, while colors show a benefit in being encoded simultaneously. See the online article for the color version of this figure.

Participants

Fifty unique U.S.-based participants from Amazon's Mechanical Turk were in the final dataset of each of the three across-participant experimental groups (total 150 participants; all with $\geq 95\%$ acceptance rates). An additional 10 participants were excluded in the nonspatial sequential condition; 15 excluded in the spatial-sequential condition; and 13 excluded in the simultaneous condition. Exclusion criteria at the subject level were performance below chance or 10% of trials excluded; trials were excluded based on the same rules as previously (reaction times < 150 ms; $> 5{,}000$ ms).

Stimuli and Procedure

In all conditions, participants performed 120 trials, 60 with color and 60 with real-world objects. Stimuli and procedure for each trial were identical to Experiment 1 with the following exceptions: Encoding type was varied across participants, such that Group 1 only performed trials in which items were presented sequentially (200 ms each with a 200 ms ISI) at the center of the screen; Group 2 only performed trials in which the items were presented sequentially at distinct spatial locations (200 ms each with a 200 ms ISI); and Group 3 performed trials in which the items were presented simultaneously. The simultaneous presentation used a longer encoding time than in Experiment 1 (2,000 ms), since we reasoned it is possible participants continue to process the stimulus during the ISI period in the sequential condition, an advantage they would not have with 1,200 ms encoding in the simultaneous condition. The first condition thus removed spatial cues from encoding, a new condition in this experiment, and the third condition examined whether objects continue to benefit from sequential encoding even with a longer encoding time in the simultaneous condition. Participants performed the same verbal suppression task as in Experiment 1, but in this study, unlike Experiment 1, compliance was not monitored continuously by the experimenter.

Results

Our results in the sequential (spatial) and simultaneous conditions replicate the results of Experiment 1 (see Figure 3). We also find that the sequential nonspatial condition, where items were all presented at the center of the screen, was similar to the results from the sequential (spatial) condition.

In particular, in all three conditions taken individually, there was a reliable and large object benefit: $t(49) = 12.17$, $p < .001$, dz = 1.72 (sequential, center); $t(49) = 11.43$, $p < .001$, dz = 1.62 (sequential, spatial); $t(49) = 5.36$, $p < .001$, dz = .76 (simultaneous).

We also replicated the dissociable effects of sequential and simultaneous encoding for objects versus colors observed in Experiment 1, even with the longer encoding time in the simultaneous condition. In particular, colors were better remembered when presented simultaneously than sequentially (spatially), $t(98) = 2.89$, $p = .005$, Cohen's d = .58, and better remembered when presented simultaneously than sequentially (at the center), $t(98) = 3.36$, $p = .001$, d = .67. By contrast, objects were remembered better when presented sequentially (spatially) than simultaneously, $t(98) = 3.02$, $p = .003$, d = .60, and also better sequentially (central) than simultaneously, $t(98) = 4.43$, $p < .001$, d = .89. The difference in performance between objects in the two sequential conditions was not statistically significant ($t(98) = 1.30$, $p = .198$, d = .26).

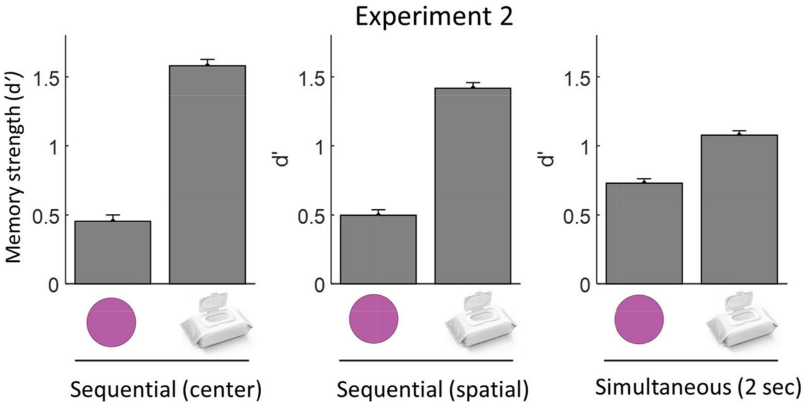
Thus, taken together, the data from Experiment 2 suggest that regardless of spatial location availability, objects show an extremely large benefit over colors when items are encoded sequentially, and a smaller (although still large) benefit when items are encoded simultaneously. They also replicate the qualitative difference in encoding between the two stimuli types: Even with longer encoding time in the simultaneous condition than sequential conditions, objects benefit more from sequential whereas colors benefit more from simultaneous encoding.

Experiment 3: Is the Overall Object Benefit Purely From Stimulus Complexity, or a Result of Knowledge/Familiarity? Lightly Scrambled Objects

Experiments 3 and 4 address two questions. First, they examine the role of semantics in the object benefit from Experiments 1 and 2 and Brady et al. (2016) and Brady and Störmer (2020). One possibility is that the object benefit is simply a "complex" stimuli benefit—that is, stimuli that are more complex are better remembered regardless of the meaningfulness of the stimulus. This could arise if, in contrast to our suggestion of high-level, meaningful features being recruited for meaningful objects, instead there are simply fixed pools of resources for each basic feature (e.g., color,

Figure 3

Results of Experiment 2



Note. We replicate reliable object benefits in all conditions. We also replicate the cross-over effect found in Experiment 1, where objects are better remembered in sequential encoding conditions whereas colors are best remembered when presented simultaneously. See the online article for the color version of this figure.

orientation, spatial frequency, etc.), and real objects benefit from recruiting multiple such pools of resources whereas colors cannot. Such a hypothesis is superficially at odds with work showing worse performance for complex but meaningless stimuli than simple stimuli (e.g., Alvarez & Cavanagh, 2004; Brady & Alvarez, 2015b); and work using perceptually matched but nonmeaningful stimuli (e.g., Asp et al., in press; Sahar et al., 2020; Stojanoski et al., 2019), but it is important to address this directly.

In addition, the following experiments also address the question of sequential versus simultaneous encoding in complex, meaningless stimuli. In particular, Experiments 3 and 4 ask whether the sequential processing benefit found in Experiment 1 is unique to objects, or dependent on how much meaningful information can be extracted from the stimuli. We hypothesized that stimuli that provide significant semantic information when processed more deeply may benefit from deeper processing—since additional informative item-specific features can be extracted—and that stimuli that are semantically meaningless (just "colored blobs") would not benefit from such processing. This would be consistent with our hypothesis that stimuli made up solely of meaningless bundles of color and orientation may be relatively unique in their affordance of parallel processing/ensembles/grouping—continuing the theme from Experiments 1 and 2 that such stimuli are perhaps not a good case study of memory as they are (uniquely) supported by feature-based attention (Treisman & Gelade, 1980).

To test this, we used two different levels of scrambling in these two experiments: Experiment 3 uses light scrambling, which slightly impairs how much the stimuli can be recognized (one side is vertically flipped) but preserves most of the meaningful information in them. Experiment 4 uses fully scrambled versions of these objects, which are effectively just colored blobs without meaning but contain many different and complex visual features. We validated the effects of these manipulations on meaningfulness in a separate pair of experiments (see Appendix).

The light scrambling we use in Experiment 3 is purposefully an extremely subtle manipulation of the objects. It has been shown to distort the meaningfulness and familiarity of objects to some extent, as well as affect memory performance (Shoval & Makovski, 2019), but preserves a fair amount of the meaning of the objects (see Appendix). By contrast, our fully scrambled objects in Experiment 4 used diffeomorphic scrambling to remove effectively all ability to recognize the objects (Stojanoski & Cusack, 2014), massively reducing their meaningfulness, while still maintaining visual complexity (see Appendix).

Note that no scrambling can perfectly match low- and midlevel features while eliminating meaning: To have exactly the same features requires having the exact same images. However, this set of experiments does nonetheless provide information about whether the presence of physically complex visual stimuli per *SE* is sufficient to reach the level of performance of realistic objects.

Overall, in Experiments 3 and 4, we predicted that like real objects lightly scrambled but still meaningful objects would benefit from deeper processing, but that fully scrambled objects, which are effectively just colored blobs with no deeper processing possible, would benefit from parallel processing like simple features. We also expected a general memory benefit for real-world objects relative to both the lightly and fully scrambled versions of these objects.

Method

Materials and data are available at: https://osf.io/va2te/.

Participants

The final dataset consists of 30 participants tested in person at UC San Diego. Data from three additional participants were excluded per the same rules as the previous experiment. We had planned 50 participants to match Experiment. 1, but data collection was interrupted by COVID-19.

Power

We hoped to power the current study to detect a main effect in performance between meaningful and scrambled objects, as well as an interaction in encoding benefits for sequential versus simultaneous if present. Using the results from Experiment 1 revealed that even with only the dataset of 30 participants, if the difference between meaningful and nonmeaningful objects was half as large as that between meaningful objects and colors, we had > 99% power to detect such an effect with this sample, as well as 77% power to detect an interaction of the size observed in Experiment 1. Thus, we analyzed the data with this sample.

Stimuli and Procedure

The experimental set-up was the same as Experiment 1 with the only exception that we used lightly scrambled objects instead of colors. Thus, the experiment consisted of four conditions: real-world objects simultaneous; lightly scrambled objects simultaneous; real-world objects sequential; and lightly scrambled objects sequential. For the lightly-scrambled-object conditions, the same object database was used as for the intact objects, but either the left or right half of the object was flipped vertically, making it more difficult to recognize the object (Shoval & Makovski, 2019), while simultaneously keeping the objects nearly identical in their visual complexity and visual features.

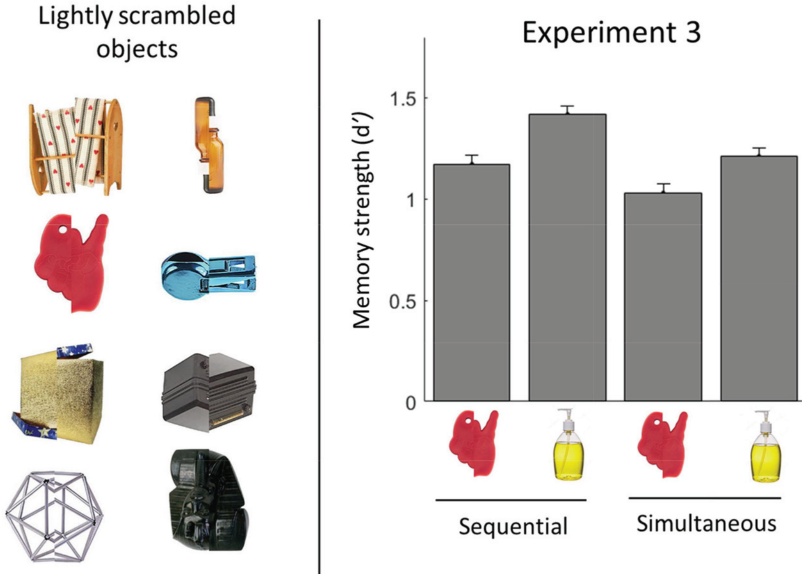
For each participant we randomly assigned which objects would be seen as lightly scrambled and which objects would be seen as intact, such that no object was used across both conditions for an individual subject. The test foils were chosen to be categorically and visually dissimilar, just as in Experiment 1; that is, we used the same object pairs as Brady et al. (2016), with or without scrambling both items. Furthermore, participants again concurrently performed a verbal interference task throughout the experiment: articulatory suppression by saying "the" out loud for the entire duration of the study that was continuously monitored by an experimenter.

Results

We observed a main effect of stimulus type, such that intact objects showed higher memory performance than scrambled versions of these objects, $F(1, 29) = 15.722$, $p = .0004$; $\eta_p^2 = .35$, replicating a meaningful object benefit in working memory (see Figure 4). Furthermore, we found a benefit of a sequential encoding for both, $F(1, 29) = 12.755$, $p = .0013$; $\eta_p^2 = .31$, and no interaction ($F(1, 29) = .595$, $p = .45$, $\eta_p^2 = .02$). Follow-up pairwise comparisons confirmed that the sequential encoding benefit was reliable for each stimulus type: For intact objects, sequential encoding resulted in higher memory performance relative to simultaneous encoding ($t(29) = 3.43$, $p = .002$, dz = .63), and for scrambled objects, memory performance was also higher for sequential relative to simultaneous encoding ($t(29) = 2.07$, $p = .048$, dz = .38).

These data are consistent with our predictions that to some extent any meaningful stimuli—that is, real-world objects and also their lightly-scrambled counterparts—benefit from sequential encoding and thus from deeper and more serial processing. Furthermore, these results replicate the general advantage for real-world objects (Brady et al., 2016), even compared to extremely visually similar stimuli. Thus, they also provide a conceptual replication of the results from Asp et al. (in press), which show enhanced memory performance and enhanced active neural storage (via the CDA) for meaningful stimuli compared to

Figure 4
Stimuli and Results of Experiment 3



Note. Objects were lightly scrambled by flipping vertically one half of the object. Real-world objects resulted in overall higher memory performance relative to lightly-scrambled objects, and both lightly scrambled objects and intact objects showed higher memory performance when presented sequentially relative to simultaneously. See the online article for the color version of this figure.

perceptually matched stimuli that are not subjectively seen as meaningful (Mooney faces), as well as other similar work (e.g., Stojanoski et al., 2019).

Experiment 4: Is the Overall Object Benefit Purely From Stimulus Complexity, or a Result of Knowledge/Familiarity? Fully Scrambled Objects

In Experiment 4, we compared real-world objects to fully scrambled objects, which were effectively just colored blobs without meaning (see Appendix for experimental validation of their lack of meaningfulness). In contrast to Experiment 3, we predicted that fully scrambled objects would benefit from parallel processing like simple features, rather than from deeper processing like the lightly scrambled objects in Experiment 3. We thus predicted they would be better remembered in simultaneous than sequential conditions.

Because these objects are similarly complex and contain many features, like real objects, this experiment thus provides a useful test of whether the meaningfulness of the stimuli is the critical aspect that leads to enhanced object memory with sequential presentations. If instead some aspect of visual complexity itself is responsible for the sequential benefit—that is, perhaps eliminating visual crowding (Whitney & Levi, 2011) is more important for real objects than simple stimuli—then there should be a sequential benefit even for these almost totally meaningless objects because of their visual complexity.

Method

Materials and data are available at: https://osf.io/va2te/.

Participants

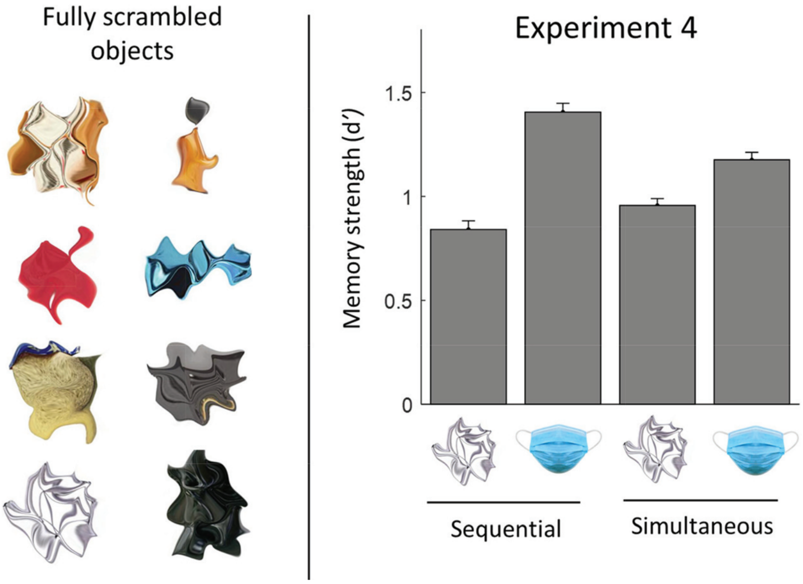
The final dataset consists of 50 participants. This experiment was run on UCSD undergraduates, but the data were collected online due to COVID-19. Six additional participants were excluded according to the same rules as the previous experiments.

Stimuli and Procedure

The experimental set-up was the same as Experiment 3 with the only exception that we used fully scrambled objects instead of lightly scrambled objects. Thus, the experiment consisted of four conditions: real-world objects simultaneous; fully scrambled objects simultaneous; real-world objects sequential; and fully scrambled objects sequential (Figure 5). For the fully-scrambled-object conditions, the same object database was used as for the intact objects, but used diffeomorphic scrambling to remove effectively all ability to recognize the objects (Stojanoski & Cusack, 2014; see Appendix for recognition data). Diffeomorphic scrambling is done by repeatedly applying a 2D-flow field to the images, with random phase and amplitude, effectively distorting the image without changing any major features of the overall distribution of pixels. It is designed to match the visual system response between the original and scrambled images in early stages of visual processing, while removing the ability for participants to recognize the images. It has no duplication of removal of parts, meaning it, for example, preserves the topography of the image, such that if there

Figure 5

Stimuli and Results of Experiment 4



Note. Objects were fully scrambled to remove the ability to recognize them while preserving relevant low and midlevel image statistics. Real-world objects resulted in overall higher memory performance relative to fully scrambled objects. Like colors, fully scrambled objects were better remembered when presented simultaneously, whereas intact objects showed higher memory performance when presented sequentially relative to simultaneously. See the online article for the color version of this figure.

are a certain number of islands of pixels surrounded by white in the original images, the same is true on the transformed images (Stojanoski & Cusack, 2014).

As in Experiment 3, for each participant we randomly assigned which objects would be seen as scrambled and which objects would be seen as intact, such that no object was used across both conditions for an individual subject. The test foils were chosen to be categorically and visually dissimilar, just as in Experiment 1. Furthermore, participants again concurrently performed a verbal interference task throughout the experiment: articulatory suppression by saying "the" out loud for the entire duration of the study.

Results

We observed a main effect of stimulus type, such that intact objects showed higher memory performance than fully scrambled versions of these objects, $F(1, 49) = 57.85$, $p < .0001$; $\eta_p^2 = .54$, replicating a meaningful object benefit in working memory. There was no main effect of sequential encoding ($F(1, 49) = 2.196$, $p = .14$, $\eta_p^2 = .04$), but there was a crossover interaction ($F(1, 49) = 16.94$, $p = .0001$; $\eta_p^2 = .26$), such that real objects were better remembered with sequential encoding and very scrambled objects were better remembered with simultaneous encoding (Figure 5). Follow-up pairwise comparisons confirmed that the differences in encoding were reliable for each stimulus type: For intact objects, sequential encoding resulted in higher memory performance relative to simultaneous encoding ($t(49) = 3.86$, $p < .001$, dz = .55), and for fully scrambled objects, memory performance was higher for simultaneous relative to sequential encoding ($t(49) = 2.13$, $p = .039$, dz = .30).

Taken together, Experiments 3 and 4 provide strong evidence that the potential for extracting meaning from a stimulus determines whether it benefits from sequential processing. On the one extreme, features like single colors have an incredible affordance of parallel processing; at the opposite extreme, real objects have significant additional semantic information that can be extracted when they are processed as individuals, improving memory in sequential conditions. Experiments 3 and 4 show there are also gradations: both kinds of reasonably meaningful stimuli—that is, real-world objects and also their lightly-scrambled counterparts—benefit from sequential encoding and thus from deeper and more serial processing, whereas fully scrambled, unrecognizable stimuli and simple features like color benefit most from simultaneous processing. Experiment 4 shows that the sequential benefit is not solely due to visual complexity but is related to the meaningfulness of the stimuli.

Overall, these experiments suggest that results from simple features at high set sizes should not be generalized to real-world objects, or even to more complex but still meaningful stimuli, like lightly scrambled versions of objects. Using simple, unidimensional stimuli like colored circles or oriented bars may not just chronically underestimate working memory performance but also provide qualitatively incorrect conclusions about how memories are best formed and stored in more realistic situations.

General Discussion

Across a series of experiments, we found higher working memory performance for real-world objects relative to simple colors and scrambled objects. This was true even in situations that are suboptimal for this object benefit to emerge, like the simultaneous presentation of all stimuli. Such conditions resulted in only moderate, but nonetheless robust, benefits for meaningful objects in comparison to simple colors and scrambled stimuli. Conditions that encouraged "deeper" item-specific processing—that is, the sequential presentation of stimuli—resulted in reliably larger benefits for meaningful stimuli, whereas such conditions uniquely disadvantaged simple colors and meaningless stimuli compared to conditions that facilitated parallel encoding (i.e., simultaneous conditions). Objects were better remembered when people were forced to focus on them one at a time, even though in the simultaneous condition we used relatively long encoding times (>1 s, and up to 2s), to match how much time each item could be processed across both conditions, thus allowing the possibility of participants' serially moving attention during the encoding period. With extremely short encoding times, which even more strongly encourage a single "snapshot" of the display rather than individualized encoding, this difference would be expected to be even further exacerbated.

Altogether, these data support a model of working memory where capacity depends on the type of information that is being remembered and how it is encoded. They suggest that using simple stimuli may be causing us to chronically underestimate working memory capacity, as well as promote the study of mechanisms of working memory storage that are unique to simple stimuli and not generalizable to complex, meaningful stimuli that we actually encounter in the world. In particular, stimuli consisting solely of meaningless visual features like color may encourage global, feature-based attention encoding strategies that are not nearly as useful as in-depth processing for realistic objects. The current data also suggest that memory is enhanced for meaningful objects especially when they can be focused on as individuals. Why might this occur? One possibility is that item-specific processing may be necessary to extract meaningful, high-level features of the stimuli (e.g., to encode faces with respect to face-specific features like eye position and nose angle, rather than simply in terms of low-level shapes). An account where the meaningfulness advantage arises from encoding "additional features" of the meaningful objects is consistent with the fact that adding features to simple objects allows more information to be encoded per object (e.g., Fougnie et al., 2013; Luck & Vogel, 1997). However, it uniquely asserts a role for meaning, since additional high-level features are only available for meaningful stimuli: Scrambled stimuli, for example, do not offer the opportunity to make use of higher-level features (e.g., like face pose; or, for a tree, branch thickness) in the same way realistic, meaningful stimuli do, despite being complex physically. Such an explanation for the meaningful object benefit is also consistent with the enhanced neural delay activity observed for meaningful stimuli (e.g., Asp et al., in press; Brady et al., 2016), as this shows that more information is being stored about meaningful objects. Finally, an "additional features" account of the meaningful object benefit also predicts that memory for realistic objects is improved only when additional meaningful features can be used in the memory probe that is given. For example, such an account would also be consistent with the fact that there is little or no benefit to remembering arbitrary colors that happen to be the colors of realistic objects (e.g., Brady et al., 2013; where capacity is similar to standard simple color square tasks), or the arbitrary spatial locations that real-world objects are in (e.g., Lam & Sprague, 2020).

Comparing Working Memory Across Stimulus Sets on Common Ground

Why have some other studies not found a working memory benefit for objects relative to colors? We believe there are a number of sources that likely underlie the discrepancies across studies. First, in another recent article we demonstrated the importance of comparing memory performance across stimulus sets on common ground by using comparable target/foil similarity at test (Brady & Störmer, 2020). Earlier work has shown that the similarity between target and foil critically determines performance (e.g., Awh et al., 2007; although see Brady & Alvarez, 2015b), and this issue has most recently been quantified in the Target Confusability Competition model of working memory (Schurgin et al., 2020), which shows that even seemingly large differences between items and foils in a given stimulus space may not be sufficient to avoid underestimating working memory capacity, as confusability between items is not at floor even for large target/foil changes. Thus, it is of tremendous importance to maximize the dissimilarity of the memory foils for both stimulus sets when comparing working memory capacity.

By contrast, some previous studies that compared working memory capacity for colors and real-world objects used maximally dissimilar colors (i.e., 180° on the color wheel), but then chose objects randomly from a database of many objects, with varying levels of visual and semantic similarity (Li et al., 2020; Quirk et al., 2020). However, this does not in any way maximize dissimilarity for objects, effectively disadvantaging them relative to colors and—at least in large part—explaining why there were no object advantages over colors at long encoding times in those studies (Brady & Störmer, 2020). In the current work, where we chose object foils to be extremely dissimilar from targets, just like colors, we repeatedly replicated Brady et al.'s (2016) finding that objects are better remembered than colors in long simultaneous encoding conditions, strongly contrasting with Li et al. (2020) and Quirk et al. (2020). Thus, one major reason why some labs have not found object benefits is that they did not make target/foil similarity comparable across their stimulus sets.

Different Encoding Strategies Benefit Objects and Meaningless Stimuli Differently

A second source of heterogeneity in object benefits—and the main focus of the current study—is how these stimuli are processed at encoding. Specifically, we demonstrate that serial and focused item-based encoding enhances working memory benefits for real-world objects, while distributed parallel encoding facilitates memory for color displays (Experiments 1 and 2). We also found a serial benefit—albeit smaller—for lightly scrambled objects that still maintained some meaningful semantic information (Experiment 3), but a simultaneous presentation benefit for fully scrambled, meaningless objects (Experiment 4). Experiment 4 suggests that the benefit of sequential encoding is not a result of their visual complexity or number of visual features; instead, sequential benefits arise only for stimuli where meaningful information can be extracted when items are individually processed.

Overall, then, the benefit for meaningful stimuli is strongest when participants can encode each item individually, recognize its identity, and connect it to existing knowledge. This is consistent with previous explanations of the object benefit: such benefits to working memory storage do not seem to arise because of more complex visual features, but rather their meaningfulness (Asp et al., in press; Brady et al., 2016). Thus, effectively visually identical stimuli are better remembered when they can be processed in a meaningful way (e.g., Alvarez & Cavanagh, 2004; Asp et al., in press; Ngiam et al., 2019), and semantic knowledge about objects in particular is critical (Starr et al., 2020).

Under this logic, any conditions that enable deeper processing of items should be particularly beneficial for real-world objects. In other work, we have used long encoding times to facilitate focused item-based encoding, and also found a selective benefit for objects in working memory at long encoding times (Asp et al., in press; Brady et al., 2016;). In the current work, we show this benefit is enhanced by independent item processing rather than simultaneous encoding of many items.

What, then, is the role of encoding versus maintenance in visual working memory limits? This question is particularly acute given that two recent studies reported benefits of long encoding for both objects and colors (Li et al., 2020; Quirk et al., 2020), seemingly at odds with the selective object benefit we previously reported as well as significantly deviating from important claims over the last 20+ years that have repeatedly argued that working memory for simple stimuli quickly "fills up" and reaches a fixed capacity limit (e.g., Alvarez & Cavanagh, 2004; Luck & Vogel, 1997). We, too, have found that color performance does improve with time (e.g., in Schurgin et al., 2020), although as shown in the current work and Brady and Störmer (2020), never reaching the level of object performance. The current work is thus consistent with a growing literature suggesting that encoding differences—either encoding strategy or encoding times—change working memory performance, raising fundamental questions about the purity of the putative fixed-capacity "memory" limits claimed in earlier work.

This is particularly important if there is significant variance in encoding strategies. While varying encoding time likely taps into different encoding strategies, such that shorter encoding time, on average, encourages parallel processing and longer encoding time encourages serial, item-based encoding, it is not clear that this is always the case, as it is an indirect way to influence how items are being processed. Participants might be able to encode solely the low-level visual details of the display, ignoring semantic and higher-level visual information, even when items are presented for a long encoding time. Indeed, individuals are known to vary in their propensity to take on more holistic encoding strategies (Babic et al., 2019; Cusack et al., 2009; Linke et al., 2011), resulting in incorrect estimates of their apparent visual working memory capacity and working memory capacity's relationship to fluid intelligence (Babic et al., 2019; Cusack et al., 2009) based only on encoding strategy differences. Such holistic strategies, as we show in the present study, are advantageous when remembering simple features, and it is thus conceivable that such strategies would be used by participants that perform color and object tasks intermixed, or in cases where that parallel "take-a-snapshot" strategy feels subjectively less effortful. Thus, longer encoding time alone may not always result in a selective object benefit but only when such encoding times are successful in promoting a deeper processing of these items. Importantly, even for simple stimuli, such a strong role for encoding points to the difficulties facing fixed-capacity models of memory performance.

If encoding strategy or encoding times change working memory performance in a relatively smooth way, even for simple stimuli, this raises important questions for all fixed capacity models of visual working memory, including those that assume fixed object limits (e.g., Adam et al., 2017), and those that assume other fixed limits can explain memory performance across set sizes (e.g., divisive normalization-based limits; Bays, 2014). For example, Bays (2014) fit data across all set sizes in a Model of working memory for simple features by assuming a single resource limit for these features (in terms of spikes) that is allocated differently across set sizes, and used this to argue for a fixed resource limit, at least within a single feature dimension (e.g., Bays, 2015). However, they focus on only a single encoding situation: a particular encoding time for simultaneously presented stimuli. How to reconcile such fixed limits with the smooth variability across encoding situations we and others observe thus remains a difficult question, which may require rethinking the assumptions of a single fixed capacity even within a single feature dimension.

In contrast, models of visual working memory that argue that all that is being assessed, even in continuous report paradigms, is effectively the signal-to-noise ratio of the memory trace (d'; Schurgin et al., 2020) take a much more fluid approach to the concept of fixed capacity. Schurgin et al. (2020), for example, point out that it would not be surprising for memories to become noisier when more items need to be encoded, or to become noisier as delay increases, and that combining multiple sources of change in signal strength (e.g., splitting attention at encoding; changing encoding time) with multiple sources of change in noise accumulation (e.g., splitting maintenance capacity; increased delay) is unlikely to result in a single fixed capacity being observed in memory performance, even if there is, deep down, some underlying limited resource in some of the components that affect signal and noise. Schurgin et al. (2020) thus suggest that while memories vary in strength continuously, the strength decrease with increasing set size is not the result of a single resource limit, but a result of many combined factors, including encoding strength, consistent with the current work.

Contribution of "Long"-Term Memory

Many studies have shown that existing knowledge or familiarity with a stimulus improves the ability to maintain information not only in long-term memory, but also over short delays, often termed long-term working memory (Ericsson & Kintsch, 1995). However, these other forms of working memory, which are thought to be, possibly, nonactive, are sometimes not considered to be core elements of working memory capacity (Awh & Vogel, 2020), even though they play major roles in most cognitive theories of working memory and are likely critical to performing everyday tasks (e.g., Cowan, 2005; Ericsson & Kintsch, 1995). Instead, the active component of working memory (sometimes referred to as the "focus of attention") is often considered particularly important (e.g., Cowan, 2005). One common question is thus the extent to which real-world object benefits arise from changes in active storage in working memory per *SE*, or from the usage of "long-term" memory systems or other forms of more passive storage that can be utilized in the short-term maintenance of information, like "activated long-term memory," or even the extent to which these concepts are truly dissociable.

We do not address this directly in the present set of studies. However, Experiment 2 does provide some indirect evidence because of the fact that the presence of distinct spatial locations versus all items being presented at the same spatial location does not modulate the object benefit very much. This provides some hints that the storage system used may be mostly based on the active component of working memory storage, because past research has shown that when items are presented at the same location sequentially, there tends to be more proactive interference across trials, whereas when items are presented at spatially distinct locations, there appears to be little lingering information trial to trial (Makovski, 2016). Given the retinotopic nature of the visual system (e.g., Golomb & Kanwisher, 2012) and the way people seem to use the visual system to hold items active in working memory (e.g., Serences, 2016), this prior work could thus be taken to indicate that in the absence of spatial cues, people may be less likely to actively hold items in mind and more likely to rely on more durable working memory traces, like activated long-term memory traces. Thus, if object benefits arose only when spatial locations were not distinct, this could be evidence that they do not arise from active storage in working memory. However, we found large and comparable benefits when spatial locations were and were not present: Regardless of whether items were presented at the same or different locations, meaningful objects were much better remembered than colors (Experiment 2), which is not in line with this possibility.

Another general concern when measuring working memory for meaningful stimuli is the contribution of not only these different more durable memory systems, but also the possibility that participants may take advantage of long encoding times to recode stimuli verbally, as verbal encoding would clearly result in fundamentally different memory traces. In the current work, to reduce the possibility of participants relying on verbal encoding, we had them perform a concurrent articulatory suppression task to hinder them from using verbal labels, and we actively monitored performance on this task in Experiments 1 and 3. Notably, the effects in the experiments with the verbal interference task actively monitored by experimenters and the replications without active monitoring of performance of this task—and thus where participants may be less inclined to perform it continuously—are similar, suggesting that verbal encoding and rehearsal strategies may not play a significant role in any of these studies.

Active Neural Representation

The present data are consistent with the results obtained in previous studies that used neural measures to assess how much information was actively held in working memory. In particular, the CDA is an EEG marker that has been associated with how much information is actively held in mind (Vogel & Machizawa, 2004). Brady et al. (2016) found that real-world objects resulted in higher memory performance than simple colors, and this performance increase was accompanied by an increase in the CDA component, suggesting it was supported by active storage in working memory. Similarly, Asp et al. (in press) showed that when participants remembered ambiguous Mooney face stimuli, memory performance and CDA were increased when participants recognized the stimuli as meaningful (i.e., a face) relative to when they just saw them as meaningless black and white shapes (see also Xie & Zhang, 2017). Critically, by asking participants whether they perceived the stimuli as a face or not, Asp et al. (in press) also showed that enhancements in active maintenance of items in visual working memory are due to the subjective perception of the stimulus as meaningful, and are not driven by physical properties of the stimulus. Evidence from fMRI is also consistent with these results: For example, Veldsman et al. (2017) found evidence of richer representations in critical working memory regions in the parietal cortex for meaningful rather than perceptually matched nonmeaningful stimuli, and Stojanoski et al. (2019) found evidence that meaningful stimuli were processed in more high-level ventral regions in preparation for visual working memory storage than perceptually matched nonmeaningful stimuli.

In contrast to this significant literature, one paper has claimed—without actually finding behavioral object benefits—that such benefits may derive from nonactive storage (Quirk et al., 2020). It is unclear what the origin of the difference between Quirk et al. (2020) and Asp et al. (in press); Brady et al. (2016); Starr et al. (2020); Stojanoski et al. (2019); and others is, although the current work suggests some possibilities, including the usage of more holistic encoding strategies in their participants and nonmatched target/foil similarity across their stimulus sets (see Brady & Störmer, 2020).

While they have taken on an outsize importance in this particular subfield, it is also not clear that thinking solely about neural measures like the CDA, rather than thinking about the cognitive operations people actually do to perform tasks (as in the origin of the term "working memory"), is actually of critical relevance when considering working memory capacity limits and how they differ for different stimuli. In theory, neither behavioral evidence based on the role of spatial locations in proactive interference nor neural markers of working memory storage like the CDA provide definitive proof that items are held actively in mind in working memory. Interpretations of the CDA, for example, are open to issues of reverse inference and circularity—the CDA is claimed to measure working memory because it tracks behavioral performance in some circumstances (e.g., Vogel & Machizawa, 2004). Then, if the CDA does not track behavior in a given instance, is this because people are not using working memory for all the information (e.g., Quirk et al., 2020), or because the CDA does not provide a pure index of working memory, but instead something more like attention (e.g., Berggren & Eimer, 2016) or does not index active maintenance memory signals that are known to be present but are not purely lateralized (Robitaille et al., 2010)? Clearly, such issues cannot be settled straightforwardly. Nevertheless, we think the evidence favors a view where behavioral benefits for meaningful stimuli are interpreted straightforwardly: People are better able to remember objects and other meaningful stimuli over short delays, and in many circumstances, this is tracked by increased neural activity consistent with working memory usage; there is evidence that stimulus-specific brain regions are engaged for working memory for meaningful stimuli alone (e.g., Druzgal & D'Esposito, 2001; Galvez-Pol et al., 2018); and such effects seem relatively unaffected by verbal encoding or diversity in spatial locations that affect proactive interference. This is altogether indicative of enhanced active storage in working memory for meaningful stimuli.

Why Real-World Objects Have a Higher Working Memory Capacity

In the current work, we investigated what conditions lead to the strongest benefits for meaningful stimuli. We found large advantages for meaningful objects in all conditions, but also found that real-world objects—and to a lesser degree lightly scrambled versions of the objects—benefit from the sequential encoding and thus deeper, focused-on-individual-items processing, while colors and nonmeaningful fully scrambled objects do not. Our results suggest that meaningless, and particularly single-feature objects, may be outliers in their affordance of parallel, quick processing, and that in more realistic memory situations, visual working memory likely relies upon representations resulting from in-depth processing of objects rather than solely being represented in terms of their low-level features. In particular, people may not solely maintain information in visual working memory in terms of colors and shapes and other "basic" visual dimensions in low-level visual cortex (e.g., Serences, 2016), but also maintain active representations of the stimuli in higher-level visual regions (e.g., FFA for face stimuli, Druzgal & D'Esposito, 2001; somatosensory regions for hand images, Galvez-Pol et al., 2018), resulting in stronger memories for these items (e.g., Asp et al., in press; Brady et al., 2016; Stojanoski et al., 2019)—at least when such meaningful features are relevant to the memory test. This may not only provide more potential sites of storage, but may also limit interference between the neural populations that must be held active, producing more distinct memories for different objects (e.g., Cohen et al., 2014; Wyble et al., 2016). Overall, then, we suggest that the working memory system can capitalize on knowledge—and connections to knowledge, enhanced by deeper processing—thereby building stronger and more robust memory representations for meaningful stimuli.

References

Adam, K. C., 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

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

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

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

Awh, E., & Vogel, E. K. (2020). Online and off-line memory states in the human brain. In Poeppel, Mangun, & Gazzaniga (Eds.), *The cognitive neurosciences* (6th ed., pp. 347–356). MIT Press.

Babic, Z., Schurgin, M. W., & Brady, T. F. (2019). Is short-term storage correlated with fluid intelligence? Strategy use explains the apparent relationship between "number of remembered items" and fluid intelligence. *PsyArXiv*. https://doi.org/10.31234/osf.io/83ch4

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience*, *7*(1), 66–80. https://doi.org/10.1162/jocn.1995.7.1.66

Bays, P. M. (2014). Noise in neural populations accounts for errors in working memory. *Journal of Neuroscience*, *34*(10), 3632–3645.

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), Article 7. https://doi.org/10.1167/9.10.7

Bays, P. M., Gorgoraptis, N., Wee, N., Marshall, L., & Husain, M. (2011). Temporal dynamics of encoding, storage, and reallocation of visual working memory. *Journal of Vision*, *11*(10), Article 6. https://doi.org/10.1167/11.10.6

Berggren, N., & Eimer, M. (2016). Does contralateral delay activity reflect working memory storage or the current focus of spatial attention within visual working memory? *Journal of Cognitive Neuroscience*, *28*(12), 2003–2020. https://doi.org/10.1162/jocn_a_01019

Brady, T. F., & Alvarez, G. A. (2015a). Contextual effects in visual working memory reveal hierarchically structured memory representations. *Journal of Vision*, *15*(15), Article 6. https://doi.org/10.1167/15.15.6

Brady, T. F., & Alvarez, G. A. (2015b). No evidence for a fixed object limit in working memory: Spatial ensemble representations inflate estimates of working memory capacity for complex objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(3), 921–929. https://doi.org/10.1037/xlm0000075

Brady, T. F., Konkle, T., & Alvarez, G. A. (2009). Compression in visual working memory: Using statistical regularities to form more efficient memory representations. *Journal of Experimental Psychology: General*, *138*(4), 487–502. https://doi.org/10.1037/a0016797

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., Gill, J., Oliva, A., & Alvarez, G. A. (2013). Visual long-term memory has the same limit on fidelity as visual working memory. *Psychological Science*, *24*(6), 981–990. https://doi.org/10.1177/0956797612465439

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., & 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

Chunharas, C., Rademaker, R. L., Brady, T. F., & Serences, J. (2019). Adaptive distortions in visual working memory. *PsyArXiv*. https://doi.org/10.31234/osf.io/e3m5a

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. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, *104*(2), 163.

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. (2005). *Working memory capacity*. Psychology Press.

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

Cusack, R., Lehmann, M., Veldsman, M., & Mitchell, D. J. (2009). Encoding strategy and not visual working memory capacity correlates with intelligence. *Psychonomic Bulletin & Review*, *16*(4), 641–647. https://doi.org/10.3758/PBR.16.4.641

Ding, S., Cueva, C. J., Tsodyks, M., & Qian, N. (2017). Visual perception as retrospective Bayesian decoding from high- to low-level features. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(43), E9115–E9124. https://doi.org/10.1073/pnas.1706906114

Druzgal, T. J., & D'Esposito, M. (2001). Activity in fusiform face area modulated as a function of working memory load. *Cognitive Brain Research*, *10*(3), 355–364. https://doi.org/10.1016/S0926-6410(00)00056-2

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*(2), 211–245. https://doi.org/10.1037/0033-295X.102.2.211

Fougnie, D., Cormiea, S. M., & Alvarez, G. A. (2013). Object-based benefits without object-based representations. *Journal of Experimental Psychology: General*, *142*(3), 621.

Galvez-Pol, A., Calvo-Merino, B., Capilla, A., & Forster, B. (2018). Persistent recruitment of somatosensory cortex during active maintenance of hand images in working memory. *NeuroImage*, *174*, 153–163. https://doi.org/10.1016/j.neuroimage.2018.03.024

Golomb, J. D., & Kanwisher, N. (2012). Higher level visual cortex represents retinotopic, not spatiotopic, object location. *Cerebral Cortex*, *22*(12), 2794–2810. https://doi.org/10.1093/cercor/bhr357

Hayhoe, M. M., Shrivastava, A., Mruczek, R., & Pelz, J. B. (2003). Visual memory and motor planning in a natural task. *Journal of Vision*, *3*(1), Article 6. https://doi.org/10.1167/3.1.6

Hollingworth, A., Richard, A. M., & Luck, S. J. (2008). Understanding the function of visual short-term memory: Transsaccadic memory, object correspondence, and gaze correction. *Journal of Experimental Psychology: General*, *137*(1), 163–181. https://doi.org/10.1037/0096-3445.137.1.163

Jackson, M. C., & Raymond, J. E. (2008). Familiarity enhances visual working memory for faces. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(3), 556–568. https://doi.org/10.1037/0096-1523.34.3.556

Lam, K., & Sprague, T. (2020). Spatial working memory performance is similar for simple stimuli and real world objects. *Journal of Vision*, *20*(11), Article 1618. https://doi.org/10.1167/jov.20.11.1618

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. https://doi.org/10.1037/xlm0000847

Lin, P. H., & Luck, S. J. (2012). Proactive interference does not meaningfully distort visual working memory capacity estimates in the canonical change detection task. *Frontiers in Psychology*, *3*, Article 42. https://doi.org/10.3389/fpsyg.2012.00042

Linke, A. C., Vicente-Grabovetsky, A., Mitchell, D. J., & Cusack, R. (2011). Encoding strategy accounts for individual differences in change detection measures of VSTM. *Neuropsychologia*, *49*(6), 1476–1486. https://doi.org/10.1016/j.neuropsychologia.2010.11.034

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

Makovski, T. (2016). Does proactive interference play a significant role in visual working memory tasks? *Journal of Experimental Psychology:*

Learning, Memory, and Cognition, *42*(10), 1664–1672. https://doi.org/10.1037/xlm0000262

Mundy, M. E., Honey, R. C., & Dwyer, D. M. (2007). Simultaneous presentation of similar stimuli produces perceptual learning in human picture processing. *Journal of Experimental Psychology: Animal Behavior Processes*, *33*(2), 124–138. https://doi.org/10.1037/0097-7403.33.2.124

Mundy, M. E., Honey, R. C., & Dwyer, D. M. (2009). Short article: Superior discrimination between similar stimuli after simultaneous exposure. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *62*(1), 18–25. https://doi.org/10.1080/17470210802240614

Nassar, M. R., Helmers, J. C., & Frank, M. J. (2018). Chunking as a rational strategy for lossy data compression in visual working memory. *Psychological Review*, *125*(4), 486–511. https://doi.org/10.1037/rev0000101

Ngiam, W. X., Khaw, K. L., Holcombe, A. O., & Goodbourn, P. T. (2019). Visual working memory for letters varies with familiarity but not complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(10), 1761–1775. https://doi.org/10.1037/xlm0000682

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

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

Robitaille, N., Marois, R., Todd, J., Grimault, S., Cheyne, D., & Jolicoeur, P. (2010). Distinguishing between lateralized and nonlateralized brain activity associated with visual short-term memory: fMRI, MEG, and EEG evidence from the same observers. *NeuroImage*, *53*(4), 1334–1345. https://doi.org/10.1016/j.neuroimage.2010.07.027

Rousselet, G. A., Thorpe, S. J., & Fabre-Thorpe, M. (2004). How parallel is visual processing in the ventral pathway? *Trends in Cognitive Sciences*, *8*(8), 363–370. https://doi.org/10.1016/j.tics.2004.06.003

Sahar, T., Sidi, Y., & Makovski, T. (2020). A metacognitive perspective of visual working memory with rich complex objects. *Frontiers in Psychology*, *11*, Article 179. https://doi.org/10.3389/fpsyg.2020.00179

Salahub, C., Lockhart, H. A., Dube, B., Al-Aidroos, N., & Emrich, S. M. (2019). Electrophysiological correlates of the flexible allocation of visual working memory resources. *Scientific Reports*, *9*(1), 1–11. https://doi.org/10.1038/s41598-019-55948-4

Salmela, V. R., Ölander, K., Muukkonen, I., & Bays, P. M. (2019). Recall of facial expressions and simple orientations reveals competition for resources at multiple levels of the visual hierarchy. *Journal of Vision*, *19*(3), Article 8. https://doi.org/10.1167/19.3.8

Schurgin, M. W., Wixted, J. T., & Brady, T. F. (2020). Psychophysical scaling reveals a unified theory of visual memory strength. *Nature Human Behaviour*, *4*, 1156–1172.

Serences, J. T. (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

Shoval, R., & Makovski, T. (2019, November). *The problem of meaning: Meaningful stimuli alter visual working memory performance* [Paper presentation]. Paper presented at Object Perception, Attention and Memory (OPAM), Montreal, Canada.

Starr, A., Srinivasan, M., & Bunge, S. (2020). Semantic knowledge influences visual working memory in children and adults. *PLOS ONE*, *15*(11), Article e0241110. https://doi.org/10.1371/journal.pone.0241110

Stojanoski, B., & Cusack, R. (2014). Time to wave good-bye to phase scrambling: Creating controlled scrambled images using diffeomorphic transformations. *Journal of Vision*, *14*(12), 6.

Stojanoski, B., Emrich, S. M., & Cusack, R. (2019). Representation of semantic information in ventral areas during encoding is associated with improved visual short-term memory. *bioRxiv*. https://doi.org/10.1101/2019.12.13.875542

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

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

Tsubomi, H., Fukuda, K., Watanabe, K., & Vogel, E. K. (2013). Neural limits to representing objects still within view. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *33*(19), 8257–8263. https://doi.org/10.1523/JNEUROSCI.5348-12.2013

Veldsman, M., Mitchell, D. J., & Cusack, R. (2017). The neural basis of precise visual short-term memory for complex recognisable objects. *NeuroImage*, *159*, 131–145. https://doi.org/10.1016/j.neuroimage.2017.07.033

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*(6984), 748–751. https://doi.org/10.1038/nature02447

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2006). The time course of consolidation in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(6), 1436–1451. https://doi.org/10.1037/0096-1523.32.6.1436

White, A. L., Runeson, E., Palmer, J., Ernst, Z. R., & Boynton, G. M. (2017). Evidence for unlimited capacity processing of simple features in visual cortex. *Journal of Vision*, *17*(6), Article 19. https://doi.org/10.1167/17.6.19

Whitney, D., & Levi, D. M. (2011). Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences*, *15*(4), 160–168. https://doi.org/10.1016/j.tics.2011.02.005

Wyble, B., Swan, G., & Callahan-Flintoft, C. (2016). Measuring visual memory in its native format. *Trends in Cognitive Sciences*, *20*(11), 790–791. https://doi.org/10.1016/j.tics.2016.08.012

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

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 follows*)

Appendix

Scrambled Stimuli Validation

To examine the relative amount of meaning that was preserved by our lightly scrambled and fully scrambled images, we ran two separate stimulus validation studies where we asked participants to recognize the images. In Experiment A2, we had participants simply freely name the images, and judged their responses. However, since previous work has clearly indicated the fully scrambled images cannot be recognized per *SE* in such conditions (i.e., participants cannot spontaneously tell that they are; Stojanoski & Cusack, 2014), we also attempted (in Experiment A1) to make a task where at least some inferences about what object it might be could be extracted even from these fully scrambled images, to allow a continuous measure of how much meaning could be extracted from the objects. In particular, in Experiment A1 we showed one image—which could be an original object image, a lightly scrambled object, or a fully scrambled object—along with two verbal labels, one of which applied to the image and one of which was instead from a different object. We asked them to, as quickly as possible, indicate which label was appropriate. We reasoned that by examining both accuracy and reaction time (RT), we can see how easily the images are recognized and how much semantic information is preserved. And, because the fully scrambled images preserve color and topology, we expected that at least some relevant information could be extracted about them in this task, allowing us to compare their overall difficulty to that of the lightly scrambled and intact objects.

Experiment A1

Method

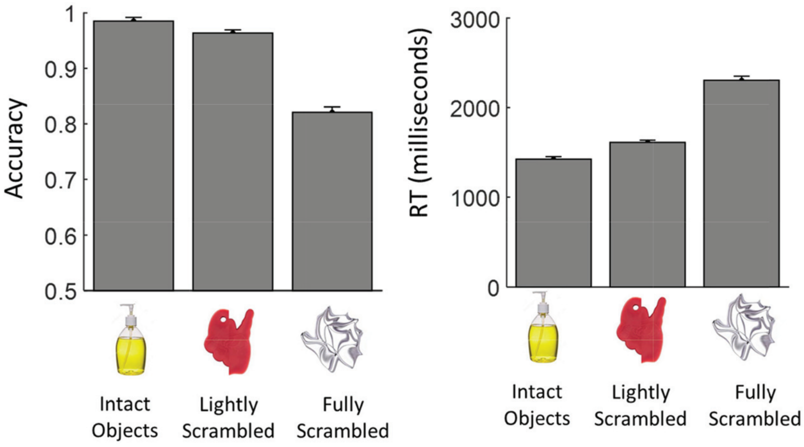
Participants. Thirty participants were recruited via Prolific. All were age 18–35, had normal or corrected-to-normal vision, and were in the United States. One participant was excluded post hoc because they were below chance at picking the appropriate label in the regular object condition, leaving 29 participants.

Procedure. On each trial, participants saw an image and two possible verbal labels (to the left and right of the image). They then indicated as quickly and accurately as possible which label was appropriate for the image using the “z” and “m” keys (for left and right, respectively). Each participant did 120 trials: 40 were original objects, 40 were lightly scrambled, and 40 were fully scrambled.

We chose the 40 objects per condition as follows: In our memory experiments, we had 240 pairs of objects that served as the tested items in each study (i.e., in the memory studies, the tested study item and foil were from a given pair). In this stimulus validation experiment, we used only 240 total object images, selecting just 1 from each pair. For each participant, we then took these 240 images and used half of them as the images that participants would be probed on, and the other 120 were given the foil labels that would be presented on each trial. This ensured the foil labels were not ones that were associated with images that were actually seen in the experiment. The 120

Figure A1

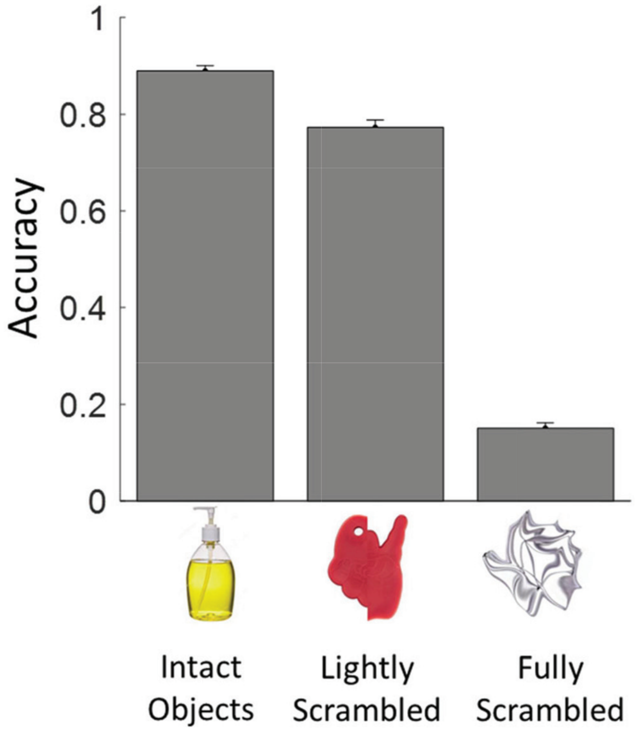
Results of Experiment A1



Note. Both accuracy and speed of matching an object with the appropriate label were strongly impacted by scrambling, with the fully scrambled objects impaired to a much greater extent than the lightly scrambled objects. See the online article for the color version of this figure.

(*Appendix continues*)

Figure A2
Results of Experiment A2



Note. Accuracy in labeling an image was strongly impacted by scrambling, with the fully scrambled objects impaired to a much greater extent than the lightly scrambled objects. See the online article for the color version of this figure.

objects that were presented to participants were then divided equally into each of the three conditions, so that the same objects would not be presented as scrambled versus intact.

Results

We found a main effect of scrambling level on accuracy, $F(2, 56) = 95.13$, $p < .0001$ (Figure A1). Intact objects were more accurately recognized than lightly scrambled objects ($t(28) = 3.13$, $p = .004$), and both were more accurately recognized than fully scrambled objects ($ps < .0001$). Similarly, in terms of median RT per participant per condition, there was a main effect of scrambling ($F(2, 56) = 120.86$, $p < .0001$), and all pairwise comparisons were significant ($ps < .0001$).

Both accuracy and speed of matching an object with the appropriate label were strongly impacted by scrambling, with the fully scrambled objects impaired to a much greater extent than the lightly scrambled objects.

Overall, the data show that the correct label for intact objects was quickly and accurately recognized. By comparison, lightly scrambled objects were recognized slightly more slowly and slightly less accurately. However, responses for fully scrambled objects were quite inaccurate and quite slow, even with only two choices, and even though the full scramble preserves color and topography as well as most low-level features.

Experiment A2

Method

Participants. Twenty participants were recruited via Prolific. All were age 18–35, had normal or corrected-to-normal vision, and were in the United States.

Procedure. On each trial, participants saw an image for only 1 second, and then typed a free response for what they believed the object to be. We emphasized that they needed to give a label for what the object is, not what it looked like (e.g., "Christmas tree," not "green"). Each participant did 120 trials: 40 were original objects, 40 were lightly scrambled, and 40 were fully scrambled. We chose the 40 objects per condition for each subject in the same way as in Experiment A1.

We graded the results by hand, completely blind to condition. For each image, we showed the responses from all subjects across all conditions, with the conditions not labeled, and marked which were accurate descriptions of the object. We judged these leniently but required an object label rather than a visual description (e.g., we accepted tree or plant for a Christmas tree). We only unblinded the conditions after grading all images for all subjects.

Results

We found a main effect of scrambling level on accuracy, $F(2, 38) = 663.7$, $p < .0001$ (Figure A2). Intact objects were more accurately recognized than lightly scrambled objects ($t(19) = 4.75$, $p < .001$, dz = 1.06), and both were more accurately recognized than fully scrambled objects ($ps < .0001$).

Discussion

In general, the data from both stimulus validation experiments suggest that lightly scrambled objects were much closer to intact objects than to fully scrambled objects, validating the idea that they retain significant amounts of meaningful information. Previous work has shown the fully scrambled objects retain little in the way of semantic information (Stojanoski & Cusack, 2014), and our results are consistent with that: Even with an extremely straightforward task of picking between two labels, where in many cases color alone may be sufficient, participants were quite inaccurate and quite slow with these fully scrambled images. When forced to name them directly, they showed extremely poor performance.

Received September 24, 2020
Revision received December 21, 2020
Accepted December 23, 2020 ■