Feature identity determines representation structure in working memory

Ricker, Timothy; Souza, Alessandra; Vergauwe, Evie

Feature identity determines representation structure in working memory

Timothy J. Ricker

University of South Dakota

Alessandra S. Souza

Center for Psychology, Faculty of Psychology and Education Sciences, University of Porto

&

Evie Vergauwe

University of Geneva

Word Count: 8361 words in the main text (including tables, figure labels, and captions).

Author's Note: All data first reported in this work and all code for our quantitative analyses are available on the Open Science Framework at https://osf.io/bv6fh/. Study materials are available upon request. An early version of this manuscript was disseminated as an Open Science Framework preprint. Some of the data and ideas presented in this manuscript were previously presented at the Distributed Working Memory Series, as well as the 2021 and 2022 Annual Psychonomic Society Conference. A. S. Souza was supported by a grant from the Swiss National Foundation (100019_169302) and from national funding from the Portuguese Foundation for Science and Technology (UIDB/00050/2020).

The authors wish to thank William Xiang Quan Ngiam for the independent verification of the reproducibility of the results of Experiment 1 and for suggesting the inclusion of a README file on our OSF page. Thanks to Liya Thomas and Mabel Etuknwa for their help with data collection in Experiments 1, 3, and 4. Thanks to Danielle Aylmer, and William Frickenstein for their help with data collection in Experiment 2.

Send correspondence to Timothy J. Ricker, University of South Dakota, 414 E. Clark St., Vermillion, SD, 57069, USA. Email: Timothy.Ricker@usd.edu

Abstract

Visual working memory maintains both continuous-perceptual information and discrete-categorical information about memory items. Both types of information are represented in working memory, but the representation structure remains unknown. Continuous and categorical information about a single stimulus could be represented separately, in two different representations. Alternatively, continuous and categorical information could be represented jointly as a single representation. To investigate this, we fitted two different computational models to delayed estimation data assuming either separate or joint representations of continuous and categorical information in working memory, for three different, commonly used features (orientation, color, and shape). Across a set of 9 experiments, model fits clearly show that feature identity drives the representation structure, with a joint representation structure for orientation, but a separate representations structure for color and shape. This pattern was remarkably invariant across a variety of task contexts. Existing models miss this distinction, leading to mischaracterization of memory precision.

Keywords: visual working memory; short-term memory; computational modeling; delayed estimation; mental representation

Public Significance Statement

Mental representations are the building blocks of human thought. Accordingly, understanding the structure of these representations is a critical step for the accurate modeling of human thought processes. In the present work, we show that the build-up of visual representations differs for three common features: orientation, color, and shape. Our results indicate that orientation memories combine perceptual (i.e., fine-grained details) and categorical (i.e., prototypical or gist-like information) into a single representation, whereas for color and shape these sorts of information are stored and used separately. By combining both sorts of information into a single unit, orientation memories are necessarily biased by categorical knowledge (e.g., cardinal orientations such as top, bottom, left and right), impacting the fidelity or precision of this information in mind.

Feature identity determines representation structure in working memory

To remember an image over a brief delay requires representing it in working memory. Research over the last two decades has increasingly focused on differentiating the number of items stored in memory and the precision with which those items are maintained (Adam et al., 2017; Bays et al., 2009; Bays & Husain, 2008; Forsberg et al., 2020; Ma et al., 2014; Ovalle-Fresa et al., 2021; Pratte et al., 2017; Rouder et al., 2008; van den Berg et al., 2014; Zhang & Luck, 2008). While studying memory precision has resulted in many new insights about visual working memory, our understanding of how information is represented in visual working memory is still quite rudimentary. This presents problems for any study relying on measures of memory precision. The precision of a memory cannot be reasonably estimated without knowledge of the structure of the involved representations. The present work aims to advance our understanding of what is stored in memory by testing whether continuous-perceptual and discrete-categorical aspects of visual features such as orientation, color and shape are represented separately or jointly in visual working memory.

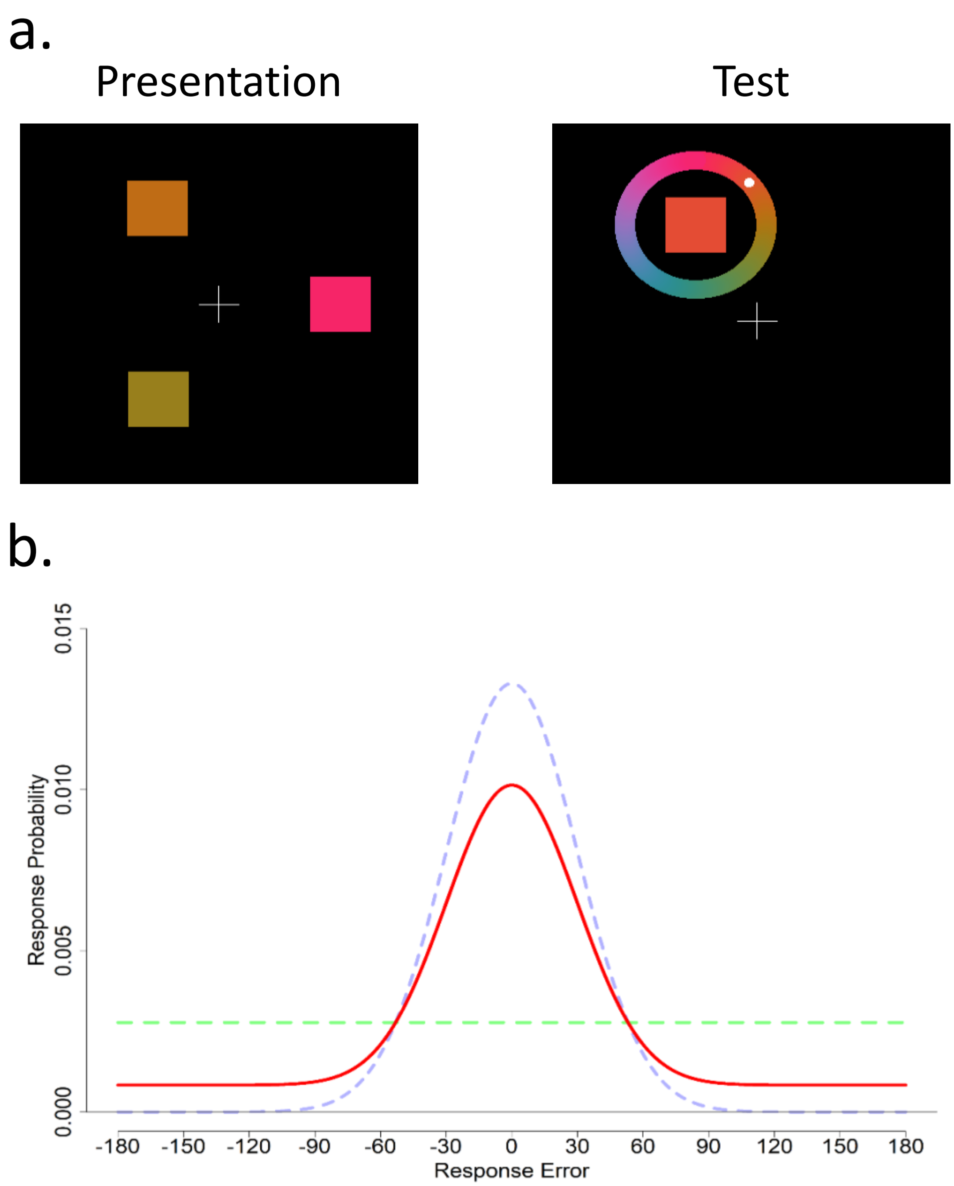
Delayed Estimation and Working Memory Representations

The delayed estimation task has become a popular experimental paradigm to examine the properties of visual representations in working memory (Bays et al., 2009; Wilken & Ma, 2004; Zhang & Luck, 2008). In this task, individuals see a set of memory items varying in a continuous visual feature, typically colors, and must reproduce this feature within a circular feature space after a brief delay (see Figure 1a). Zhang and Luck (2008) developed a computational model of performance on this task that characterizes responses as memory-based or guessing (see Figure 1b). Memory-based responses were assumed to occur with some probability, P_m. These responses were assumed to be centered on the feature of the presented item with some standard deviation (δ), which were assumed to characterize the variance in responses around the presented feature, and hence the precision of the memory representation. Guessing responses were assumed to occur with probability $1\text{-}P_m$ and were characterized as a distribution of uniform responses throughout the feature space. The assumption that memory-based responses are centered on the presented feature implies that the memory representation is of the exact continuous feature value, not of a gist or categorical representation of the visual feature.

Zhang and Luck (2008)'s central argument was not about the nature of the representation, but rather the nature of capacity limits. They argued in favor of a discrete item limit governing visual working memory capacity based upon the finding that changes in the set size lead primarily to changes in the probability an item was in memory. The primary challenge to the discrete item-based model of delayed estimation performance has been from the distributed resource model of Bays and Husain (2008) who argue that capacity is limited by a flexible resource that determines the precision of representations. According to Bays and Husain, all items are represented in memory, but the memory capacity limits how precise each of these items are represented. If more of the resource is devoted to an item, then that item should show a higher level of precision and this should lead to lower precision for other memory items. Much work has since debated whether a discrete item limit or a flexible-resource limits working memory capacity (e.g., Adam et al., 2017; Bays et al., 2009; Ma et al., 2014; Pratte et al., 2017; van den Berg et al., 2014). This is not surprising as a debate between fixed-discrete capacity or variable-continuous memory capacity limits has preceded the computational modeling of delayed estimation performance (Baddeley et al., 1975; Brown et al., 2007; Cowan, 2001; Crowder, 1976; Kintsch, 1967; Rouder et al., 2008; Wilken & Ma, 2004).

Figure 1

(a.) A typical delayed estimation task for color. (b.) In the delayed estimation model proposed by Zhang and Luck (2008), errors for memory-based responses (in blue) are assumed to be centered at zero and distributed with some standard deviation representing memory precision. Guesses (in green) are assumed to be equally distributed across all possible feature values. A weighted combination of the memory-based and guessing responses produces each individual's overall response distribution (in red).



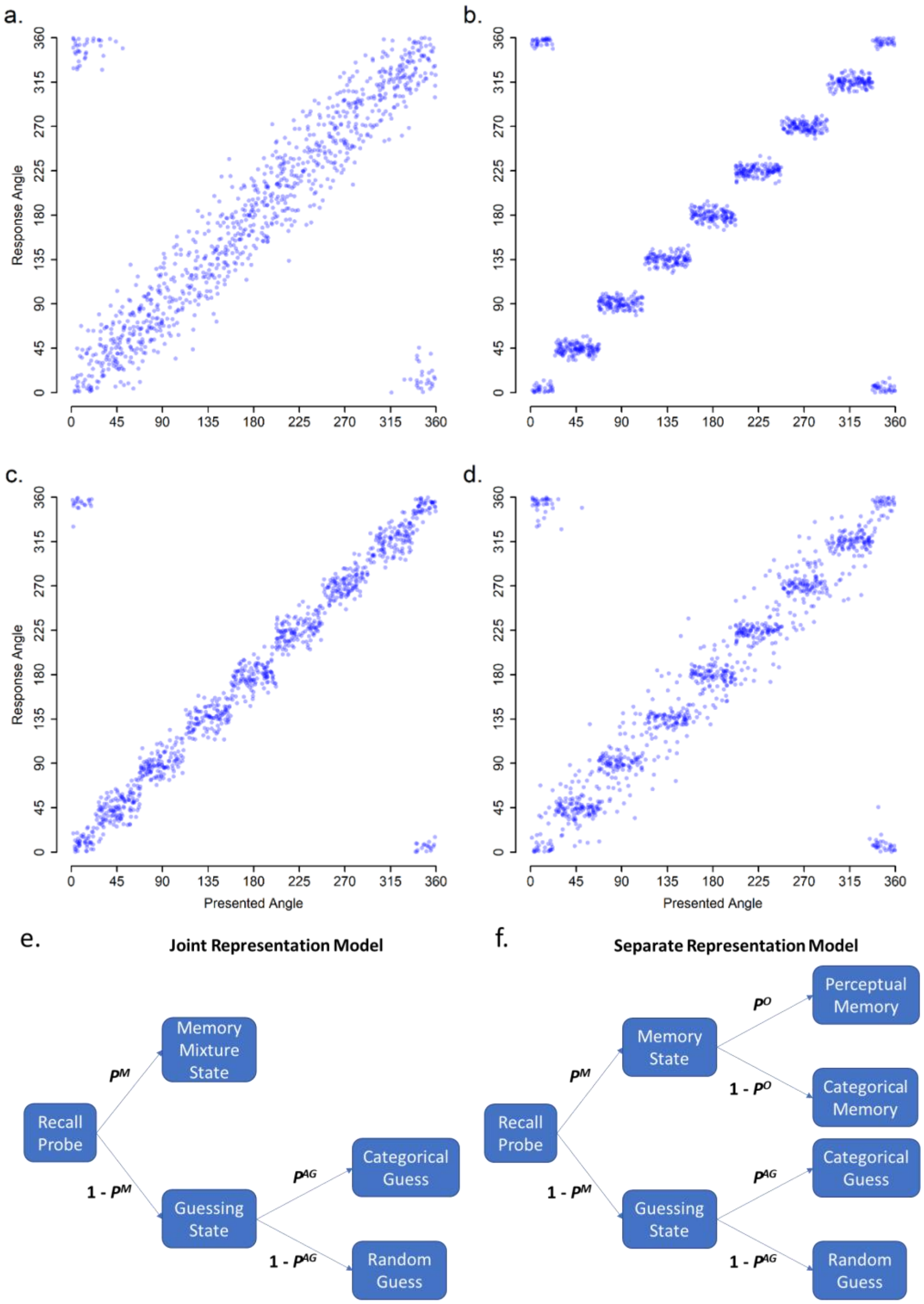
The implicit assumption underlying both approaches to computational modeling of delayed estimation data is that memories are representations of the continuous feature value of the target item and not a verbal label or gist representation. This assumption is what allows the uniform response distribution to be mapped onto guessing rates and the standard deviation of the error distribution to be mapped onto memory precision. Bae et al. (2014) questioned this assumption by asking whether individuals reproduce all perceived and remembered colors equally well. If not, then some process is causing the mental representation to deviate systematically from the color of the presented item. They found consistent biases both in a perceptual estimation task and a delayed working memory estimation task indicating that individuals systematically underproduce some color hues and overproduce others. Participants in their studies tended to reproduce some colors more frequently than others, with the clusters of frequent responses coinciding with color categories, and infrequent responses moving away from color-category boundaries.

Modeling Continuous and Categorical Representations in Working Memory

Bae et al. (2015) expanded upon the notion that participants were remembering color categories by providing a computational model of delayed estimation similar to the Zhang and Luck model but incorporating color categories into the memory response. The traditional approach had been to model the response error, namely the distance between the reported and presented feature value, which disregards what was the specific feature value studied. Rather than modeling response errors, Bae et al. modeled response color as a function of the studied color. This approach allowed them to see evidence for categorical representations, gist representations such as the concept “red” or “left”, that are unobservable in error distributions collapsed across stimulus locations. Categorical representations can be thought of as separate from continuous representations of the exact feature value which have been the more typical target of research. Figure 2 shows a scatterplot of response values as a function of the presented feature value in several scenarios. Panel a shows responses when there is only memory for the continuous feature value. Panel b shows responses when there is only memory for the item category. Whereas continuous memories of the exact feature produce a smooth pattern along the diagonal in Figure 2a, categorical responses result in a staircase pattern as shown in Figure 2b whereby the responses for a range of different stimulus values cluster together around the same response value. Modeling of error distributions, as shown in Figure 1b, cannot differentiate the two response patterns shown separately in Figures 2a and 2b. Error distributions become blind to color category signatures because all stimulus values are pooled into a single error distribution, making the category-based offset from the diagonal appear to be reduced memory precision. Changes in estimates of precision inevitably alter the estimates of guessing rates as both parameters are estimated jointly. The lack of accurate model specification in earlier models brings any findings regarding parameter values into question.

Figure 2

Response angle resulting from successful memory retention as a function of the presented angle for (a.) continuous representations and (b.) categorical representations. Simulated response patterns produced by successful memory retention in the (c.) joint representation model and (d.) separate representation model. Tree model representations of the (e.) joint representation variant and (f.) separate representation variant. In the joint representation model P^O represents the memory-state mixture weighting. Larger values mean more weight on the continuous perceptual memory element. Smaller values mean more weight on the categorical memory element.



Hardman et al. (2017) developed a model of color representation similar to Bae et al. (2015), but with differences in how color categories are determined and in how categorical representations contribute to the representation held in visual working memory. Whereas Bae et al.'s approach had individuals label their internal color categories, Hardman et al. estimated the category locations from the data itself (see also Pratte et al., 2017), thereby allowing the category locations to vary across participants and trials. Both Bae et al. (2015) and Hardman et al. (2017) agree that categorical representations are necessary to accurately model delayed estimation data. They differ in how categorical information is incorporated into the overall mental representation of the stimulus. In Bae et al.'s approach the categorical representation is implemented as a bias upon the precise continuous representation. The category and the continuous information are combined to form a mental representation that is different from either one alone. We refer to this model as the joint representation model (see Figure 2c). This representation is conceptually equivalent to remembering the exact shade of red presented, but with a bias towards the individual's conception of a generic red color. The resulting memory is more stereotypical than the presented color, with the deviation from the presented color being larger as the presented shade gets farther from the generic category value. Within this model, responses are drawn from a single distribution that is a mixture of the categorical and continuous response distributions. Memory-based responses form a characteristic pattern on the stimulus-response plot in Figure 2c, zigzagging above and below the diagonal across the stimulus space.

Hardman et al., on the other hand, implemented categorical and continuous information as two separate mental representations. This second approach assumes that the exact hue and the general color category can both be remembered, but that these representations are maintained and used separately. We refer to this model as the separate representation model (see Figure 2d).

Memory-based responses on the stimulus-response plot in Figure 2d show the individual response patterns from Figures 2a and 2b superimposed upon one another.

The two methods of incorporating categorical representations into delayed estimation models were tested and contrasted by Hardman et al. (2017) to establish whether individuals maintained joint or separate representations of the color category and the continuous color hue presented. Hardman et al. referred to these as the within and between model variants, but joint and separate representation models are perhaps more intuitive names for differentiating the two variants. In two experiments, Hardman and colleagues found that the separate representation model outperformed the joint representation model, indicating that individuals represent categorical and continuous information as separate mental representations when remembering colors for brief periods of time.

Studies of visual working memory using the delayed estimation paradigm continue to be abundant with a strong focus on interpreting computational model parameter estimates (e.g., Forsberg et al., 2020; Huang, 2020; Long et al., 2020; Ovalle-Fresa et al., 2021; Pratte, 2019; Rhodes et al., 2020; Schneegans et al., 2021; Son et al., 2020). For these parameter values to be meaningful, an accurate model of the structure of the mental representation involved must be used for their estimation. This includes an accurate characterization of how categorical information is used in combination with continuous perceptual detail in visual working memory. While Hardman et al. (2017) found a superior fit of the separate representation model over the joint representation model (see also Souza & Skóra, 2017), this has not been tested with any features beyond color. That is the goal of the present study.

The Present Study

Here, we examined whether representations of continuous detail and categorical gist are always stored separately or whether, instead, the representation structure varies across feature identity, physical characteristics of the stimuli, the number or features per stimulus, and task set-ups. First, we report four new experiments that were conducted to examine the structure of the working memory representation of orientation across different task set-ups (Experiments 1-4). Next, we report the re-analysis of three published experiments in which the memoranda were multi-feature objects, allowing us to study the structure of the working memory representation of color, orientation, and shape when more than one feature was maintained (Experiments 1a-b and 2 of Overkott & Souza, in press). Finally, we report the re-analysis of two published experiments in which shape needed to be remembered, allowing us to study the structure of the representation for yet another feature in visual working memory, again across task situations (Experiments of Souza et al., 2021, and Li et al., 2022). Across these three sets of analyses, we found that not all memory representations use the same structural format. Feature identity (orientation, color, or shape) drives the structure of the memory representation irrespective of the type of the physical characteristics of the stimuli, the number or features per stimulus, and task set-up.

Representing Orientation: Experiments 1-4

In a first step, four experiments were conducted to test whether the mental representation of orientation maintains continuous and categorical information separately or jointly. In these experiments, a variable number of orientation stimuli were presented. Once all stimuli were presented, they were immediately probed and reproduced by the participants. Orientations varied from 1-360 degrees by increments of 1 degree. Participant response angles were modeled using both the joint (see Figure 2c) and separate (see Figure 2d) model variants developed by Hardman et al. (2017) to determine the structure of the mental representation. In the following paragraphs we explain the basic model structure for each variant, as illustrated in Figure 2e for the joint variant and 2f for the separate variant. For formal model specifications see the original work by Hardman et al.

The separate-representations model variant (see Figure 2f) assumes that there is some probability, P^M, that a probed memory item will be available in mind at test, otherwise the participant responds by guessing, with probability $1- P^M$. If a response is informed by memory, then there is some probability, P^O, that the response is based on a continuous representation (similar to the distribution shown in Figure 2a). In this case, the participant's response distribution will be centered at the location of the presented feature with continuous imprecision parameter, σ^O, indexing memory variability. Alternatively, with probability $1- P^O$, the memory response is based on a categorical representation of the item (as illustrated in Figure 2b). In that case, the response will be centered on values that represent the stereotypical representation of a category, such as 180 degrees for down or 90 degrees for right. Response variability of categorical responses is characterized by parameter, σ^A. Hence, in the separate-representation model variant, for each memory item on a given trial, the participant's response is based on either a continuous representation or a categorical representation of the item, but not both. The number and center of the categories are estimated from the data for each participant with variability assumed across trials, allowing the model to account for variability in categorization schemes across participants and throughout an experimental session. Categories are estimated without fixing the distance between category centers, allowing both narrow categories and categories that stretch across a wide stimulus space. Category estimation details are presented in

Hardman et al. (2017). When participants guess, they respond by selecting a random category, with probability P^{AG}, or by selecting a random value from a uniform distribution with probability 1 - P^{AG}.

The joint-representation model variant (see Figure 2e) uses the same basic structure as the separate representations model, with one major change. In this model there is only one type of memory representation that drives all memory-based responses. This memory representation has both a continuous perceptual component and a categorical component. Thus, in contrast to the separate representation model, here, for each memory item on a given trial, the participant's response is based on a mixture of continuous and categorical information of the item. Parameter P^O serves as the continuous distribution mixing weight in this model variant and can be thought of as the proportion of the representation of the item in question that is based upon the continuous perceptual representation. When $P^O = 1$, then participants only use the continuous information for responding, with no categorical component (see Figure 2a). When $P^O = 0$, the response is instead purely based on the categorical information (see Figure 2b). Intermediate values determine the relative strength of each component that makes up the joint memory representation and that drives the participant's response. Figure 2c shows the pattern of memory-based responses produced when $P^O = 0.6$.

In Experiments 1-4, we test whether the separate representation or the joint representation model captures best how orientation is represented in visual working memory. It is important to note that our approach at this step was to determine via model comparison which model variant (separate or joint representation) best described the response patterns across task set-ups. This contrast will allow us to make inferences about the basic structure of the mental representation used in working memory, something that changes in parameter values cannot reveal. In the General Discussion, we will discuss how the model variant impacts parameter values.

Method

Participants

One-hundred nighty-nine individuals (ages 18-52 years, M = 20.0; 139 female, 60 male)[1] with normal or corrected-to-normal vision participated in Experiments 1 to 4 (47 in Experiment 1, 46 in Experiment 2, 61 in Experiment 3, 45 in Experiment 4). The recruitment strategy was to collect data from at least 40 participants in each experiment. Participants signed up several weeks in advance for each session. Enrollment was reviewed weekly to determine whether more data needed to be collected. This resulted in sample sizes larger than the minimum target number. Participants were recruited from the psychology department participant pool at the College of Staten Island and received partial course credit for participation. The study was approved by the City University of New York Integrated Institutional Review Board, project #2015-1156. All participants gave informed consent prior to participation.

Methods and Procedure

All four experiments used a variation of the same basic design and procedure shown in Figure 3. Within each experiment the memory set size (1, 2, 3, or 5 items) was varied randomly across trials. All other variables were held constant across trials within an experiment. We next explain the basic structure of a trial in all experiments and then concisely summarize the differences across experiments.

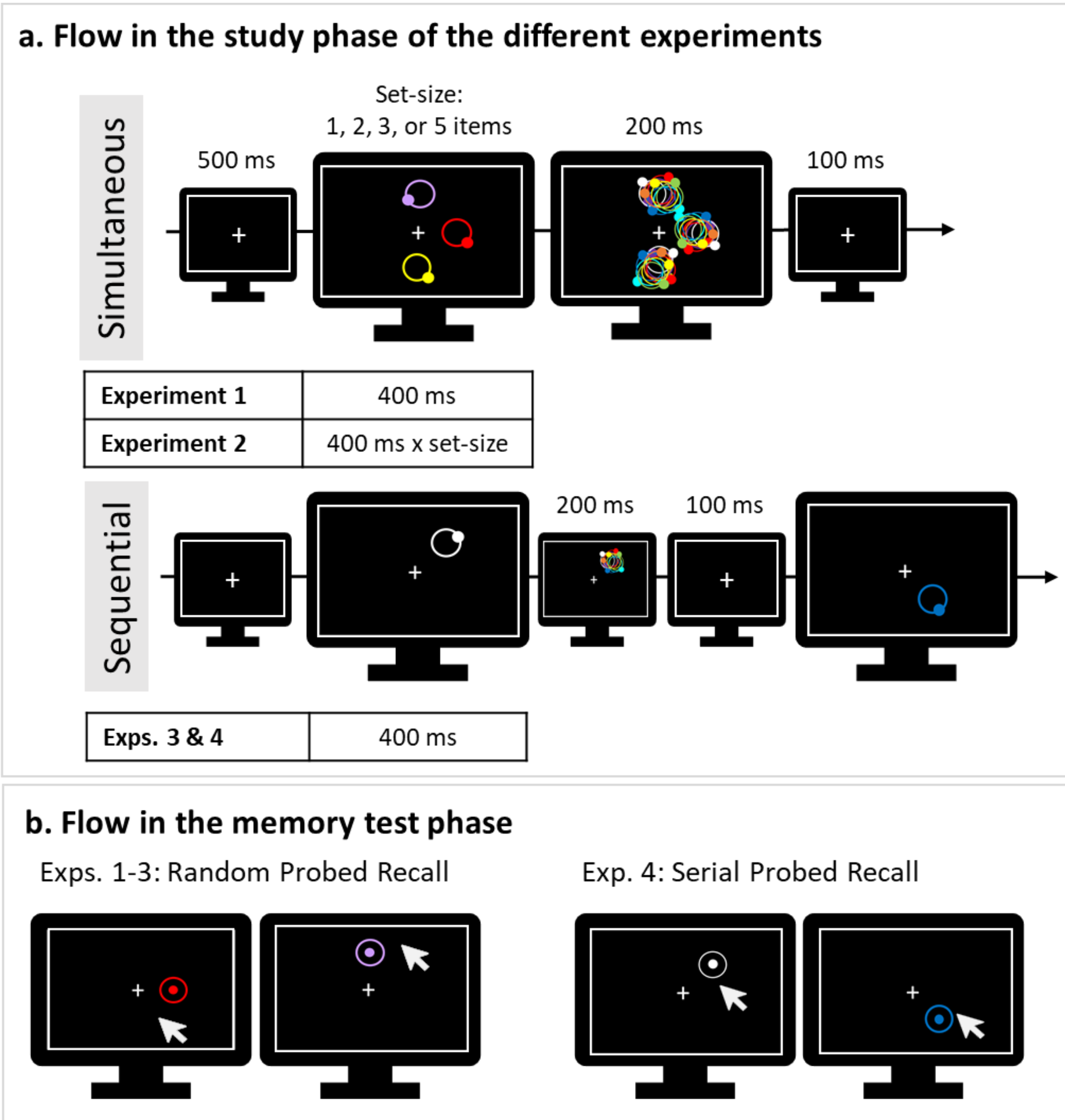
Participants initiated each trial by pressing the space bar. A fixation cross was presented on the screen for 500 ms, followed by memory item presentation. Memory items were composed

[1] Participant gender was collected by a demographic question asking participants to “Select Your Gender” with two options “Male” and “Female”.

of a ring with a dot somewhere along its perimeter (360 possible locations, equally spaced). The task was to remember the orientation of each dot on the ring. The ring diameter was in 36 mm in Experiments 1 and 2, and 23 mm in Experiments 3 and 4. Dot diameter was 5 mm in Experiments 1 and 2, and 3 mm in Experiments 3 and 4. The rings appeared in 1 of 8 equally-spaced locations along an invisible circle, centered on the middle of the screen. Each presentation location was at a distance from the center of the screen of 67 mm in Experiments 1 and 2, and 42 mm in Experiments 3 and 4. The difference in size across experiments was due to a change in computer monitors. Each location was associated with a specific color. Whenever a ring was shown it appeared in the color associated with that location (colors: red, grey, blue, yellow, violet, orange, lime, cyan). Ring color was irrelevant to the task and intended as a redundant retrieval cue in addition to the item location. The background color was black.

Figure 3

Example of the flow of events for Experiments 1-4, the new experiments to examine representation structure of orientations. Panel (a.) shows the item study phase. Panel (b.) shows the item recall phase.



In Experiments 1 and 2, all memory items were presented concurrently as a memory array. In Experiment 1 the entire array was presented for 400 ms. In Experiment 2 the array was presented for 400 ms x the set-size (i.e., 400, 800, 1200, or 2000 ms, for 1, 2, 3, or 5 memory items, respectively). Experiments 3 and 4 used serial presentation of the memory items with each item presented for 400 ms.

After presentation of each memory item a mask appeared at the location of that memory item for 200 ms. The mask was composed of 8 rings, each slightly displaced from the location of the memory item, and eight dots placed randomly within the area that the rings were drawn. Each of the rings and dots were a unique color matching the colors associated with each of the eight locations. After each mask was a 100 ms delay containing only a fixation cross before the next stimulus appeared.

Once the item presentation sequence was completed for all items, memory was probed for all items. Memory probes were presented for each item individually. The probe sequence was in random order in Experiments 1, 2, and 3. The probe sequence was in the order of presentation in Experiment 4.

After each trial, feedback was given. The feedback screen showed all rings with the correct location marked by a white dot and the participants response marked by a colored dot matching the color of the ring. If mean response error was less than or equal to 30 degrees a happy tone sequence played. If mean error was greater than 30 and less than or equal to 60 degrees a neutral tone sequence played. If mean error was greater than 60 degrees a sad tone sequence played.

There were 8 practice trials in all experiments. In Experiments 1, 3, and 4, there were 6 blocks of 35 experimental trials in. In Experiment 2, there were 8 blocks of 40 trials. A summary of the experimental manipulations in each experiment is given in Table 1.

Table 1
Summary of Experimental Manipulations in Experiments 1-4

Experiment	Presentation Type	Presentation Time	Recall Order
1	Simultaneous	400ms	Random
2	Simultaneous	400ms x Set Size	Random
3	Serial	400ms per item	Random
4	Serial	400ms per item	Serial

Analysis

We fit the two versions of our model to individual response angles with Bayesian Markov-Chain Monte-Carlo sampling using the "*CatContModel*" package (Hardman, 2017) for the R statistical computing language. Formal model specification can be found in Hardman et al. (2017). We ran 11,000 iterations and discarded the first 1,000 as burn in. Model selection was performed using Watanabe-Akaike Information Criterion (WAIC), a version of AIC appropriate for hierarchical Bayesian model selection that accounts for effective number of parameters in the penalty term. For details on WAIC calculation and interpretation see Gelman et al. (2014).

Transparency and Openness

All data first reported in this work and all code for our quantitative analyses are available on the Open Science Framework at https://osf.io/bv6fh/. This study was not preregistered.

Results

Mean error and key parameter estimates for Experiments 1-4 are provided as a function of Set Size and Experiment in Figure 4. As can be seen in Table 2, the joint model variant outperformed the separate model variant by a large margin in all four experiments, with a WAIC advantage of 261 points in Experiment 1, 586 points in Experiment 2, 693 points in Experiment 3, and 514 points in Experiment 4.

Figure 4

Error and parameter values as a function of condition and experiment. (a.) Mean error, error bars are standard error of the mean. (b.) Probability an item is in Memory (c.) Proportion of the mixture that is from continuous perceptual information. (d.) Imprecision of the Continuous distribution. Error bars in panels b-d represent the 95% credible interval of the posterior of the estimated parameter.

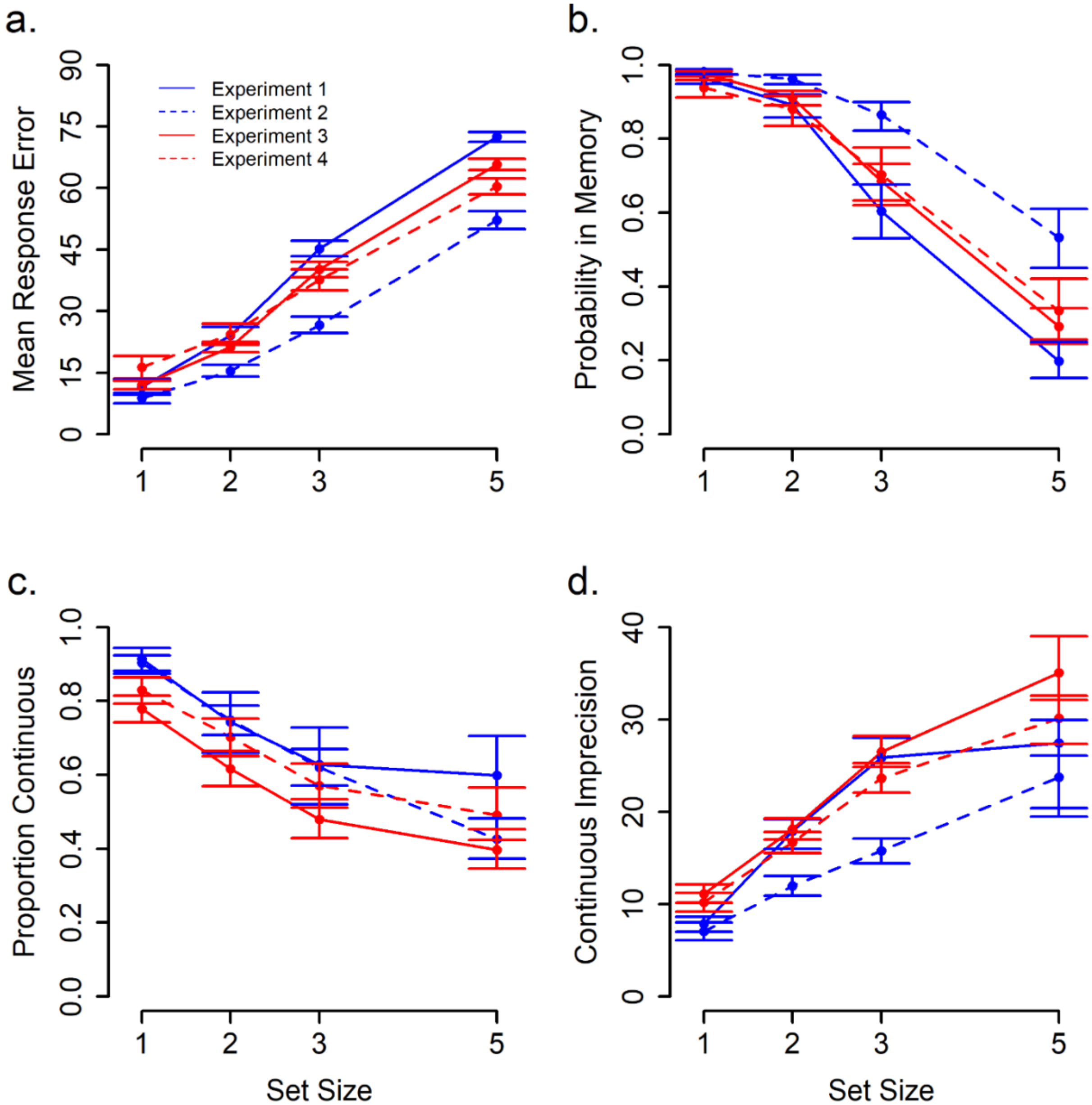


Table 2
WAIC fits for Orientation Experiments 1-4 by Experiment for each Model Variant

Experiment	Model	WAIC	WAIC difference from best model
1	Separate	286,257	261
	Joint	**285,996**	**0**
2	Separate	350,331	586
	Joint	**349,745**	**0**
3	Separate	359,737	693
	Joint	**359,044**	**0**
4	Separate	261,285	514
	Joint	**260,771**	**0**

**Bold rows indicate the winning model for each experiment.*

Discussion

The results from all four experiments produced the same pattern of findings. The joint representation model fit participant response-angle data for orientations better than the separate representation model by a clear margin irrespective of the task set-up. This is in contrast to the findings of Hardman et al. (2017) who found that the color data from their two experiments was better fit by the separate representation model. The fact that both sets of observations are consistently observed across the respective experiments indicates that orientation and color use different representation structures within working memory to complete similar tasks. In the following sets of experiments, we assess whether this conclusion is warranted by replicating our current findings for orientation and those of Hardman et al. for colors in several published data sets. We also test whether a third feature, that of continuous shape, is represented using separate or joint representations. Modeling data from different studies using different memory materials and task contexts allowed us to test whether stimulus feature identity drives the representation structure in a way that is task independent.

Representing Orientation and Color in Multi-feature Objects

Overkott and Souza (in press) reported three experiments in which participants recalled the color and orientation of multi-feature objects using a continuous feature space. We fit the separate and joint representation variants of Hardman et al. (2017)'s model to this data to assess whether multiple features of an object can be stored in differing representation structures concurrently. If the stimulus feature drives the representation structure, then we should see that orientation is best modeled as a joint representation of continuous and categorical information while color is best modeled as separate continuous and categorical memory representations, even in multi-feature objects. Alternatively, it could be that the representation structure is determined at the object level and that all features of a single object are encoded with the same type of representation. If this alternative prediction is true, then our model-based analysis should show that orientation and color data both favor a common representation structure when they are part of the same object (i.e., either both using a separate representations structure or both using a joint representation structure). To be clear, this is not a test of whether object-level representations exist, there is ample evidence that they do (Cowan, 2001; Donkin et al., 2013; Ngiam et al., 2022; Rouder et al., 2008). Here we test whether differing features of the same object can use different representation structures or if they are constrained to using a common representation structure.

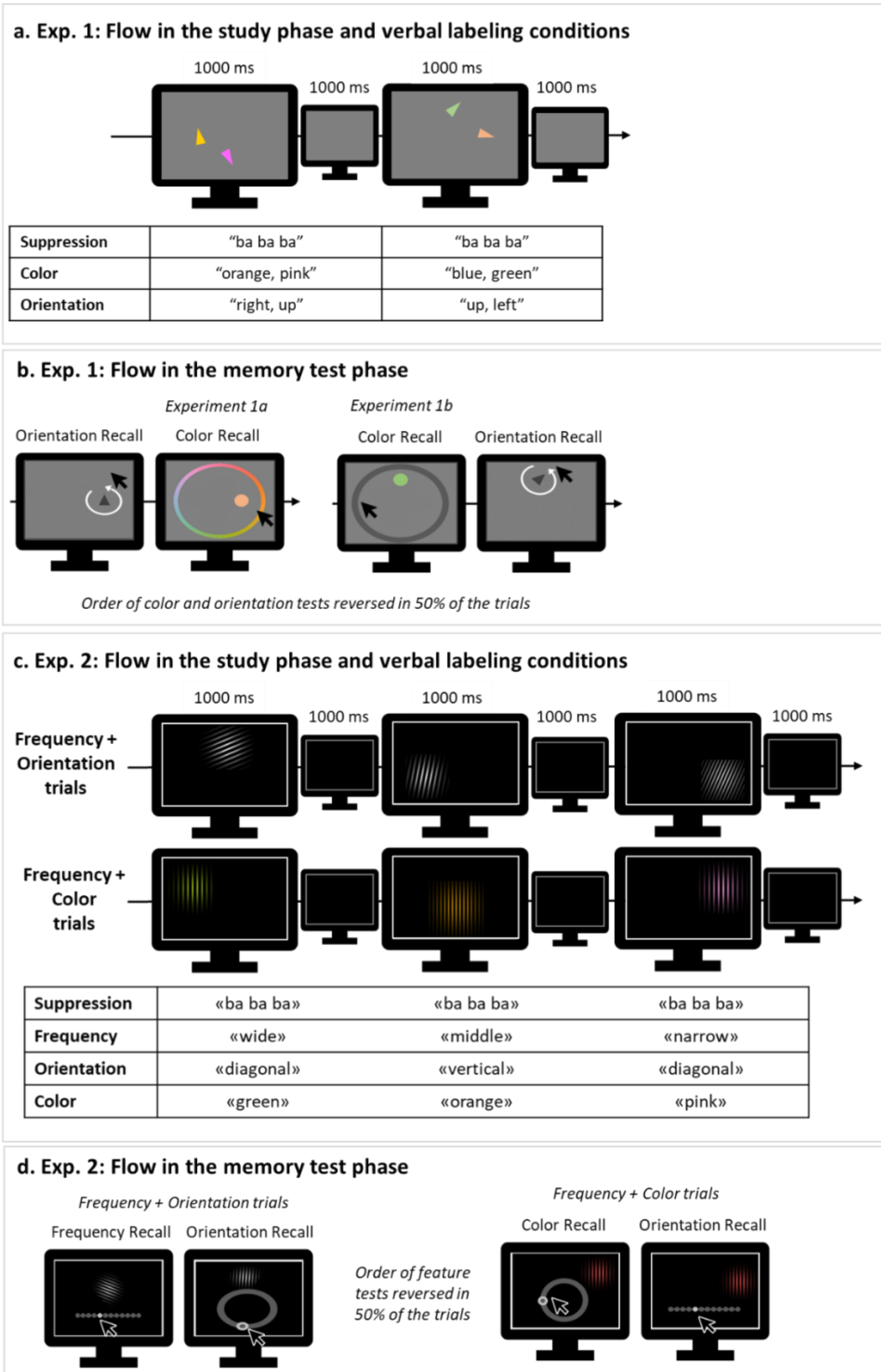
Procedure of Overkott and Souza (in press)

Participants remembered several multi-featured objects and then recalled them on a circular feature space. In the first two experiments (Experiments 1a and 1b, N = 36 and 59, respectively), participants remembered triangles that varied in their orientation and color to recall after a brief delay (see Figures 5a and 5b). In the third experiment, participants (N = 51) saw Gabor patches with varying orientation, color, and frequency for later recall (see Figures 5c and 5d). Depending on the condition, participants in the third experiment were asked to either remember the color and grating frequency (orientation was held constant at 90°) or the orientation and grating frequency (color was held constant at white). The grating frequency took only a limited number of discrete values, so we were not able to model the format of the frequency representation. In all three experiments, participants were asked to remember two features of the same item and were probed to reproduce both features of a randomly selected item.

In all experiments, each memory presentation lasted for 1000 ms with a 1000 ms blank inter-stimulus interval. In Experiments 1a and 1b, two sets of two-colored triangles were presented on each trial, for a total set-size of four items. In Experiment 2, three Gabor patches were presented sequentially. In all experiments, participants were instructed to either engage in articulatory suppression or name the value of each feature as the items were presented (see Figure 5a for labeling examples). After all items were presented, both remembered features of a single (randomly selected) item were recalled by reproducing their feature value within a circular feature space (for orientation or color) or linear space (for frequency). The order of feature recall was random.

Figure 5

Flow of events in Overkott and Souza (in press)(a.) Experiments 1a and b's study phase, (b.) Experiment 1a and 1b's test phase, (c.) Experiments 2's study phase, and (d.) Experiment 2's test phase.



In all experiments, color recall was accomplished by presenting a probe in one of the memory locations along with a color wheel (Experiment 1a) or grey wheel (Experiment 1b and 2). Participants used the mouse to change the probe color. The grey wheel covered the color wheel to minimize color interference: only one color at a time was shown in the probe location as participants moved the mouse around the grey wheel (see Figures 5b and 5d for a visual depiction of the response procedure). Orientation recall in Experiments 1a and 1b was accomplished by presenting a randomly rotated triangle in dark grey at the tested location and allowing participants to adjust the orientation of the probe triangle with the mouse. Orientation recall in Experiment 2 was accomplished by representing the memory item at a random orientation and allowing participants to move a dot along a grey wheel. Moving the dot caused the Gabor patch to rotate.

Results

Model fits for the orientation and color responses in Overkott and Souza (in press)'s Experiments 1a, 1b, and 2 are provided in Table 3. For color, the separate-representation variant was consistently found to be the preferred model, like in Hardman et al. (2017). For orientation, the joint-representation variant was consistently found to be the preferred model, like in the current Experiments 1-4. For color, the separate model WAIC advantage was 221 points in Experiment 1a, 266 points in Experiment 1b, and 339 points in Experiment 2. For orientation, the joint model WAIC advantage was 41 points in Experiment 1a, 294 points in Experiment 1b, and 423 points in Experiment 2.

Table 3
WAIC fits to ***Overkott and Souza (in press)****'s Data by Experiment and Feature Type for each Model Variant*

Experiment & Feature	Model	WAIC	WAIC difference from best model
1a Color	**Separate**	**89,418**	**0**
	Joint	89,639	221
1b Color	**Separate**	**140,980**	**0**
	Joint	141,246	266
2 Color	**Separate**	**121,291**	**0**
	Joint	121,630	339
1a Orientation	Separate	88,904	41
	Joint	**88,863**	**0**
1b Orientation	Separate	141,414	294
	Joint	**141,120**	**0**
2 Orientation	Separate	111,750	423
	Joint	**111,327**	**0**

**Bold rows indicate the winning model for each experiment.*

Discussion

The results of our model-based analysis are consistent across Overkott and Souza (in press)'s experiments. Color was always represented with separate representations for continuous and categorical information while orientation was always represented with a joint continuous and categorical representation. These findings agree with our orientation findings in Experiments 1–4 and the color findings by Hardman et al. (2017).

Maintaining representations of multiple features for the same object did not coerce the representations into a common structure. It was not the object that defined the representation structure, but the stimulus feature-identity. This suggests that it is not a strategic choice to hold the items in mind using different structures, at least by default, but rather a property arising from the natural mental representation of the feature itself. In our final set of analyses, we assessed a third feature, namely shape, across different task set-ups.

Representing Shape

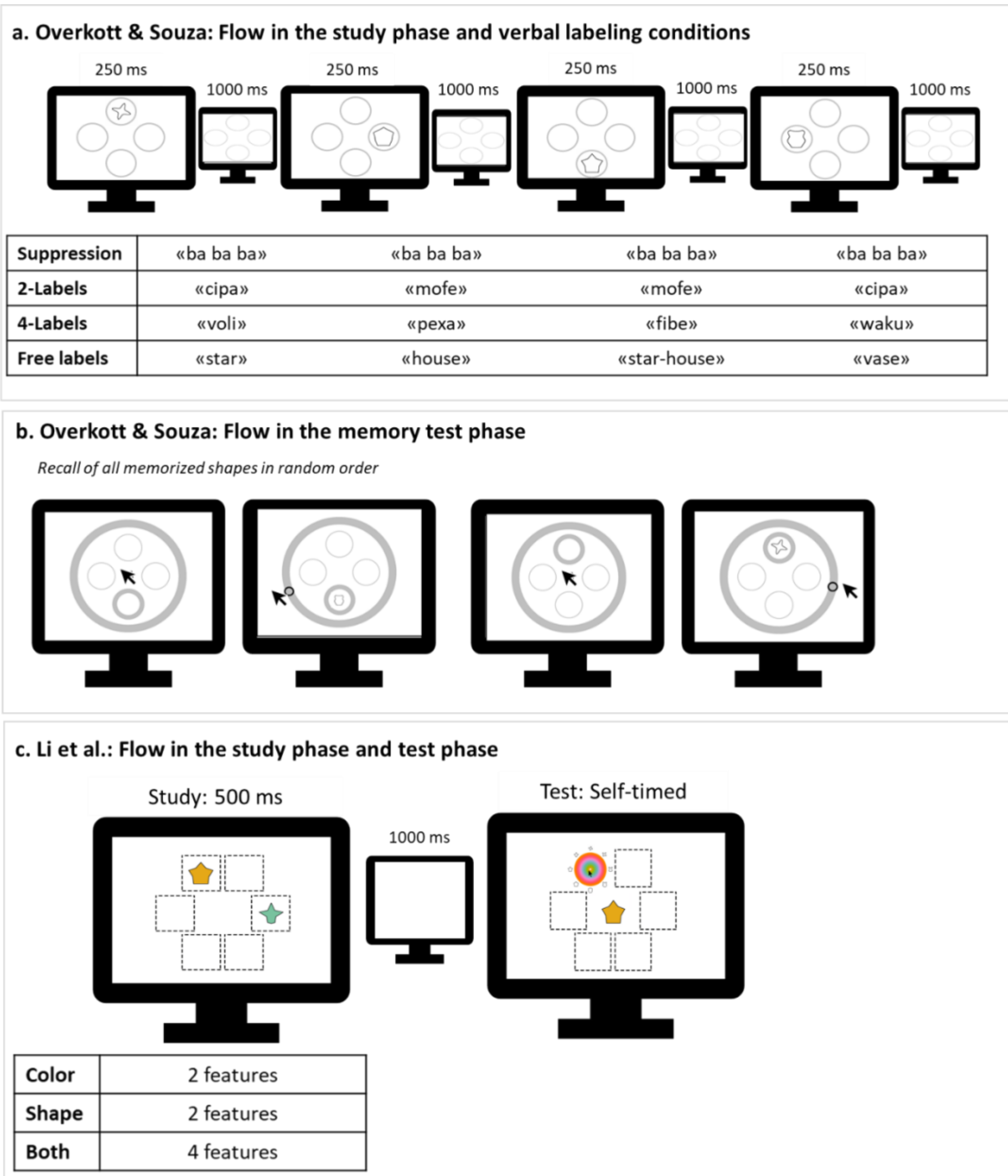
Based on our analyses of response data for orientations and colors, it appears that the structure of the visual representations in working memory depends on feature identity, with categorical and continuous information being represented jointly for orientations but separately for colors. Our final set of modeling analyses explore the mental representation in working memory of yet a third visual feature, shape. Abstract shapes can be gradually morphed from one to another to create a circular shape space with similar properties to continuous orientation or color space. We reanalyzed the data of two experiments which collected continuous response data using the circular shape space developed by Li et al. (2020). Data reported by Souza et al. (2021, Experiment 2) were modeled to identify the structure of the mental representation of shape information (separate vs. joint), and to test whether, like for orientation and color, the revealed structure is invariant across different task set-ups (verbal labeling conditions in this case). Data reported by Li et al. (2022) were modeled to confirm the findings based on the Souza et al. data set in a task set-up without verbal labeling instructions and when multi-featured objects were maintained in mind.

Procedures of Souza et al. (2021) and Li et al. (2022)

In both experiments, continuous shapes were presented and then immediately recalled using a continuous response wheel. Souza et al. (2021, Experiment 2) presented four shapes sequentially, each for 250 ms with a 1000 ms blank screen following each item. Participants completed this task under four conditions: Suppression, 2-Labels, 4-Labels, and Free Labels. Before the experimental trials, Souza et al. trained participants (*N* = 31) in the 2-Label and 4-Label conditions to verbally labels all possible shapes using either 2 or 4 arbitrary labels (non-words randomly assigned to sections of the wheel). During the presentation phase of the experiment (see Figure 6a), participants were instructed to use the learned verbal labels (in the 2-Label and 4-Label conditions), apply their own verbal labels to the shapes (in the free labeling condition), or to engage in articulatory suppression. Condition was blocked and varied within-subjects, and block order was counterbalanced. In the recall phase (see Figure 6b), participants were probed in random order to reproduce the shapes of all memorized items by moving the mouse around a grey wheel. The mouse location determined the shape that appeared in the probe location. The shape varied continuously across 360 different values.

Figure 6

Flow of events in Souza et al. (2021)'s (a.) study phase, (b.) test phase, and (c.) in Li et al. (2022)'s Experiment 1.



Li et al. (2022, Experiment 1) presented 1 or 2 colored shapes during a single 500 ms presentation screen, followed by a 1000 ms blank screen. The task was to remember, depending on condition, either the shape, the color, or both the shape and color for reproduction at test. Li et al. instructed people (N = 29) to avoid verbal labels entirely whenever possible. See Figure 6c for a graphic depiction of the procedure. Li et al. (2022, Experiment 1) cued the item to be recalled by placing a combined shape and color wheel at the location to be recalled.

Results

Model fits for Souza et al. (2021, Experiment 2) are provided for each labeling condition in Table 4. Like for orientation and for color, we found the preferred model to be highly consistent across task conditions. For shape, in all cases, including the suppression condition, the separate-representation model variant was the preferred model. The separate model WAIC advantage was 90 points in the suppression condition, 538 points in two-label condition, 194 points in four-label condition, and 237 points in free labeling condition.

Table 4
WAIC fits to Souza et al. (2021, Experiment 2) Shape Data by Condition for each Model Variant

Condition	Model	WAIC	WAIC difference from best model
Suppression	**Separate**	**66,224**	**0**
	Joint	66,314	90
Two-Labels	**Separate**	**66,682**	**0**
	Joint	67,220	538
Four-Labels	**Separate**	**61,133**	**0**
	Joint	61,327	194
Free Labeling	**Separate**	**62,260**	**0**
	Joint	62,497	237

**Bold rows indicate the winning model for each experiment.*

Model fits for Li et al. (2022, Experiment 1) are provided separately for the color and shape data in the single and dual feature conditions in Table 5. In all conditions and for both features, the separate model variant was the preferred model. The separate model WAIC advantage was 137 points in the color single-feature condition, 60 points in the color dual-feature condition, 194 points in shape single-feature condition, and 221 points in the shape dual-feature condition.

Discussion

Our modeling analysis of Souza et al. (2021, Experiment 2) and Li et al. (2022, Experiment 1) show that shape memory, like color memory, is supported by separate representations of continuous information and categorical information. As in the previous analyses, feature identity drove the representation structure. Articulation condition, labeling, and concurrent feature maintenance for each object did not change the structure used to represent shape in working memory.

Table 5

WAIC fits to Li et al. (2022, Experiment 1) by Feature Type and Recall Condition for each Model Variant

Condition	Model	*W*AIC	*W*AIC Diff*erence from best model*
Color: Single Feature	**Separate**	**67,626**	**0**
	Joint	67,763	137
Color: Dual Feature	**Separate**	**69,764**	**0**
	Joint	69,824	60
Shape: Single Feature	**Separate**	**64,926**	**0**
	Joint	65,120	194
Shape: Dual Feature	**Separate**	**66,613**	**0**
	Joint	66,834	221

**Bold rows indicate the winning model for each experiment.*

General Discussion

In nine experiments we explored whether continuous perceptual information and categorical gist information for a memory item are represented separately or as a single joint representation. We found that the type of structure used to represent a visual feature in working memory is entirely and consistently driven by the identity of the feature represented. Orientation memories are composed of a joint representation of continuous and categorical information, but color and shape memories are composed of separate continuous and categorical representations. In our four new experiments we varied several procedural variables to assess whether we could alter the representation structure through alternative means. These variables included sequential/simultaneous array presentation, presentation duration, set size, and random/serial response order. None of these factors altered the representation structure of orientation. In our reanalysis of previously reported data, we explored whether several additional variables could alter the representation structure, specifically, maintaining single or multi-feature objects, the identity of concurrently held features within an object (i.e., orientation and color, orientation and spatial frequency, color and spatial frequency, color and shape), articulatory labeling, set size, testing procedure (whole report vs. single probe), and the surface features of the perceptual stimulus (i.e., whether orientation was through a dot on a circle, a triangle, or a Gabor patch). None of these factors altered the representation structure of color, orientation, or shape. Together, this makes a strong argument that feature identity drives the structure of the mental representation, that different features utilize fundamentally different representations, and that the structure of these representations is invariant across task context.

Visual working memory representations vary based upon stimulus-specific features. Analysis techniques that aggregate data across stimulus-specific variance, using measures such as mean condition performance and error distribution comparisons, will mischaracterize the structure of the underlying representation. While a coarse level of resolution may be sufficient for predicting typical overall performance on a task, detailed models that characterize trial-to-trial variation in stimulus properties are necessary to understand the structure of individual memories. The importance of this issue should not be underestimated. Many studies have explored whether individual memory representations rely upon discrete or continuous resources using the delayed estimation task (Adam et al., 2017; Bays et al., 2009; Bays & Husain, 2008; Donkin et al., 2015; Keshvari et al., 2013; Pratte, 2019; van den Berg et al., 2014; Zhang & Luck, 2008). The norm in this task is to model individual memory items as purely continuous representations by characterizing only a single memory representation and guessing distribution, sometimes with the inclusion of swap errors.

To assess how using the wrong model impacts memory performance we compared parameter estimates across the separate, joint, and a standard ZL model[2] based upon the approach of Zhang and Luck (2008). We again found a remarkably consistent pattern with only a few examples of deviation from the pattern shown in Figure 7 (see the Supplementary Figures for similar parameter comparisons of all experiments analyzed in the paper). In most experiments we found that all three models arrived at similar estimates of guessing rates. Memory precision differed dramatically across the three model variants while the separate and joint models produced modestly different estimates of P^O. The failure to account for categorical representations or the use of the wrong continuous-categorical structure alters estimates of memory precision and gives a minor misestimate of the contribution of categorical representations. Hardman et al. (2017) also found that use of the wrong model impacted conclusions about whether parameters changed across set sizes. Without the correct model, experimental analyses and the accompanying theoretical conclusion relating to memory precision and categorical elements may not be valid.

One could argue that all three model variants (joint, separate, and ZL) all show the same pattern of parameter value change across conditions so all model variants would result in the

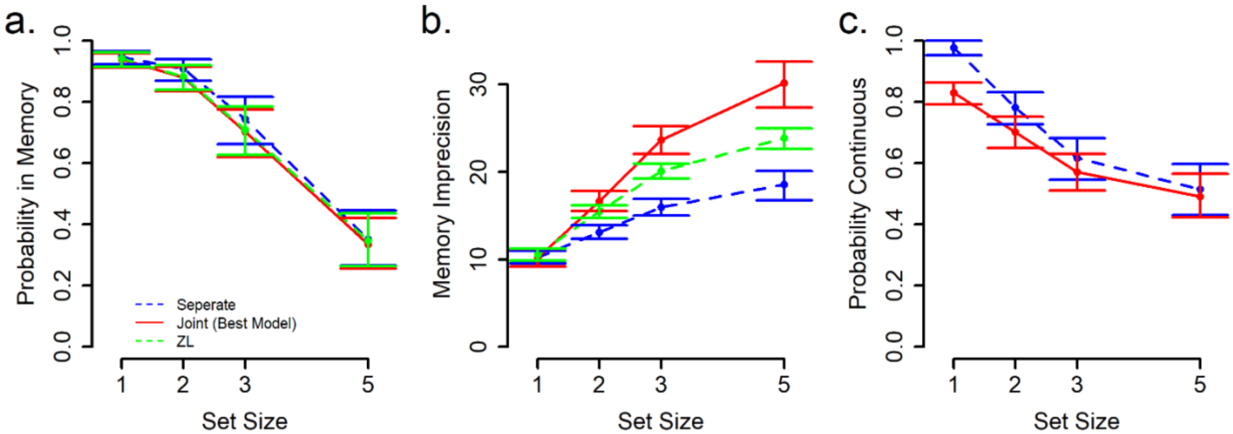
---

[2] We implemented the ZL model using the hierarchical structure of Hardman et al. (2017) with the probability continuous set to 0.

same pattern of results across experimental conditions and thus, ultimately, in the same theoretical conclusions. While it is certainly true that Figure 7b shows clear loss of precision as set size increases across all model variants, the magnitude of change across the different set sizes is clearly different between the different models (e.g., response error increases from about 10 degrees for set size 1 to a bit over 30 degrees for set size 5 for the joint model, but it only increases to a bit under 20 degrees for set size 5 for the ZL model). This dramatic difference between the models in the magnitude of change of memory precision across different set sizes could lead to a failure to detect a change across conditions or an artificially inflated effect across conditions when using an incorrect model. It is also clear that the use of the incorrect model would introduce problems in any context where the parameter value does matter or in any case of comparing precision across estimates derived from different models.

Figure 7

Mean parameter estimates produced by 3 model variants in Experiment 4. Error bars represent 95% credible intervals. Panel (a) shows the probability an item was in memory. Panel (b) shows the memory imprecision. Panel (c) shows the proportion of memory that was continuous. In the separate model this means the probability an item in memory was represented continuously. In the joint model this means the relative weight of the continuous memory on the joint representation.



Mental Representations of Color, Shape, and Orientation in Working Memory

Two studies had previously assessed whether continuous and categorical information are maintained through separate or joint representations. Hardman et al. (2017) and Souza and Skóra (2017) found evidence in favor of separate representations of continuous and categorical features for colors. The current study concurs with this conclusion for color memory, replicating the findings of Hardman et al. and Souza and Skóra by analyzing four additional experiments requiring color memory reproduction using differing experimental procedures, memoranda (colored triangles, colored gratings, colored shapes), and task instructions (suppression, labeling, single vs. dual-feature storage). Other researchers have previously advocated for the use of color categories in memory representation (Bae et al., 2015; Bae et al., 2014; Donkin et al., 2015; Olsson & Poom, 2005; Smith et al., 2020), but here, we not only argue that color categories are used, but that they are maintained in mind alongside and distinct from continuous perceptual information about color, resulting in separate representations.

Previous work by Souza et al. (2021) demonstrated clear use of categories in memory for abstract shapes. In reanalysis of their experiment and of Li et al (2022), we found that shape used a separate representation structure for continuous and categorical memory. Across these experiments many variables were manipulated including set size, verbal labeling, and whether the objects were single or multi-featured. In all cases a consistent separate representation structure was clearly observed for shape, matching the findings from color memory experiments.

In four new experiments and reanalysis of three published experiments we examined the structure of orientation memory. We again found consistent results across many task manipulations including, set size, presentation time, recall order, task set-up, single versus multi-featured items, and verbal labeling conditions. This time the opposite structure was observed. Orientation showed a clear pattern of joint representation of continuous and categorical memory. This representation structure occurred across three different types of orientation stimuli, rings, triangles, and Gabor patches (and they seem to occur also for orientation information conveyed by clockface stimuli, see Ngiam et al., 2022), indicating that the particular memory items used were not the cause of the joint representation structure for orientation. Taken together, the current study (1) confirmed that colors use separate representations of continuous and categorical features, (2) revealed that shapes also use separate representations of continuous and categorical features, and (3) uncovered that orientations use joint representations of continuous and categorical features.

Different Representation Structure for Different Features

When comparing the results across the three features, it becomes clear that two different representation structures exist in visual working memory, and that the structure depends on feature identity. The existence of two differing representations structures in visual working memory for the storage of continuous and categorical information indicates that individuals are not simply attending to and retaining the mental image of a stimulus when performing a visual working memory task. Other laboratories exploring the structure of mental representations have come to similar conclusions while focusing on different structural elements of the representation than those we focus on here. Studies investigating memory for object sets have found strong evidence that memory representations contain ensemble statistics reflecting features of the entire memory set or a predictable subset (Brady & Alvarez, 2011; Lew & Vul, 2015; Son et al., 2020; Utochkin & Brady, 2020). For example, Brady and Alvarez (2011) asked participants to reproduce the size of a single circle within a memory set. They found that the reproduced size of the circle was biased toward the average size of all circles that shared a color with the target. These findings support a hierarchical memory representation that includes representation of context or basic summary statistics of relevant set features. Mathy and colleagues have argued for a similar representation resulting from stimulus compression. When stimuli undergo data compression, commonalities across memory stimuli are represented by a single shared representation, lowering the amount of data the system must maintain (Chekaf et al., 2016; Mathy & Feldman, 2012). Converging results emphasize that working memory makes use of several levels of representation rather than being an isolated verbal or visual trace held within a segregated buffer.

Our finding of differing representations across differing feature types also speaks to the larger debate about domain-specific versus domain-general working memory systems (Baddeley, 1986; Berry et al., 2019; Cocchini et al., 2002; Cowan, 1995; Morey, 2018; Vergauwe et al., 2010). Differing feature types are not represented by a common abstracted representation within working memory, but rather are driven by the feature type itself. This does not necessarily mean that all mechanisms acting on those representations are domain- or feature-specific. It is possible that separate and joint representations may be acted upon by the common domain-general mechanisms. The present findings are consistent with domain-specific models that propose fine-grained segmentation of maintenance structures (Logie & Pearson, 1997; Wang et al., 2017), but also with domain-general systems that operate upon feature-specific representations (Cowan, 1995; Postle, 2006).

Although consistent with both domain-specific and domain-general models of memory, models of memory differentiating visual and spatial storage buffers (Logie & Pearson, 1997) provide a potential explanation of why the representational structure of orientation is different from that of color and shape. One could consider orientation a spatial feature and color and shape visual features. If the spatial buffer uses a joint-representation structure and the visual buffer uses a separate-representation structure this could explain the difference in representation structure across features (see also work on multiple object tracking observing a similar dissociation between tracking visual identity of objects vs. the location of objects; e.g., Pylyshyn, 2004). Future work testing the representational structure of a larger range of both visual and spatial features is needed to fully grasp the broader implications of the current findings. It is important to know whether orientation is an outlier or if other features also use joint representation structures. Understanding the representational structure of a wider range of features could also provide evidence for or against a domain-based fractionation of visuo-spatial working memory maintenance into separate, domain-specific buffers (Logie & Pearson, 1997) .

Differences in representation structure across visual features also have implications for the theoretical interpretation of many experimental findings. For example, many approaches assume that verbal rehearsal or labeling effects act by enhancing only a semantic/categorical representation and not the exact visual representation (Alogna et al., 2014; Donkin et al., 2015; Schooler & Engstler-Schooler, 1990; Sense et al., 2017). This may be consistent with our model of color and shape representations, but orientation does not appear to have a separate (potentially verbal) trace different from the perceptual-visual image. Instead with orientation memory, categorical information acts as an anchor which is then biased by the perceptual-visual information. Verbal rehearsal or labeling in this context may strengthen the anchor but would never be truly independent of the perceptual-visual image itself. Accordingly, recent work assessing the impact of verbal labeling of color and orientation memories observed that labeling affected color and orientation memories differently. Labeling added categorical and continuous information to color memories, but only continuous information to orientation memories (Overkott & Souza, in press). Rather than viewing manipulations as affecting a specific resource or type of representation, the present work makes clear that stimulus-specific characteristics will influence how the information is represented mentally and thus, different cognitive mechanisms may operate differently on them.

Better understanding of representation structure is also needed to verify the validity of previous work using memory models assuming no categorical representation (Bays et al., 2009; Fougnie et al., 2012; Zhang & Luck, 2008) . In our experiments, the model used to estimate memory parameters had only a minor impact on guessing parameters but had a strong impact on memory precision estimates (see Figure 7 and supplemental materials). Findings relying on memory precision estimates may be invalid when estimated with the wrong model of categorical memory or without estimating categorical memory contributions. Going forward it is important to recognize that continuous, categorical, ensemble characteristics, and other complex qualities of mental representation all contribute to the working memories we utilize in day-to-day cognition. The present work provides a crucial step toward understanding the basic representations that underly human thought.

Constraints on Generality

Our findings are based upon data collected in three different countries (The United States, Canada, and Switzerland). In Experiments 1-4 the sample was collected from an urban community college participant pool composed of individuals reflecting the demographics of the general population in the area (New York City). This sample had wide diversity, including a high number of immigrants from around the globe, first-generation Americans, and a majority of non-white individuals. The samples from Switzerland come from the Zurich area, which is quite cosmopolitan, including people from different backgrounds. Despite this diversity, all our participants did reside in North America or Europe at the time of testing, they were attending college, and most, but not all, were young adults. It is possible that individuals with different educational backgrounds may use different representation structures, although we view it as unlikely. If representation structures reflect brain architecture or optimal data formats, then there is no compelling argument to support a demographics-based difference in representation structure. Our research does not rule out the alternative possibility that cultural artifacts of some sort drive how we structure basic visual representations. Research from a large variety of different cultures and socioeconomic backgrounds would be necessary to test this later prediction more thoroughly.

Author Contributions Statement

T.R. – Conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, visualization, writing: original draft. A.S. – Conceptualization, data curation, visualization, writing: review & editing. E.V. – Conceptualization, methodology, writing: review & editing.

References

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

Alogna, V. K., Attaya, M. K., Aucoin, P., Bahník, Š., Birch, S., Birt, A. R., . . . Zwaan, R. A. (2014). Registered Replication Report:Schooler and Engstler-Schooler (1990). *Perspectives on Psychological Science*, *9*(5), 556-578. https://doi.org/10.1177/1745691614545653

Baddeley, A. D. (1986). *Working Memory*. Oxford University Press.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of verbal learning and verbal behavior*, *14*(6), 575-589. https://doi.org/https://doi.org/10.1016/S0022-5371(75)80045-4

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), 744-763. https://doi.org/10.1037/xge0000076

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

Bays, P. M., Catalao, R. F., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, *9*(10).

Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science*, *321*(5890), 851-854.

Berry, E. D. J., Allen, R. J., Waterman, A. H., & Logie, R. H. (2019). The effect of a verbal concurrent task on visual precision in working memory. *Experimental Psychology*, *66*(1), 77-85. https://doi.org/10.1027/1618-3169/a000428

Brady, T. F., & Alvarez, G. A. (2011). Hierarchical Encoding in Visual Working Memory:Ensemble Statistics Bias Memory for Individual Items. *Psychological science*, *22*(3), 384-392. https://doi.org/10.1177/0956797610397956

Brown, G. D., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological review*, *114*(3), 539-576. https://doi.org/10.1037/0033-295X.114.3.539

Chekaf, M., Cowan, N., & Mathy, F. (2016). Chunk formation in immediate memory and how it relates to data compression. *Cognition*, *155*, 96-107. https://doi.org/https://doi.org/10.1016/j.cognition.2016.05.024

Cocchini, G., Logie, R. H., Della Sala, S., MacPherson, S. E., & Baddeley, A. D. (2002). Concurrent performance of two memory tasks: Evidence for domain-specific working memory systems. *Memory & Cognition*, *30*(7), 1086-1095.

Cowan, N. (1995). In *Attention and memory: An integrated Framework.* (Vol. 26). Oxford University Press.

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

Crowder, R. G. (1976). *Principles of learning and memory*. Erlbaum.

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. https://doi.org/10.3758/s13423-014-0675-5

Donkin, C., Nosofsky, R. M., Gold, J. M., & Shiffrin, R. M. (2013). Discrete-slots models of visual working-memory response times. *Psychological review*, *120*, 873-902. https://doi.org/10.1037/a0034247

Forsberg, A., Johnson, W., & Logie, R. H. (2020). Cognitive aging and verbal labeling in continuous visual memory. *Memory & Cognition*, *48*(7), 1196-1213. https://doi.org/10.3758/s13421-020-01043-3

Fougnie, D., Suchow, J. W., & Alvarez, G. A. (2012). Variability in the quality of visual working memory. *Nature communications*, *3*, 1229.

Gelman, A., Hwang, J., & Vehtari, A. (2014). Understanding predictive information criteria for Bayesian models. *Statistics and Computing, 24*(6), 997-1016. https://doi.org/10.1007/s11222-013-9416-2

Hardman, K. O. (2017). *CatContModel: Categorical and Continuous Working Memory Models for Delayed Estimation Tasks (Version 0.8.0) [Computer software]. Retrieved from https://github.com/hardmanko/CatContModel/releases/tag/v0.8.0*. In

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. https://doi.org/10.1037/xhp0000290

Huang, L. (2020). Distinguishing target biases and strategic guesses in visual working memory. *Attention, Perception, & Psychophysics*, *82*(3), 1258-1270. https://doi.org/10.3758/s13414-019-01913-2

Keshvari, S., van den Berg, R., & Ma, W. J. (2013). No evidence for an item limit in change detection. *PLoS computational biology*, *9*(2), e1002927.

Kintsch, W. (1967). Memory and decision aspects of recognition learning. *Psychological review*, *74*(6), 496-504. https://doi.org/10.1037/h0025127

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

Li, A. Y., Fukuda, K., & Barense, M. D. (2022). Independent features form integrated objects: Using a novel shape-color "conjunction task" to reconstruct memory resolution for multiple object features simultaneously. *Cognition*, *223*, 105024. https://doi.org/https://doi.org/10.1016/j.cognition.2022.105024

Li, A. Y., Liang, J. C., Lee, A. C. H., & Barense, M. D. (2020). The validated circular shape space: Quantifying the visual similarity of shape. *Journal of Experimental Psychology: General*, *149*(5), 949-966. https://doi.org/10.1037/xge0000693

Logie, R. H., & Pearson, D. G. (1997). The Inner Eye and the Inner Scribe of Visuo-spatial Working Memory: Evidence from Developmental Fractionation. *European Journal of Cognitive Psychology*, *9*(3), 241-257. https://doi.org/10.1080/713752559

Long, F., Ye, C., Li, Z., Tian, Y., & Liu, Q. (2020). Negative emotional state modulates visual working memory in the late consolidation phase. *Cognition and Emotion*, *34*(8), 1646-1663. https://doi.org/10.1080/02699931.2020.1795626

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature neuroscience*, *17*(3), 347-356.

Mathy, F., & Feldman, J. (2012). What's magic about magic numbers? Chunking and data compression in short-term memory. *Cognition*, *122*(3), 346-362. https://doi.org/https://doi.org/10.1016/j.cognition.2011.11.003

Morey, C. C. (2018). The case against specialized visual-spatial short-term memory. *Psychological bulletin*, *144*(8), 849-883. https://doi.org/10.1037/bul0000155

Ngiam, W. X. Q., Foster, J. J., Adam, K. C. S., & Awh, E. (2022). Distinguishing guesses from fuzzy memories: Further evidence for item limits in visual working memory. *Attention, Perception, & Psychophysics*. https://doi.org/10.3758/s13414-022-02631-y

Olsson, H., & Poom, L. (2005). Visual memory needs categories. *Proceedings of the National Academy of Sciences*, *102*(24), 8776-8780. https://doi.org/doi:10.1073/pnas.0500810102

Ovalle-Fresa, R., Uslu, A. S., & Rothen, N. (2021). Levels of Processing Affect Perceptual Features in Visual Associative Memory. *Psychological science*, *32*(2), 267-279. https://doi.org/10.1177/0956797620965519

Overkott, C., & Souza, A. (in press). The fate of labeled and non-labeled visual features in working memory. *Journal of Experimental Psychology: Human Perception and Performance*. https://doi.org/https://doi.org/10.1037/xhp0001089

Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience*, *139*(1), 23-38. https://doi.org/https://doi.org/10.1016/j.neuroscience.2005.06.005

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

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

Pylyshyn, Z. (2004). Some puzzling findings in multiple object tracking: I. Tracking without keeping track of object identities. *Visual cognition*, *11*(7), 801-822. https://doi.org/10.1080/13506280344000518

Rhodes, S., Abbene, E. E., Meierhofer, A. M., & Naveh-Benjamin, M. (2020). Age differences in the precision of memory at short and long delays. *Psychology and Aging*, *35*(8), 1073-1089. https://doi.org/10.1037/pag0000565

Rouder, J. N., Morey, R. D., Cowan, N., Zwilling, C. E., Morey, C. C., & Pratte, M. S. (2008). An assessment of fixed-capacity models of visual working memory. *Proceedings of the National Academy of Sciences*, *105*(16), 5975-5979.

Schneegans, S., Harrison, W. J., & Bays, P. M. (2021). Location-independent feature binding in visual working memory for sequentially presented objects. *Attention, Perception, & Psychophysics*, *83*(6), 2377-2393. https://doi.org/10.3758/s13414-021-02245-w

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

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

Smith, P. L., Saber, S., Corbett, E. A., & Lilburn, S. D. (2020). Modeling continuous outcome color decisions with the circular diffusion model: Metric and categorical properties. *Psychological review*, *127*(4), 562-590. https://doi.org/10.1037/rev0000185

Son, G., Oh, B.-I., Kang, M.-S., & Chong, S. C. (2020). Similarity-based clusters are representational units of visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(1), 46-59. https://doi.org/10.1037/xlm0000722

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/https://doi.org/10.1016/j.jml.2021.104242

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

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

van den Berg, R., Awh, E., & Ma, W. J. (2014). Factorial comparison of working memory models. *Psychological review*, *121*(1), 124-149. https://doi.org/10.1037/a0035234

Vergauwe, E., Barrouillet, P., & Camos, V. (2010). Do mental processes share a domain-general resource? *Psychological science*, *21*(3), 384-390. https://doi.org/10.1177/0956797610361340

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

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision*, *4*(12).

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233-235.