Author final-version accepted for publication in the *Journal of Experimental Psychology: Human Perception and Performance*

The Fate of Labeled and Non-Labeled Visual Features in Working Memory

Clara Overkott[1] and Alessandra S. Souza[1,2]

[1] University of Zurich, [2]University of Porto

Author Note

Clara Overkott, Department of Psychology, University of Zurich, Switzerland (https://orcid.org/0000-0002-8954-7381). Alessandra S. Souza, Department of Psychology, University of Zurich, Switzerland, and Faculty of Psychology and Education Sciences, University of Porto, Portugal (https://orcid.org/0000-0002-1057-8426). This research was support by a grant from the Swiss National Science Foundation (100019_169302) awarded to A. S. Souza. We thank Gian-Luca Gubler, Natascha Zimmerli, Vanessa Vallesi, Jasmine Kerr, Olenka Dworakowski, Elena Throm, and Valentina Vogel for helping with data collection. Materials, data, and analysis scripts of all experiments, as well as preregistration of Experiment 2, can be found at: https://osf.io/z3yp3/?view_only=1fee8ba2857948c9996cc68620572732.

Correspondence concerning this article should be addressed to Alessandra S. Souza, Faculty of Psychology and Education Sciences, University of Porto, Rua Alfredo Allen, 4200-135 Porto, Portugal. E-mail: alessandra@fpce.up.pt.

Abstract

Visual objects often contain several features. Previous studies showed that verbally labeling a visual feature boosts its retention in a continuous format in visual working memory. Yet, the fate of non-labeled visual features remained unexplored. One hypothesis is that labeling induces tradeoffs in the allocation of working memory capacity across labeled and non-labeled features. To test this, we asked participants to memorize multi-feature objects (varying in color, orientation, and spatial frequency), while labeling either its (a) color, (b) orientation, or (c) spatial frequency. To inhibit labeling, they repeated “bababa” aloud in a control condition. At test, labeled and non-labeled features were reproduced using a continuous scale. Across four experiments, labeling increased continuous memory for the labeled feature, even when labels were arbitrary. Labeling aftereffects on non-labeled features were mixed: only sometimes guessing increased. These findings are inconsistent with the hypothesis that labeling induces a capacity-allocation tradeoff. Rather, costs to non-labeled features accrued when the labeling task was attentionally demanding (e.g., using less familiar or arbitrary labels). We conclude that labeling activates conceptual knowledge, thereby protecting and boosting continuous memory of the labeled feature; yet the attentional demands imposed by labeling itself can lead to the forgetting of non-labeled features.

Keywords: visual working memory, labeling, visual features, involuntary filtering, mixture modeling

Statement of Public Significance

We often verbally label the visual world around us. Previous studies showed that labeling a visual feature (e.g., color) promoted the immediate retention of fine-grained information about the labeled feature (i.e., its exact color hue). Visual objects are, however, usually composed of multiple features – for example, color, shape, orientation, size, etc – and not all features may be labeled at once. The present study showed that labeling only one of the relevant features of an object always boosted immediate memory of the labeled feature, increasing how much fine-grained information was retained in mind, while sometimes also producing costs to the non-labeled features. These costs were generated as an aftereffect of the demand of the labeling task: when labeling was hard, it consumed attentional resources that would be used to encode non-labeled features.

The Fate of Labeled and Non-Labeled Visual Features in Working Memory

Visual working memory (VWM) is the system that holds and maintains visual information available for immediate processing. VWM capacity is severely limited (see Oberauer et al., 2016 for a review). Verbal labeling can counteract these limitations: previous research demonstrated that describing visual objects, for example by saying *"this is a blue building"*, can help to retrieve detailed visual information regarding the color of the building a moment later (Overkott & Souza, 2021; Souza et al., 2021; Souza & Skóra, 2017). But, does this come at the expense of remembering other non-labeled features of the object (e.g., its size, material, shape, orientation)? Visual objects usually contain multiple features, and the more features of an object one retains in VWM, the worse the memory for each individual feature becomes (Fougnie et al., 2010; Hardman & Cowan, 2015; Oberauer & Eichenberger, 2013; Olson & Jiang, 2002; Palmer et al., 2015; Quak et al., 2018; Swan et al., 2016). This indicates that there is a common capacity limit for storing different features in VWM. Accordingly, it is possible that boosting the precision of one visual feature in VWM may come at the expense of the retention of other memory features. The main aim of the present study was to investigate whether describing one feature of a visual object can induce a tradeoff on the retention of the labeled and the non-labeled features.

What Happens When Visual Information is Labeled?

Recent research has shown that verbal labeling improves VWM (Forsberg et al., 2019, 2020; Overkott & Souza, 2021; Souza et al., 2021; Souza & Skóra, 2017). Souza and Skóra (2017) presented four colored discs sequentially for study and asked participants to either label the colors or repeatedly say “bababa” aloud to prevent labeling (aka suppression condition). During test, the colors of all four discs were reproduced on a color wheel. Color labeling improved memory performance in contrast to suppression. But what was the source of this benefit? To answer this question, the data was analyzed through computational models that can separate the differential sources of influences on the responses. Mixture modeling has been used to estimate the proportion of responses that included information retrieved from VWM (as opposed to guessing) along with the fidelity or precision of the information in memory (Bays et al., 2009; Zhang & Luck, 2008). Recently, these models have been extended to account for categorical biases (Bae et al., 2015; Hardman et al., 2017; Pratte et al., 2017). These models assume that memory responses are based either on (a) categorical information, such as memory that an item belonged to the blue or left category, or (b) continuous information, e.g., the exact hue of blue or exact direction the item pointed to. Mixture modeling in Souza and Skóra (2017) revealed that the labeling benefit accrued from two sources: (1) participants retained categorical information about more items, and (2) they also retained more continuous information: they either had a greater probability of remembering the exact color hue of a larger proportion of the memory items or they stored this continuous information more precisely.

This finding has been replicated in different task set-ups. For example, Overkott and Souza (2021) presented three colored objects sequentially and asked participants to label the color (e.g., “blue”), the object (e.g., “dog”), or the color+object binding (e.g., “blue dog”); and they contrasted these conditions to suppression. At test, participants reproduced the color of the object with the use of a color wheel. Labeling the color or the color+object binding improved performance in contrast to suppression. Mixture modeling revealed that color labeling increased the accessibility of representations (categorical and continuous) and memory precision for the continuous information, whereas color+object labeling increased the probability of storage of continuous information. Critically, in this study, labeling only the object produced costs compared to the suppression condition. This result suggests that labeling can also lead to costs depending on which feature is labeled and which feature needs to be retrieved from VWM.

To summarize, recent research suggests that verbal labeling is especially helpful for the retention of detailed information in VWM: labeling increases the probability of storing continuous information about more items or the precision with which this information is stored. Some initial evidence also points to tradeoffs between labeled and non-labeled features: labeling one feature could lead to the loss of the non-labeled feature (Overkott & Souza, 2021). However, investigating these tradeoffs were not the main goal of Overkott and Souza (2021), and it is not clear whether this happened because the labeled feature (the object's shape) was less relevant in the task given that it only served as a retrieval cue. Investigating this further will be the main goal of the present study.

Hypotheses of the Labeling Benefit

Researchers have commonly assumed that labeling would only provide categorical knowledge about the visual trace (Alogna et al., 2014; Donkin et al., 2015; Lupyan, 2008; Schooler & Engstler-Schooler, 1990; Sense et al., 2017). In the worst-case scenario, the verbal label would overshadow the visual input leading to the loss of the visual trace, and hence to less precise memory (a hypothesis known as *verbal recoding*, see Souza & Skóra, 2017). In the best-case scenario, the label would just add another source of information (i.e., categorical knowledge) with no change on the visual trace (i.e., a *dual-trace* hypothesis). This latter scenario would be consist with the idea of the multi-component model of working memory (Baddeley, 2012; Baddeley & Hitch, 1974; Logie, 2011): the label and the visual feature would be stored across different components of the system (i.e., the phonological loop and the visuospatial sketchpad). This would predict only an increase in categorical knowledge in the presence of the label with continuous information remaining constant. Neither of these hypotheses can account for the labeling benefits for the retention of continuous information described above.

To account for these findings, Souza and Skóra (2017) proposed a *categorical visual long-term memory* (VLTM) hypothesis that assumes that verbal labeling activates categorical visual information in VLTM (see also the label feedback hypothesis, Lupyan, 2012), and that this categorical knowledge facilitates the retention of continuous information in VWM. This hypothesis assumes that two visual memory traces are created in memory: one consisting of the visual representation of the presented item and the other one of the visual information that was activated in VLTM by the verbal label. As the feature category is activated in VLTM, it is possible that this either facilitates the encoding and consolidation of the visual memory trace in VWM permitting the storage of more precise continuous information regarding the labeled visual feature, or the categorical activation protects the continuous information from interference during maintenance or retrieval, sustaining it in a more robust state.

There are other hypotheses of the labeling effect, for example, that labeling would act as a cue to focus attention on the labeled feature (Kelly & Heit, 2017a). According to the *attentional-cue* hypothesis proposed by Kelly and Heit, the verbal label would be helpful when it directs attention to an item's feature making it relevant. The authors found that labeling the color of an object reduced color bias towards a color prototype in a surprise VLTM test compared to study conditions involving animacy judgement or preference rating. They argued that this was due to the label directing attention to the color feature during the study phase, which helped performance when this feature suddenly became relevant during the delayed memory test. However, this labeling effect vanished once participants were made aware of the memory test already before the study phase. This led them to argue that labeling benefits in VLTM occur when the labeled feature is incidentally encoded, but not when encoding is intentional. This hypothesis as formulated does not account for the continuous labeling benefits observed in VWM (Forsberg et al., 2020; Overkott & Souza, 2021; Souza et al., 2021; Souza & Skóra, 2017), but it can account for costs of labeling on non-labeled features (Overkott & Souza, 2021): As a consequence of directing attention to one feature, other non-attended features would be suppressed, leading to a cost.

Overkott and Souza (2021) argued that a revised version of the attentional-cue hypothesis could account for both benefits and costs of labeling: Participants may not uniformly pay attention during the memory trials - e.g., their minds may wonder (Adam et al., 2015; Adam & Vogel, 2017; Arnicane et al., 2021). Labeling could increase the amount of attention or how much time attention dwells on the labeled feature, boosting its encoding and retention because a larger share of VWM resources is assigned to the attended feature. According to this view, increasing attention to the labeled feature is coupled with the suppression of non-labeled features, and hence we should observe a cost for recall of the non-labeled features.

So far, previous studies have not evaluated whether this revised version of the *attentional-cue hypothesis* could provide a better account of the labeling effect compared to the *categorical visual long-term memory hypothesis*. One way to distinguish between these hypotheses is to assess the fate of labeled and non-labeled features: the attentional-cue hypothesis predicts that labeling benefits go hand-in-hand with costs to non-labeled features. The categorical visual long-term memory hypothesis does not assume a tradeoff: labeling simply boosts the labeled feature. Testing these predictions will be the main goal of the present study.

Next, we will briefly review the literature on the storage of multi-feature objects in VWM and the selective encoding of some features over others.

Multi-Feature Objects in Visual Working Memory

VWM performance decreases as the number of visual features stored for a given object increases (Fougnie et al., 2010; Hardman & Cowan, 2015; Oberauer & Eichenberger, 2013; Olson & Jiang, 2002; Palmer et al., 2015; Quak et al., 2018; Swan et al., 2016). This does not mean that all object's features are stored together: when participants report multiple features of the same object, errors are usually uncorrelated, indicating that visual features can be stored or retrieved independently of each other (Bays et al., 2011; Fougnie & Alvarez, 2011; Schneegans & Bays, 2017; Shin & Ma, 2017).

Several studies have shown that selective attention can be used to gate the entrance of relevant features in VWM (Bocincova et al., 2017; Bocincova & Johnson, 2019; Chen & Wyble, 2015; Maniglia & Souza, 2020; Rock et al., 1992; Serences et al., 2009; Swan et al., 2016; Yu & Shim, 2017). For example, Swan et al. (2016) displayed colored arrows and asked participants to memorize only color (see also, Shin & Ma, 2016). Hence participants had the incentive to voluntarily filter out orientation, if they could. After half of the trials, a surprise memory test on the irrelevant orientation feature followed (see also, Rock et al., 1992). Thereafter participants were asked at random to either recall the orientation or color feature – meaning that both features were now relevant for the memory test. In the surprise trial, orientation memory was poor and clearly worse than color memory on the previous trials, indicating that participants were voluntarily filtering out this feature during encoding. Mixture modeling revealed that this cost was particularly evident in the memory precision parameter, as precision of the orientation memory was low. Once participants were aware of both feature tests, orientation performance improved compared to when it was irrelevant, suggesting that once orientation was relevant participants voluntarily encoded this feature.

In sum, recent evidence is mounting in support of the view that features can be voluntarily filtered out. What is not known yet is whether this filtering process could occur involuntarily: that is, participants may inadvertently suppress or forget information that is task-relevant and they intend to maintain in mind. This could happen, for example, when a person labels only one feature of the object. Arguably, labeling directs attention to this feature thereby boosting its encoding, consolidation, or maintenance in VWM. Could this come at the expense of the other relevant yet non-labeled features of the memory item? If labeling one feature leads to the involuntary forgetting of non-labeled features, this would indicate that the labeling boost involves suppression of non-labeled information to more effectively gate the entrance of the labeled feature in VWM, indicating that labeling induces tradeoffs on how capacity is allocated in VWM.

The Present Study

Our main goal was to assess the fate of labeled and non-labeled features in VWM. In particular, we aimed to assess whether labeling of one feature would lead to the involuntary forgetting of the non-labeled features. We investigated this question across three visual features: color, orientation, and spatial frequency.

The attentional-cue hypothesis predicts that labeling of one feature will increase attention to this feature at the expense of the processing of the other feature dimension or perhaps with its active suppression, leading to costs to the non-labeled feature. To test for this possibility, we conducted four experiments (Experiments 1a, 1b, 2, and 3). The general procedure across these experiments consisted of a delayed estimation task for memory items varying on two features (colored triangles; Experiments 1 and 3) or two out of three feature combinations (Gabor patches with color, orientation, and spatial frequency; Experiment 2). During item presentation, participants were asked to label only one of the item's features, for example "green" for color, "left" for orientation, or "wide" for spatial frequency. Performance in these conditions was contrasted to a control condition, in which participants were asked to repeatedly say "bababa" aloud, thereby inhibiting verbal labeling. Finally, participants were then tested on both features of a randomly probed item. This testing procedure guaranteed that both features were relevant to the task, although only one of the features was labeled during study. This permitted us to test the degree in which labeling benefits would be coupled with corresponding costs to the non-labeled features of an object. If this is the case, then the labeling benefit will be more likely to be explained as a by-product of attentional processing: attention boosts the attended feature, while at the same time suppressing the non-attended information.

Experiment 1

In Experiments 1a and 1b, participants were presented colored triangles and they were asked to remember the color and orientation of all objects. They completed this task under suppression and under two labeling conditions that required them to either label only the colors or only the orientations of the objects. At the memory test, the color and the orientation of a single (randomly selected) object was reproduced. In Experiment 1a, a color wheel was presented for color reproduction. We replaced the color wheel with a grey wheel (i.e., a hidden color wheel) in Experiment 1b to rule out that the results in Experiment 1a could be explained by color wheel interference (Souza et al., 2016).

We modeled our data to assess how labeling affected the labeled and non-labeled features. Replicating previous research, we expected that labeling colors would increase its retention in a continuous format, and we expected to extend this finding to orientation memory. Critically, with regards to the non-labeled feature, we expected guessing to increase due to this representation being involuntarily filtered out.

To foreshadow our results, labeling improved recall of the labeled features in contrast to suppression, and mixture modeling showed that this benefit was due to the retention of more continuous information about the labeled features. Effects on the non-labeled features depended on the feature type: color was more likely to be forgotten when orientation was labeled, whereas orientation memory remained unchanged when color was labeled. This mixed pattern does not fit well the predictions of the attentional-cue hypothesis.

Methods

Transparency and Openness

All materials, analysis scripts, and data underlying all experiments reported here are available at: https://osf.io/z3yp3/?view_only=1fee8ba2857948c9996cc68620572732

Experiment 1 was not preregistered, whereas Experiment 2 was preregistered and can be found here: https://osf.io/myp8b/.

Participants

In total, 102 students from the University of Zurich were tested across Experiments 1a and 1b. Experiment 1a originally included a sample of 42 participants (*M* = 23.71; *SD* = 4.42; 25 women). Of these, five participants were excluded as they failed to follow the labeling instructions by either not labeling anything or giving the wrong type of label for the condition. Another person was excluded as the verbal label recording did not work. In total, the data of 36 participants were retained for the final analysis. Sample-size decision was as follows. We started with testing 30 participants. We then decided to increase the number of participants because the evidence obtained for the contrast of some of our conditions of interest (i.e., evidence for labeling effect on the labeled and non-labeled features) was ambiguous (i.e., Bayes Factor, BF, was between 0.33 and 3). To obtain clearer evidence for these contrasts, we added a second batch of 12 people (considering our counterbalancing across 6 participants).

Experiment 1b included a sample of 60 students (*M* = 22.87; *SD* = 3.91; 49 women). First, we obtained data of 36 participants to match the final dataset of Experiment 1a. As there was ambiguous evidence regarding a cost for the non-labeled feature in orientation recall, we increased the sample size to a total of 60 participants. One participant was excluded for not following the labeling instructions, leaving a total of 59 valid data-sets in Experiment 1b.

Only participants with German (or Swiss-German) mother tongue, aged between 18-35 years, and reporting normal color vision and normal or corrected-to-normal visual acuity could take part in the experiment. Participants had to sign an informed consent form prior to the study and were debriefed at the end. The experimental protocol was in accordance with the Institutional Review Board of the Psychology Institute from the University of Zurich and it did not require special approval.

Materials

All experiments were run in MATLAB using the Psychophysics Toolbox 3 (Brainard, 1997; Pelli, 1997). Memory items were colored isosceles triangles (see Figure 1). Colors were sampled from 360 values that varied on a continuous color circle defined in CIELAB color space, with L= 70, a= 20, b= 38, and radius= 60 (Zhang & Luck, 2008).[1] The vertex of the isosceles triangle (30° angle) pointed in directions that varied in 360 degrees. Hereafter we will refer to it as the orientation feature. The length of the side of the triangle was set to 120 pixels. Memory objects were presented within an imaginary circle with a radius of 200 pixels. The positions of the objects were determined as follows: The position of the first object was randomly selected from 360° (e.g., 45°). The remaining three objects were presented within a distance of 90°, 180° and 270° from the first one (e.g., 135°, 225°, and 315°), thereby evenly spacing the memoranda, but varying their relative position from trial to trial. The objects were presented in a sequence of two

[1] Bae et al (2014) observed that these color wheel parameters tend to generate differences between the intended and the actual rendered color values on the screen, mainly because some of the RGB values fell outside of the monitor Gamut. These could create perceptual distortions. We note, however, that all of our manipulations are within-subjects and do not depend on the presentation of specific stimuli. Hence, we believe that these differences are inconsequential to our main findings. Yet, future studies should consider using a color wheel specification that is completely within the monitor Gamut.

displays containing adjacent objects (distance of 90 degrees). The memory objects were presented against a grey background (RGB 128 128 128). In Experiment 1a, a color wheel was used for the memory test, whereas in in Experiment 1b, the color wheel was replaced by grey wheel (RGB 96 96 96).

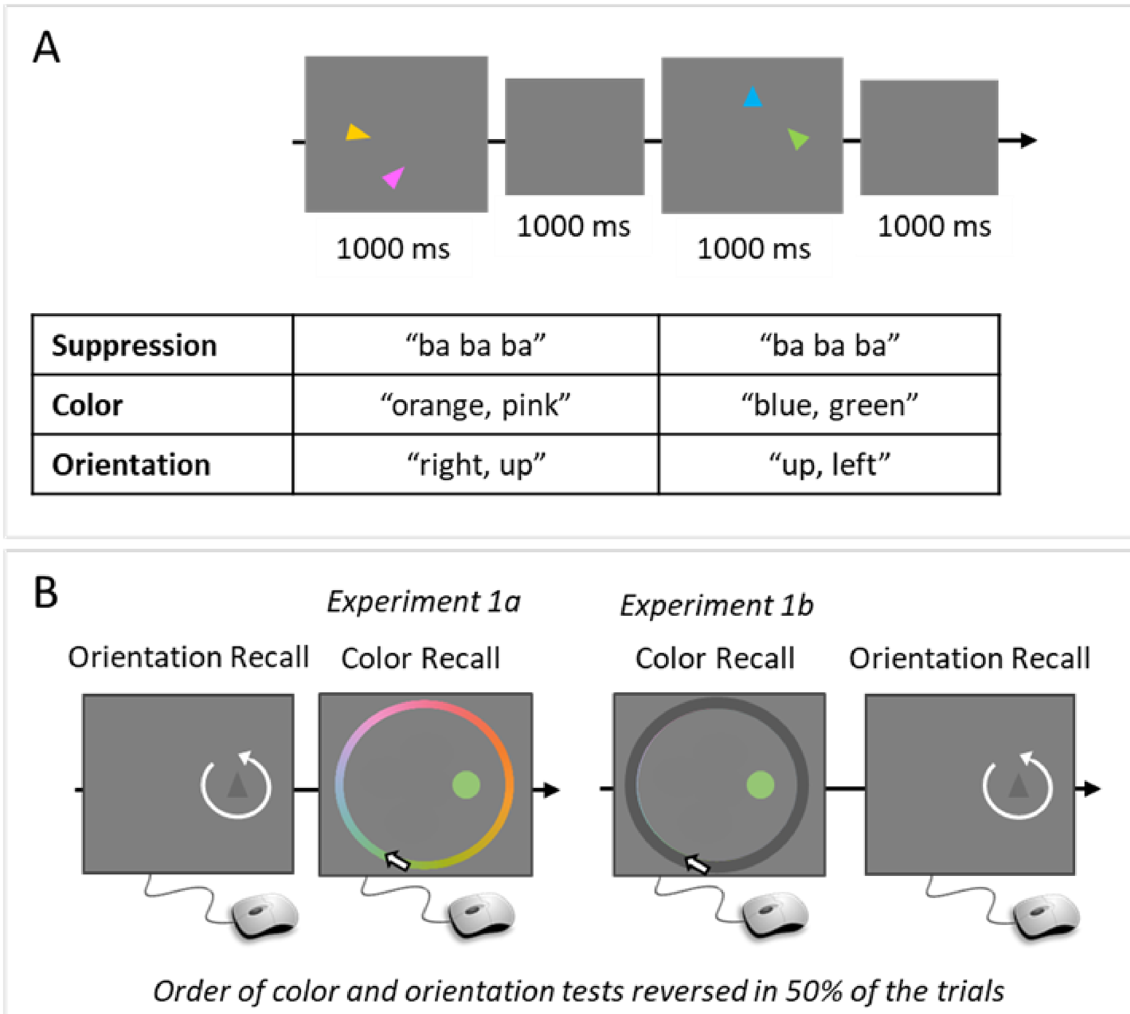
In all experiments reported here, participants were tested in individual booths where they sat approximately 50 cm from the computer screen (viewing distance was unconstrained). Participants wore a headset and were informed that their speech would be recorded in order to check for compliance with the experimental instructions.

Procedure

At the beginning of every trial, a white (RGB 255 255 255) fixation cross was presented in the center of the screen for 1000 ms. Each display was then presented for 1000 ms, followed by a blank interval of 1000 ms (see Figure 1A). In total, four memory objects were presented, two objects for each display. During memory presentation, three labeling conditions were applied: participants were asked to either (a) label the color, (b) label the orientation, or (c) say “bababa” aloud (suppression condition). Participants were free to label the features with any term they wanted. In other words, they were only told they should label the colors or orientations, and they were free to use whatever labels they preferred; there were no experimental constrains except that they should label them overtly. Their verbal output was recorded for offline check of compliance with the labeling instructions. Note that given the time constrains (2 s per display), participants only labeled the features of the last two presented items. We did not observe attempts to rehearse previously labeled items.

Figure 1

Illustration of the Experimental Procedure in Experiment 1 along with the Labeling Conditions in Effect during Study.



Note. Panel A illustrates the flow of events in the study phase and Panel B shows the recall test display for color and orientation recall (order randomly determined). Note that the labels displayed here are only illustrative (participants were free to use any terms they wanted), and that Experiments 1a and 1b differ only regarding whether the colors of the color wheel were visible or covered by a grey wheel.

For the memory test (see Figure 1B), the orientation and the color of one item had to be recalled with the use of a continuous scale. The order in which the two features were tested was counterbalanced across trials. The test item (hereafter target) was randomly selected from the first or the second display for an equal number of trials, and a memory probe was shown at the target's location. For the orientation test, a dark-grey (RGB 112 112 112) probe-triangle appeared at the location of the tested item with the mouse-cursor located in the probe's center. The initial orientation of the probe-triangle was randomly chosen from the 360 possible angles. The probe rotated once participants started moving the mouse-cursor around, with the vertex of the triangle pointing in the same direction as the mouse cursor.

For the color test, a dark-grey dot appeared at the location of the tested item together with a color wheel (in Experiment 1a) or a grey wheel (in Experiment 1b) and the mouse cursor. To change the color of the probe, participants moved the mouse cursor along the wheel which prompted the change in the probe's color. When the color wheel was visible, participants were exposed to all possible colors in one incidence, thereby creating interference. In the orientation recall test, participants only saw a single orientation at a time whilst adjusting the position of the probe. To equate the color and orientation tests, in Experiment 1b we covered the color wheel with a grey wheel: as participants moved the mouse over the grey wheel, the probe changed its color to the single one at the current mouse position. This equated the amount of interference produced at the color and orientation tests.

Participants were instructed to adjust the orientation or the color of the probe as accurately as possible, and to click with the mouse to confirm their selection. After responding to both memory tests, a message indicating to press the spacebar to initiate the next trial appeared along with a reminder of the current labeling condition (e.g., say "ba ba ba" out loud now; label the colors; label the orientations).

Labeling conditions were completed in three separate blocks, and the order of these three blocks was counterbalanced across participants. In total, 4 practice trials and 80 test trials were completed in each block, resulting in 240 test trials.

Data Analysis

Recall Performance. Recall performance was assessed by calculating the deviation between the given response and the true value of the studied item in degrees. The absolute value of the deviation can be taken as a model-free index of performance, referred to as recall error. We submitted this data to a Bayesian ANOVA (hereafter BANOVA). The advantage of a Bayesian analysis is that it quantifies the strength of the evidence for both the null and the alternative hypothesis. One commonly employed measure is the Bayes Factor (BF). The BF is the strength of evidence for one hypothesis (e.g., the alternative) over another hypothesis (e.g., the Null), given the observed data. BFs should be interpreted as a continuous index of the strength of evidence in the data in support of one model over the other and provides the factor by which the ratio of our prior beliefs should be updated in light of the data. The BF can be reported in favor of the alternative (BF_{10}) or the null (BF_{01}), where $BF_{01} = (1/BF_{10})$. A BF_{10} larger than 1 gives evidence for the alternative hypothesis (i.e., for the presence of an effect), a BF_{10} lower than 1 provides evidence against an effect. A BF_{10} of 10 indicates that the alternative hypothesis is 10 times more likely than the null. Usually, BFs > 3 are regarded as providing substantial evidence for one hypothesis over the other, BFs > 10 as providing strong evidence, and BFs > 100 as providing decisive evidence (Jeffreys, 1961; Wetzels & Wagenmakers, 2012). We computed the BFs in line with Rouder et al. (2012) and Rouder et al. (2009) by using the BayesFactor package with default prior settings (Morey

& Rouder, 2015) implemented in R (R Core Team, 2014). We also computed effect-sizes (Cohen's *d*) for the contrast of each labeling condition to suppression using the package BEST (Kruschke & Meredith, 2021), which used the following equation: $(\mu - compVal)/\delta$. μ represents the within-subjects difference between conditions, and δ the variance of these values. CompVal represents the comparison value, which was always set to 0. We report the mean and 95% credible interval of d.

Categorical-Continuous Mixture Modeling. We modeled participants' responses using the Bayesian hierarchical categorical-continuous mixture model of Hardman et al. (2017). This model is illustrated in Figure 2 applied to the recall of orientations. The model assumes that responses about a stimulus (S) are either informed by memory (M; P^M) or reflect guessing (G; 1- P^M). Responses informed by memory can further be divided into categorical (1 - P^O) or continuous (P^O) information about the memory item. Panels 2A an 2B illustrate these two types of information by plotting the relation between the studied feature value against the recalled feature value. In Figure 2A, responses cluster around four canonical values (up, left, down, right) along the feature space representing only categorical knowledge. Figure 2B shows responses that vary continuously with the studied feature (thereby falling on a diagonal indicating covariation between the two). Continuous responses can be more or less fine-grained, reflecting its continuous imprecision (σ^O). Very fine-grained responses are reflected by a dense diagonal line as depicted in Figure 2B, whereas less fine-grained responses lie around a broad diagonal line. When participants guess (G), they can do so by randomly selecting among the categories (P^{AG}), as illustrated in Figure 2C, or by randomly sampling one of the continuous values (1-P^{AG}), as shown in Figure 2D. The main model parameters (i.e., P^{M}, P^{O}, σ^{O}) are allowed to vary across experimental conditions.

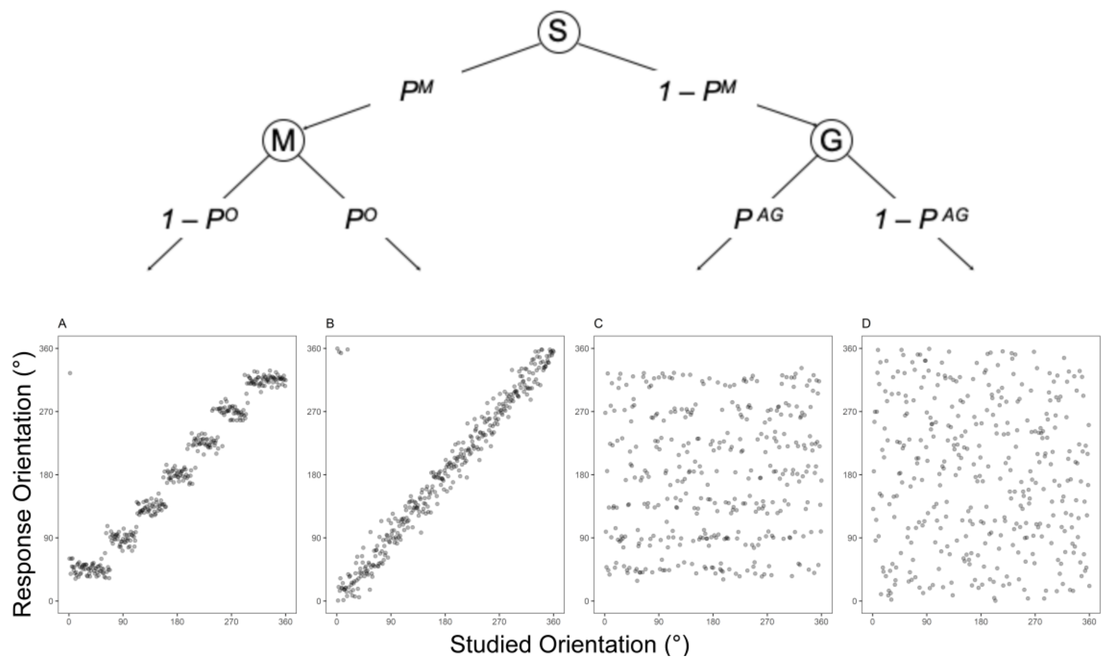
In this mixture model, the number and the locations of categories are allowed to vary between participants with a distribution describing the group-level behavior. Each category is represented by a category mean and standard deviation. Accordingly, the model does not impose an *apriori* category structure; it estimates them based on the data (the posterior of the category centers is presented in the Supplementary Materials available at https://osf.io/z3yp3/). There are parameters in the model to describe (a) categorical selectivity, namely how rapidly the responses transition from one category to the next, and (b) categorical imprecision, which reflects imprecision in selecting the category center. These parameters are freely estimated by the model, but are assumed to be constant across conditions modeled simultaneously. Given that we modeled our conditions simultaneously, the model estimated one single parameter to described the number of categories, their location (μ), the categorical imprecision, and also the categorical selectivity across all labeling conditions. For a more detailed description of the model parameters please refer to Hardman et al. (2017).

For all experiments, we fitted the between-item model of the CatContModel package (Hardman, 2016) implemented in R. In this model variant, both categorical and continuous information relative to a memory item can be held in memory at the same time. At test, however, the response is based on either categorical or continuous information. The within-item model variant, in contrast, assumes that both categorical and continuous information are integrated to inform response selection, but it has been reported to have worse model fit to the data of this task (Hardman et al., 2017; Souza &

Skóra, 2017). Hierarchical models view the parameters of individual participants in a given condition as samples from a population-level distribution. The parameter values and distributional probabilities were determined through Markov chain Monte Carlo (MCMC) sampling techniques.

Figure 2

Categorical-Continuous Model by Hardman et al. (2017) Exemplified for the Orientation Feature Space.



Note. The upper part shows the model tree, where responses of a stimulus (S) can either be informed by memory (M; with probability P^M) or reflect guessing (G; with probability 1-P^M). Memory responses are either informed by categorical representations (with probability 1-P^O) or continuous representations (with probability P^O), which can be more or less fine-grained (σ^O). Guessing is divided into categorical (P^{AG}) or continuous guessing (1-P^{AG}). The four bottom panels show the response plotted against the studied feature value. Panel A depicts examples of categorical responses, here regarding four categories (up, right, down, left). Panel B shows continuous responses which align along the diagonal, with a denser line reflecting more precise responses. Panel C reflects categorical guessing, distributed randomly along the orientation categories. Panel D shows uniform guessing.

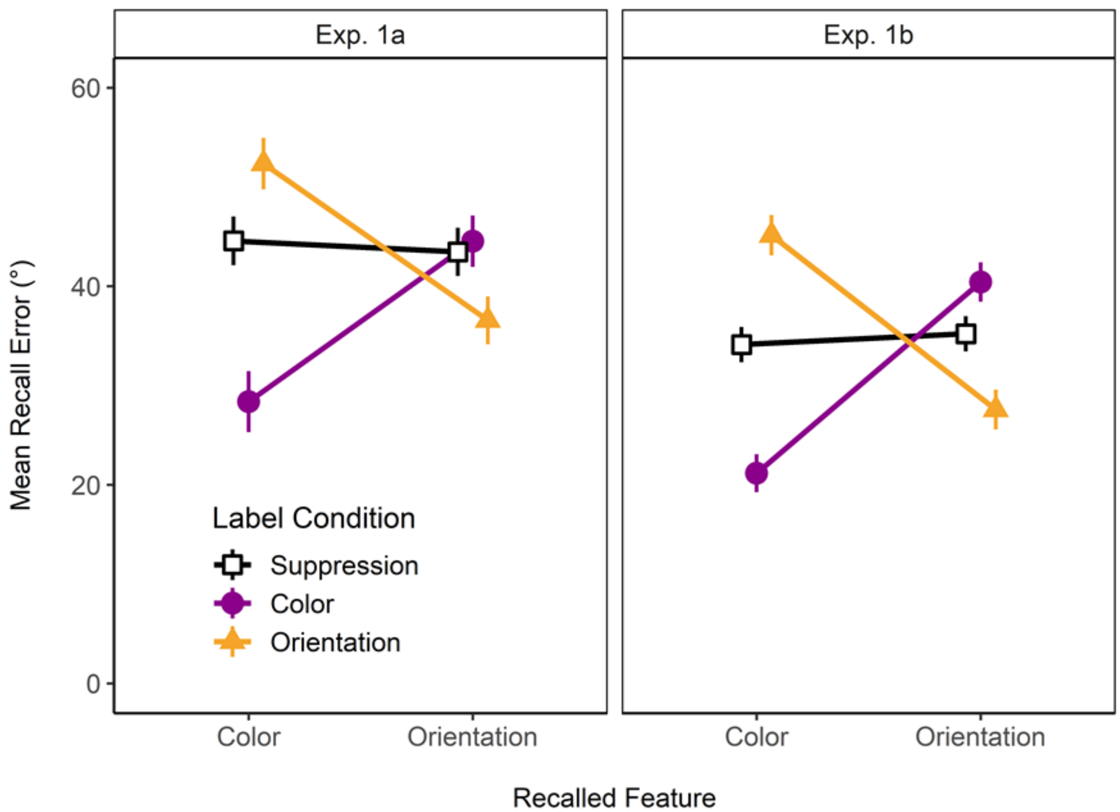
Results

Recall Performance

The aim of Experiment 1 was to investigate whether labeling one visual feature of an item boosts memory for this feature at the expense of the non-labeled features, indicating that non-labeled features were involuntarily lost. Figure 3 shows the recall error as a function of which feature was being probed (*x*-axis), with each line representing one of the labeling conditions in effect during study.

Figure 3

Mean Recall Error in Experiment 1a and Experiment 1b as a Function of the Recalled Feature and the Label Condition During Study. It Shows Lower Recall Error When the Recalled Feature Was Labeled Compared to Suppression. For Color Recall, Labeling Orientation Increased the Error Compared to Suppression in both Experiments.



Note. The error bars represent the 95% within-subjects confidence interval (Morey, 2008).

For both experiments, color recall improved (showing a smaller recall error) when color was labeled, and it showed an impairment when orientation was labeled (showing a higher recall error), in contrast to suppression. Likewise, orientation recall improved when orientation was labeled compared to suppression; whereas when color was labeled, orientation memory remained either unaffected (Experiment 1a) or a small but credible cost was observed (Experiment 1b).

Experiment 1a. To estimate the labeling effect on recall error we first ran a two-way BANOVA with labeling condition (suppression vs. color label vs. orientation label) and recalled feature (color vs. orientation) as predictors. Table S1 in the Online Supplementary Materials presents the BF of each model. The best model of the data included both main effects and their interaction ($BF_{10} = 2.32 \times 10^{20}$), and there was overwhelming evidence to keep the interaction term in the model ($BF_{10} = 1.4 \times 10^{18}$). The interaction shows that labeling has different effects depending on whether the labeled or non-labeled feature is recalled.

To estimate the labeling benefits and costs, we further ran Bayesian *t*-tests separately contrasting each labeling condition to suppression. For color recall, there was evidence for a color labeling benefit, $d = 1.44$ [95% CI: .90, 2.02], $BF_{10} = 9.27 \times 10^6$, and an orientation labeling cost, $d = -.81$ [-1.24, -.42], $BF_{10} = 815.19$, in contrast to suppression. For orientation recall, there was an orientation labeling benefit, $d = .77$ [.35, 1.21], $BF_{10} = 153.47$, and evidence against a color labeling effect, $d = -.08$ [-.43, .26], $BF_{10} = 0.21$.

Experiment 1b. In Experiment 1b, we replaced the color wheel by a grey wheel to assess whether the higher susceptibility of the color memory to labeling-induced loss observed in Experiment 1a was due to color interference produced by the color wheel at the memory test. By using the grey wheel, we equated the test interference between the color and orientation recall procedures. The same BANOVA applied to the data of Experiment 1b (see Table S1) also revealed that the best model included both main effects and their interaction (BF_{10} = 1.43 x 10^{47}) with overwhelming support for including the interaction term (BF_{10} = 2.87 x 10^{46}). Again, the presence of the interaction indicates that labeling one feature has different effects depending on whether the labeled or non-labeled feature is recalled.

We then estimated the labeling benefits and costs with Bayesian *t*-tests. For color recall, there was evidence for a color labeling benefit, d = 1.38 [.97, 1.76], $BF_{10} = 7.30 \times 10^{11}$ and for an orientation labeling cost, d = -1.15 [-1.56, -.75], $BF_{10} = 2.86 \times 10^{8}$, in contrast to suppression. For orientation recall, there was evidence for an orientation labeling benefit, d = .75 [.46, 1.05], $BF_{10} = 4.77 \times 10^{4}$, and for a color labeling cost, d = -.53 [-.83, -.24], BF_{10} = 123.83, in contrast to suppression. The finding of a color labeling cost for orientation memory stands in contrast to the result of Experiment 1a, where we found some evidence against a color labeling cost. The difference between the two experiments was that in Experiment 1b, the color wheel was hidden under a grey wheel to reduce color interference. Indeed, removal of the color wheel was associated with overall better performance particularly in the suppression condition (compare Experiment 1a to Experiment 1b). Results of Experiment 1b therefore show that the higher susceptibility of color memory to involuntary labeling-induced loss cannot be accounted by this feature being more prone to interference from the test situation. If anything, this modification made orientation memory more susceptible to labeling costs.

Categorical-Continuous Mixture Modeling

We modeled the data of the three labeling conditions (i.e., suppression, color labeling, and orientation labeling) simultaneously in each experiment. However, we ran separate models for recall of color and orientation given that categorical biases in these two feature domains can differ. For each model, we ran 10,000 iterations, of which 2,000 were regarded as burn-in. Scatterplots presenting the recalled feature against the studied feature for all experiments reported here are available in the Online Supplementary Materials. These figures allow to visualize the changes in performance produced by labeling that the modeling is characterizing, namely changes in guessing, continuous and categorical memory.

We proceeded with model comparison to determine which model parameters were affected by labeling. First, we ran a full model allowing for an effect of labeling condition on all three main parameters of the model (i.e., P^M, P^O, and σ^O). In subsequent steps, we removed the effect of labeling from each of the parameters and combinations of parameters as indicated in Table 1. For each model, we computed the Watanabe-Akaike Information Criterion (WAIC). The WAIC is used to assess the model fit based on the model predictive accuracy and includes a correction for the number of parameters used in the model (Gelman et al., 2014). This penalty term was helpful here as we constrained our models in regard to the number of parameters. The model with the smallest WAIC is the one that best explains the data.

As shown in Table 1, the best model for color recall in Experiment 1a included an effect of labeling on all three parameters of the model. For orientation recall, the best model did not include an effect of labeling condition on the continuous imprecision parameter. In Experiment 1b, for both color and orientation recall, the best model did not include an effect of labeling condition on continuous imprecision. However, for both recall tests, this model was only favored by a WAIC difference of 1 in comparison to the full model including all fixed predictors, which may indicate ambiguous evidence to exclude the effect of labeling on continuous imprecision.

We then assessed the group-level posterior estimates in each condition. We report the posterior estimates of the model including an effect of labeling on all parameters, even when this was not the best model because this allowed us to see the variability in the posterior of the parameters. Here, we report four parameters that were allowed to vary between conditions: categorical and continuous memory, guessing, as well as continuous imprecision. To assess continuous memory, we calculated $P^M \times P^O$ – this value reflects the proportion of responses informed by continuous memory representations. Likewise, categorical memory was assessed by calculating $P^M \times (1\text{-}P^O)$ – reflecting the remaining proportion of memory responses that were informed by categorical information. Finally, guessing was computed as $1\text{-}P^M$. To illustrate this, suppose that the model estimates that $P^M = 0.80$, and $P^O = 0.30$. This indicates that continuous memory representations informed 0.24 of the responses, categorical memory informed 0.56 of the responses, whereas the remaining 0.20 of the responses reflected guessing. The continuous imprecision parameter (σ^O) was reported as outputted by the model, and it reflects the imprecision of the continuous memory. All reported models fitted the obtained data well (see Online Supplementary Materials). Table S2 in the Online Supplementary Materials presents group-level estimates for continuous memory, categorical memory, and memory imprecision in each labeling condition.

Table 1

WAIC Comparison for all Fitted Mixture Models to the Data of Experiments 1a, 1b, and 2

		Color Model Labeling Effect on:					Orientation Model Labeling Effect on:				
		P^M	P^O	σ^O			P^M	P^O	σ^O		
Exp.	Model				WAIC	ΔWAIC				WAIC	ΔWAIC
E1a	1	✓	✓	✓	**89396**	**0**	✓	✓	✓	88838	11
	2		✓	✓	89747	351		✓	✓	88845	18
	3	✓		✓	89745	349	✓		✓	88882	55
	4	✓	✓		89400	4	✓	✓		**88827**	**0**
E1b	1	✓	✓	✓	140903	1	✓	✓	✓	141283	1
	2		✓	✓	141615	712		✓	✓	141442	160
	3	✓		✓	140919	16	✓		✓	141314	32
	4	✓	✓		**140902**	**0**	✓	✓		**141282**	**0**
E2	1	✓	✓	✓	**121250**	**0**	✓	✓	✓	**111647**	**0**
	2		✓	✓	121708	458		✓	✓	111724	74
	3	✓		✓	121250	0	✓		✓	111732	82
	4	✓	✓		121268	18	✓	✓		111650	3

Note. Δ indicates the difference score for this particular model in comparison to the best model. P^M = probability that information is in memory, P^O = probability of continuous information, σ^O = continuous imprecision.

Figure 4 presents posterior differences between each labeling condition in contrast to the suppression condition, thereby indicating how labeling modulated the retention of the labeled and non-labeled features. The posterior differences are presented with their mean (point) along with its highest density interval (HDI). The HDI reflects the range of values that covers 95% of the posterior. The zero represents no difference between the labeling condition and the suppression condition. When the HDI does not include zero, the labeling condition credibly differs from the suppression condition. Values above zero indicate a labeling benefit and values below 0 a labeling cost.

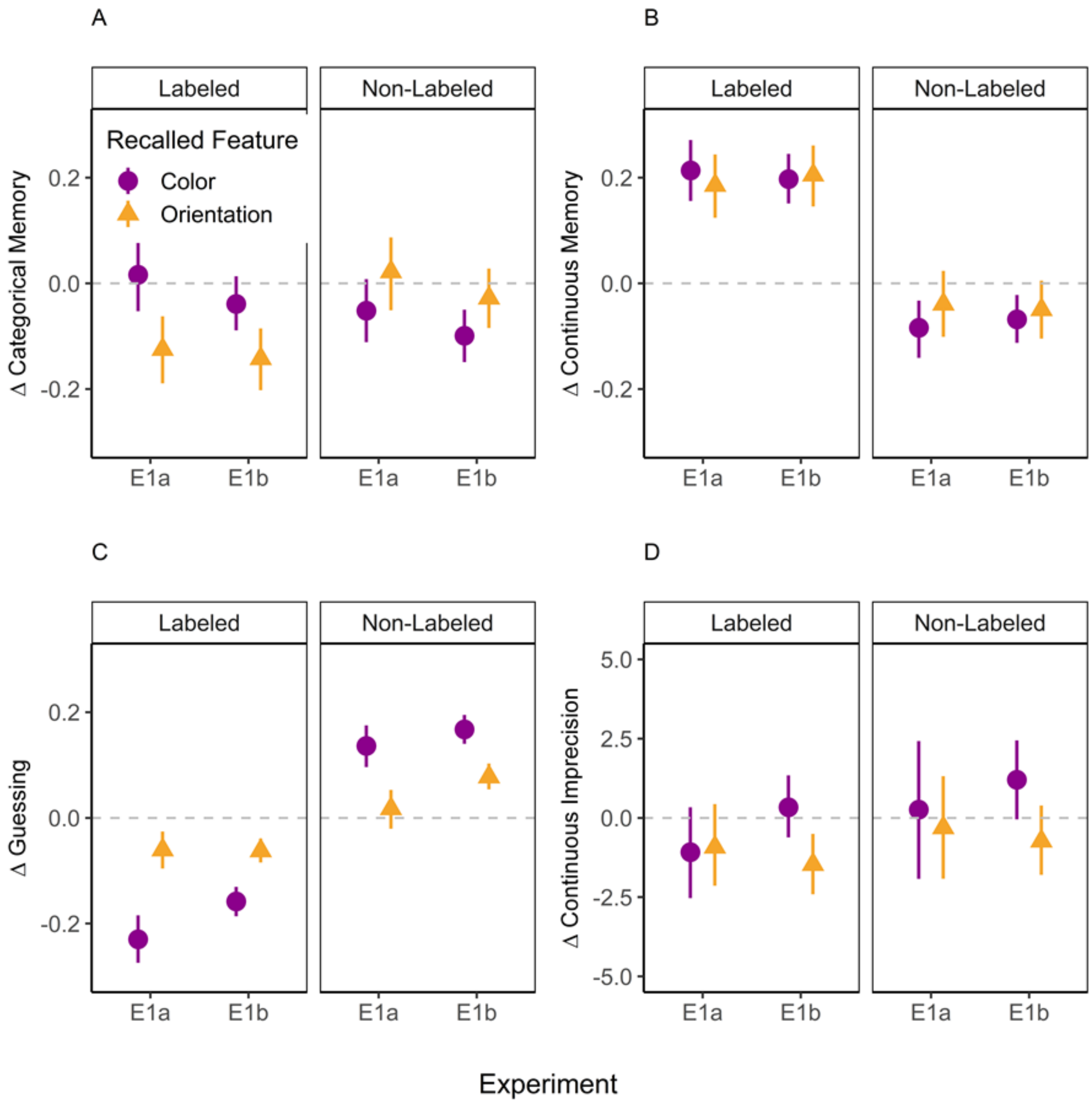
Figure 4A presents changes in categorical memory for the labeled and non-labeled features. With regards to the labeled feature, there was no change in categorical memory for color, but a reduction for orientation. For the non-labeled feature, there was no credible change in comparison to suppression for orientation. For color, categorical memory tended to get reduced, but this effect was only credible in Experiment 1b.

Figure 4B presents changes in continuous memory. Continuous memory for the labeled feature always improved, whereas continuous memory for the non-labeled feature tended to decrease. Yet, the latter was only credible when the recalled feature was color.

Figure 4C presents changes in guessing. Guessing decreased for the recall of the labeled features. This shows that although categorical memory for the labeled orientation was reduced, this was a by-product of labeling boosting orientation continuous memory. For the non-labeled feature, guessing increased for color but not orientation, corroborating that only color memory suffered when the other feature was labeled.

Figure 4

Changes in the Mixture Model Parameters for the Labeled and Non-Labeled Features (Different Sub-Panels) in Experiments 1a and 1b. In Sum, When the Recalled Feature was Labeled, Continuous Memory (Panel B) Credibly Increased, Whereas Guessing (Panel C) Credibly Decreased for Both Features. When the Recalled Feature was Non-Labeled, Decreases in Categorical (Panel A) and Continuous Memory (Panel B), as well as an Increase in Guessing (Panel C) Were Observed for Color Memory Only.



Note. Dots depict the mean difference of the posterior distributions in the labeling conditions (color or orientation labeling) in relation to the suppression baseline. The error bars depict the 95% HDI of the difference. Panel A reflects the changes in the probability of retrieving categorical information, and Panel B of continuous information. Panel C presents changes in guessing, and Panel D changes in the estimates of continuous memory imprecision.

Finally, Figure 4D presents changes in continuous imprecision. In line with the WAIC analysis, there were hardly any credible differences for continuous imprecision across the labeling conditions when compared to the suppression. Only for Experiment 1b, labeling orientation credibly reduced orientation memory imprecision.

Discussion

In Experiment 1 we showed that labeling one feature of a multi-feature object can have two types of outcomes: (1) it can benefit the labeled feature by decreasing guessing while increasing the probability that participants make a response based on a continuous memory representation about the labeled feature, and (2) it can produce costs for the non-labeled feature: categorical and continuous information about the non-labeled feature may get lost, increasing guessing. This was the case for color memory, but not credibly for orientation memory. More specifically, we found that labeling orientation led to some cost for categorical color memory (see Exp. 1b) and a very credible cost in continuous color memory (Exps. 1a and 1b) - indicating that fine-grained information about the color feature was involuntarily lost when orientation was labeled. In contrast, color labeling did not lead to a credible cost for either categorical or continuous information about orientation as revealed by the parameters in the mixture model. There was, however, a credible, albeit small, cost for color labeling on orientation recall error (a raw index of performance) in Experiment 1b. These results suggest that orientation memory was less likely to be involuntarily lost than color memory. We also showed that this higher cost for color memory was not explained by color memory becoming more susceptible to color interference at test, as reducing color interference at test by using a grey wheel in Experiment 1b did not change the pattern of results.

These asymmetric effects of color and orientation labeling on recall performance indicate that some visual features may be involuntarily lost when other features are labeled. So far, however, we do not know whether involuntary loss of visual features is the norm and orientation is an exception, or whether the reverse is more likely with color information being particularly vulnerable to involuntary forgetting.

Experiment 2

The goal of Experiment 2 was to test the impact of labeling other visual features besides color and orientation. This experiment used Gabor patches that could vary in three visual dimensions (i.e., spatial frequency, color, and orientation). By adding the third visual feature (namely spatial frequency), we aimed to test whether labeling this feature could boost its retention in memory at the expense of other visual features, and whether labeling the other features (color or orientation) would lead to a cost for spatial frequency memory. We aimed to assess the likelihood of two hypotheses: (H1) = Labeling enhances memory for the labeled feature at the expense of the non-labeled feature; (H2) = labeling enhances memory for the labeled feature with no costs for the non-labeled features. Or, a mix of the two depending on the visual feature. We also hoped this would provide further insight regarding to what types of visual features are susceptible to a labeling cost.

We preregistered our hypotheses (https://osf.io/2spwt/) [2]: we expected to replicate the results obtained in Experiment 1, namely that labeling spatial frequency

[2] In the preregistration, we mentioned that the results of Experiment 1a and 1b were similar. This was based on N = 36 participants tested in Experiment 1b, but we decided to later test up to 60 participants to determine whether the cost for orientation memory in recall error was credible.

improves memory for this feature, it impairs memory for color, and it has no impact on memory for orientation.

Methods

Participants

In total, a new sample of 61 students of the University of Zurich were tested in Experiment 2 (M = 23.97; SD = 4.18; 45 women). Participants completed two 1-hour sessions under the same conditions as in Experiment 1. In the preregistration, we mentioned to start data collection with a sample of 24 participants, and that we would add more participants until we reached BFs ≥ 10 for comparison of our conditions of interest or that we would stop data collection once we have collected a total of 60 participants. The latter was our key determinant to stop data collection.

Seven participants were excluded due to not following the labeling instructions [3], one for not attending the second session, one for aborting the experiment in the second session, and one because they admitted to the experimenter after the experiment was over that they did not understood the instructions for labeling and just repeated the terms appearing in the instructions. Thus, the final data set submitted to the analysis consisted of 51 participants.

Participants fulfilled the same inclusion criteria as in Experiment 1, except for two participants who did not inform us prior to the study that their mother tongue was not German. As their German was sufficiently good to label the features, we included their

[3] Five of these participants additionally labeled the other feature during the frequency labeling condition, and the other two participants labeled the correct feature on less than 70% of occasions.

data into the analysis. Participants were exposed to the same experimental protocol as in the previous experiments.

Materials and Procedure

In Experiment 2 (see Figure 5A), three Gabor patches were presented sequentially against a black background (RGB 0 0 0), with each Gabor remaining onscreen for 1000 ms, followed by an inter-item interval of 1000 ms. The envelope of each Gabor patch was defined with a size of 61 pixels and a radius of 61 pixels. The Gabor standard deviation (sigma) was set to 10.17. The background of the Gabor was black (RGBA 0 0 0 0) with a pre-contrast multiplier of 1.0. The three Gabor patches were presented equally spaced in an imaginary circle centered in the middle of the screen. The exact locations of the items varied from trial to trial. The Gabor spatial frequencies ranged from 12 pixels/cycle to 24 pixels/cycle (0.19° and 0.39°) in 13 steps.

Participants completed two sessions. In one session, they were presented with Gabor patches that varied in spatial frequency (1 out of 13 values) and orientation (0-180°), whereas color was fixed at a single value (e.g., white). In the other session, the Gabor patches varied in spatial frequency and color (0-360 colors as defined in Experiment 1) and orientation was fixed (0°).[4] Hence although the items contained three features, only two features were varied at a time.

As for the critical labeling manipulation, participants were asked to perform the task under suppression, or they were required to label the spatial frequency aloud, or the other feature of the item (either color or orientation). As in Experiment 1, participants

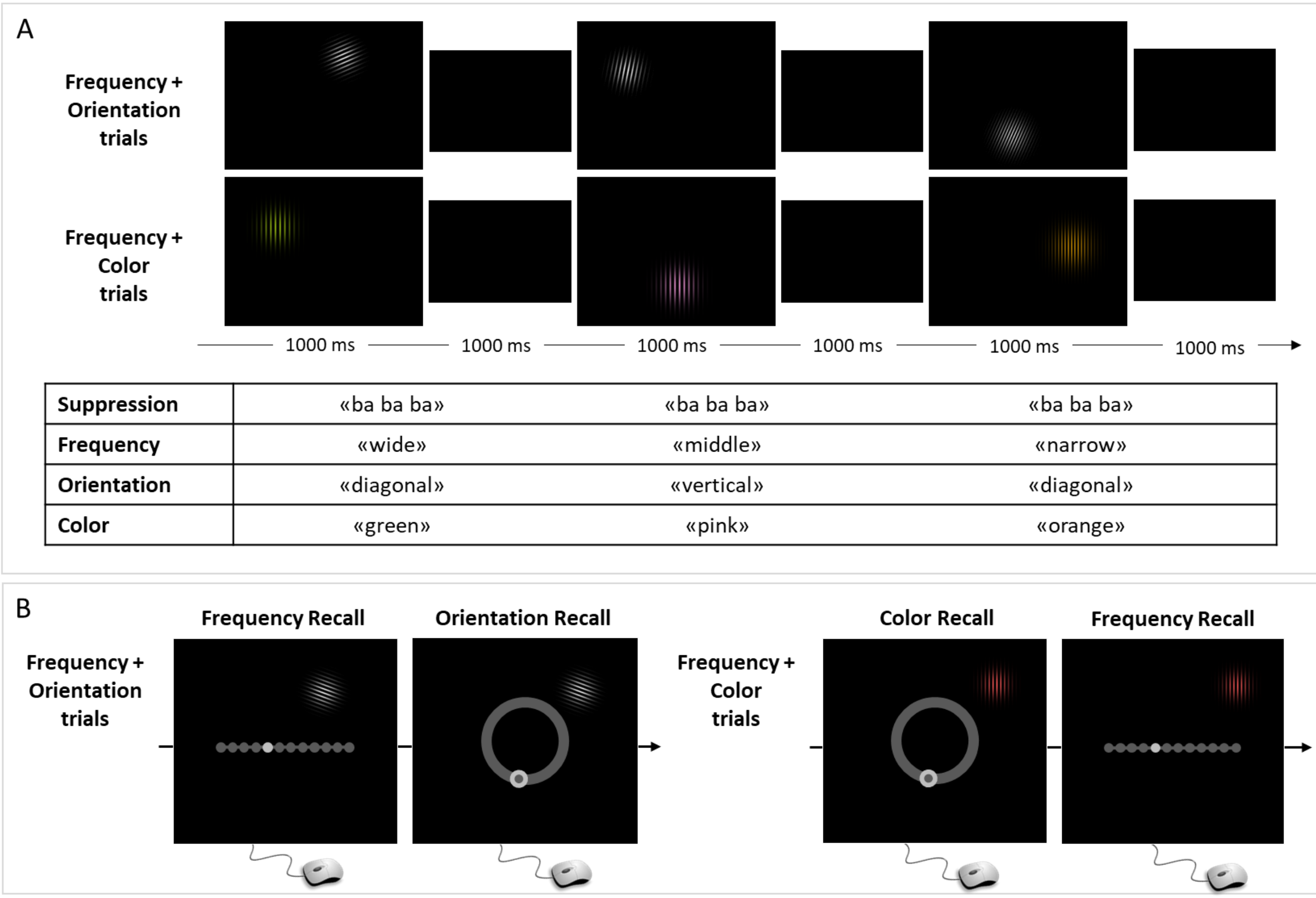
[4] We did not conduct a color calibration of the monitor. Hence it is possible that the color hue varied slightly within the Garbor.

were free to label the features with any term they wanted; the only requirement being that they did it aloud.

We did not assess the impact of orientation labeling on color memory and vice versa because we had addressed this issue in Experiments 1a and 1b, and we aimed to maximize the number of trials on the new conditions that addressed the new experimental questions posed in Experiment 2. The labeling conditions were implemented in three blocks whose order was counterbalanced across participants. Each block consisted of 78 trials, resulting in 234 trials in total for one session. Before the start of each block, participants completed three practice trials. Participants started each trial by pressing the space bar. Before pressing the space bar, they were reminded of the current labeling condition (e.g., say “ba ba ba” out loud now; label the spatial frequencies, label the colors, or label the orientations). The verbal output during the study phase was recorded for offline check of compliance with the labeling instructions. In this experiment, we coded each of the verbal responses to assess which labels were applied to which memoranda.

Figure 5

Example of the Flow of Events in Experiment 2 for the Frequency+Orientation and the Frequency+Color Sessions



Note. Panel A illustrates the flow of events in the study phase for the two types of session trials (frequency+orientation and frequency+color) along with the labeling conditions. Panel B illustrates the memory tests for these two types of sessions. Note that labels displayed here are only illustrative: participants were free to use any term they wanted.

At test (see Figure 5B), one item was randomly chosen as the test target and a memory probe was shown at the target's location. The probe was shown with a randomly selected spatial frequency. For the session in which orientation was varied, a random orientation was selected for the probe, and color was always white. For the session in which color was varied, a random color was selected for the probe, and orientation was fixed at 0°. Participants were requested to reconstruct both relevant features of the target item. The order in which the two features were tested was counterbalanced across trials. Participants adjusted the spatial frequency of the probe by moving a dot (RGB 150 150 150) on a dark grey slider (RGB 96 96 96) presented in the middle of the screen. They adjusted the orientation or color of the probe by moving a dot on a grey wheel presented in the middle of the screen. When participants were satisfied with the adjusted feature, they confirmed their response by a left-mouse click.

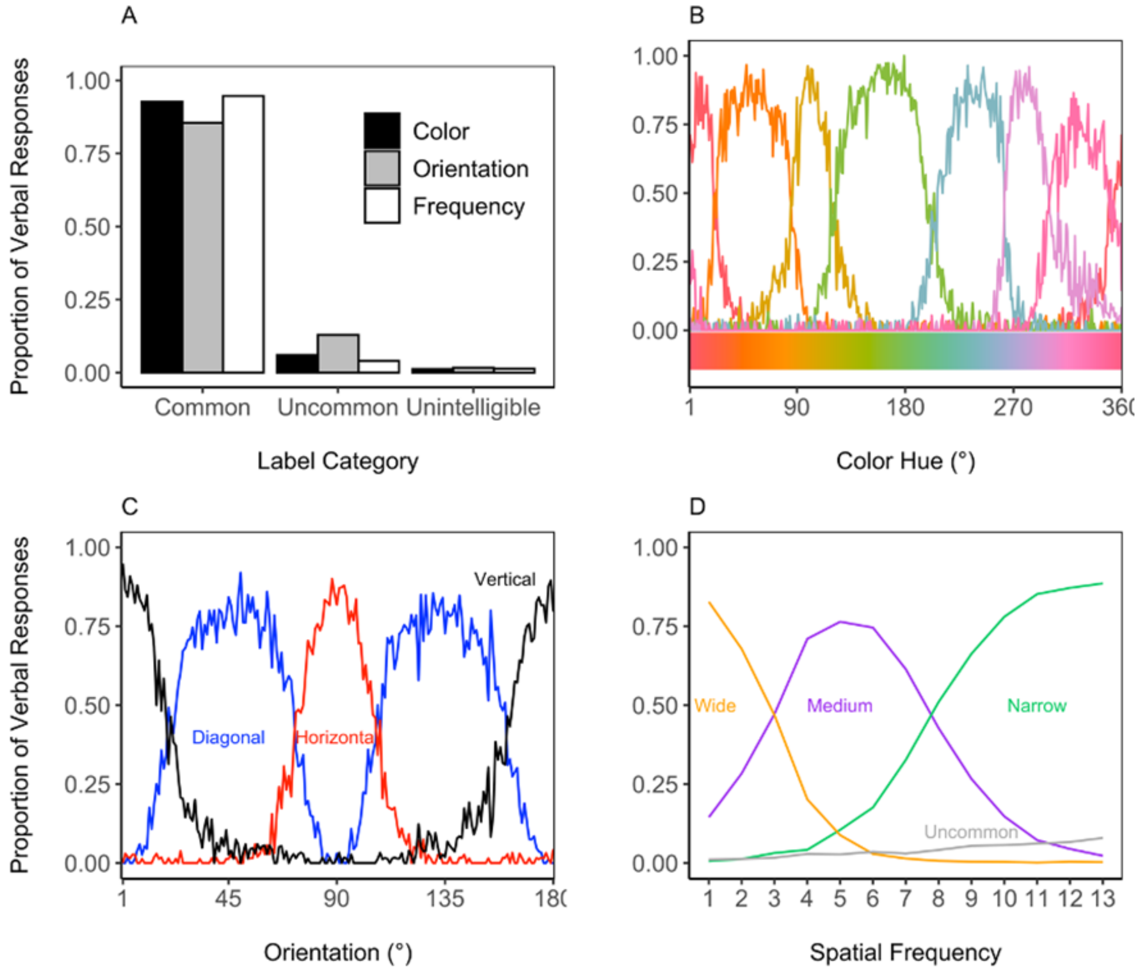
Results

Verbal Labeling Output

We recorded the verbal responses during the study phase of the working memory task in all our experiments. However, in Experiment 1, two memory items were simultaneously presented, which did not allow for a direct inference of which label was assigned to which memory item. In Experiment 2, memory items were sequentially presented, which allowed us to further analyze the verbal labeling output data to assess the variety of labels applied to the colors, orientations, and spatial frequency of each Gabor patch, and which feature values the labels were applied to.

Figure 6

Analysis of the Labels used by the Participants in Experiment 2 Showing the Range of Labels Used by Participants to Label each Type of Feature Dimension.



Note. Panel A shows the proportion of occasions in which common, uncommon, and unintelligible labels were applied for color, spatial frequency, and orientation. Panel B shows the proportion of times one of the seven common color labels was used to refer to a given color on the wheel (as shown in the x-axis). A proportion of 1 indicates that the x color on the wheel was labeled with the same label by all participants. The lower the proportion, the less often participants used that label to refer to that given color. Each color term is represented by the line with its prototypical color. Similarly, Panel C shows the proportion of times one of the three common orientation labels was used by the participants to refer to the different orientations. Panel D shows the proportion of times each of the three common spatial frequency labels (and the uncommon labels) were applied to each of the spatial frequency values used in the study.

Participants used a total of 90 different color labels, 93 orientation labels, and 31 frequency labels. The majority of the color labels belonged to a set of seven basic color categories (e.g., red, orange, yellow, green, blue, purple, and pink) - hereafter referred as common category, as opposed to the usage of more uncommon labels (e.g., turquoise, yellow-green, dark orange, blueish), or unintelligible responses. Likewise, three terms were commonly used for orientation (e.g., diagonal, horizontal, vertical), and three terms were used for spatial frequency (e.g., wide, thin, medium). The proportion of occasions in which these sets of 7 color terms, 3 orientation terms, and 3 spatial frequency terms were used (hereafter common category) is depicted in Figure 6A. Other terms that did not belong to this set were classified as uncommon, and we also coded for unintelligible responses (output was not understandable or the participants remained silent). Based on this classification, it is clear that more labels were assigned to the color space than to the orientation and frequency space. However, overall participants used the common category on the majority of trials, and this was similar across the features.

Figure 6B presents the proportion of occasions one of the seven basic color labels was used (across all participants) to refer to the 360 colors. There was high agreement between participants regarding the labeling of the colors during the memory trials. The same approach was used for the orientation labels, which ranged from 1 to 180 values.

Figure 6C shows that for orientation, three broad labels were used across the orientation space. Figure 6D shows that three broad labels were used for frequency labeling. We further plotted the proportion of the uncommon labels, which is distributed in close proximity to the x-axis, showing that these labels were not systematically applied to a section of the spatial frequency continuum.

Recall Performance

For all features, we computed a measure of recall error by computing the absolute difference between the true feature value of the item and the participant's response. For spatial frequency, this measure ranged between 1 and 12 pixels/cycle. For orientation, this measure ranged between 0° and 90°, and for color between 0° and 180°. We evaluated the effect of the labeling condition (e.g., suppression, labeling the reported feature, labeling the other feature) upon each of these measures separately.

Figure 7A shows the error in recalling spatial frequency and Figure 7B the error for recalling orientation and color. Figure 7A shows that labeling spatial frequency reduced the error in recalling this feature compared to suppression, whereas labeling the color or the orientation of the Gabor had no credible impact. This shows that labeling boosted spatial frequency memory, and that information about spatial frequency was not lost when participants labeled the other features.

For orientation memory (Figure 7B), orientation labeling reduced recall error, whereas spatial frequency labeling increased recall error compared to suppression. Likewise, for color memory, color labeling reduced recall error, whereas spatial frequency labeling increased recall error compared to suppression. These results indicate that orientation and color information were lost when spatial frequency was labeled.

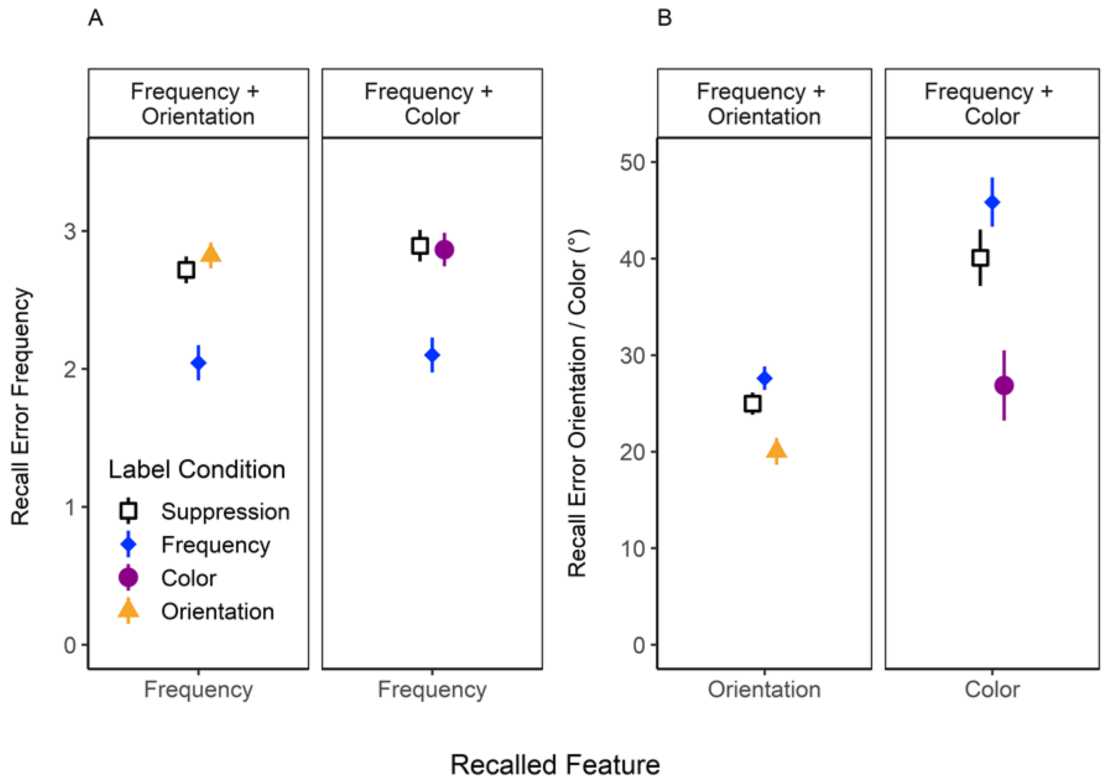
We contrasted recall error in each labeling condition to the one observed in the respective suppression condition using Bayesian *t*-tests [5]. Table 2 presents the evidence for the Alternative hypothesis that the two conditions differ (BF_{10}). The BF is neutral

[5] The Bayesian *t*-tests were not preregistered.

regarding the direction of the effect (i.e., whether recall error is larger or smaller in the labeling condition compared to suppression). To provide information regarding the direction of the difference, we entered in the table a B (benefit) for the observation of a smaller recall error and a C (cost) for the observation of larger recall error in the labeling than in the suppression condition as well as the Cohen's *d* of the difference.

Figure 7

Mean Recall Error in Experiment 2 as a Function of Labeling Condition (Different Dots) and Recalled Feature (x-axis). Panel A Shows that the Error in Reproducing Frequency Decreased when Frequency was Labeled Compared to Suppression, and it Remained Unchanged when Orientation or Color was Labeled. Panel B Shows that Recall Error for Orientation and Color Decreased when these Features were Labeled Compared to Suppression, and it Increased when Frequency was Labeled.



Note. Panel A shows the mean recall error for frequency recall. The error is presented for the two sessions, frequency + color and frequency + orientation recall test along with the labeling conditions. Panel B shows mean recall error for orientation (first subpanel) and color (second subpanel) as a function of labeling condition, for their respective sessions. Note that recall error varied from 1-90° for orientation recall and from 1-180° for color recall. Error bars represent the 95% within-subjects confidence interval (Morey, 2008).

Table 2

BFs in Favor (i.e., BF_{10}) of a Labeling Benefit (B) or Cost (C) in Experiment 2, and Cohen's d.

Labeling	Recall			
	Frequency + Orientation Session		Frequency + Color Session	
	Frequency	Orientation	Frequency	Color
Frequency	B = 5.47×10^6 *d* = 1.58 [.90, 2.39]	**C = 20.40** *d* = -.49 [-.81, -.17]	B = 5.19×10^9 *d* = 1.45 [.92, 2.01]	**C = 25.78** *d* = -.52 [-.87, -.21]
Orientation	0.70 *d* = -.24 [-.53, .06]	B = 8.62×10^3 *d* = .76 [.43, 1.09]	---	---
Color	---	---	0.16 *d* = .02 [-.26, .32]	B = 7.68×10^3 *d* = .85 [.45, 1.24]

Note. Green font and a B indicate substantial evidence (BF > 3) for a labeling benefit. The red, bold font and a C indicate substantial evidence for a labeling cost. We present the mean effect-size and its 95% credible interval.

Verbal labeling benefitted recall of the labeled feature in all three labeling conditions in contrast to suppression. For spatial frequency memory, there was inconclusive evidence whether labeling orientation led to a cost, and evidence against a cost when color was labeled (BF_{10} = 0.16 which is equivalent to a BF_{01} = 6.25). In contrast, spatial-frequency labeling led to a cost for the recall of color and orientation.

In the preregistration we mentioned to submit the data to a BANOVA. For this purpose, we calculated *z*-score values to directly compare the three different types of feature recall using the same scale. The *z*-scores were computed for each of the four recall tests depicted in Figure 7 by subtracting the mean recall error for that type of test averaged across all labeling conditions, divided by the standard deviation.

We first ran a 3-way BANOVA on the *z*-scored recall error with labeling (color, orientation, frequency), test condition (frequency-orientation, frequency-color) and tested feature (orientation, color, frequency) as fixed predictors, and subject as random predictor. The best model included labeling condition and tested feature as well as their interaction into the model, $BF_{10} = 1.86 \times 10^{47}$. This model was favored against the second-best model including condition, test condition, tested feature and the interaction of condition × tested feature by a $BF_{10} = 6.04$.

To estimate more closely the effect of verbal labeling on the recalled feature we ran two independent BANOVAs for the frequency-orientation and frequency-color sessions having labeling condition and tested feature as predictors (see Table S3 in the Online Supplementary Materials). For both BANOVAs, the best model included both main effects and their interaction (BFs = 2.37×10^{23}; 4.34×10^{24}) and the inclusion of the interaction was clearly favored (BFs = 4.11×10^{23}; 4.46×10^{23}).

Categorical-Continuous Mixture Modeling

For the continuous feature dimensions used in Experiment 2, namely, color and orientation, we submitted the data to the CatCont mixture model, as done in Experiment 1 [6]. We could not apply the mixture model to the spatial frequency data because this is not a continuous circular space. Note that, as in the previous experiments, we separately modeled recall of color and orientation given that categorical biases are different in these feature dimensions, but modeled all labeling conditions simultaneously. For each model, we ran 10,000 iterations, of which 2,000 were regarded as burn-in. We first ran the full model, containing the fixed effect of labeling condition on all parameters. We then constrained this full model as done in Experiment 1. Table 1 presents all models, alongside their WAICs, and their relative comparison. For color recall, two models yielded the same WAIC. One of them was the full model, and we decided in favor of this

[6] The mixture modeling analysis was not preregistered.

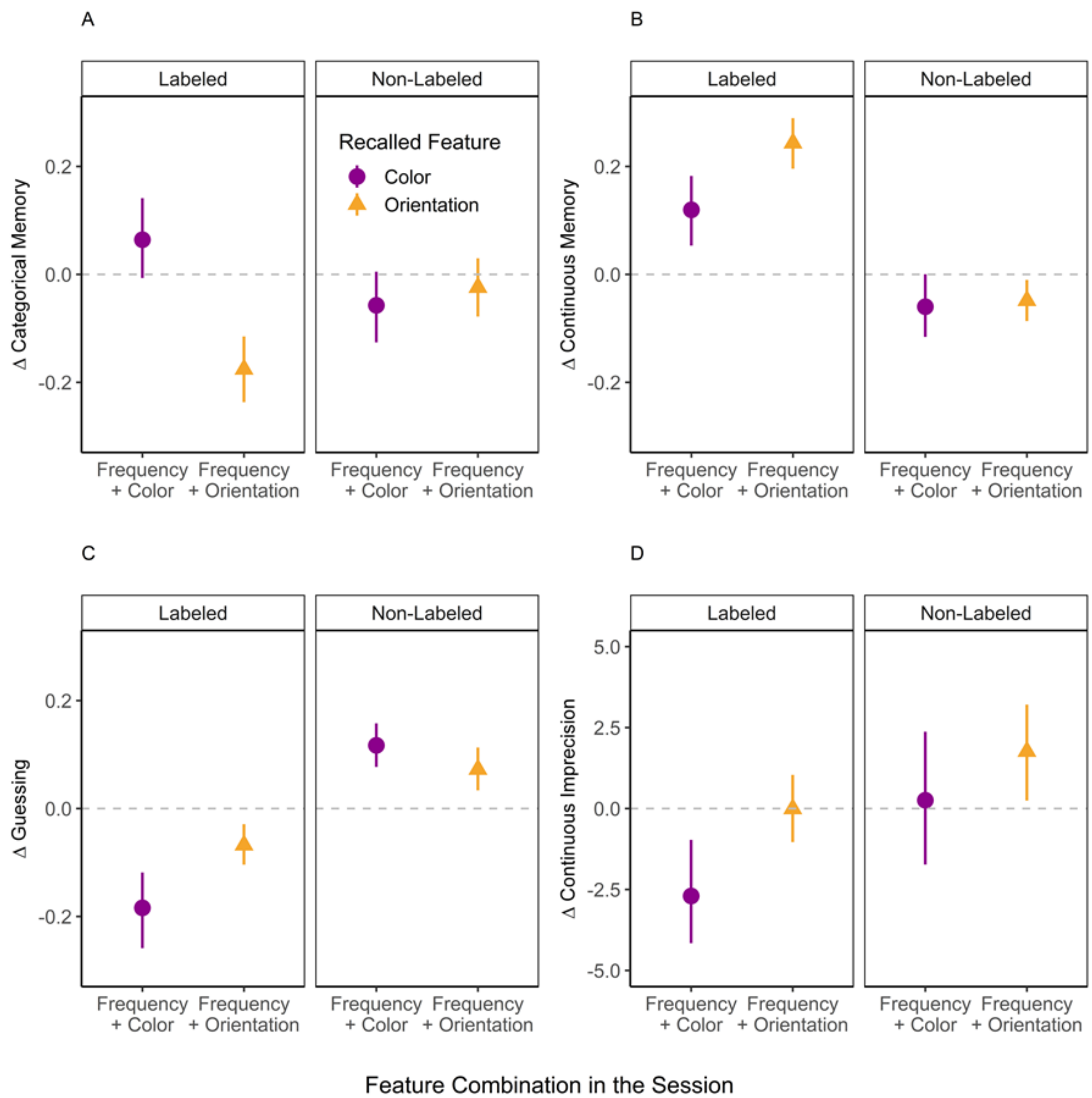
model to be more conservative. For orientation recall, the full model was the best model. Note that the model for orientation recall fitted the obtained data less well (see Online Supplementary Materials).

As done in Experiment 1, based on P^M and P^O we calculated the probability of retrieving categorical and continuous memory, and guessing (1- P^M), and these estimates are presented in Table S2 along with the group-level estimates for continuous imprecision. We then computed posterior differences between the labeling conditions compared to the suppression condition, which are displayed in Figure 12. To recapitulate, values above zero indicate a labeling benefit and values below 0 a labeling cost.

Figure 8A shows the changes in categorical memory as a function of labeling. For the labeled feature, color categorical memory tended to increase (although non-credibly) when color was labeled, whereas orientation categorical memory decreased when orientation was labeled. There was no credible change in either color and orientation categorical memory when they were the non-labeled features (i.e., when frequency was labeled). Figure 8B shows changes in continuous memory. Continuous memory increased for color and orientation when they were the labeled features, and decreased when they were the non-labeled features, although this decrease was not fully credible for color. Figure 8C shows changes in guessing. When color and orientation were labeled, guessing decreased, whereas when they were the non-labeled features, guessing increased credibly. Finally, Figure 8D shows that continuous color imprecision decreased when color was labeled, but not when orientation was labeled. When color was the non-labeled feature, there was no credible change in memory imprecision, yet when orientation was the non-labeled feature, memory imprecision credible increased.

Figure 8

Changes in the Mixture Model Parameters for the Recall of the Labeled and Non-Labeled Features in the two Sessions of Experiment 2 (i.e., Frequency+Color and Frequency+Orientation). In Sum, Labeling the Tested Feature Consistently Increased the Storage of Continuous Features (Panel B) and Reduced Guessing (Panel C) for the Recall of Color and Orientation. When Color and Orientation Were the Non-Labeled Features, Guessing Increased Credibly.



Note. Dots depict the mean difference of the posterior distributions in the labeling condition in comparison to the suppression baseline. The error bars depict the 95% HDI of the difference. The dotted line represents no change in relation to the suppression condition. Panel A reflects the changes in the probability of retrieving categorical information, and Panel B of continuous information. Panel C presents changes in guessing, and Panel D changes in the estimates of continuous memory imprecision.

Discussion

In Experiment 2 we assessed the impact of labeling other visual features besides color and orientation. We used Gabor patches that could vary in three dimensions (namely, spatial frequency, color, and orientation). Replicating and extending our previous research, we again showed that the labeled features were retained better, now demonstrating that this effect extends to spatial frequency. Likewise, mixture modeling again showed that this benefit was consistently observed in the probability of storing a continuous representation of the labeled feature.

Critically, Experiment 2 was designed to assess the likelihood that labeling enhanced memory for the labeled feature at the expense of the non-labeled feature (H1) or whether there was no cost for the non-labeled feature (H2). We found evidence for both hypotheses, indicating a complex picture: the presence or not of costs still varied depending on the feature combinations.

Experiment 2 also showed that the types of labels applied to the different features varied: Participants used far more labels for colors than for orientation and spatial frequency. One possible explanation for the pattern of costs to the non-labeled features we observed is related to how much attention is engaged by the cognitive process of labeling. Color labeling is probably a much more routine activity than orientation and spatial frequency labeling. Therefore, it is possible that color labeling consumes fewer attentional resources. Since working memory is highly sensitive to the attentional demands imposed by dual-tasks, known as the *cognitive load effect* (Barrouillet et al., 2007, 2011; Oberauer et al., 2018), the non-labeled feature may be suffering due to the cognitive load imposed by the labeling activity, whereas the labeled feature remains protected from it. The goal of Experiment 3 was to address this possibility.

Experiment 3

In Experiments 1 and 2, color memory was always hindered when it was the non-labeled feature, whereas when orientation was the non-labeled features it was less likely to show a cost. One potential explanation is that color labeling was less demanding than orientation labeling (and spatial frequency labeling), and it is the attentional demands of the labeling activity *per se* that generated costs. If this is the case, then reducing the attentional demands of orientation labeling while increasing the demands of color labeling would, respectively, reduce the costs of orientation labeling on color memory and increase the costs of color labeling on orientation memory. Hence, the attentional demands of color and orientation labeling would be equalized.

The goal of Experiment 3 was to test these predictions. In Experiment 3, we trained participants to use an intuitive orientation labeling system: to label the direction of the triangles as if they were the hours in a clockface. Our reasoning was that if this more intuitive labeling system is less demanding, the costs on color memory might get reduced. We trained participants on using this labeling system, and required them to use it across two sessions. With this, we expected participants to get more and more used to these labels, thereby reducing the costs of orientation labeling on color memory.

Additionally, we manipulated the type of color labels participants used across the two sessions. In the first session, we allowed participants to label the colors using their preferred color terms, as done in Experiments 1 and 2. In the second session, we trained participants in an arbitrary color labeling system. The color wheel was arbitrarily partitioned in six equal sections (values randomly determined for each participant), and each section was associated to a German non-word. Participants learned how to use these arbitrary labels to categorize the colors in each section in a short practice block, and then applied these labels during the VWM task. Our reasoning was that this color labeling system would be much more attentionally demanding than the more familiar color labels. Accordingly, if the labeling demand is what determines the observation of costs to the non-labeled features, then arbitrary color labeling would produce costs to orientation memory, whereas familiar color labeling would be cost-free as in the previous experiments.

Method

Participants

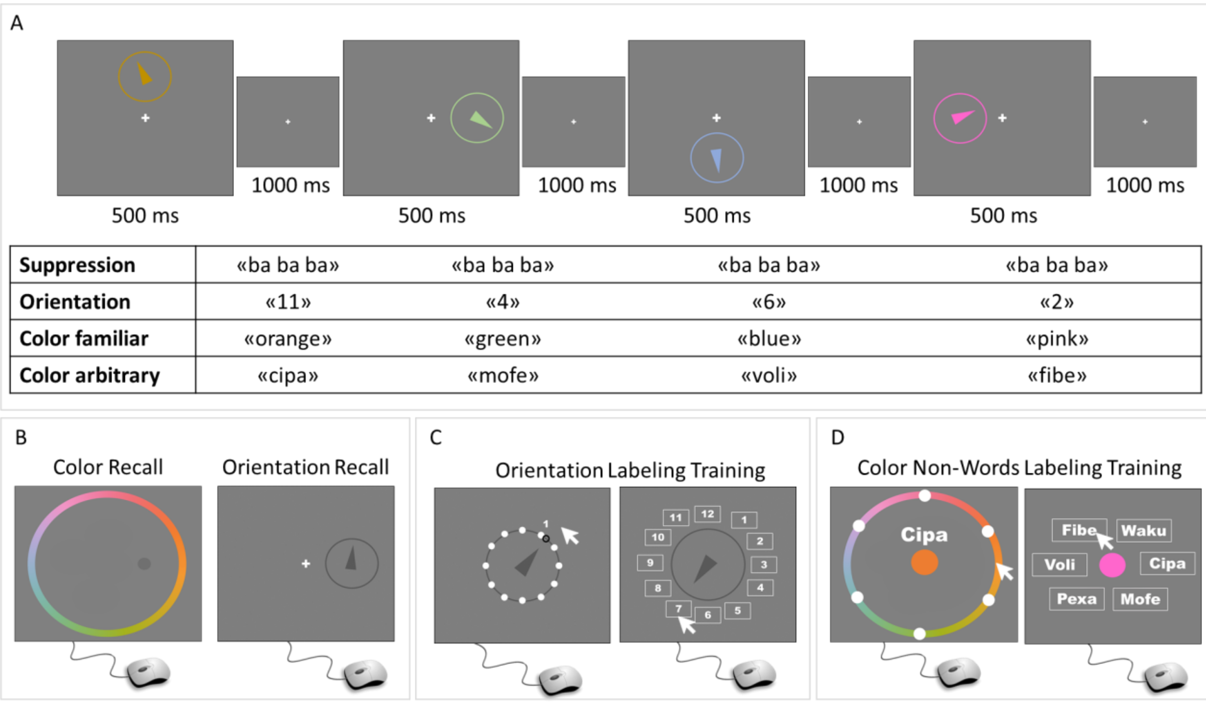
We collected data of 30 German-speaking students of the University of Zurich (*M* = 22.96; *SD* = 3.07; 6 men). Two additional participants started the experiment but had to be excluded: one because they failed to learn the arbitrary color labels in the second session and one because they never showed for the second session. Participants completed two 1-hour sessions under the same laboratory conditions as in Experiment 1.

Materials and Procedure

The same materials used for Experiment 1 were used in Experiment 3. The only difference was that we included a circle frame around the memory stimuli to make it more similar to a clockface.

Figure 9

Flow of Events in Experiment 3. Panel A Illustrates the Study Phase, Panel B the Recall Procedure. Panel C Illustrates the Orientation Labeling Training, and Panel D the Training with the Arbitrary Non-Words Assigned to Colors.



The main experimental task consisted of the presentation of a sequence of four colored triangles, each surrounded by a circle-frame. We included the frames to make the stimuli resemble a clockface. The stimuli appeared in four fixed locations (top, right, down, and left) in clockwise order as shown in Figure 9A. Each memory item was presented for 500 ms, with an inter-stimulus interval of 1000 ms. Next, a randomly selected item was probed, and participants had to reproduce its color and orientation (see Figure 9B). In a random 50% of the trials, color was probed first followed by orientation. In the remaining 50% of the trials, the order was reversed.

Participants completed two sessions. In each session, there were three labeling blocks whose order was counterbalanced across participants. In the *suppression* block, participants had to say *"ba ba ba"* aloud continuously during the study phase, akin to the previous experiments. In the *orientation labeling* block, participants were instructed to label the orientations of the presented stimuli using the numbers 1-12, as if they would be naming the hours on a clock. We selected these labels because this is a labeling system that participants are likely to be familiar with and that can be easily applied to improve memory of orientations. To make sure participants were able to accurately label the orientations with the clock-system, participants completed an orientation training block in Session 1 (see details below). Participants were prompted to use the same orientation labeling system in both sessions.

Finally, participants completed a *color labeling* block, where they had to label the colors of the presented stimuli. The type of color label was manipulated between the two sessions. In Session 1, participants were instructed to use any color label they wanted, thereby replicating the color labeling conditions used in the previous experiments. Thereafter we will refer to this session as the *Familiar Color Labels* condition. In Session 2, participants were required to label the colors with one of six non-words that they were pre-trained to assign it to randomly defined sections of the color wheel. This procedure was inspired by previous work from our lab (Souza et al., 2021), in which we demonstrate that arbitrary non-word labels can be used to improve memory of continuous shapes. Here we extended this procedure to colors, and expanded the usage to create newly defined color categories. We will refer to this session as the *Arbitrary Color Labels* condition. Participants also completed a color labeling training phase right before this condition.

Orientation Labeling Training. This training block occurred right before the VWM trials in the orientation labeling block in Session 1. The training consisted of three parts. In the first part, participants were shown a grey triangle inside a circle frame that was divided into 12 sections. As participants moved the mouse around, the position of the triangle changed (Figure 13C). As they rotated the triangle, the number corresponding to the hour the triangle pointed to appeared onscreen. Participants were told to explore the label mapping and press continue when ready. Next, part two started. Participants completed trials in which they were presented with a triangle surrounded by boxes with the numbers 1-12 positioned around the stimulus (see Figure 9C). They had to click on the number label that best represented the orientation of the triangle. If they clicked on the correct label, the box turned green for 500 ms; if they clicked on the wrong box, it turned red and the correct response turned green simultaneously for 500 ms, and the next trial started. Participants practiced this task for a minimum of 48 trials. After that, accuracy of classification was verified, and if they reached a criterion of 80% correct responses, part two was finished. If accuracy was below this level, they completed another set of 12 trials. Then accuracy was again evaluated. Training was finished when the 80% criterion was reached or a maximum of 240 trials were completed. If participants failed to reach this criterion, participation in the experiment was terminated. All participants learned the labels within this limit. Finally, in part three, participants were presented an orientation for 500 ms, and they had another 1000 ms to say the label associated with the orientation aloud. Then the correct label and the stimulus was presented onscreen and participants had to indicate by mouse click if they said the correct label aloud or not (self-scored accuracy) by clicking in one of two boxes. Participants completed a minimum of 48 trials practicing labeling the stimulus within this time limit. They had to achieve a minimum of 80% self-scored correct responses, or the training continued until a maximum of 240 trials.

Color Labeling Training. As with the orientation labeling training, this training consisted of three parts. In the first part, participants were first shown a colored patch with a non-word on top of it and a color wheel that was randomly divided into six parts (see Figure 9D). As participants moved the mouse around the color wheel, the color of the central patch changed and so did the label atop. Each section was associated to one of six German non-words: voli, pexa, mofe, cipa, waku, and fibe (Souza et al., 2021). Participants were told to explore the label mapping and press continue when ready. Next, part two started: participants completed trials in which they were presented with a color surrounded by six boxes with the non-words (see Figure 9D). They had to click on the label associated with the presented color. If they clicked on the right label, the box turned green for 500 ms. If they clicked on the wrong box, it turned red and simultaneously the correct label turned green for 500 ms. Thereafter the next trial started. They practiced this task for a minimum of 48 trials, and a maximum of 240 trials, until they reached an 80% accuracy criterion (as in the orientation training phase). If participants failed to reach this criterion, participation was terminated. Only one participant was excluded for failing to learn the labels. Finally, in part three of the training, participants were presented a color patch for 500 ms, and they had another 1000 ms to say the non-word label associated with the color aloud. Then the correct label and the stimulus were presented onscreen and participants had to indicate by a mouse click if they said the correct label aloud or not. Participants completed a minimum of 48 trials practicing labeling the stimulus within this time limit. As in phase 2, they had to achieve a minimum of 80% correct responses, or the training continued until a maximum of 240 trials.

Results

Recall Performance

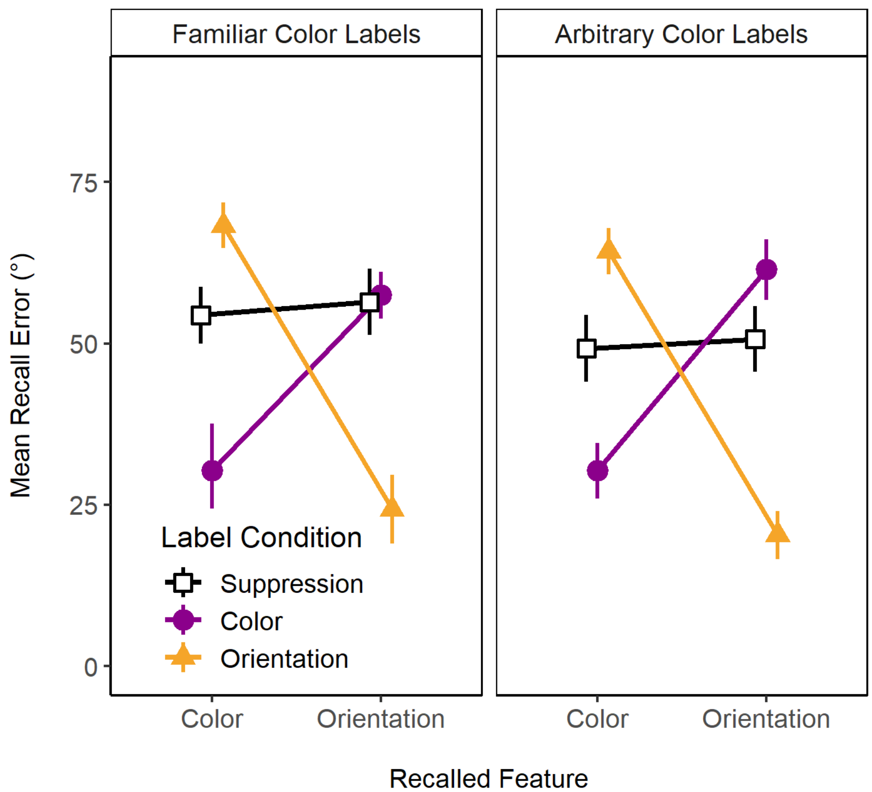
Figure 10 shows recall error in Experiment 3 as a function of the probed feature (color or orientation), with each line representing one of the labeling conditions during study. The data of the two sessions in which the color labels were manipulated (i.e., familiar vs. arbitrary labels) are presented in different panels. Table S4 in the Online Supplementary Materials presents the evidence for the effects of labeling condition and recalled feature for each session. Replicating the previous experiments, the best model of the data in each session was always the full model ($BF_{10} = 2.83 \times 10^{28}$; $BF_{10} = 1.91 \times 10^{24}$), indicating that the effect of labeling depended on whether the probed feature was labeled or non-labeled.

To estimate the labeling benefits and costs for each recalled feature in each session, we ran Bayesian *t*-tests separately contrasting each labeling condition to suppression. For color recall, there was evidence for a color labeling benefit irrespective of whether the labels were familiar, $d_{fam} = 1.00$ [.53, 1.50], $BF_{10} = 1727.8$, or arbitrary, $d_{arb} = 1.15$ [.67, 1.72], $BF_{10} = 1.4 \times 10^4$. These results show for the first time that even when participants apply a completely new and arbitrarily defined color labeling system to the memoranda, labeling benefits are observed.

In contrast, color recall suffered when orientation was labeled in the familiar color labeling session (Session 1), $d_{fam} = -.93$ [-1.40, -.48], $BF_{10} = 1005.9$, as well as the arbitrary color labeling session (Session 2), $d_{arb} = -.92$ [-1.34, -.44], $BF_{10} = 1110.1$. In both sessions, the orientation labels used were the same, but practice with using these labels increased. Yet, the size of the cost was the same across the two sessions, indicating that this training was not sufficient to reduce the costs of orientation labeling on color memory.

Figure 10

Mean Recall Error in Experiment 3 as a Function of the Recalled Feature and the Label Condition During Study for Session 1 in Which Familiar Color Labels Were Used and Session 2 in which Arbitrary Color Labels were Used. It Shows Lower Recall Error When the Recalled Feature Was Labeled Compared to Suppression. For Color Recall, Labeling Orientation Increased the Error Compared to Suppression in both Sessions. For Orientation Recall, Color Labeling Only Produced a Cost when Color Labels Were Arbitrary.



Note. Error-bars represent 95% within-subjects confidence intervals.

For orientation recall, there was an orientation labeling benefit in both sessions, d_{fam} = 1.65 [1.03, 2.29], $BF_{10} = 1.1 \times 10^7$; d_{arb} = 1.95 [1.29, 2.59], $BF_{10} = 6.6 \times 10^8$. Replicating the previous experiments, color labeling did not induce a cost when the color labels were familiar, d_{fam} = -.16 [-.58, .25], BF_{10} = 0.21, but a substantial color labeling cost was observed when the color labels were arbitrary, d_{fam} = -.69 [-1.19, -.19], BF_{10} = 11.64. These findings are in line with the assumption that the costs to the non-labeled features vary depending on the attentional demands of using the labels.

Categorical-Continuous Mixture Model

We fitted the data of each session and feature (color and orientation) separately. We fitted the data with 10000 iterations, and excluded the first 2000 as burn-in. Table S5 presents the posteriors for each of the relevant parameters in Experiment 3. Figure 11 presents the changes in parameters of the mixture model for the labeled and non-labeled features (in the labeling conditions) compared to the suppression baseline.

Figure 11A shows changes in categorical memory. When color was the labeled feature, categorical memory increased, whereas the opposite was observed for orientation. The reduction in categorical memory when orientation was labeled replicates the previous experiments. As it will become clear from the analysis of continuous memory and guessing, it mostly reflects a tradeoff with a change from categorical to continuous memory reliance. For the non-labeled features, there was only evidence for a credible reduction of color memory when the color labels were arbitrary.

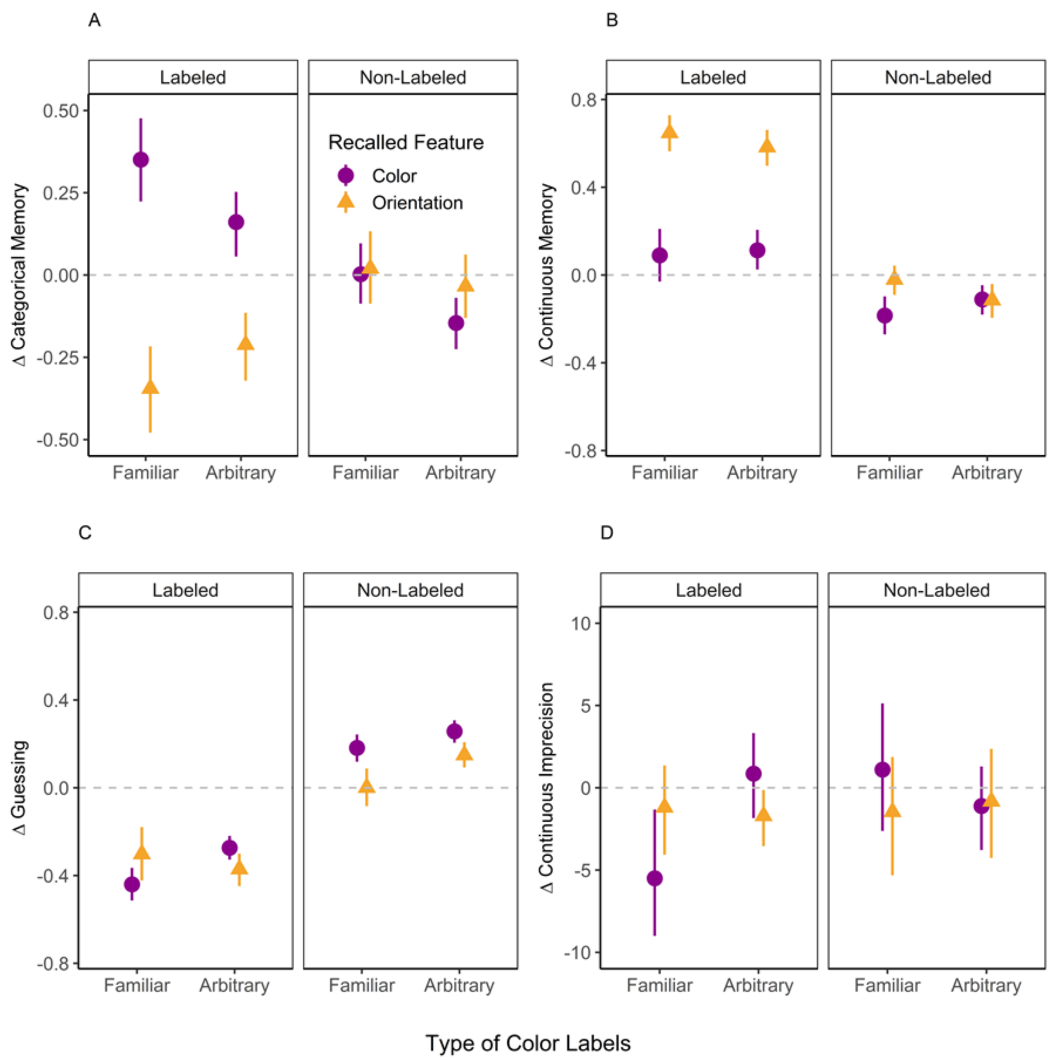
Figure 11B shows changes in continuous memory. Continuous memory increased credibly for the labeled features, except for color memory in the session with familiar color labels. The gain in continuous memory was quite large for orientation memory (larger than in our previous experiments), indicating that the clockface labeling system was quite helpful for storing a larger proportion of continuous orientations. In contrast, continuous memory decreased for the non-labeled features, except for orientation memory in the familiar color labels session.

Figure 11C shows changes in guessing. In line with the effects for categorical and continuous memory, guessing for the labeled features was reduced, whereas guessing for the non-labeled features increased. The only exception was orientation memory when the color labels were familiar. Overall, the pattern observed when orientation memory was the non-labeled feature shows that it only suffered when the color labels were arbitrary, in line with our prediction that the demands of labeling impose a cognitive load that hinders memory for the non-labeled features.

Figure 11D presents changes in continuous imprecision. There were only two credible changes in this parameter. In the familiar color label session, continuous imprecision was reduced when the color feature was labeled. Hence although the increase in the probability of continuous memory during familiar color labeling was not fully credible, continuous memory improved by other means, namely by the fidelity of feature storage in memory. The second credible memory imprecision reduction was in the arbitrary color labels session during orientation labeling. As the manipulation of color labels had nothing to do with orientation labeling, this finding is likely related to practicing the clockface labeling system (given this was the second session).

Figure 11

Changes in the Mixture Model Parameters for the Recall of the Labeled and Non-Labeled Features in the two Sessions of Experiment 3. In Sum, Labeling the Tested Feature Tended to Increase Continuous Memory (Panel B) and Reduced Guessing (Panel C) for the Recall of Color and Orientation. When Color Was the Non-Labeled Feature, Guessing Increased Credibly Irrespectively of the Type of Color Labels (Panel C). When Orientation was the Non-Labeled Feature, Guessing Only Increased When Color Labels Were Arbitrary.



Note. Dots depict the mean difference of the posterior distributions in the labeling condition in comparison to the suppression baseline. The error bars depict the 95% HDI of the difference. The dotted line represents no change in relation to the suppression condition. Panel A reflects the changes in the probability of retrieving categorical information, and Panel B of continuous information. Panel C presents changes in guessing, and Panel D changes in the estimates of continuous memory imprecision.

Discussion

In Experiment 3 we trained participants in using a more intuitive orientation labeling system: naming the orientations as they if they were reporting hours. We even increased the similarity of our stimuli to a clockface to make this labeling system more intuitive. Using these labels yielded substantial benefits for orientation memory: continuous memory increased, whereas guessing decreased. Yet, the aftereffects of orientation labeling on color memory remained detrimental: continuous color memory decreased and guessing increased when color was the non-labeled feature. This result shows that we were unable to reduce the costs of orientation labeling by training participants in using a more favorable orientation labeling system.

At first sight these results go against our hypothesis that the costs of the labeling activity itself would be the reason behind the pattern of costs observed for the non-labeled features. Yet, it is worth pointing out that we did not have a direct measure of the attentional demands of this new orientation labeling system. Hence it is perfectly possible that the use of this more common labeling system was not less attentionally demanding after all. Additionally, this strategy involved using 12 labels, which could also contribute to the demands of using it. This might explain why people do not spontaneously use it to label orientations, although they are somewhat familiar with it. We do have some training in reading clocks, but perhaps not as much as it would be required to make this a costfree labeling strategy.

Crucially, Experiment 3 also included an experimental manipulation of the costs of color labeling. Experiments 1 and 2 already showed that the familiar color labeling system did not generate costs for orientation memory, hence our approach was to increase the costs of color labeling in Experiment 3 by training participants in using a completely arbitrary color labeling system. Participants had to apply non-words to arbitrarily defined sections of the wheel. Hence this labeling system, which was learned within a 20-min training session, required them to inhibit their familiar color labeling tendencies in order to use the new arbitrary system. We showed that color memory similarly benefited from this arbitrary labeling system: continuous memory increased, whereas guessing was decreased. This pattern was similar to the one observed when color labels were familiar. This shows that even the use of newly learned categories that do not match our well-learned familiar categories can benefit visual working memory (see also Souza et al., 2021). Critically, our manipulation of color labels had the expected effects on the non-labeled feature: the usage of familiar color labels did not harm orientation memory, but the use of arbitrary color labels did. This is exactly what is predicted from the hypothesis that labeling *per se* is an attentionally demanding activity. The labeled feature remains protected from the cognitive load imposed by this activity. However, non-labeled features do not have any support from the semantic network, and therefore they suffer when the cognitive load is high.

Summary of Findings Across Experiments 1-3

Table 3 presents a summary of the changes (as reflected in Cohen's *d*) in recall error observed for the labeled and non-labeled features across all experiments. As shown in this table, recall of the labeled feature always improved (lower recall error), and this effect was large (Cohen's d varying from .75 to 1.95 across features and experiments). This result extends previous findings regarding the benefits of labeling to new features such as orientation and spatial frequency. Experiment 3 also showed that using completely arbitrary labels (German non-words) to classify randomly defined sections of the color wheel produced large benefits for the labeled feature. Hence benefits of verbal labeling seem to generalize across visual features, and do not depend on using very well-established labels or categories (see also Souza et al, 2021).

Table 3

Summary of the Labeling Effect (Mean Cohen's d) on Recall Error in Experiments 1a, 1b, 2, and 3. Labeling the Recalled Feature Always Yielded Benefits. Recall of a Non-Labeled Feature Produced Costs in Some of the Comparisons.

Labeled Feature	Recalled Feature		
	Color	Orientation	Frequency
Color	**Benefit for Labeled** **E1a: 1.44 [.90, 2.02]** **E1b: 1.38 [.97, 1.76]** **$E2_{color}$: .85 [.45, 1.24]** **$E3_{fam}$: 1.00 [.53, 1.50]** **$E3_{arb}$: 1.15 [.67, 1.72]**	**Mixed** E1a: -.08 [-.43, .26] **E1b: -.53 [-.83, -.24]** $E3_{fam}$: -.16 [-.58, .25] **$E3_{arb}$: -.69 [-1.19, -.19]**	No Effect $E2_{color}$: .02 [-.26, .32]
Orientation	**Cost to Non-Labeled** **E1a: -.81 [-1.24, -.42]** **E1b: -1.15 [-1.56, -.75]** **$E3_{fam}$: -.93 [-1.40, -.48]** **$E3_{arb}$: -.92 [-1.34, -.44]**	**Benefit for Labeled** **E1a: .77 [.35, 1.21]** **E1b: .75 [.46, 1.05]** **$E2_{orient}$: .76 [.43, 1.09]** **$E3_{fam}$: 1.65 [1.03, 2.29]** **$E3_{arb}$: 1.95 [1.29, 2.59]**	No Effect $E2_{orient}$: -.24 [-.53, .06]
Frequency	**Cost to Non-Labeled** **$E2_{color}$: -.52 [-.87, -.21]**	**Cost to Non-Labeled** **$E2_{orient}$: -.49 [-.81, -.17]**	**Benefit for Labeled** **$E2_{orient}$: 1.58 [.90, 2.39]** **$E2_{color}$: 1.45 [.92,2.01]**

Note. The labeling effect was calculated as a difference in performance in the labeling conditions compared to the suppression baseline. Mean and 95% credible intervals of Cohen's *d* are presented. Credible values are presented in bold-face. Positive values indicate that the labeling condition produced a lower error than the suppression condition, whereas negative values indicate an increase in the recall error. $E2_{color}$ and $E2_{orient}$ = Data from the two separate sessions in Experiment 2 in which color and orientation, respectively, were encoded with spatial frequency. $E3_{fam}$ and $E3_{arb}$: Data of the two sessions in Experiment 3 in which color labels were familiar and arbitrary, respectively.

Recall of the non-labeled feature was sometimes hindered (increase in recall error) and sometimes remained unchanged. Table 3 shows that this effect was also large, although not as large as the benefits to the labeled feature. Non-labeled costs were consistently observed for color memory. Orientation memory suffered when spatial frequency was labeled and when color labeling was arbitrary (apart from a cost in Experiment 1b that never replicated). Spatial frequency was not harmed by color or orientation labeling. In Experiment 3 we reasoned that these costs for non-labeled features might arise due to how much labeling demanded attention, and observed that color labeling was not costly for orientation memory when participants could rely on familiar color labels, but a sizeable cost (effect size between medium and large) was observed when color labels were arbitrary.

We also summarized the effects of labeling on the mixture model parameters (i.e., changes in model parameters compared to suppression). Table 4 presents a summary of the effects of labeling for the labeled and non-labeled features separately for color and orientation. For the labeled feature, we consistently observed a reduction in guessing, and an increase in continuous memory. The decrease in guessing was larger for color than for orientation, except for Experiment 3, in which the benefits were similar across the two features. The increase in continuous memory was similar across the two features in most experiments, again with Experiment 3 being one exception in which orientation memory was much more continuous. This might be related to the orientation labeling system in Experiment 3 which afforded much more precise labels. This dovetails with our prior work in which we showed that gains in precision depend on the number of labels used (Souza et al., 2021). It is worth noting that when orientation was the labeled feature, there was a credible reduction in the reliance of categorical memory in all experiments, indicating that labeling created a tradeoff toward more continuous memory at the expense of retaining a coarse, categorical memory of the stimulus. There was no consistent effect on continuous imprecision across our experiments, but if anything, sometimes benefits were observed for the labeled feature with a reduction on memory imprecision, and only once a cost for the non-labeled feature was observed in this parameter.

Overall, Table 4 shows that labeling a feature had consistent effects on memory of the labeled feature. Costs for the non-labeled feature also accrued in similar parameters, yet costs were not always observed. Experiment 3 suggested that these costs may occur when labeling diverts attention away from the encoding of the relevant, yet non-labeled, features.

Table 4

Change in Mixture Model Parameters (Posterior Mean and 95% Credible Interval) for the Labeled and Non-Labeled Features (i.e. Performance in the Labeling Conditions) Compared to Performance in the Suppression Baseline.

Parameter	Exp.	Labeled		Non-Labeled	
		Color	Orientation	Color	Orientation
Categorical Memory	1a	= .02 [-.05, .08]	↓ **-.12 [-.19, -.06]**	= -.05 [-.11, .01]	= .02 [-.05, .09]
	1b	= -.03 [-.09, 0.01]	↓ **-.14 [-.20, -.09]**	↓ **-.09 [-.15, -.05]**	= -.03 [-.08, .03]
	2	= .06 [-.01, .14]	↓ **-.18 [-.24, -.12]**	= -.06 [-.12, .01]	= -.02 [-.08, .03]
	3_{fam}	↑ **.35 [.22, .48]**	↓ **-.34 [-.48, -.22]**	= .00 [-.09, .10]	= -.02 [-.09, .13]
	3_{arb}	↑ **.16 [.06, .25]**	↓ **-.21 [-.32, -.11]**	↓ **-.15 [-.23 -.07]**	= -.03 [-.13, .06]
Continuous Memory	1a	↑ **.21 [.16, .27]**	↑ **.19 [.12, .24]**	↓ **-.08 [-.14, -.03]**	= -.04 [-.10, .02]
	1b	↑ **.19 [.15, .24]**	↑ **.20 [.15, .26]**	↓ **-.06 [-.11, -.02]**	= -.05 [-.10, .01]
	2	↑ **.12 [.05, .18]**	↑ **.24 [.20, .29]**	= -.06 [-.12, .00]	↓ **-.05 [-.09, -.01]**
	3_{fam}	= .09 [-.03, .21]	↑ **.65 [.56, .73]**	↓ **-.18 [-.27, -.10]**	= -.02 [-.09, .04]
	3_{arb}	↑ **.11 [.03, .21]**	↑ **.58 [.50, .66]**	↓ **-.11 [-.18, -.05]**	↓ **-.12 [-.20, -.04]**
Guessing	1a	↓ **-.23 [-.27, -18]**	↓ **-.06 [-.10, -.03]**	↑ **.14 [.09, .17]**	= .02 [-.02, .05]
	1b	↓ **-.16 [-.19, -.13]**	↓ **-.06 [-.08, -.04]**	↑ **.17 [.14, .19]**	= .01 [-.02, .05]
	2	↓ **-.18 [-.26, -.12]**	↓ **-.07 [-.10, -.03]**	↑ **.12 [.08, .16]**	↑ **.07 [.03 .11]**
	3_{fam}	↓ **-.44 [-.51, -.37]**	↓ **-.30 [-.42, -.18]**	↑ **.18 [.12, .24]**	= .00 [-.08, .09]
	3_{arb}	↓ **-.27 [-.33, -.22]**	↓ **-.37 [-.45, -.30]**	↑ **.26 [.21, .31]**	↑ **.15 [.09, .21]**
Continuous Imprecision	1a	= -1.08 [-2.53, .33]	= -.92 [-2.14, .43]	= .26 [-1.92, 2.42]	= -.31 [-.192, 1.31]
	1b	= .33 [-062, 1.34]	↓ **-1.47 [-2.41, -.51]**	= 1.20 [-.05, 2.44]	= -.73 [-1.79, .39]
	2	↓ **-2.70 [-4.16, -.97]**	= -.01 [-1.03, 1.03]	= .26 [-1.73, 2.37]	↑ **1.76 [.25, 3.02]**
	3_{fam}	↓ **-5.50 [-9.00, -1.32]**	= -1.19 [-4.07, 1.35]	= 1.10 [-2.62, 5.13]	= -1.46 [-5.32, 1.87]
	3_{arb}	= .86 [-1.84, 3.12]	↓ **-1.71 [-3.55, -.14]**	= -1.11 [-3.79, 1.29]	= -.83 [-4.26, 2.36]

Note. (=) no credible change; (↓) credible decrease; (↑) credible increase.

General Discussion

Verbal labeling of continuous colors and shapes in delayed estimation tasks have been found to produce benefits for VWM (Forsberg et al., 2020; Overkott & Souza, 2021; Souza et al., 2020; Souza & Skóra, 2017a). Here we extended this finding to two additional visual features, namely orientation and spatial frequency, and for the use of arbitrary labels to colors. We always observed a substantial reduction in recall error for the labeled feature. Mixture modeling indicated that this labeling benefit originated from an increase in continuous memory for the labeled feature, meaning that verbal labeling increased the retention of fine-grained information. This result stands in contrast to prior assumptions that labeling would only provide categorical information about an item (Donkin et al., 2015; Hardman et al., 2017), as implied in versions of the multi-component model of working memory (Baddeley, 2012; Baddeley & Hitch, 1974; Logie, 2011). This hypothesis would predict a gain in categorical memory with no change in continuous memory, which does not match our pattern of findings. Our results also defy claims that verbal labeling overshadows the visual input (Alogna et al., 2014; Schooler & Engstler-Schooler, 1990) leading to less precise memories. This hypothesis predicts an increase in categorical memory with a reduction in continuous memory, again a mismatching prediction to our findings.

The benefits of labeling for the retention of continuous memory for the labeled feature in VWM have been proposed to arise from the activation of visual categorical knowledge which would help consolidating and maintaining fine-grained information about the presented stimulus (Forsberg et al., 2020; Overkott & Souza, 2021; Souza et al., 2021; Souza & Skóra, 2017). The present findings are in line with the predictions of this hypothesis. Yet, these findings could also be explained by an attentional cue hypothesis assuming that labeling focuses attention on the labeling feature, more efficiently gating the entrance of the labeled feature in VWM, while suppression other types of information and distractions. So far, studies have not attempted to distinguish between these hypotheses.

Critically, in the present work we reasoned that although the categorical visual long-term memory hypothesis was silent about the potential impact of labeling on non-labeled object features, the attentional cue hypothesis made specific predictions regarding the balance between labeling benefits and costs, offering an opportunity to distinguish between them. According to the attentional cue hypothesis, labeling would gate the entrance of the labeled feature in VWM because attention is directed to this feature, whereas the non-labeled feature would be suppressed. This would explain why more continuous memory is available for the labeled feature: more VWM resources would be deployed for the labeled feature. This would however come at the expense of the representations of other information, and non-labeled features would be blocked from VWM. Hence, in the present study, we set-up to test whether labeling would boost memory for the labeled feature at expense of the non-labeled feature.

Across Experiments 1-3, we consistently observed that labeling color, orientation, and spatial frequency always benefited the labeled feature. Mixture modeling showed that the gains derived from labeling accrued from similar sources across features, namely a reduced guessing associated with an increase in continuous memory for the labeled feature. Yet, the labeling costs for the non-labeled features were not uniform across the features. Costs were consistently observed when color was the non-labeled feature, but not orientation and spatial frequency. This finding cannot be readily explained by the attentional-cue hypothesis. This hypothesis would predict that the labeling gains go hand-in-and with the costs: As one feature receives privileged access to VWM, the other feature would be suppressed, thereby explaining labeling benefits as a differential allocation of VWM resources. However, this hypothesis was challenged here in several occasions by the observation of labeling benefits without any costs. Our general data pattern is therefore inconsistent with the predictions of the revised attentional cue hypothesis (Kelly & Heit, 2017b; Overkott & Souza, 2021).

Alternatively, the pattern of findings we obtained seems more readily explained as the combination of two effects. First, for the labeled feature, the activation of categorical information helped the consolidation of continuous information about the labeled feature in VWM, protecting it from forgetting, as predicted by the *categorical visual long-term memory* hypothesis advanced by Souza and Skóra (2017). Second, the non-labeled feature lacked this protection, and hence it remained susceptible to interference, as the one produced by dual-tasks demands. When the demands imposed by overt labeling were substantial, for example, because participants had to quickly (in ca. 1 s) classify the visual information presented using a non-habitual labeling system, the attentional resources consumed by this task imposed a cognitive load that diverted relevant attentional resources that would be used to encode or consolidate the non-labeled feature. Cognitive load is an important drive of working memory performance (Barrouillet et al., 2007, 2011), being considered one of the benchmarks findings in working memory research (Oberauer et al., 2018).

This hypothesis fits well with the data of Experiment 3: when familiar color labels were used to label orientations, the attentional demands of labeling were minimal, and hence orientation memory did not suffer. However, when we increased the demands of color labeling by requiring the use of a non-habitual, arbitrary color-labeling system, orientation memory suffered, akin to the costs observed for color memory that were incurred by labeling the orientations.

In sum, our results suggest that labeling protects the labeled feature, while leaving non-labeled features intact. Yet, this does not mean that non-labeled features will not show a cost. These features do not benefit from the protective activation generated by labeling, and hence they are labile to suffer from any potential interference, including the one created by the act of labeling itself. If the demands on labeling are high, labeling will create a cost. This is quite clear for the use of arbitrary color labels, but it seems also sensible to assume that orientation and spatial frequency labeling could be similarly challenging. Participants probably struggled to generate labels to these features during the experiment, thereby demanding a large share of attention to complete the task. Conversely, if labeling becomes automatic, as it is the case for the use of familiar color terms, then labeling would not draw attentional resources away from encoding and consolidation and the non-labeled feature will remain intact.

All in all, these results continue to support the *categorical visual long-term memory hypothesis* as the best explanation of the labeling effect (Forsberg et al., 2020; Overkott & Souza, 2021; Souza et al., 2021; Souza & Skóra, 2017b), and help rule out the attentional-cue hypothesis as a candidate for accounting for the interaction of verbal and visual inputs in working memory.

Conclusion

Verbally labeling only one of the relevant features of an object always boosted continuous working memory of the labeled feature. Labeling was sometimes inconsequential to non-labeled features, but sometimes it increased its forgetting. These costs were an aftereffect of the attentional demands of the labeling task: when labeling was hard, attention was diverted away from non-labeled features, leading to their forgetting.

References

Adam, K. C. S., Mance, I., Fukuda, K., & Vogel, E. K. (2015). The Contribution of Attentional Lapses to Individual Differences in Visual Working Memory Capacity. *Journal of Cognitive Neuroscience*, *27*(8), 1601–1616. https://doi.org/10.1162/jocn_a_00811

Adam, K. C. S., & Vogel, E. K. (2017). Confident failures: Lapses of working memory reveal a metacognitive blind spot. *Attention, Perception, & Psychophysics*, 1–18. https://doi.org/10.3758/s13414-017-1331-8

Alogna, V. K., Attaya, M. K., Aucoin, P., Bahník, Š., Birch, S., Birt, A. R., Bornstein, B. H., Bouwmeester, S., Brandimonte, M. A., & Brown, C. (2014). Registered replication report: Schooler and engstler-schooler (1990). *Perspectives on Psychological Science*, *9*(5), 556–578.

Arnicane, A., Oberauer, K., & Souza, A. S. (2021). Validity of attention self-reports in younger and older adults. *Cognition*, *206*, 104482. https://doi.org/10.1016/j.cognition.2020.104482

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

Baddeley, A., & Hitch, G. J. (1974). Working memory. *Psychology of Learning and Motivation*, *8*, 47–89.

Bae, G. Y., Olkkonen, M., Allred, S. R., Wilson, C., & Flombaum, J. I. (2014). Stimulus-specific variability in color working memory with delayed estimation. *Journal of Vision*, *14*(4), 7–7. https://doi.org/10.1167/14.4.7

Bae, G.-Y., & Flombaum, J. I. (2013). Two items remembered as precisely as one: How integral features can improve visual working memory. *Psychological Science*, *24*(10), 2038–2047.

Bae, G.-Y., Olkkonen, M., Allred, S. R., & Flombaum, J. I. (2015). Why some colors appear more memorable than others: A model combining categories and particulars in color working memory. *Journal of Experimental Psychology: General*, *144*(4), 744.

Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(3), 570–585. https://doi.org/10.1037/0278-7393.33.3.570

Barrouillet, P., Portrat, S., & Camos, V. (2011). On the law relating processing to storage in working memory. *Psychological Review*, *118*(2), 175–192. https://doi.org/10.1037/a0022324

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

Bays, P. M., Wu, E. Y., & Husain, M. (2011). Storage and binding of object features in visual working memory. *Neuropsychologia*, *49*(6), 1622–1631.

Bocincova, A., & Johnson, J. S. (2019). The time course of encoding and maintenance of task-relevant versus irrelevant object features in working memory. *Cortex*, *111*, 196–209. https://doi.org/10.1016/j.cortex.2018.10.013

Bocincova, A., van Lamsweerde, A. E., & Johnson, J. S. (2017). The role of top-down suppression in mitigating the disruptive effects of task-irrelevant feature changes in visual working memory. *Memory & Cognition*, *45*(8), 1411–1422.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Chen, H., & Wyble, B. (2015). The location but not the attributes of visual cues are automatically encoded into working memory. *Vision Research*, *107*, 76–85.

Donkin, C., Nosofsky, R., Gold, J., & Shiffrin, R. (2015). Verbal labeling, gradual decay, and sudden death in visual short-term memory. *Psychonomic Bulletin & Review*, *22*(1), 170–178.

Forsberg, A., Johnson, W., & Logie, R. H. (2019). Aging and feature-binding in visual working memory: The role of verbal rehearsal. *Psychology and Aging*, *34*(7), 933–953. https://doi.org/10.1037/pag0000391

Forsberg, A., Johnson, W., & Logie, R. H. (2020). Cognitive aging and verbal labeling in continuous visual memory. *Memory & Cognition*.

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

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

Gelman, A., Hwang, J., & Vehtari, A. (2014). Understanding predictive information criteria for Bayesian models. *Statistics and Computing*, *24*(6), 997–1016.

Hardman, K. O. (2016). *CatContModel: Categorical and Continuous Working Memory Models for Delayed Estimation Tasks* (0.7.5) [Computer Software]. https://github.com/hardmanko/CatContModel/releases/tag/v0.7.5

Hardman, K. O., & Cowan, N. (2015). Remembering Complex Objects in Visual Working Memory: Do Capacity Limits Restrict Objects or Features? *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *41*(2), 325–347. https://doi.org/10.1037/xlm0000031

Hardman, K. O., Vergauwe, E., & Ricker, T. J. (2017). Categorical working memory representations are used in delayed estimation of continuous colors. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(1), 30–54. https://doi.org/10.1037/xhp0000290

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

Kelly, L. J., & Heit, E. (2017a). Recognition memory for hue: Prototypical bias and the role of labeling. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(6), 955.

Kelly, L. J., & Heit, E. (2017b). Recognition memory for hue: Prototypical bias and the role of labeling. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(6), 955–971. https://doi.org/10.1037/xlm0000357

Kruschke, J., & Meredith, M. (2021). *BEST: Bayesian Estimation Supersedes the t-Test* (0.5.4). https://CRAN.R-project.org/package=BEST

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

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.

Lupyan, G. (2012). Linguistically modulated perception and cognition: The label-feedback hypothesis. *Frontiers in Psychology*, *3*, 54.

Maniglia, M. R., & Souza, A. S. (2020). Age Differences in the Efficiency of Filtering and Ignoring Distraction in Working Memory. *Manuscript Submitted for Publication*.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Reason*, *4*(2), 61–64.

Morey, R. D., & Rouder, J. N. (2015). *BayesFactor: Computation of Bayes Factors for Common Designs.* (R package version 0.9.12-2) [Computer software].

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

Oberauer, K., Farrell, S., Jarrold, C., & Lewandowsky, S. (2016). What limits working memory capacity? *Psychological Bulletin*, *142*(7), 758–799. https://doi.org/10.1037/bul0000046

Oberauer, K., Lewandowsky, S., Awh, E., Brown, G. D. A., Conway, A., Cowan, N., Donkin, C., Farrell, S., Hitch, G. J., Hurlstone, M. J., Ma, W. J., Morey, C. C., Nee, D. E., Schweppe, J., Vergauwe, E., & Ward, G. (2018). Benchmarks for models of short-term and working memory. *Psychological Bulletin*, *144*(9), 885–958. https://doi.org/10.1037/bul0000153

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

Overkott, C., & Souza, A. S. (2021). Verbal descriptions improve visual working memory but have limited impact on visual long-term memory. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0001084

Palmer, J., Boston, B., & Moore, C. M. (2015). Limited capacity for memory tasks with multiple features within a single object. *Attention, Perception, & Psychophysics, 77*(5), 1488–1499. https://doi.org/10.3758/s13414-015-0909-2

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442. https://doi.org/10.1163/156856897X00366

Pratte, M. S., Park, Y. E., Rademaker, R. L., & Tong, F. (2017). Accounting for stimulus-specific variation in precision reveals a discrete capacity limit in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(1), 6–17. https://doi.org/10.1037/xhp0000302

Quak, M., Langford, Z. D., London, R. E., & Talsma, D. (2018). Contralateral delay activity does not reflect behavioral feature load in visual working memory. *Biological Psychology*, *137*, 107–115.

R Core Team. (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/.

Rock, I., Linnett, C. M., Grant, P., & Mack, A. (1992). Perception without attention: Results of a new method. *Cognitive Psychology*, *24*(4), 502–534.

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374.

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

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

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.

Sense, F., Morey, C. C., Prince, M., Heathcote, A., & Morey, R. D. (2017). Opportunity for verbalization does not improve visual change detection performance: A state-trace analysis. *Behavior Research Methods*, *49*(3), 853–862. https://doi.org/10.3758/s13428-016-0741-1

Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-Specific Delay Activity in Human Primary Visual Cortex. *Psychological Science*, *20*(2), 207–214. https://doi.org/10.1111/j.1467-9280.2009.02276.x

Shin, H., & Ma, W. J. (2016). Crowdsourced single-trial probes of visual working memory for irrelevant features. *Journal of Vision*, *16*(5), 10–10. https://doi.org/10.1167/16.5.10

Shin, H., & Ma, W. J. (2017). Visual short-term memory for oriented, colored objects. *Journal of Vision*, *17*(9), 12–12. https://doi.org/10.1167/17.9.12

Souza, A. S., Overkott, C., & Matyja, M. (2020). Categorical distinctiveness constrains the labeling benefit in visual working memory. *Submitted Manuscript*.

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

Souza, A. S., Rerko, L., & Oberauer, K. (2016). Getting more from visual working memory: Retro-cues enhance retrieval and protect from visual interference. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(6), 890.

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

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

Swan, G., Collins, J., & Wyble, B. (2016). Memory for a single object has differently variable precisions for relevant and irrelevant features. *Journal of Vision*, *16*(3), 32–32. https://doi.org/10.1167/16.3.32

Wetzels, R., & Wagenmakers, E.-J. (2012). A default Bayesian hypothesis test for correlations and partial correlations. *Psychonomic Bulletin & Review*, *19*(6), 1057–1064.

Yu, Q., & Shim, W. M. (2017). Occipital, parietal, and frontal cortices selectively maintain task-relevant features of multi-feature objects in visual working memory. *NeuroImage*, *157*, 97–107. https://doi.org/10.1016/j.neuroimage.2017.05.055

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