AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: General

ISSN: 0096-3445

2025, Vol. 154, No. 9, 2432–2446
https://doi.org/10.1037/xge0001801

Limits of Verbal Labels in Cognition: Category Labels Do Not Improve Visual Working Memory Performance for Obfuscated Objects

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

Across many studies, verbal labels have been shown to significantly affect perceptual and cognitive processes. Here, we ask whether verbal labels can also improve visual working memory performance by making objects more recognizable. In a series of experiments, participants were asked to remember visual details of unrecognizable shapes derived from real-world objects. Participants were either provided with category labels prior to stimulus presentation or not. We reasoned that the labels could aid object recognition, which would then help maintain them in working memory based on recent work showing that meaningful, recognizable objects are better remembered than arbitrary visual shapes. Contrary to our hypothesis, we found no reliable increase in visual working memory performance when participants were given verbal labels. This pattern of results persisted across various experimental manipulations, including different methods of distorting the objects, testing label memory, and using within- or between-category stimulus sets. Overall, our results indicate that category labels did not readily enhance working memory performance for obfuscated shapes that are difficult to recognize, suggesting that improving visual working memory for meaningful stimuli may depend on how the visual system reorganizes incoming visual information perceptually rather than associating visual inputs with verbal knowledge. Thus, verbal labels do not seem to have as broad and general effects on cognitive functioning as previously assumed.

Public Significance Statement

Past research has shown powerful effects of language on perceptual and cognitive functioning, but it is unclear whether verbal labels can also aid storing visual objects in working memory. Based on recent work indicating that visual working memory benefits from recognizing a stimulus as a meaningful object, we investigated whether providing category labels to ease the recognition of abstract, difficult-to-recognize stimuli would strengthen their memory representations. Across multiple experiments, we found that participants who received verbal labels did not show better memory performance compared to those who did not. These results suggest that verbal labels alone do not help memory for unfamiliar or distorted objects. Our findings challenge the idea that simply adding labels can boost visual working memory, indicating that improving memory for hard-to-recognize shapes may require changes in how the incoming visual information is organized at encoding, rather than solely associating that information with category labels. This work highlights the importance of understanding how different types of information affect working memory storage and suggests that enhancing memory for less familiar objects may need more than just categorization cues.

Keywords: visual working memory, verbal working memory, semantic processing, meaningful stimuli

This article was published Online First June 30, 2025.

Ashleigh Maxcey served as action editor.

Yong Hoon Chung https://orcid.org/0009-0005-3943-7845

All data are available on the Open Science Framework (https://osf.io/nz8dh/). A subset of this work was presented at the Annual Meeting of the Vision Sciences Society in 2021 (Williams et al., 2021). The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article. Timothy F. Brady and Viola S. Störmer were supported by the National Science Foundation Directorate for Social, Behavioral, and Economic Sciences (Grant BCS 1829434).

Yong Hoon Chung played a lead role in data curation, formal analysis, writing–original draft, and writing–review and editing and an equal role in conceptualization, investigation, methodology, validation, and visualization. Lauren H. Williams played a lead role in data curation and formal analysis, a supporting role in writing–original draft, and an equal role in conceptualization, investigation, methodology, validation, and visualization. Timothy F. Brady played a lead role in funding acquisition, investigation, methodology, project administration, resources, supervision, validation, and writing–review and editing and an equal role in conceptualization. Viola S. Störmer played a lead role in funding acquisition, investigation, methodology, project administration, resources, supervision, validation, and writing–review and editing and an equal role in conceptualization.

Correspondence concerning this article should be addressed to Yong Hoon Chung, Department of Psychological and Brain Sciences, Dartmouth College, 3 Maynard Street, Hanover, NH 03755, United States. Email: yong.hoon.chung.gr@dartmouth.edu

Semantic knowledge can aid visual object recognition, speeding the identification of objects. One common and effective way semantic knowledge can be activated is through language. Previous studies have demonstrated that providing verbal labels results in faster recognition of visual objects that match the category of the label (e.g., Edmiston & Lupyan, 2015; Lupyan & Thompson-Schill, 2012), and in some cases, these effects appear to occur early on in perceptual processing, modulating visual–cortical activity within the first few hundred milliseconds of processing (Boutonnet & Lupyan, 2015; Enge et al., 2023; but also see de Leeuw et al., 2021). Thus, labels can serve as effective cues to facilitate visual object recognition. Additionally, verbal labels—and semantic knowledge more broadly—have been shown to play a crucial role in the retention of visual information in long-term memory. For example, providing labels, and thus semantic information, for images that otherwise look abstract and meaningless—for example, "droodles"—can improve later recall of these drawings (Bower et al., 1975; Nishimoto et al., 2010), though this sometimes comes at a cost of accurate memory for some semantically inconsistent visual details in favor of semantically consistent ones (e.g., Brandimonte et al., 1992; Carmichael et al., 1932; Schooler & Engstler-Schooler, 1990). Similarly, Kouststaal et al. (2003) found that ambiguous shapes were better remembered in a long-term memory task when disambiguating labels were provided that helped object recognition, though this again came with a cost in terms of higher false alarm rates for category-consistent novel stimuli. Other work by Verhaeghen et al. (2006) showed that English-speaking participants remembered novel Chinese characters better when those Chinese characters were associated with English words, making them semantically more relevant. Broadly, findings like these have been taken as evidence that verbal labels can help interpret and disambiguate equivocal visual information and that connecting visual inputs to semantic knowledge and meaning is critical for long-term retention and retrieval, even if top-down knowledge can distort memory in some ways (Bradshaw & Anderson, 1982; Brady et al., 2019; Collins & Loftus, 1975; Konkle et al., 2010; Wiseman & Neisser, 1974). This benefit of semantics on visual information storage (and in some cases verbal labels specifically) is consistent with the dual-coding framework of long-term memory that assumes stronger memories for information that can be coded in both visual and verbal formats (Paivio, 1971); an alternative explanation that has been put forward more recently is the idea that semantic or verbal information may scaffold memory for visual details in a hierarchical way, rather than being stored in separate formats (e.g., Brady et al., 2024).

Building on this prior work showing the influence of verbal labels on object recognition and the importance of semantic understanding for storing visual information in long-term memory, we here ask whether verbal labels can influence visual working memory storage, where detailed visual information is actively maintained for short periods of time. Visual working memory is often assumed to be fundamentally perceptual, reflecting primarily continued engagement of perceptual neurons in low-level visual regions like V1 (e.g., Serences, 2016), and thus is traditionally not assumed to be directly affected by top-down semantic knowledge. Similar to the dual-coding framework (Paivio, 1971) that has been dominant in long-term memory work on verbal labels, many working memory models assume that when verbal information must be stored along with visual information in working memory, verbal information and visual information are stored separately, in separate "buffers," with separate capacities (e.g., Baddeley, 1992). Thus, according to these accounts, verbal and visual information would not be stored in an integrated way but would be maintained in separate formats, potentially leading to redundant, and thus more robust, coding of information, just like dual-coding in long-term memory. Consistent with this, several studies have shown that verbal labels can improve memory for associated visual information. For instance, providing category labels for to-be-remembered simple features such as colors (e.g., saying the word "red" when encoding a red-colored circle) can improve color working memory performance compared to only relying on visual information (Souza & Skóra, 2017). This is especially true for visual features that can be easily categorized, such as colors, but more subtle for verbally ambiguous visual stimuli like orientations (Overkott & Souza, 2023). Moreover, such labeling benefits for simple features can be found across various age groups, suggesting a robust interaction between verbal and visual codes in working memory (Overkott et al., 2023).

The present study asks a separate question, namely whether verbal labels can support visual working memory for ambiguous, difficult-to-recognize objects by means of helping to recognize them. In this case, improvements in visual working memory would not be due to classic dual-coding of separate verbal and visual information, but verbal labels would help to structure and store the visual inputs in a meaningful way because they facilitate object recognition at encoding. Indeed, recent studies have shown that visual working memory performance is higher for stimuli that are recognized as meaningful relative to abstract nonmeaningful stimuli (e.g., Brady et al., 2016; Brady & Störmer, 2022, 2024; Chung, Shlipak, & Störmer, 2025; Thibeault et al., 2024; Torres et al., 2024; for reviews, see Chung, Brady, & Störmer, 2024; Hart et al., 2025). For example, Asp et al. (2021) showed that visual working memory capacity is increased for ambiguous Mooney images, two-tone patterns that can sometimes be perceived as human faces, when these ambiguous stimuli were recognized as faces and were thus meaningful to the observer (see also Williams et al., 2023). Other work has shown that once such ambiguous stimuli are recognized as meaningful, their visual representations change in various visual regions including early visual areas (Hsieh et al., 2010; Kanwisher et al., 1998). Thus, recognizing ambiguous visual inputs as meaningful fundamentally changes how they are represented in perceptual regions, which has been hypothesized to support their storage in visual working memory (Chung, Brady, & Störmer, 2024). For example, once a round red shape is recognized as an *apple*, this can activate associated semantic knowledge (e.g., apples grow on trees, they taste sweet, they are ripe in the fall), eliciting a robust and higher dimensional memory representation (Brady et al., 2024; Chung, Brady, & Störmer, 2024; Wyble et al., 2016). Importantly, these studies reveal an effect of semantic knowledge that is not due to a separate verbal subsystem because the benefit of semantic knowledge on visual working memory persists even with concurrent articulatory suppression tasks (e.g., repeating four digits throughout trials), ruling out the idea that the meaningfulness benefit arises from simple strategies of verbal rehearsal (e.g., Brady et al., 2016; Chung et al., 2023, 2025; see also Brown & Wesley, 2013).

To summarize, verbal labels have been shown to facilitate object recognition and support visual long-term memory. At the same time, recent data indicate that visual objects that connect to preexisting semantic knowledge are remembered better in visual working

memory than nonmeaningful stimuli. Based on these findings, we hypothesized that verbal labels may provide an effective way to improve performance for obfuscated visual stimuli that are not identifiable immediately but where verbal labels could help with recognizing them. This, in turn, could help participants interpret the visual input in a meaningful way and store the images not only in terms of their low-level visual features but also in terms of their semantic meaning, which, based on recent studies (e.g., Brady & Störmer, 2022), should increase memory strength. To test this, we conducted a series of experiments in which we asked participants to memorize distorted visual stimuli over a short delay period. During encoding, we either provided participants with additional verbal labels that could help them recognize the stimuli or not.

In contrast to our prediction, we consistently found that category labels did not reliably improve visual working memory performance for ambiguous visual stimuli. This shows that simply attaching verbal labels to visual stimuli that are difficult to recognize is not sufficient for improving visual working memory performance.

Transparency and Openness

The sample size, exclusion criteria, hypothesis, and analysis plan were preregistered for Experiment 1 (https://aspredicted.org/blind.php?x=LL1_NWW), Experiment 2 (https://aspredicted.org/blind.php?x=F8Z_5LF), Experiment 3 (https://aspredicted.org/blind.php?x=WVK_R4W), Experiment 3C (https://aspredicted.org/blind.php?x=GHB_R3K), and Experiment 6 (https://aspredicted.org/blind.php?x=56C_NWC). All data, materials, and example analysis codes are available online (https://osf.io/nz8dh/; Chung, Williams, et al., 2024). Data were analyzed using MATLAB (Version R2020b) and R Studio (Version 1.3.1093). All participants provided informed consent before participating in our studies, and our experimental procedures were approved by the Institutional Review Boards at the University of California San Diego (UCSD) and Dartmouth College.

Experiment 1

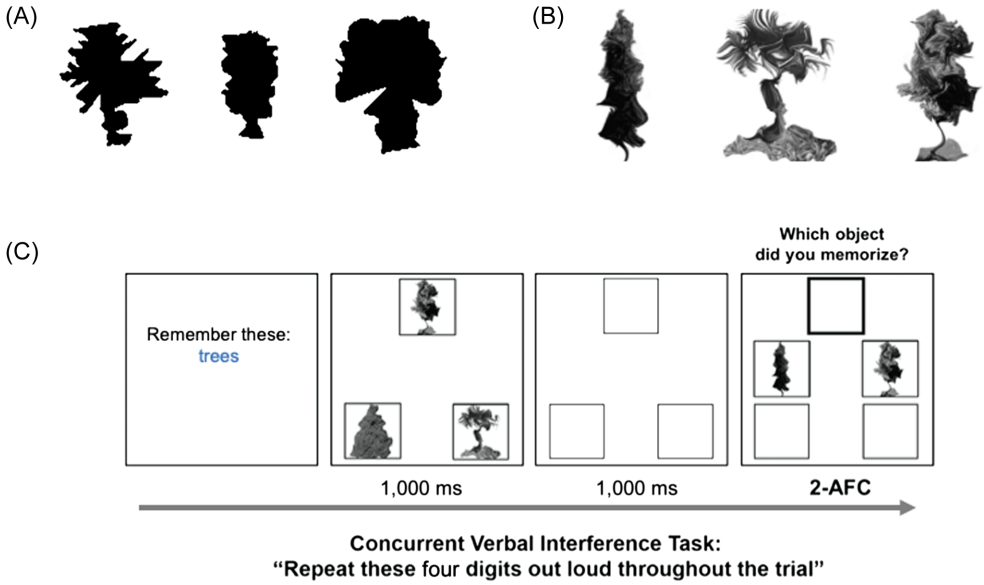
In Experiment 1, participants memorized a set of arbitrary shape silhouettes, which have been shown to be difficult to recognize when conceptual information is withheld but relatively easier to recognize when category labels are provided (Kouststaal et al., 2003; see Figure 1A for example stimuli). In both conditions, participants memorized the exact same visual stimuli. But in one condition, additional conceptual information was provided via verbal category labels (e.g., "trees"), and in the other, a generic description of them as "shapes" was provided instead. Note that as we were interested in whether semantic information would support visual memory for details, participants were always tested for their memory using a within-category foil.

Method

Participants

The final sample consisted of 200 participants (100 participants per group) recruited from the undergraduate population at the UCSD using an online recruitment platform (Sona System). Participants completed the experiment in a web browser on their own devices. Participants were excluded from the data analysis if their overall

Figure 1
Example Stimuli and Experiment Procedure



Note. (A) Tree-shaped stimuli from Kouststaal et al. (2003), used in Experiment 1. (B) Scrambled tree images created by the diffeomorphic scrambling method (Stojanoski & Cusack, 2014) used in Experiments 2–5. (C) Experimental procedure for Experiment 1. All stimulus sizes and distances are for illustration purposes only. 2-AFC = 2-alternative-forced-choice. See the online article for the color version of this figure.

memory performance (d') was below .5, if more than 10% of their trials were excluded, or if their error rate for the articulatory suppression digits task exceeded 10%. Individual trials were excluded if the response time was less than 150 ms or greater than 5,000 ms. Using these preregistered exclusion criteria, data from 72 participants were excluded and replaced.

The mean age of participants in Experiment 1 was 19.95 years (range = 18–30 years), and 196 reported as women, 63 reported as men, three reported as nonbinary, and 10 declined to report their gender.

Procedure

On each trial, half of the participants were prompted to "remember these shapes" and the other half were given a category label instead of the word "shapes" (e.g., "remember these beds," "remember these trees"). Participants were asked to click their mouse to start each trial, and the prompt stayed on the screen for 2,000 ms more after the click. Following the prompt, three ambiguous shape silhouettes (150 px height and 150 px width) were presented for 1,000 ms. The memory items for each trial were randomly drawn from a set of 10 objects from the same object category. After a 1,000-ms delay period, one object location was probed and participants were shown two test items on either side of the fixation cross for a 2-alternative-forced-choice (AFC) discrimination task. One of the items was from the memory set (target), and the other item was a novel object from the same category as the target (foil). Participants were asked to indicate which of the two items was one of the memory set items by pressing the left or right arrow key. To ensure that participants did not rely on verbal strategies to perform the visual working memory task, on every trial, participants were asked to rehearse a series of four digits out loud and report them at the end of each trial. See Figure 1C for illustration. There were 10 practice trials, followed by 125 experimental trials. Each object category was repeated five times throughout the experiment. All experimental procedures were approved by the Institutional Review Boards at the UCSD and Dartmouth College.

Statistical Analysis

As a measure of memory performance, d' was calculated for each condition using the standard formula for a 2-AFC task: $(\text{zHit} - \text{zFalseAlarm})/\sqrt{2}$. We hypothesized that the group that received the category labels would have better working memory performance for the within-category discrimination task than the group that did not receive category labels. To test this prediction, we analyzed the data using a between-participants t test. In addition to frequentist statistics, we also added Bayesian analyses to assess evidence for the null effect (Bayes factor [BF]) using a default prior of .707 (Morey et al., 2015; Rouder et al., 2009).

Results

Contrary to our prediction, working memory performance did not significantly differ between the group that received the category labels and the group that did not receive the category labels (Figure 2A), $t(198) = 0.90$, $p = .37$, Cohen's $d_z = 0.13$, 95% confidence interval (CI) [−0.15, 0.05], $BF_{01} = 4.44$. The average d' was 1.13 for the label group and 1.17 for the no-label group.

Experiment 2

In Experiment 1, providing category labels to support object recognition did not help participants remember sets of black silhouettes that have previously been shown to be ambiguous and better recognized and remembered in long-term memory with verbal labels (Kouststaal et al., 2003). However, it is possible that these simple shape silhouettes lacked critical visual features, such as texture or more complex contours, that may be important to elicit differences in memory performance between recognizable and unrecognizable objects. Thus, in Experiment 2, we used a different stimulus set of warped photo-realistic objects, which were difficult to recognize by themselves (Figure 1B).

Method

Participants

The final sample consisted of 200 participants (100 participants per group) recruited from the UCSD. Using the same preregistered exclusion criteria as in Experiment 1, data from 60 participants were excluded and replaced.

The mean age of participants in Experiment 2 was 20.47 years (range = 18–36 years), and 166 reported as women, 85 reported as men, one reported as nonbinary, and eight declined to report their gender.

Procedure and Stimuli

We used the diffeomorphic scrambling technique (Stojanoski & Cusack, 2014) to create a set of warped images that are difficult to recognize. The warped images of objects were created from a subset of Konkle et al.'s (2010) stimulus set. First, each real-world object was converted to a grayscale image. Next, the objects were warped until the point at which object recognition became difficult (37.5% of the maximum warp; see Figure 1B for example stimuli) similar to prior studies (e.g., Chung et al., 2023; Thibeault et al., 2024). Because this approach provided a larger stimulus set than Experiment 1, no object categories were repeated in the experiment. Aside from the stimulus set, Experiment 2 was identical to Experiment 1.

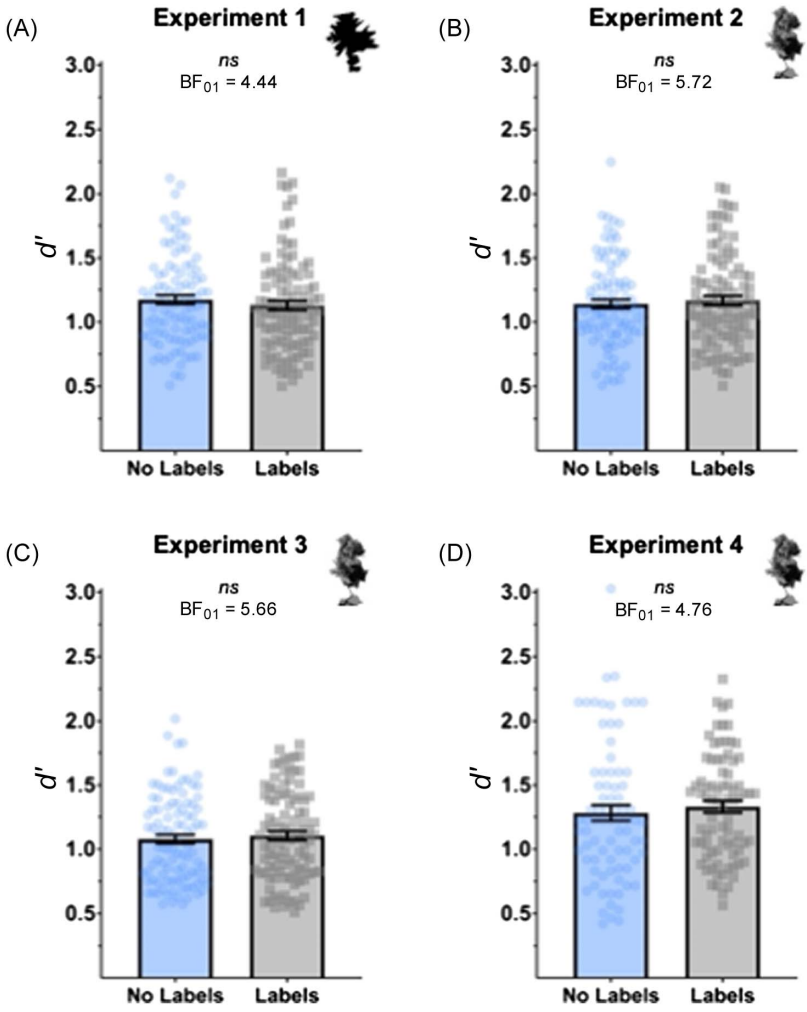
Statistical Analysis

Analysis was identical to Experiment 1.

Results

Contrary to our prediction, working memory performance again did not significantly differ between the group that received the category labels and the group that did not receive the category labels (Figure 2B), $t(198) = 0.52$, $p = .60$, Cohen's $d_z = 0.07$, 95% CI [−0.07, 0.12], $BF_{01} = 5.72$. The average d' was 1.17 for the label group and 1.14 for the no-label group.

Figure 2
Results of Experiments 1–4



Note. Mean *d*′s are plotted separately for the no-label (blue) and label (gray) conditions. Individual dots represent individual subjects' data. (A) Experiment 1: between-subjects design with the silhouette stimulus set (shown in Figure 1A). (B) Experiment 2: between-subjects design with the scrambled stimulus set (shown in Figure 1B). (C) Experiment 3: between-subjects design with the scrambled stimulus set and without the verbal interference task. (D) Experiment 4: within-subjects design with the scrambled stimulus set and the additional label 2-alternative-forced-choice task. Across all four experiments, no reliable improvements were found when participants were provided with verbal category labels to remember the distorted shapes. BF = Bayes factor; ns = nonsignificant. See the online article for the color version of this figure.

Experiment 3

In Experiments 1 and 2, participants were asked to rehearse four digits out loud throughout the trial to ensure they did not rely solely on verbal rehearsal strategies to perform the visual working memory task, as we were interested in how visual working memory per se was affected by the potential recognition of the stimuli, rather than dual-coding across different working memory subsystems. However, the concurrent verbal interference task may have prevented or discouraged participants from using the verbal category labels to improve object recognition. In Experiment 3, we thus removed the verbal interference task to address this possibility.

Method

Participants

The final sample consisted of 200 participants (100 participants per group) recruited from the UCSD. Using the same preregistered exclusion criteria as in Experiment 1, data from 47 participants were excluded and replaced.

No specific demographic information was collected in Experiment 3, but the sample is representative of the UCSD demographics reported in other experiments.

Procedure and Stimuli

Aside from the exclusion of the verbal interference task, Experiment 3 was identical to Experiment 2.

Statistical Analysis

Analysis was identical to Experiment 1.

Results

Working memory performance did not significantly differ between the group that received the category labels and the group that did not receive the category labels (Figure 2C), $t(198) = 0.55$, $p = .59$, Cohen's $d_z = 0.08$, 95% CI [−0.07, 0.12], $BF_{01} = 5.66$. The average d' was 1.10 for the label group and 1.08 for the no-label group.

We also ran two other versions of Experiment 3: one with longer encoding time (2,000 ms; Experiment 3B) and one using a phase-scrambled stimulus set (1,000 ms encoding; Experiment 3C) to further test whether the pattern of no differences would replicate with slightly different experimental parameters and stimuli. In both of these additional experiments, there were no significant differences across the label conditions—Experiment 3B: $t(38) = 1.21$, $p = .24$, $BF_{01} = 1.82$; Experiment 3C: $t(49) = 0.60$, $p = .55$, $BF_{01} = 5.56$ (see Appendix for more details).

Experiment 4

In the first three experiments, participants were provided with labels for ambiguous stimuli at the beginning of each trial. However, there was no test of whether participants remembered and used the labels throughout the memory trial. For instance, it could be that using a verbal label in a visual working memory task adds an extra cognitive burden for the participants, which may discourage them from encoding and remembering the labels. To ensure that participants were reading and memorizing the labels, in Experiment 4, we added a label memory task at the end of each trial.

Method

Participants

This was a within-subjects design where each participant experienced both label and no-label condition blocks. The final sample consisted of 80 participants recruited from the UCSD. Each label trial where the participant got the label memory task incorrect was excluded from analysis. Afterward, using the identical exclusion criteria as previous experiments, data from 26 participants were excluded and replaced. As Experiment 4 used a within-subjects design where each participant was exposed to both experimental conditions, we determined a smaller final sample size compared to previous experiments.

The mean age of participants in Experiment 4 was 20.71 years (range = 18–35 years), and 80 reported as women, 27 reported as men, and one reported as nonbinary.

Procedure and Stimuli

Participants remembered a set of warped images identical to Experiment 2. Images were presented for 2,000 ms, followed by a 1,000-ms delay and the within-category 2-AFC. Each participant was exposed to both label and no-label condition trials (each 62 trials, intermixed). On every category label condition trial, after the 2-AFC for the visual memory task, participants were prompted with a label memory task in which they were given two label options to choose. The foil label was chosen from a separate novel label list.

Statistical Analysis

Every label condition trial with an incorrect label memory response was excluded from analysis. Because Experiment 4 was a within-subjects design, a paired t test was used to compare the performance between the label condition and no-label condition. All other analysis procedures were identical to Experiment 1.

Results

Average label memory performance was 97% correct with a standard deviation of 3% after exclusion based on overall visual working memory d' and response times following previous criteria. Thus, only 3% of trials were overall excluded from the final analyses due to accuracy in the label 2-AFC task. Working memory performance on warped images did not differ significantly between the condition that received the category labels and the condition that did not receive the category labels (Figure 2D), $t(79) = 1.07$, $p = .29$, Cohen's $d_z = 0.09$, 95% CI [−0.04, 0.14], $BF_{01} = 4.76$. The average d' was 1.33 for the label condition and 1.28 for the no-label condition. Thus, even after confirming that participants know the labels and can remember them well, there was no benefit of the labels for visual working memory performance.

Experiment 5

While the ambiguous shapes used in Experiment 1 and warped real-world objects used in Experiments 2–4 were difficult to clearly recognize on their own, it is still possible that participants could have extracted some level of semantic information from these stimuli that allowed them to know the category of a given stimulus display, even in the conditions without labels. If this were the case, it may be that semantic knowledge is activated automatically during encoding (e.g., Logie, 2011) regardless of whether category labels were provided or not, leading to no reliable difference in performance between the two groups. To test this, Experiment 5 used the original, nonwarped real-world objects as stimuli, providing an upper bound of working memory performance for clearly recognizable stimuli as used in previous studies (e.g., Thibeault et al., 2024). If participants adequately recognized the obfuscated objects, or at a minimum extracted semantic meaning (even if incorrect), there should be no reliable difference in performance between warped objects (Experiment 3) and intact objects (Experiment 5).

Method

Participants

The final sample consisted of 200 participants (100 participants per group) recruited from the UCSD. Using the identical exclusion criteria as previous experiments, data from 57 participants were excluded and replaced.

No specific demographic information was collected in Experiment 5, but the sample is representative of the UCSD demographics reported in other experiments.

Procedure and Stimuli

Instead of distorted objects, intact real-world objects sampled from Konkle et al. (2010) were used in Experiment 5 (see Figure 3A for example stimuli). Aside from the stimuli, Experiment 5 was identical to Experiment 3.

Statistical Analysis

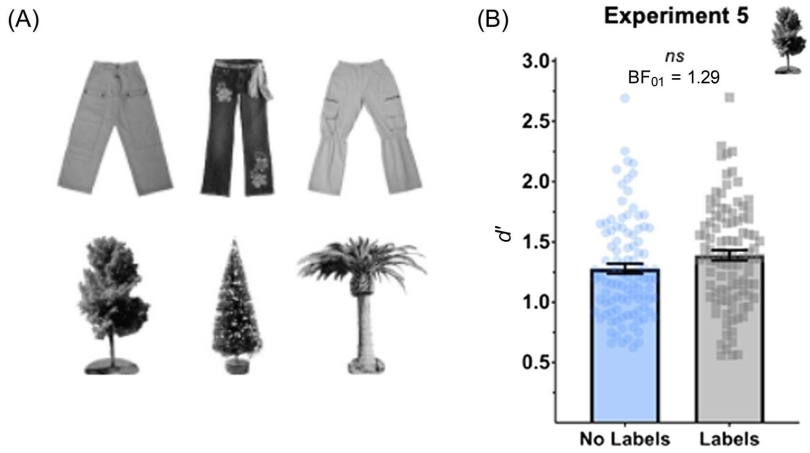
Our main interest was whether working memory performance would differ between nonwarped intact objects (Experiment 5) and the warped objects (Experiment 3). To test this, we conducted a 2 × 2 analysis of variance with stimulus types (warped vs. intact) and label groups (label vs. no-label) as between-subjects factors across Experiments 3 and 5. Otherwise, the analysis of Experiment 5 data comparing the label group and the no-label group was identical to Experiment 1.

Results

The 2 × 2 analysis of variance across Experiments 3 and 5 yielded a significant main effect of stimulus types, $F(1, 396) = 39.48$, $p < .001$, $\eta^2 = 0.089$, $BF_{01} < 0.001$, such that intact objects were better remembered compared to warped objects. There was no significant main effect of label conditions, $F(1, 396) = 3.20$, $p = .07$, $\eta^2 = 0.007$, $BF_{01} = 2.22$, and no significant interaction, $F(1, 396) = 1.21$, $p = .27$, $\eta^2 = 0.003$, $BF_{01} = 3.62$. Overall, this replicates previous findings that recognizable real-world objects can boost visual working memory performance (e.g., Brady & Störmer, 2022; Chung et al., 2023; Thibeault et al., 2024) and provides strong evidence that participants indeed could not adequately recognize the obfuscated objects that were also used in the other experiments. Moreover, this result extends beyond previous studies by demonstrating that working memory performance is improved when remembering details of real-world objects, even if all objects on a given trial are drawn from the same category.

Following all other experiments, we also directly tested the effect of labels on intact objects and found that working memory performance did not differ significantly between the group that received the category labels and the group that did not receive the category labels (Figure 3B), $t(198) = 1.86$, $p = .06$, Cohen's $d_z = 0.26$, 95% CI [−0.006, 0.23], $BF_{01} = 1.29$. However, although not reliable, we observed a numerical trend of the label group performance being higher than the no-label group in Experiment 5 ($d' = 1.39$ for the label group and $d' = 1.28$ for the no-label group). Although additional experiments will be necessary to replicate and, if reliable, fully understand this effect, the observed trend may suggest that presenting labels prior to easily recognizable objects

Figure 3
Example Stimuli and Results of Experiment 5



Note. (A) Grayscale real-world objects sampled from Konkle et al. (2010). Three objects from the same category were presented on a given trial, as in all other experiments. (B) Results of Experiment 5: between-subjects design with the intact real-world object stimulus set. There was no reliable difference between the no-label (blue) and label (gray) groups. Importantly, working memory performance was higher for these intact objects compared to the matched scrambled objects in Experiment 3 (see Figure 2). BF = Bayes factor; ns = nonsignificant. See the online article for the color version of this figure.

might prepare observers to recognize the objects more efficiently and extract relevant visual details more effectively. However, even so, this had a modest effect on visual working memory performance (Cohen's $d_z = 0.26$ when comparing the intact label group vs. the intact no-label group), especially in comparison to the performance enhancement for visually nonrecognizable versus visually recognizable objects (Cohen's $d_z = 0.53$ when comparing the intact no-label group vs. the warped no-label group).

Experiment 6

Across the first four main experiments (and two additional experiments described in more detail in the Appendix), providing verbal category labels did not improve working memory performance for distorted sets of objects in a within-category 2-AFC task. We used this within-category test because we were interested in whether the labels could improve memory for detailed visual features of the objects, rather than which category of an object participants encoded, in which case they could rely on just the labels at test. However, one possibility is that broad category labels are not useful when the task is a within-category 2-AFC that requires detailed object discrimination. In other words, the labels may be helpful for categorizing the objects as a group, but not for individuating different objects from the same category to support a fine-grained, within-category discrimination task. Thus, in Experiment 6, we tested whether using objects from three different categories in each memory display and providing distinct category labels for each item would improve working memory performance, relative to a no-label condition.

Method

Participants

The final sample consisted of 200 participants (100 participants per group) recruited from the UCSD. Using the preregistered exclusion criteria, data from 68 participants were excluded and replaced.

The mean age of participants in Experiment 6 was 20.72 years (range = 18–44 years), and 180 reported as women, 75 reported as men, three reported as nonbinary, and 10 declined to report their gender.

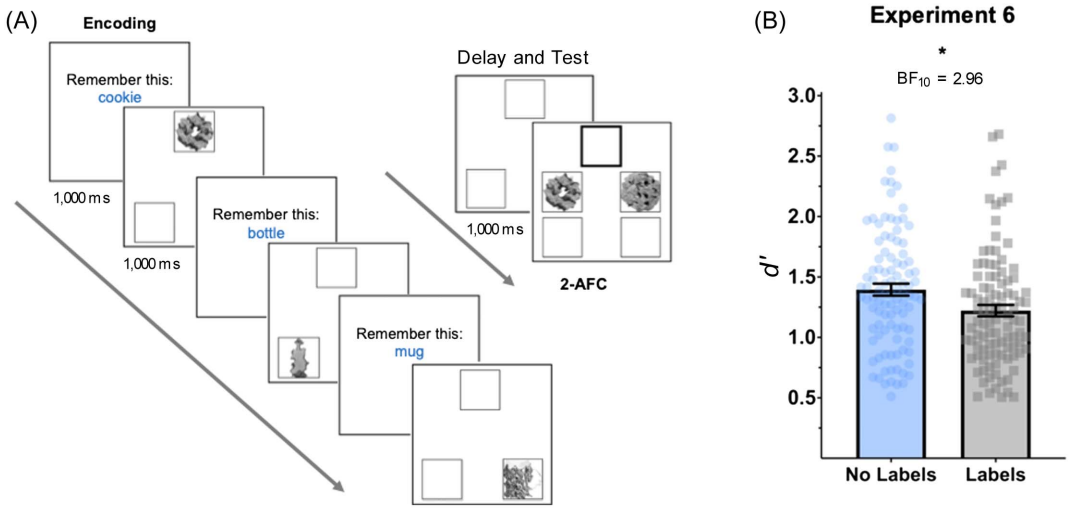
Procedure and Stimuli

Participants were asked to remember a set of distorted objects (identical to Experiment 3)—presented sequentially for 1,000 ms each over a brief delay. The memory items for each trial were randomly drawn from a set of distorted objects from different object categories. Half of the participants were prompted to "remember this shape" prior to each item, and the other half were given a category label instead (e.g., "remember this bed"). Each label was presented for 1,000 ms. After a 1,000-ms delay period, participants completed the 2-AFC task with target and foil from the same object category, identical to previous experiments. See Figure 4A for illustration.

Statistical Analysis

Analysis was identical to Experiment 1.

Figure 4
Procedure and Results for Experiment 6



Note. (A) Each to-be-remembered stimulus was sampled from different categories, and images were presented sequentially, with each image being preceded by a unique category label for the labeling group. All stimulus sizes and distances are for illustration purposes only. (B) Results of Experiment 6 depicting the mean d' separately for the no-label (blue) and label (gray) groups; individual dots represent individual subjects' data. When memory stimuli were sampled from different categories, providing category labels resulted in a significant drop in performance. 2-AFC = 2-alternative-forced-choice; BF = Bayes factor. See the online article for the color version of this figure.
* $p < .05$.

Results

Unexpectedly, working performance was better for the group that did not receive the category labels than the group that did receive the category labels (Figure 4B), $t(198) = 2.53$, $p = .01$, Cohen's $d_z = 0.35$, 95% CI [−0.31, −0.04], $BF_{01} = 0.34$ ($BF_{10} = 2.96$). The average d' was 1.22 for the label group and 1.39 for the no-label group.

What could cause this reversed result? Based on previous work showing that verbal labels can sometimes cause memory representations to shift toward the category prototype or lead to an overreliance on verbal labels rather than the visual details of the individual objects (e.g., Lupyan, 2008), it is plausible that a kind of verbal-overshadowing effect occurred, leading to performance in the label group being lower in the visual 2-AFC task, which required a fine-grained discrimination in this experiment (see Figure 4A: discriminating between two cookies requires detailed visual information about the encoded objects).

General Discussion

We tested whether providing category labels for distorted, difficult-to-recognize objects can benefit visual working memory for their shapes. Across a series of experiments, participants were either given category labels or not prior to the presentation of a set of obfuscated objects that were difficult to recognize. After a short delay period, they performed a 2-AFC task with a foil that was from the same object category as the target object. We hypothesized that even though the detailed memory test would prevent people from only relying on the broad category information provided by the verbal label (as in typical dual-coding label studies), if category labels can successfully help disambiguate the visual stimuli at encoding, this would allow participants to recognize them as meaningful objects, connect them to prior semantic knowledge, extract relevant visual features, and ultimately support storing them in visual working memory. For instance, if an observer can recognize an abstract shape as a tree with the help of the category label "trees," they could extract relevant features (i.e., shapes of branches, size of crowns, length of trunk), which would support the encoding and storage of these features, and thus the exemplar, in working memory. This should then result in better performance in a 2-AFC task where they need to remember the specific visual details of the shape, similar to what has been shown for recognizable real-world objects (e.g., Brady et al., 2016). Contrary to our prediction, however, we did not find reliable benefits of category labels in visual working memory for ambiguous objects. This was true for different types of stimuli (i.e., distorted silhouettes, warped object images, or noisy object images), long or short encoding times, and whether articulatory suppression was implemented or not. Experiment 4 further showed that participants successfully remembered the verbal labels on each trial; thus, despite them knowing the label, this knowledge did not facilitate storing the associated detailed visual information. Experiment 5 ruled out the possibility that participants recognized the obfuscated objects even without labels, just like intact real-world objects, by showing that the nonwarped, intact real-world objects resulted in reliably higher visual working memory performance compared to the warped objects. Thus, it was not the case that the warped objects were already recognized sufficiently to begin with, such that the labels could not provide additional support. Moreover, we found that providing distinct category labels for individual objects at encoding actually resulted in significantly lower working memory performance when these stimuli were sampled from different categories (while the 2-AFC test required a within-category discrimination, which again prevented a simple dual-coding benefit via remembering the label itself). Thus, we overall did not find any evidence that providing category labels improves visual working memory for difficult-to-recognize stimuli.

Verbal Labels in Visual Working Memory

Traditionally, working memory is thought to be modular in a way that verbal storage is separate from visuospatial storage (Baddeley & Hitch, 1974). Experimental results corroborate this as concurrent articulatory suppression tasks typically do not impair visual working memory performance (e.g., Luck & Vogel, 1997), and instead verbal and visual information can be stored largely in parallel (e.g., Baddeley, 1992). Other studies have shown that usage of category labels can aid visual working memory, especially for simple features such as colors that can be clearly labeled and where this verbal information, even if stored separately, can be used to improve performance through redundant coding mechanisms (e.g., Overkott & Souza, 2022, 2023). For instance, Souza and Skóra (2017) found that labeling colored circles (e.g., saying "blue" when encoding a blue circle) leads to improved memory recall compared to the control condition with no labeling. These results suggested that verbal labels can activate categorical information that can sharpen feature representation in visual working memory (Souza & Skóra, 2017), as well as providing a redundant representation of the approximate color of the stimulus (e.g., dual-coding).

In contrast, here, we found that category labels—that could not be used, on their own, to provide information about the memory test questions—did not reliably improve visual working memory performance for ambiguous, difficult-to-recognize objects. We hypothesized that by providing category labels, participants' ability to recognize the obscure shapes (e.g., a scrambled silhouette of a tree) as meaningful objects (e.g., a pine tree) would then allow them to pull out relevant visual and semantic features, strengthening their memory representation. In this case, it is critical that participants can actually utilize the category labels to *perceive* the ambiguous shapes as meaningful objects, effectively changing the representational states in the visual system, in order to improve visual memory performance. For instance, recognizing ambiguous Mooney images as meaningful can alter neural representations of the same stimuli in early visual cortices (Hsieh et al., 2010). Such perceptual reorganization, arising from the recognition of the objects as meaningful, can then support stronger memory representations in both short- (Asp et al., 2021) and long-term (Brady et al., 2019) memory tasks. Related work shows that blurry photos can be recognized much faster once the observers have already seen what the original photos are, as the participants can continue to hold on to the correct perceptual organization of the stimuli in their visual system (or can be impaired if they hold onto the wrong organization; e.g., Bruner & Potter, 1964). This principle of perceptual reorganization of initially arbitrary-looking stimuli might have not occurred with the current manipulation of providing labels, in contrast to cases like Mooney images, where participants organize the ambiguous visual information naturally themselves (Asp et al., 2021; Brady et al., 2019; Wiseman & Neisser, 1974) or with the help of "visual labels"

(i.e., previews of the stimuli; see Figure 5 for illustration; Hsieh et al., 2010) and thus vary in whether they perceive (or do not perceive) a meaningful stimulus. This would suggest that the meaningfulness benefit in visual working memory only arises when object features are reorganized in a perceptual way and are interpreted differently by the visual system and not solely by associating a label or abstract semantic knowledge with the visual input. Results from Experiment 5 further corroborate this hypothesis, showing enhanced working memory performance for easily recognizable real-world objects relative to the scrambled objects, revealing that visual recognizability facilitates detailed memory representations even for objects within the same category.

While the present study did not find a reliable benefit of category labels for obfuscated objects, future studies can incorporate other ways to more directly aid visual recognition of unfamiliar stimuli through perceptual reorganization. For instance, objects that are scrambled to a lesser degree would likely make it easier to extract meaningful features and may also be more likely to benefit from verbal labels. More generally, a more straightforward visual reference for unfamiliar objects might be helpful in facilitating visual recognition. Alternatively, previous studies showed that training participants to connect category labels with novel visual stimuli can lead to label benefits in visual recognition (Lupyan & Thompson-Schill, 2012; Souza et al., 2021). Such extensive training on the category labels and unfamiliar stimuli might ultimately lead to label benefits in visual working memory tasks as well.

Verbal Strategies in Visual Working Memory

Another aspect to consider is to what extent participants used—or tried to use—the verbal labels in the present task. Visual working memory is often thought to be closely related to perceptual processes involving early sensory regions (Adam et al., 2022; Postle, 2016; Scimeca et al., 2018; Serences, 2016), but studies have shown that nonvisual strategies can be engaged as well (Pearson & Keogh, 2019). For instance, when completing a spatial working memory task, participants can use either a verbal strategy (i.e., "thinking of a word or a phrase to describe the stimuli") or a visual strategy (i.e., "keeping the image of the spatial array as is in mind") and such difference in strategy usage is reflected in different areas of the brain engaged during the task (Sanfratello et al., 2014). Another study showed that using verbal strategies to semantically recode abstract black and white visual patterns can enhance visual working memory performance, especially when the patterns are hard to recognize as something meaningful (Brown & Wesley, 2013). However, a more recent study by Hart and Nicholls (2024) also found that simply instructing participants to engage in semantic strategies (i.e., telling participants to try to see meaningful shapes out of abstract visual patterns) did not change their overall working memory performance. Thus, it may be the case that using verbal strategies to aid visual working memory is not straightforward for participants, and just providing participants with categorical labels as done here was not sufficient to encourage the usage of those labels in the present study, even when participants were asked to explicitly remember the labels as in Experiment 4. Furthermore, in the current experiment, participants performed 2-AFC with foils that were from the same original object category as the target, requiring them to remember visual details of the shapes. This way participants could not rely on a simple strategy of remembering only the provided labels to later recall the target shape during 2-AFC but had to remember the visual details. Thus, it is possible that a more explicit way of promoting the usage of labels to recognize the abstract shapes is required to better promote working memory benefits of category labels (e.g., by allowing the labels themselves to be informative about the memory tests, like in previous dual-coding experiments but unlike in the current experiments).

Figure 5

Examples of Stimuli Used in the Present Study (Left Top and Bottom; Warped Images in Experiments 2–4 and 6 and Phase-Scrambled Images in Experiment 3C) and Mooney Face Images Used in Other Studies (Right Top and Bottom; e.g., Asp et al., 2021)

Note. Mooney face images are sampled from "Mooney Face Classification and Prediction by Learning Across Tone," T.-W. Ke, S. X. Yu, and D. Whitney, *2017 IEEE International Conference on Image Processing (ICIP) (p. 2027)*, 2017, Institute of Electrical and Electronics Engineers (https://doi.org/10.1109/ICIP.2017.8296637). Copyright 2017 by Institute of Electrical and Electronics Engineers, for illustration purposes. Asp et al. (2021) showed that once Mooney images are recognized as faces, observers can start using semantic features (i.e., age, emotional expression, gender) to increase visual working memory capacity. While the ambiguous shape stimuli were generated from real-world objects, recognizing them as meaningful objects and extracting their semantic features may be much harder to do than for Mooney images. See the online article for the color version of this figure.

Category Labels Can Sometimes Hurt Visual Working Memory

In Experiments 1–5, participants were presented and tested with objects from the same category in each trial (i.e., three different types of trees) and only given the broad category label ("trees"). The idea was that helping participants to recognize the images as trees would then more easily reveal to them the specific kinds, such as a maple, oak, and palm tree; this, in turn, would help participants maintain the specific trees and separate their representations in working memory (cf. Asp et al., 2021; Chung, Brady, & Störmer, 2024). Interestingly, when objects were from different categories during encoding and tested with a within-category foil in Experiment 6, the category labels actually impaired visual working memory performance. One possibility could be that processing the verbal labels in the label group introduced an additional cognitive load in the task; however, this seems unlikely given that all other experiments did not result in such impairment in working memory performance in the label group, especially even in Experiment 4 where participants were explicitly asked to remember the labels. Then, what could explain these results? Previous studies have shown that memory recall of abstract visual stimuli such as line drawings can be impaired when verbal labels are provided as participants rely too heavily on semantic knowledge induced by the labels and lose the visual details (e.g., Brandimonte et al., 1992; Carmichael et al., 1932; Schooler & Engstler-Schooler, 1990). A different study showed that category labels can shift representations of objects to be more prototypical of the object category (Lupyan, 2008), again suggesting that labels can change details about specific objects. Furthermore, learning categories of arbitrary stimuli can induce participants to have more frequent false alarms to novel stimuli that are consistent with category characteristics (De Brigard et al., 2017; Kouststaal et al., 2003; O'Neill et al., 2022; Yin et al., 2019). Lastly, Souza et al. (2021) showed that broad category labels can reduce the precision of color working memory, while categorically distinct labels increased the precision. These results collectively suggest that category labels for objects from different categories in Experiment 6 may have pushed memory representation to be more categorical, ultimately hurting memory performance for visual details: As we always tested participants in a within-category 2-AFC task, requiring people to remember the details of the objects (which exact cookie they saw; see Figure 4), relying mostly on the category label ("cookie") is not sufficient or could even be detrimental for memory performance as tested here.

Conclusion

This study tested whether verbal labels can improve working memory for visual details. Based on prior studies showing that (a) labels help visual object recognition and (b) recognizable objects are better remembered than unrecognizable objects, we hypothesized that associating verbal labels with ambiguous visual objects would benefit visual working memory performance. Across several high-powered studies, we found no evidence for such a benefit. Prior work has shown that visual processing and object recognition can be strongly influenced by the mere presentation of labels prior to the onset of the stimuli (e.g., Boutonnet & Lupyan, 2015; Edmiston & Lupyan, 2015); however, the current data suggest that this is not always the case, at least with respect to the processes that are critical for supporting visual working memory for objects. Thus, we conclude that providing verbal category labels alone is not sufficient to support working memory for visual details. Instead, it appears that the critical factor for improved visual working memory due to top-down knowledge is a change in the visual representations themselves due to the perceptual reorganization of the incoming visual inputs (e.g., Asp et al., 2021). Overall, this suggests that semantic understanding may not simply be linked to incoming visual information via labels but needs to be integrated and extended to early perceptual regions to change the representational format.

Constraints on Generality

We recruited young adults from the UCSD and the Prolific participant pool (ages 18 and older). We did not collect detailed demographic data for all of our experiments, but these subject pools usually consist of young participants, between the ages of 18 and 22 for UCSD participants and a little older (18–35 years) for Prolific participants. While working memory is generally known to decline throughout aging (e.g., Brockmole & Logie, 2013; Park & Reuter-Lorenz, 2009; Salthouse, 1994), and thus overall performances may vary across the adult lifespan, we expect the overall pattern of results to be comparable for younger or older adults. Participants from the UCSD were representative of the local population in terms of race, culture, and ethnicity. Although we would not expect large differences in visual working memory processes across cultures, the effect of labels and/or meaningfulness of objects could vary depending on which stimuli are used. In the present study, most stimuli were objects that are common in many societies (i.e., trees, mugs, bottles); however, there may still be some cultural differences in the ability to recognize the specific categories of objects used in the present study, as some of the objects are only found in modern environments (i.e., airplanes). This could lead to differences in results in other participant groups in the world for the specific stimuli used in the study.

References

Adam, K. C., Rademaker, R. L., & Serences, J. T. (2022). Evidence for, and challenges to, sensory recruitment models of visual working memory. In T. F. Brady & W. A. Bainbridge (Eds.), *Visual memory* (pp. 5–25). Routledge.

Asp, I. E., Störmer, V. S., & Brady, T. F. (2021). Greater visual working memory capacity for visually matched stimuli when they are perceived as meaningful. *Journal of Cognitive Neuroscience*, *33*(5), 902–918. https://doi.org/10.1162/jocn_a_01693

Baddeley, A. D. (1992). Working memory. *Science*, *255*(5044), 556–559. https://doi.org/10.1126/science.1736359

Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation*, *8*, 47–89. https://doi.org/10.1016/s0079-7421(08)60452-1

Boutonnet, B., & Lupyan, G. (2015). Words jump-start vision: A label advantage in object recognition. *The Journal of Neuroscience*, *35*(25), 9329–9335. https://doi.org/10.1523/JNEUROSCI.5111-14.2015

Bower, G. H., Karlin, M. B., & Dueck, A. (1975). Comprehension and memory for pictures. *Memory & Cognition*, *3*(2), 216–220. https://doi.org/10.3758/BF03212900

Bradshaw, G. L., & Anderson, J. R. (1982). Elaborative encoding as an explanation of levels of processing. *Journal of Verbal Learning and Verbal Behavior*, *21*(2), 165–174. https://doi.org/10.1016/S0022-5371(82)90531-X

Brady, T. F., Alvarez, G. A., & Störmer, V. S. (2019). The role of meaning in visual memory: Face-selective brain activity predicts memory for ambiguous face stimuli. *The Journal of Neuroscience*, *39*(6), 1100–1108. https://doi.org/10.1523/JNEUROSCI.1693-18.2018

Brady, T. F., Robinson, M. M., & Williams, J. R. (2024). Noisy and hierarchical visual memory across timescales. *Nature Reviews Psychology*, *3*(3), 147–163. https://doi.org/10.1038/s44159-024-00276-2

Brady, T. F., & Störmer, V. S. (2022). 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*, *48*(7), 942–958. https://doi.org/10.1037/xlm0001014

Brady, T. F., & Störmer, V. S. (2024). Comparing memory capacity across stimuli requires maximally dissimilar foils: Using deep convolutional neural networks to understand visual working memory capacity for real-world objects. *Memory & Cognition*, *52*(3), 595–609. https://doi.org/10.3758/s13421-023-01485-5

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

Brandimonte, M. A., Hitch, G. J., & Bishop, D. V. (1992). Influence of short-term memory codes on visual image processing: Evidence from image transformation tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(1), 157–165. https://doi.org/10.1037/0278-7393.18.1.157

Brockmole, J. R., & Logie, R. H. (2013). Age-related change in visual working memory: A study of 55,753 participants aged 8–75. *Frontiers in Psychology*, *4*, Article 12. https://doi.org/10.3389/fpsyg.2013.00012

Brown, L. A., & Wesley, R. W. (2013). Visual working memory is enhanced by mixed strategy use and semantic coding. *Journal of Cognitive Psychology*, *25*(3), 328–338. https://doi.org/10.1080/20445911.2013.773004

Bruner, J. S., & Potter, M. C. (1964). Interference in visual recognition. *Science*, *144*(3617), 424–425. https://doi.org/10.1126/science.144.3617.424

Carmichael, L., Hogan, H. P., & Walter, A. A. (1932). An experimental study of the effect of language on the reproduction of visually perceived form. *Journal of Experimental Psychology*, *15*(1), 73–86. https://doi.org/10.1037/h0072671

Chung, Y. H., Brady, T. F., & Störmer, V. S. (2023). No fixed limit for storing simple visual features: Realistic objects provide an efficient scaffold for holding features in mind. *Psychological Science*, *34*(7), 784–793. https://doi.org/10.1177/09567976231171339

Chung, Y. H., Brady, T. F., & Störmer, V. S. (2024). Meaningfulness and familiarity expand visual working memory capacity. *Current Directions in Psychological Science*, *33*(5), 275–282. https://doi.org/10.1177/09637214241262334

Chung, Y. H., Shlipak, K. K., & Störmer, V. S. (2025). Fluid intelligence correlates with working memory capacity for both real-world objects and simple-feature stimuli. *Journal of Memory and Language*, *143*, Article 104648. https://doi.org/10.31234/osf.io/pvbj7

Chung, Y. H., Tam, J., Wyble, B., & Störmer, V. S. (2025). Conceptual information of meaningful objects is stored incidentally. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *51*(1), 82–96. https://doi.org/10.1037/xlm0001339

Chung, Y. H., Williams, L., Brady, T. F., & Störmer, V. S. (2024). *Limits of verbal labels in cognition: Category labels do not improve visual working memory performance for obfuscated objects*. https://osf.io/nz8dh

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407–428. https://doi.org/10.1037/0033-295X.82.6.407

De Brigard, F., Brady, T. F., Ruzic, L., & Schacter, D. L. (2017). Tracking the emergence of memories: A category-learning paradigm to explore schema-driven recognition. *Memory & Cognition*, *45*(1), 105–120. https://doi.org/10.3758/s13421-016-0643-6

de Leeuw, J. R., Andrews, J., Barney, L., Bigler, M., Bruna, P. J., Chen, Y., Cherry, R., Dowie, D. R., Forbes, E., Haffey, B., Hu, X., Jaklitsch, M., Leopold, N., Lewis, C., MacDonald, D., McShaffrey, C., Monroy-Spangenberg, E., Nakayama, K., Olstad, W., … Zhang, L. (2021). Words may jump-start meaning more than vision: A non-replication of early ERP effects in Boutonnet and Lupyan (2015). *Collabra: Psychology*, *7*(1), Article 29763. https://doi.org/10.1525/collabra.29763

Edmiston, P., & Lupyan, G. (2015). What makes words special? Words as unmotivated cues. *Cognition*, *143*, 93–100. https://doi.org/10.1016/j.cognition.2015.06.008

Enge, A., Süß, F., & Abdel Rahman, R. (2023). Instant effects of semantic information on visual perception. *The Journal of Neuroscience*, *43*(26), 4896–4906. https://doi.org/10.1523/JNEUROSCI.2038-22.2023

Hart, R., Logie, R. H., & Brown Nicholls, L. A. (2025). Towards theoretically understanding how long-term memory semantics can support working memory performance. *Quarterly Journal of Experimental Psychology*, *78*(2), 370–390. https://doi.org/10.1177/17470218241284414

Hart, R., & Nicholls, L. A. B. (2024). *A semantic strategy instruction intervention aimed at boosting young and older adults' visual working memory capacity*. PsyArXiv. https://doi.org/10.31234/osf.io/qf832

Hsieh, P. J., Vul, E., & Kanwisher, N. (2010). Recognition alters the spatial pattern of FMRI activation in early retinotopic cortex. *Journal of Neurophysiology*, *103*(3), 1501–1507. https://doi.org/10.1152/jn.00812.2009

Kanwisher, N., Tong, F., & Nakayama, K. (1998). The effect of face inversion on the human fusiform face area. *Cognition*, *68*(1), B1–B11. https://doi.org/10.1016/S0010-0277(98)00035-3

Ke, T.-W., Yu, S. X., & Whitney, D. (2017). Mooney face classification and prediction by learning across tone. *2017 IEEE International Conference on Image Processing (ICIP)* (pp. 2025–2029). Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/ICIP.2017.8296637

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

Kouststaal, W., Reddy, C., Jackson, E. M., Prince, S., Cendan, D. L., & Schacter, D. L. (2003). False recognition of abstract versus common objects in older and younger adults: Testing the semantic categorization account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(4), 499–510. https://doi.org/10.1037/0278-7393.29.4.499

Logie, R. H. (2011). The functional organization and capacity limits of working memory. *Current Directions in Psychological Science*, *20*(4), 240–245. https://doi.org/10.1177/0963721411415340

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

Lupyan, G. (2008). From chair to "chair": A representational shift account of object labeling effects on memory. *Journal of Experimental Psychology: General*, *137*(2), 348–369. https://doi.org/10.1037/0096-3445.137.2.348

Lupyan, G., & Thompson-Schill, S. L. (2012). The evocative power of words: Activation of concepts by verbal and nonverbal means. *Journal of Experimental Psychology: General*, *141*(1), 170–186. https://doi.org/10.1037/a0024904

Morey, R. D., Rouder, J. N., Jamil, T., & Morey, M. R. D. (2015). *Package "bayesfactor"*. https://cran.r-project.org/web/packages/BayesFactor/BayesFactor.pdf

Nishimoto, T., Ueda, T., Miyawaki, K., Une, Y., & Takahashi, M. (2010). A normative set of 98 pairs of nonsensical pictures (droodles). *Behavior Research Methods*, *42*(3), 685–691. https://doi.org/10.3758/BRM.42.3.685

O'Neill, K., Liu, A., Yin, S., Brady, T., & De Brigard, F. (2022). Effects of category learning strategies on recognition memory. *Memory & Cognition*, *50*(3), 512–526. https://doi.org/10.3758/s13421-021-01207-9

Overkott, C., & Souza, A. S. (2022). Verbal descriptions improve visual working memory but have limited impact on visual long-term memory. *Journal of Experimental Psychology: General*, *151*(2), 321–347. https://doi.org/10.1037/xge0001084

Overkott, C., & Souza, A. S. (2023). The fate of labeled and nonlabeled visual features in working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *49*(3), 384–407. https://doi.org/10.1037/xhp0001089

Overkott, C., Souza, A. S., & Morey, C. C. (2023). The developing impact of verbal labels on visual memories in children. *Journal of Experimental Psychology: General*, *152*(3), 825–838. https://doi.org/10.1037/xge0001305

Paivio, A. (1971). Imagery and language. In S. J. Segal (Ed.), *Imagery: Current cognitive approaches* (pp. 7–32). Academic Press. https://doi.org/10.1016/B978-0-12-635450-8.50008-X

Park, D. C., & Reuter-Lorenz, P. (2009). The adaptive brain: Aging and neurocognitive scaffolding. *Annual Review of Psychology*, *60*(1), 173–196. https://doi.org/10.1146/annurev.psych.59.103006.093656

Pearson, J., & Keogh, R. (2019). Redefining visual working memory: A cognitive-strategy, brain-region approach. *Current Directions in Psychological Science*, *28*(3), 266–273. https://doi.org/10.1177/0963721419835210

Postle, B. R. (2016). How does the brain keep information "in mind"? *Current Directions in Psychological Science*, *25*(3), 151–156. https://doi.org/10.1177/0963721416643063

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian *t* tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225–237. https://doi.org/10.3758/PBR.16.2.225

Salthouse, T. A. (1994). The aging of working memory. *Neuropsychology*, *8*(4), 535–543. https://doi.org/10.1037/0894-4105.8.4.535

Sanfratello, L., Caprihan, A., Stephen, J. M., Knoefel, J. E., Adair, J. C., Qualls, C., Lundy, S. L., & Aine, C. J. (2014). Same task, different strategies: How brain networks can be influenced by memory strategy. *Human Brain Mapping*, *35*(10), 5127–5140. https://doi.org/10.1002/hbm.22538

Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology*, *22*(1), 36–71. https://doi.org/10.1016/0010-0285(90)90003-M

Scimeca, J. M., Kiyonaga, A., & D'Esposito, M. (2018). Reaffirming the sensory recruitment account of working memory. *Trends in Cognitive Sciences*, *22*(3), 190–192. https://doi.org/10.1016/j.tics.2017.12.007

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

Souza, A. S., Overkott, C., & Matyja, M. (2021). Categorical distinctiveness constrains the labeling benefit in visual working memory. *Journal of Memory and Language*, *119*, Article 104242. https://doi.org/10.1016/j.jml.2021.104242

Souza, A. S., & Skóra, Z. (2017). The interplay of language and visual perception in working memory. *Cognition*, *166*, 277–297. https://doi.org/10.1016/j.cognition.2017.05.038

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), Article 6. https://doi.org/10.1167/14.12.6

Thibeault, A. M. L., Stojanoski, B., & Emrich, S. M. (2024). Investigating the effects of perceptual complexity versus conceptual meaning on the object benefit in visual working memory. *Cognitive, Affective, & Behavioral Neuroscience*, *24*(3), 453–468. https://doi.org/10.3758/s13415-024-01158-z

Torres, R. E., Duprey, M., Campbell, K. L., & Emrich, S. (2024). Not all objects are created equal: The object benefit in visual working memory is supported by greater recollection, but only for some objects. *Memory & Cognition*. Advance online publication. https://doi.org/10.3758/s13421-024-01655-z

Verhaeghen, P., Palfai, T., & Johnson, M. P. (2006). Verbal labeling as an assimilation mnemonic for abstract visual stimuli: The sample case of recognition memory for Chinese characters. *Memory & Cognition*, *34*(4), 795–803. https://doi.org/10.3758/BF03193427

Williams, L. H., Brady, T. F., & Störmer, V. S. (2021). Category labels do not improve working memory performance for ambiguous shapes. *Journal of Vision*, *21*(9), Article 2778. https://doi.org/10.1167/jov.21.9.2778

Williams, L. H., Störmer, V. S., & Brady, T. F. (2023). *The working memory advantage for meaningful stimuli persists under high levels of proactive interference*. PsyArXiv. https://doi.org/10.31234/osf.io/e5fkz

Wiseman, S., & Neisser, U. (1974). Perceptual organization as a determinant of visual recognition memory. *The American Journal of Psychology*, *87*(4), 675–681. https://doi.org/10.2307/1421974

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

Yin, S., O'Neill, K., Brady, T. F., & De Brigard, F. (2019). The effect for category learning on recognition memory: A signal detection theory analysis. *CogSci* (pp. 3165–3171).

(*Appendix follows*)

Appendix

Additional Follow-Up Experiments to Experiment 3

Experiment 3B

In Experiments 1–3, encoding time for memory stimuli was 1,000 ms. While this is comparable to previous studies on visual working memory for meaningful objects (e.g., Brady & Störmer, 2022, 2024), it is possible that this was not enough time for participants to sufficiently use category labels to help object recognition. Thus, in Experiment 3B, the encoding time was increased to 2,000 ms in order to test whether additional processing time would lead to better working memory performance for the category labels group.

Method

Participants

The final sample consisted of 40 participants (20 participants per group) recruited from the Prolific online recruitment platform. Using the exclusion criteria from the previous experiments, data from five participants were excluded and replaced. As Experiment 3B was a follow-up to Experiment 3, a smaller sample size was used than the main experiments.

No specific demographic information was collected from the Prolific participant pool.

Procedure

In Experiment 3B, the encoding time was increased from 1,000 to 2,000 ms. Otherwise, Experiment 3B was identical to Experiment 3.

Statistical Analysis

Analysis was identical to Experiment 1.

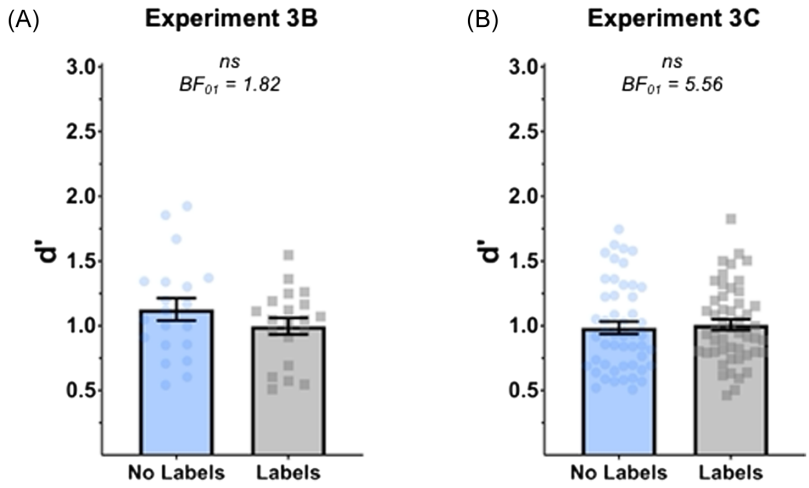
Results

Working memory performance did not significantly differ between the group that received the category labels and the group that did not receive the category labels (Figure A1A), $t(38) = 1.21$, $p = .24$, Cohen's $d_z = 0.38$, 95% confidence interval [−0.35, 0.09], $BF_{01} = 1.82$. The average d' was 0.99 for the label group and 1.13 for the no-label group.

Experiment 3C

To test whether the finding generalizes to a different stimuli set, in Experiment 3C, phase-scrambled object images were used instead of warped images.

Figure A1
Results of Experiments 3B and 3C



Note. (A) Experiment 3B: 2,000 ms encoding time. (B) Experiment 3C: phase-scrambled stimuli. BF = Bayes factor; ns = nonsignificant. See the online article for the color version of this figure.

(*Appendix continues*)

Method

Participants

The final sample consisted of 50 participants recruited from the University of California San Diego. Using the preregistered exclusion criteria, data from 20 participants were excluded and replaced. As Experiment 3C was another follow-up experiment to Experiment 3 and Experiment 3B and used a within-subjects design, a smaller sample size was used than the main experiments.

The mean age of participants in Experiment 3C was 20.68 years (range = 18–32 years), and 48 reported as women, 20 reported as men, one reported as nonbinary, and one declined to report their gender.

Procedure and Stimuli

Participants memorized a set of phase-scrambled real-world images (Figure A2). Experiment 3C used a within-subjects design, such that participants received a verbal category label (e.g., "remember these cookies") on 50% of trials (125 trials each condition, intermixed). Otherwise, Experiment 3C was identical to Experiment 3, including the encoding time of 1,000 ms.

Statistical Analysis

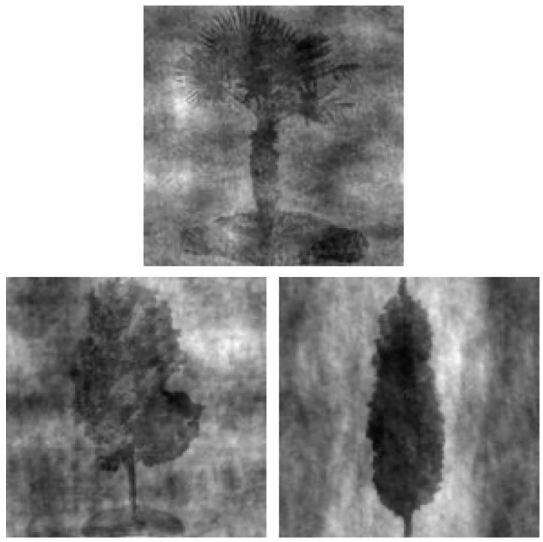
A paired *t* test was used to compare the performance between the label condition and no-label condition. All other analysis procedures were identical to Experiment 1.

Results

Working memory performance did not significantly differ between the label condition that provided the category labels and the no-label condition that did not provide the category labels (Figure A1B), $t(49) = 0.60$, $p = .55$, Cohen's $d_z = 0.08$, 95% confidence interval [−0.05, 0.09], $BF_{01} = 5.56$. The average d' was 1.01 for the label condition and 0.98 for the no-label condition.

Figure A2

Example Phase-Scrambled Stimuli Used in Experiment 3C



Received October 17, 2024
Revision received April 9, 2025
Accepted May 3, 2025 ■