Ensemble clustering in visual working memory biases location memories and reduces the Weber noise of relative positions

Timothy F. Lew
Department of Psychology, University of California, San Diego, CA, USA

Edward Vul
Department of Psychology, University of California, San Diego, CA, USA

People seem to compute the ensemble statistics of objects and use this information to support the recall of individual objects in visual working memory. However, there are many different ways that hierarchical structure might be encoded. We examined the format of structured memories by asking subjects to recall the locations of objects arranged in different spatial clustering structures. Consistent with previous investigations of structured visual memory, subjects recalled objects biased toward the center of their clusters. Subjects also recalled locations more accurately when they were arranged in fewer clusters containing more objects, suggesting that subjects used the clustering structure of objects to aid recall. Furthermore, subjects had more difficulty recalling larger relative distances, consistent with subjects encoding the positions of objects relative to clusters and recalling them with magnitude-proportional (Weber) noise. Our results suggest that clustering improved the fidelity of recall by biasing the recall of locations toward cluster centers to compensate for uncertainty and by reducing the magnitude of encoded relative distances.

Introduction

Our visual working memory is limited in its ability to remember objects. In addition to remembering the individual elements of scenes, people may also extract the higher-order structure of an image, such as the elements' average size (e.g., Ariely, 2001) or average location (e.g., Alvarez & Oliva, 2009). People can then use that statistical structure to help remember objects (Brady & Alvarez, 2011; Brady, Konkle, & Alvarez, 2009; Sims, Jacobs, & Knill, 2012). Knowing that your papers are scattered in a pile around your desk, for example, constrains their possible locations (e.g., it is unlikely they are in the bathroom) and can help you remember where individual papers are. Given that people appear to encode and utilize not only individual objects but also the higher-order structure of objects, what is the format of structured memories?

In contrast to the traditional assumption that objects in visual working memory are encoded independently (Anderson, Vogel, & Awh, 2011; Bays & Husain, 2008; Zhang & Luck, 2008; for review, see Ma, Husain, & Bays, 2014), recent studies have demonstrated that memory exploits the statistical structure of scenes. Specifically, people infer the ensemble statistics of objects (such as the average location of objects; Alvarez & Oliva, 2009; Ariely, 2001) and combine these ensemble statistics with uncertain estimates of individual object properties (Brady & Alvarez, 2011; Brady & Tenenbaum, 2013; Orhan & Jacobs, 2013). This encoding strategy can be described as reliance on a hierarchical generative model: People infer that object features are drawn from a distribution of features and make uncertain inferences accordingly. In our desk example, this would imply that if you did not know exactly where a paper was, you might recall it as closer to the center of the pile to compensate for your uncertainty; although this strategy yields some *bias* in in your estimate of the location, it decreases *variance* and thus improves overall memory fidelity.

The structure of multiple objects may also constrain the individual constituent objects more rigidly into multiobject "chunks" (Brady & Tenenbaum, 2013; Cowan, 2001; Miller, 1956). Chunking accounts tacitly assumes that an inferred chunk completely constrains its subparts (e.g., encoding "FBI" fully determines its constituent letters). Thus, chunking is classically considered to be a fixed memory structure (what we might call "hard chunking"), such that people remember only the chunk and nothing about its constituent elements. However, if this encoding strategy is softened to allow some information to be preserved about the

Citation: 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, 1–14, doi:10.1167/15.4.10.

doi: 10.1167/15.4.10 Received February 7, 2015; published September 11, 2015 ISSN 1534-7362 © 2015 ARVO

Downloaded from jov.arvojournals.org on 07/01/2019

constituent elements of a chunk ("soft chunking"), such an account is consistent with encoding a hierarchical generative model that probabilistically constrains individual elements.

Additionally, studies of spatial memory suggest that people encode the *relative* positions of objects: Rather than remember the absolute position of a paper, you may remember its position relative to your desk (e.g., the paper is one foot northwest of your desk; Hollingworth, 2007; Huttenlocher, Hedges, & Duncan, 1991). This relative encoding may be adapted to accommodate hierarchical structures via an assumption that people encode the relative discrepancy between features of individual objects and the average features of the ensemble. This relative encoding view is consistent with vector-summation models of multiobject motion parsing (Gershman, Tenenbaum & Jäkel, in press; Johansson, 1973) and spatial positions (Mutluturk & Boduroglu, 2014). Intuitively, instead of remembering the locations of your papers relative to your desk, you may remember the locations of individual papers relative to the centroid of all the papers.

Thus, the space of possible structures that people might use to encode objects can be considered along several dimensions: (a) Do people encode individual items with no information about their structure (independent encoding)? Or do they only encode the structure, losing all information about constituent elements (hard chunking)? Or something in between, such that the overarching structure informs individual object features (hierarchical generative model or soft chunking encoding)? (b) Insofar as people encode both higher-order structure and individual element features, are these both encoded in absolute terms and inform one another probabilistically (absolute encoding), or are objects in the hierarchy encoded relative to their "parent" (objects relative to their ensembles and ensembles relative to cluster groups), such that object properties are ascertained by accumulating relative offsets in the hierarchy (relative encoding)?

Here we evaluate these dimensions of visual memory structure by asking people to remember and report the locations of objects arranged in different spatial clustering structures. Subjects recalled objects more accurately when they were arranged in fewer clusters that each contained more objects separated by smaller relative distances. To directly evaluate the format of subjects' structured memories, we compared human behavior to that of three cognitive models: a hard chunking model, a hierarchical generative model, and a relative position model. The relative position model best accounted for human performance, followed closely by the hierarchical generative model, with the hard chunking model missing key aspects of human behavior. Our results demonstrate two compatible ways in which hierarchical encoding improves the fidelity of visual working memory. First, objects are biased toward their ensemble statistics to compensate for uncertainty about individual object properties. Second, objects are encoded relative to their parents in the hierarchy, and relative positions are corrupted by Weber noise,[1] such that larger relative distances yield greater errors.

Experiment

To distinguish different hierarchical encoding strategies that people may use, we asked subjects to report the positions of objects arranged in different clustering structures. Different encoding strategies yielded distinct patterns of errors across scenes that varied in the number of objects and the number of clusters in which they were arranged. Thus, we then examined if subjects' responses across different types of environments were consistent with different forms of structured encoding.

Methods

Subjects

Thirty-five students from the University of California, San Diego, participated for course credit.

Stimuli

We generated 70 environments, each containing objects arranged into different clustering structures. We selected 440 images from Brady, Konkle, Alvarez, and Oliva (2008) for the objects. Although we did not control how much objects varied perceptually and semantically, we made sure each object type was unique (e.g., there was only one bicycle, clock, etc.). The dimensions of the environments were 700×1000 pixels. Each subject saw the same environments but in a random order.

Each environment had one of seven clustering structures: four clusters each containing one object (4C1), two clusters containing two objects (2C2), 1C4, 8C1, 4C2, 2C4, 1C8 (Figure 1). We generated the locations of the clusters and objects by selecting cluster centers from a uniform distribution across the entire environment and then sampling object locations from each center using a two-dimensional isotropic normal distribution ($SD = 45$) with the restriction that objects could not overlap. There were 10 unique environments for each clustering structure for a total of 70 environments.

Downloaded from jov.arvojournals.org on 07/01/2019

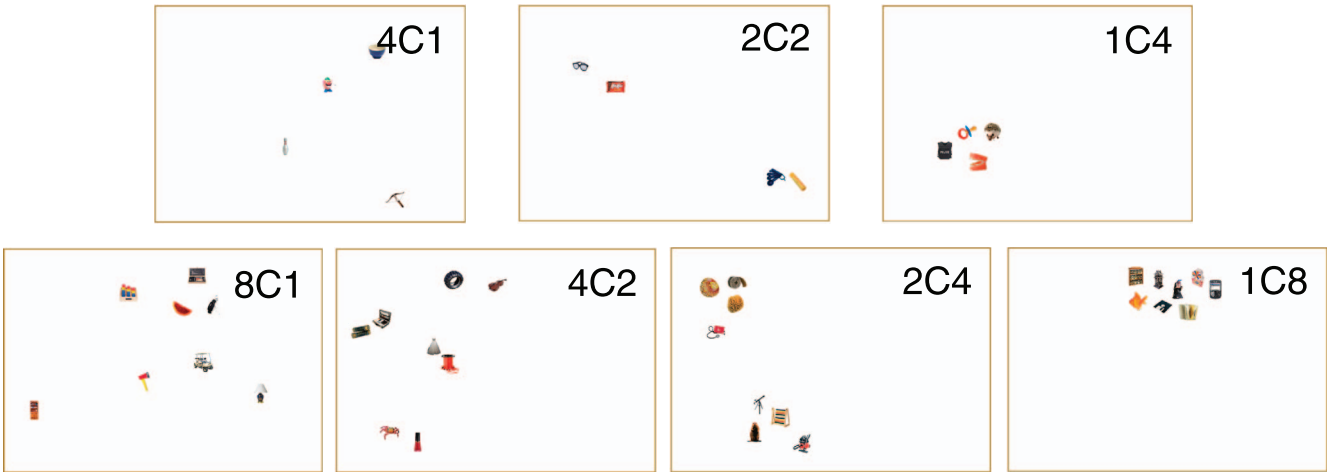

Figure 1. Examples of environments from each of the clustering structures. From left to right, each row is arranged in order of increasing clustering (clusters contain more objects). For this figure, a label indicating each environment's clustering structure is superimposed. Labels are read 4C2 = four clusters each containing two objects. Images of objects from Brady et al. (2008).

Procedure

Subjects studied the four-object environments (4C1, 2C2, and 1C4) for 4 s and the eight-object environments (8C1, 4C2, 2C4, and 1C8) for 8 s. After a 1-s pause, subjects saw an empty environment with the objects located at the bottom of the screen and had unlimited time to place the objects in their correct locations by clicking and dragging with the mouse. Our analyses focus on the reported spatial locations of all the objects in a display.

Results

Did subjects encode objects according to their clustering structure?

If subjects did encode and utilize the clustering structure of objects instead of independently encoding objects, the errors for objects in the same cluster should be more similar (in the same direction) than expected by chance. We defined the similarity of the errors (q) in reporting the locations of two objects as

$$q_{ij} = \frac{x_i x_j^T}{||x_i|| ||x_j||}$$

where $\boldsymbol{x_i}$ and $\boldsymbol{x_j}$ are vectors containing the spatial translational error of the two objects' reported locations. The numerator is the projection of the translational error vectors with positive values indicating vectors in the same direction and negative values indicating vectors in the opposite direction. The denominator normalizes the numerator, such that q falls between −1 and 1. Thus, if the recalled locations of two objects were both shifted in exactly the same direction, q would be 1; if they were shifted in orthogonal directions, q would be 0; and if they shifted in opposite directions, q would be −1.

We calculated the translational error similarity (q) of objects in the same cluster for each environment (Figure 2). We excluded environments without clustering (4C1 and 8C1) from this analysis. For all clustering structures, subjects recalled objects in the same cluster with more similar errors than expected by independent encoding, smallest t value, $t(34) = 16.05$, $p < 0.001$). Subjects did not appear to encode the objects independently and instead used the clustering structure of objects.

How did clustering structure affect recall fidelity?

If subjects encoded objects independently, then clustering structures should not have affected how accurately subjects recalled locations. We assessed the effect of clustering structure upon the fidelity of recall by calculating the root mean square error ($RMSE^2$) of subjects' responses (Figure 3). We used a mixed-effects model that included the number of objects, the number of clusters, and their interaction as fixed effects and subjects as random effects to test whether object load and clustering structure affected recall. RMSE was lower in the four-object conditions compared to the eight-object conditions, $t(241) = 12.47$, $p < 0.001$ for the linear effect of number of objects, and decreased as the number of objects in each cluster increased for both

Downloaded from jov.arvojournals.org on 07/01/2019

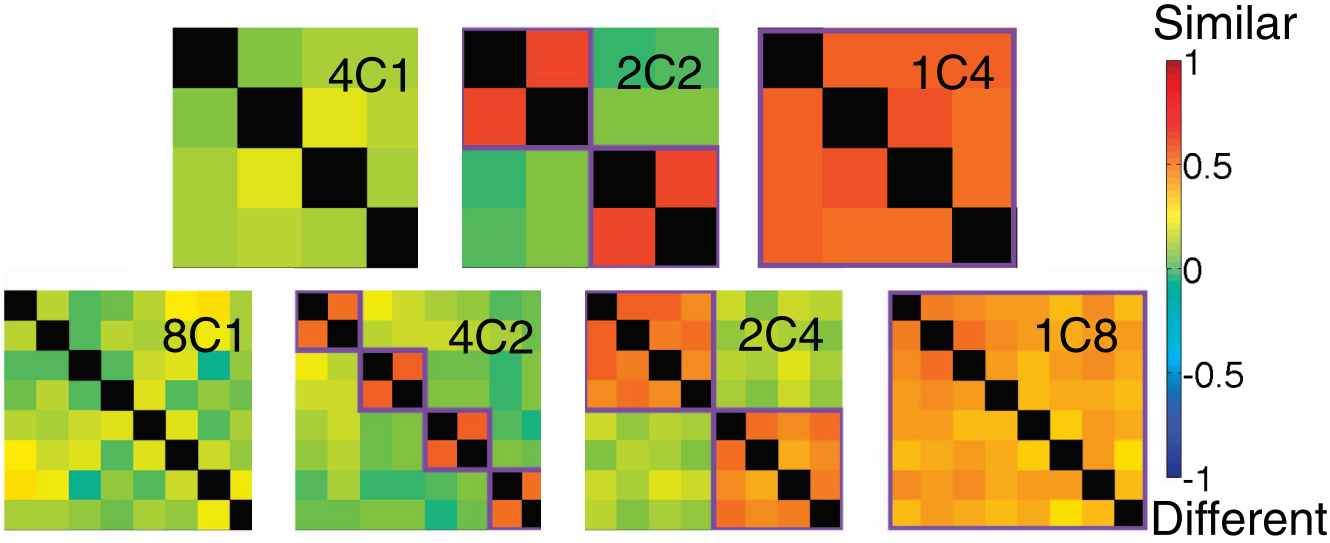

Figure 2. Error similarity heat maps with labels indicating the clustering structure superimposed. Warmer colors indicate more similar errors. Each square represents the error similarity between two different objects. Objects in the same cluster are outlined in purple. Objects in the same cluster were recalled with more similar errors.

the four-object and eight-object conditions, $t(241) = 16.95$, $p < 0.001$ for the linear effect of number of clusters. Post hoc Tukey's honest significant difference (HSD) pairwise comparisons confirmed that performance improved with every increment of cluster size in both the four-object conditions (smallest difference: 13.30, 95% confidence interval = 3.69–22.92, $p = 0.0042$) and the eight-object conditions (smallest difference: 14.71, 95% confidence interval = 3.55–25.88, $p = 0.0046$). The decrease in RMSE with increasing cluster size seems constant across the four- and eight-object conditions, $t(241) = .31$, $p = 0.76$ for the interaction of the number of objects and the number of clusters, i.e., the difference in slope of RMSE as a function of number of clusters. The effect of clustering structure on performance suggests that subjects did not encode the objects independently and that subjects used clustering to help remember objects more accurately.

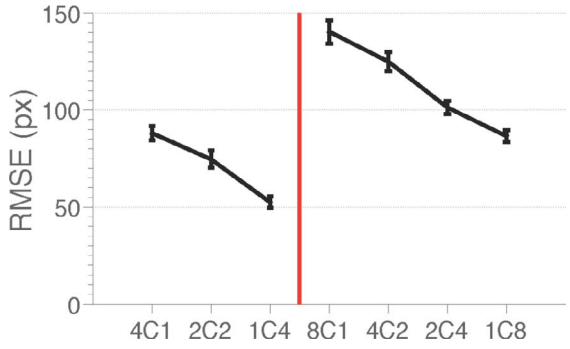

Figure 3. Raw performance measured in root mean square error (RMSE) for each of the clustering structures, arranged in order of increasing clustering. The red line separates the four-object conditions from the eight-object conditions. Error bars indicate *SEM*. Performance improved as objects were arranged in fewer clusters containing more objects.

Error model

Thus far, we have demonstrated that subjects did not encode objects independently. Given that subjects appeared to use the clustering structure of objects, how did that structure constrain the locations of objects? Did subjects encode objects using hard chunking, a hierarchical generative model, and/or a relative position tree? These encoding models predict different levels of reliance on (and bias toward) objects' hierarchical structure and different patterns of noise. To determine what type(s) of structured encoding subjects' errors were consistent with, we constructed an error model that estimates the extent of errors due to misassociations, bias, and noise.

First, subjects may have had difficulty remembering which objects were in which locations. We estimated the probability of correctly matching an object to its location, p_T, and the probability of making a misassociation between an object and another object's location, $p_M = 1 - p_T$. The probability of misassociating to a particular location then was $\frac{p_M}{n-1}$, where n is the number of locations. To determine exactly to which location each object was misassociated, we assumed a bijective mapping of objects to locations (f), such that only one object could be paired with each location. $f^{-1}(i)$ denotes the inverse mapping from locations to objects.

Downloaded from jov.arvojournals.org on 07/01/2019

Second, subjects may have been uncertain about objects' locations but used their memories of cluster locations to inform their responses. This would have resulted in objects being drawn toward their clusters. We accounted for two types of such "regularization" bias: the degree to which clusters are drawn toward the global centroid of all objects (cluster-to-global bias, β_c) and the degree to which objects are drawn toward their cluster centers (object-to-cluster bias, β_o). Here, a bias of zero indicates the object/cluster is unbiased, and a bias of one indicates the element is drawn completely toward its parent.

To parameterize how the locations of objects would be shifted by these sources of bias, we decomposed the true locations of objects, t, into their relative positions and then weighted the relative positions by the bias parameters. The decomposition of the true locations yielded a relative position tree in which the locations of objects were represented relative to their clusters (x), the locations of clusters were relative to the global centroid (c), and the global centroid (g) was the mean of the true locations (t). Conditional on the mapping $f^{-1}(i)$ of the true locations t to response locations s, the position of an object i's cluster relative to the global center was defined by

$$c_i = C_{M\left(f(i)\right)} - g$$

where $M(\,)$ maps objects to the clusters of which they are members, and C is the absolute position of the cluster center, calculated by averaging the locations of all objects in that cluster. Similarly, the positions of objects relative to their clusters were defined by

$$x_i = t_{f^{-1}(i)} - c_i - g.$$

We then weighted the relative positions of clusters and objects by the cluster-to-global bias (β_c) and the object-to-cluster bias (β_o), respectively. Thus, the biased absolute positions of an object, b_i, were

$$b_i = g + (1 - \beta_c)^* c_i + (1 - \beta_o)^* x_i.$$

Finally, subjects may have remembered locations with some imprecision. To account for this, the model includes three levels of spatial noise that might induce correlations in errors across objects: that which is shared globally across all object locations (σ_g), for locations within the same cluster (σ_c), and individual object locations (σ_o). This decomposition of object positions induces an expected correlation structure on the errors in reporting individual objects, which can be parameterized with a covariance matrix, Σ, of the form

$$\Sigma_{i,j} = \begin{cases} \sigma_g^2 & i \neq j \quad M(f(i)) \neq M(f(j)) \\ \sigma_c^2 + \sigma_g^2 & i \neq j \quad M(f(i)) = M(f(j)) \\ \sigma_o^2 + \sigma_c^2 + \sigma_g^2 & i = j \end{cases}$$

where the three conditions reflect (in order) error covariance shared by all objects, error covariance for objects in the same cluster, and error variance for individual objects.

Let Θ be the set of parameters $\{p_M, \beta_c, \beta_o, \sigma_g, \sigma_c, \sigma_o\}$. Altogether, for each environment, the likelihood of a set of responses given the targets and parameters was

$$LIK(s|t,f,\Theta) = (\mathrm{p}_T^{n_T})\left(\frac{p_M}{n-1}\right)^{n_M} N(s|b,\Sigma)$$

where s denotes the response locations, n is the number of objects, n_T is the number of objects correctly mapped to their locations by f, and n_M is the number of objects incorrectly mapped to their locations by f. We estimated these parameters (f, p_M, β_c, β_o, σ_g, σ_c, σ_o) for each environment across subjects using a Markov chain Monte Carlo algorithm (see Appendix C for more details concerning our Markov chain Monte Carlo algorithm and Appendix D for all parameter fits).

Did subjects encode objects in addition to their hierarchical structure?

Encoding objects as components of hard chunks or a hierarchical generative model should result in distinct patterns of object-to-cluster bias. If subjects encoded objects as hard chunks, they should have retained minimal information about the objects' locations and recalled the objects with a large bias toward their respective cluster centers. If subjects encoded objects in a hierarchical generative model, then they should have recalled objects with more bias toward their cluster centers when clusters contained more objects. Intuitively, subjects can more precisely estimate the centers of clusters that contain more objects and consequently should rely on those clusters more when they are uncertain about the locations of the individual objects.

The bias of objects toward clusters was consistently low (β_o: $M = .19$, $SEM = .02$, max $= .62$), suggesting that subjects remembered the locations of individual objects within their clustering structure rather than storing chunks and discarding their internal components. Additionally, contrary to the pattern of bias we expected to find if subjects encoded objects in a hierarchical generative model, as objects were arranged in fewer clusters containing more objects, the objects tended to be recalled with less bias toward their clusters (Figure 4), $t(47) = 7.14$, $p < 0.001$ for the linear effect of number of clusters on β_o in a model including fixed effects of number of objects and number of clusters). Post hoc Tukey's HSD pairwise comparison tests confirmed that objects' bias toward their clusters varied with the number of clusters for the four-object conditions (smallest difference: .099, 95% confidence interval = .060–.14, $p < 0.001$). With the exception of

Downloaded from jov.arvojournals.org on 07/01/2019

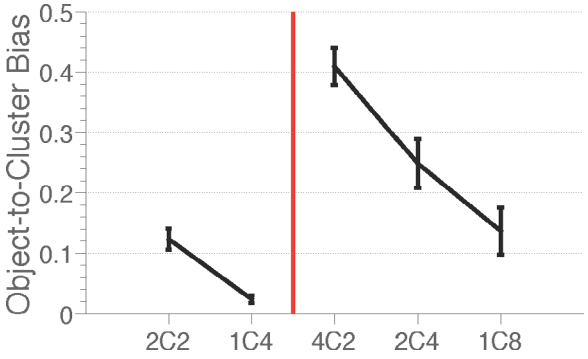

Figure 4. The extent to which objects were drawn toward their cluster centers (β_o) for each clustering structure. Larger object-to-cluster bias indicates objects are drawn more toward their clusters. Zero indicates the object is not biased toward the cluster, and one indicates an object is drawn completely to the cluster. The red line separates the four-object and eight-object structures. Error bars indicate *SEM*. Object-to-cluster bias was generally low, suggesting subjects did not solely encode chunks (thus forgetting relative object position within a cluster), and contrary to the predictions of a hierarchical generative model, the bias of objects toward their clusters decreased as clusters contained more objects. Nevertheless, in all conditions, objects were drawn toward their clusters to some degree.

the 2C4 and 1C8 conditions (difference: .11, 95% confidence interval = −.017–.24, $p = 0.098$), the bias of objects toward their clusters also varied for the eight-object conditions (smallest difference: .16, 95% confidence interval = .031–.29, $p = 0.01$). However, even though the bias of objects toward their clusters was generally low, objects were consistently recalled with *some* bias. Together, this pattern of bias suggests that subjects encoded objects in a hierarchical generative model but did not rely primarily on this form of representation.

Did subjects encode objects in a relative position tree?

Subjects may have encoded objects in a relative position tree, wherein object positions are coded as relative offsets from the cluster centers, and cluster centers are coded as relative offsets from the global center. At first glance, this is no different from encoding the objects according to their absolute position. However, if relative positions are recalled with Weber noise (Sims et al., 2012; Tudusciuc & Nieder, 2010), then larger relative distances will be more difficult to recall. Because the relative distances between objects decrease with more clustering, this could explain why subjects remembered more densely clustered objects more accurately.

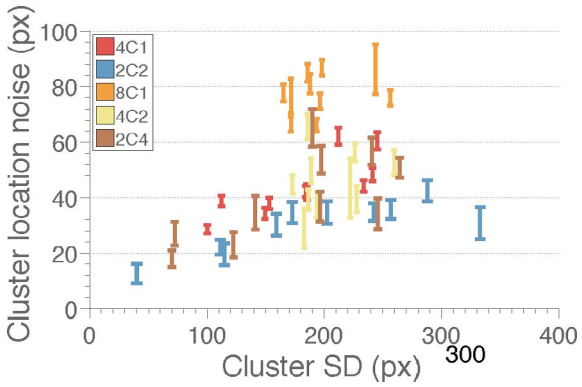

Figure 5. The noise of recalled cluster locations (σ_c) given the dispersion of clusters. Each point represents an environment estimated across subjects. Points are color-coded by clustering structure. Error bars indicate *SD* of the posterior distribution. As clusters were further apart, cluster locations were recalled less accurately.

Under such a relative encoding scheme, environments that happened to contain more dispersed clusters[3] require larger relative distances to represent positions. Consequently, as the dispersion of clusters in the environment increases, subjects should recall clusters less precisely (that is, σ_c should increase). The dispersion of clusters in an environment was significantly correlated with the precision with which subjects recalled cluster centers ($r = 0.38$ $p < 0.01$) (Figure 5), consistent with subjects encoding objects according to their relative positions and having difficulty recalling larger relative distances.

Comparing chunking, hierarchical generative, and relative position models

To directly test explicit formulations of different encoding theories, we designed three cognitive models that would encode a display and generate responses according to its biases: a hard chunking model that only remembers clusters, a hierarchical generative model that encodes absolute positions (similar to Orhan & Jacobs, 2013), and a model that encodes objects in a relative position tree and recalls relative positions with Weber noise. Each model uses a nonparametric Dirichlet process to determine the clustering of the objects (Ferguson, 1983). We evaluated how well these models could predict subject performance (measured in RMSE) in each environment.

Nonparametric Dirichlet process

We used a nonparametric Dirichlet process to determine the clustering structure of the objects

Downloaded from jov.arvojournals.org on 07/01/2019

	4C1	2C2	1C4	8C1	4C2	2C4	1C8	All
Ch	.0095	.37	−.24	.38	.44	.70*	−.21	.55**
HG	.70*	.63	.16	.43	.54	.58	−.43	.70**
RP	.73*	.85**	.80*	.63	.67*	.61	.53*	.89**

Table 1. *r* values of the correlation between subject RMSE and model RMSEs for the environments within each clustering structure (4C1–1C8) and for all environments across clustering structures (All). *Notes*: Ch: chunking model, HG: hierarchical generative model, RP: relative position model. *$p < 0.05$, **$p < 0.01$. The relative position model predicted the difficulty of environments within each clustering structure most accurately.

(Ferguson, 1983). Although we used specific clustering structures to generate the locations of objects, the actual distribution of objects in a particular display may have been consistent with a clustering structure we did not design. Such impromptu clustering is especially likely in environments without built-in clustering (4C1 and 8C1). Nonparametric Dirichlet clustering assumes that each object's location is drawn from an isotropic Gaussian cluster with some position and standard deviation. Crucially, this clustering model estimates the number of clusters, the assignment of objects to clusters, and the breadth and locations of clusters that best explain the locations of the objects.

We used a Gibbs sampler (Geman & Geman, 1984) to estimate the clustering structure of objects and a concentration parameter. The concentration parameter captures a prior on the number of clusters, and its average median value was .11 ($SD = .033$). The chunking, hierarchical generative and relative position models all use the maximum likelihood clustering structures of the environments estimated by the nonparametric Dirichlet process.

Chunking model

The hard chunking model uses solely information about the clusters and which objects belong to which clusters to recall the locations of objects. Importantly, the chunking model knows nothing about the locations of the individual objects. Instead, the model recalls the location of an object by randomly sampling from the object's cluster based on the center and standard deviation of the cluster estimated by the Dirichlet process. The model has no free parameters.

Hierarchical generative model

The hierarchical generative model uses knowledge of clusters' locations to compensate for uncertainty in the individual objects' locations. This model is similar to the Dirichlet process mixture model used by Orhan and Jacobs (2013).

The hierarchical generative model noisily encodes the absolute locations of all the objects as well as the properties of their clusters. Because the model pools memories of individual objects to determine the mean and dispersion of their respective clusters, each additional object in a cluster allows the model to estimate the position of that cluster more precisely. This model uses the same process to estimate the precision of the global center from the locations of the clusters. During recall, the model first recalls the locations of the clusters by averaging the positions of the clusters and global center, weighted by their precisions. The model then recalls the locations of individual objects by averaging the positions of the objects and their clusters, weighted by the precision of the encoded object locations and the posterior predictive spread of objects within a cluster, respectively.

This model has one free parameter: the noise with which objects are encoded. We set the noise parameter to the average object location noise (σ_o) estimated by our error model separately for the four-object and eight-object conditions.

Relative position model

The relative position model remembers the relative positions of objects and clusters with Weber noise and uses clustering to reduce the magnitude of relative positions. Using the clustering structure inferred by the Dirichlet process, the relative position model remembers the positions of objects relative to their clusters and the clusters relative to the global center. The model encodes relative positions via their distance and angle and recalls them with circular Gaussian noise on angle and proportional (Weber) noise on distance. The angular and distance noise are captured by two free parameters. We fit the model separately for the four-object and eight-object conditions.

Can the models predict the difficulty of environments?

We tested whether the models could predict the difficulty, measured in RMSE, of each of the environments across and within clustering structures (Table 1). All models were able to predict the difficulty of the

Downloaded from jov.arvojournals.org on 07/01/2019

environments across clustering structures. However, the chunking model was the worst predictor of subjects' performance ($r = .55$, 95% confidence interval = .37–.70). The relative position model fit environments across clustering structures slightly better than the hierarchical generative model (hierarchical generative: $r = 0.70$, 95% confidence interval = .56–.80; relative position: $r = 0.89$, 95% confidence interval = .82–.93). Within clustering structures, the hierarchical generative model and relative position models generally predicted the difficulty of environments accurately. Notably, however, the hierarchical generative model matched subjects' behavior particularly poorly for 1C4 and 1C8 environments. This is most likely because when all the objects are in a single cluster, the hierarchical generative model tends to recall objects excessively biased toward the cluster centers. Instead, as our analysis of the bias of objects toward their clusters demonstrated, subjects retained a lot more information about the individual objects in these one-cluster environments. This pattern and the relative position model's better ability to predict behavior suggest that relative position encoding dominated subjects' errors.

General discussion

People can encode more information about multiple objects if they exploit the objects' shared statistical structure rather than encoding them independently. We considered several ways people might use this structure when encoding objects and found that in addition to using a hierarchical generative model to *infer* object properties, people also use the hierarchy to *encode* object properties as relative offsets from the central tendency of their group. Because relative positions seem to be recalled with Weber noise, hierarchical clustering reduces the number of large distances that subjects encoded and thus increases overall accuracy.

Implications for the structure of visual working memory

We found that people encoded objects in a relative position tree (Gershman et al., in press; Mutluturk & Boduroglu, 2014), using clustering to reduce the Weber noise of relative distances. Even though the relative position model provided the best quantitative account of our data, the qualitative pattern of results is not entirely consistent with the "pure" chunking, hierarchical generative model, or relative position accounts. In contrast to the predictions of a chunking account, people retained more than just information about the hierarchical structure; they also remembered rich information about the individual object locations. Despite subjects recalling positions biased toward cluster centers in all conditions—consistent with subjects encoding positions via a hierarchical generative model (Brady & Alvarez, 2011; Brady & Tenenbaum, 2013; Orhan & Jacobs, 2013)—this bias decreased as clustering density increased, contrary to the predictions of such hierarchical encoding. Furthermore, although a relative position account could explain errors scaling with increasing relative distances, in isolation it does not predict the systematic biases toward cluster centers. Thus, our results suggest that human memory relies on some amalgamation of these structured representations. Indeed, encoding the relative positions of objects requires first determining the hierarchical clustering structure of the scene, and insofar as this is done under uncertainty, biases should be expected from such inference. Altogether, it seems that both hierarchical inference and relative encoding must play a role in human memory encoding.

The extent to which relative encoding or hierarchical inference dominates the pattern of memory errors is likely to vary across circumstances, either due to strategy switching or even from a constant strategy that incorporates both mechanisms. Insofar as clustering structure or individual object properties may be apprehended more easily with brief presentations or other task constraints, different experimental protocols may yield errors that reflect the clustering structure or the relative encoding. Similarly, stimuli designed with large variations in relative feature offsets will yield more error variability captured by Weber properties of distance encoding, and more homogeneous displays will not show such patterns. In short, although human behavior in our task was best described by the relative position model, we suspect that this result may vary with task parameters and that uncovering this task-dependent variation in error structure may reveal more fine-grained details of visual working memory mechanisms.

Implications for visual working memory capacity

Our findings that subjects remembered the locations of many objects accurately, even in environments containing eight objects, is at odds with models predicated on a fixed number of slots in visual working memory (Anderson et al., 2011; Zhang & Luck, 2008). Additionally, neither such slot models nor flexible resource models (Bays & Husain, 2008; for review, see Ma et al., 2014) capture the effect of scene structure on memory fidelity. Instead, our results are consistent with recent work suggesting that visual working memory performance is constrained by both memory capacity

Downloaded from jov.arvojournals.org on 07/01/2019

and the encoded statistical structure of objects (Brady et al., 2009; Orhan, Sims, Jacobs, & Knill, 2014; Sims et al., 2012). By decreasing the relative distances between objects, clustering may have allowed a more efficient encoding of the objects, ostensibly increasing observers' capacity.

Limitations

Although we defined chunking as subjects retaining memories of clusters but not individual objects, there are other ways subjects could have encoded objects' structure while discarding information about the individual objects. Subjects may have encoded sets of locations as familiar shapes, such as squares, triangles, etc. (Yantis, 1992). They could have then used these remembered shapes, rather than the cluster centers, to constrain the locations of objects. Under this account, no information about individual objects would be preserved over and above the "chunk," but our analysis would still yield reliable information about the relative (within cluster) positions of individual objects.

Another ambiguity of our analysis arises from the assumption that subjects computed the centers of clusters and encoded individual objects relative to those centers (and reported objects with bias toward those center). An alternative possibility is that subjects encoded the positions of objects relative to *each other* with greater bias exerted by nearby objects (e.g., such as gravity with force dropping off with distance). Unfortunately, our results cannot distinguish whether objects were biased toward each other or toward inferred cluster centers.

Although our report focuses on people's memories of object locations, our model analyses revealed that subjects sometimes recalled locations correctly but matched the wrong objects to the locations (Appendix D). Neither the relative position model nor the hierarchical generative model can account for this behavior. It is likely that subjects' real-world priors caused them to expect the locations and identities of objects to be related; subjects may have consequently sought to connect the two. Because locations and identities were independent, the conflict between subjects' priors and the lack of structure in the stimuli may have even impaired performance (Orhan et al., 2014). If the structure of locations and identities had been correlated—such as if all the objects in the same cluster were the same color or same type of animal—subjects may have used the structure of one to inform the other. Given that being able to perceptually group objects based on proximity appears to improve the ensemble encoding of other features (Im & Chong, 2014), it is possible that objects in the same spatial cluster would have even been recalled with more similar features/identities. Future studies may examine how the hierarchical encoding of objects affects binding.

Other factors may have improved subjects' apparent memory capacity in our study. Unlike many prior studies, we used distinct objects that never repeated, which may have reduced interference between objects (Endress & Potter, 2014). Furthermore, many subjects reported using verbal strategies (e.g., "the pants are above the shoes") to help remember displays. We suspect that such strategies would have been only minimally helpful, both because they seem to play a minimal role in long-term memory using comparable encoding times (e.g., Brady, Konkle, Gill, Oliva, & Alvarez, 2013)[4] and because they seem insufficient to attain the precision exhibited by visual spatial memory. Because verbally encoded spatial relations (such as "above" or "left") offer only imprecise location information, we suspect that the main benefit of such verbal encoding was to reduce misassociations between objects (Lew, Pashler, & Vul, in press) rather than encoding the locations themselves. Additionally, patterns of oculomotor movements and attentional shifts could have influenced performance by interfering with encoding in visual memory (Lawrence, Myerson, & Abrams, 2004). Although the uniform distribution of cluster centers in our study still mandates many changes of fixation, it is possible that clustering yields fewer eye movements and attentional shifts between objects in the same cluster, improving the fidelity of memories. Our presentation times were also longer than most visual working memory studies, which may have given subjects more time to encode objects. Given that performance appears to asymptote with display times shorter than those used in the current study (Bays, Gorgoraptis, Wee, Marshall, & Husain, 2011), our results may reflect how people encode stimuli when given enough time to thoroughly observe all objects. Varying the encoding time, delay time, or the environment statistics might reveal how people navigate the space of possible encoding schemes.

Finally, a relative position-encoding scheme may have been particularly well suited for exploiting the structure of spatial positions. Computing relative positions is straightforward for spatial locations and, most likely, other features with Euclidean spaces, such as size or aspect ratio. However, it is less clear how relative encoding would work in more complex, higher-dimensional spaces, such as color or texture. For well-defined but non-Euclidean features, such as hue or orientation, encoding relative positions will likely be helpful if the stimuli are constrained to a narrow range of the space (such that the space is effectively locally Euclidian), but it is not obvious what relative encoding would mean, or predict, if the features span the full range of a circular feature dimension. It is possible that for more complex object properties (such as face

Downloaded from jov.arvojournals.org on 07/01/2019

identity) people collapse those stimuli onto a small set of salient or trained dimensions (such as organizing faces according to race or gender; Hopper et al., 2014). If so, relative memory encoding for such complex objects would be possible in this low-dimensional representation; however, finding evidence of such an encoding strategy would require solving a considerably harder problem: specifying the dimensions along which such stimuli are encoded.

Conclusion

We examined how people encode and use the hierarchical structure of objects under different object loads and structures. In addition to recalling objects biased toward their ensembles, people encoded objects in a relative position tree, using clustering to reduce the Weber noise of relative positions. Our findings are consistent with previous work suggesting that people select encoding schemes that allow them to efficiently represent a given set of stimuli with high fidelity and demonstrate a novel form of encoding.

Keywords: visual working memory, ensemble encoding, chunking

Acknowledgments

We thank Kevin Smith and Drew Walker for reading through many iterations of this manuscript. This was supported by the National Science Foundation (NSF grant CPS # 1239323).

Commercial relationships: none.
Corresponding author: Timothy Franklin Lew.
Email: tflew@ucsd.edu.
Address: Department of Psychology, University of California, San Diego, CA, USA.

Footnotes

[1] In this study, Weber noise refers to errors that are normally distributed in log space.

[2] We calculated RMSE using the formula $\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$ where (x_1,y_1) and (x_2,y_2) are the true location of the object and the subject's reported location, respectively.

[3] In our study, we held the standard deviation of objects within clusters constant, preventing us from analyzing the effect of relative distance on the accuracy of objects. We predict that this relationship between relative distance and accuracy should remain true for objects within the same cluster.

[4] Although Brady et al. (2013) assessed the influence of verbal strategies in long-term visual memory, they also found that both short- and long-term visual memory rely on similar representations; thus, it seems reasonable to apply their findings to short-term memories in our experiments. Moreover, the greater precision in short-term memory would seem to make verbal encoding even less effective here than in long-term memory.

[5] The proportion of locations mismatched by object-to-location mapping function f gives similar misassociation rates.

References

Alvarez, G. A., & Oliva, A. (2009). Spatial ensemble statistics are efficient codes that can be represented with reduced attention. *Proceedings of the National Academy of Sciences, USA, 106*(18), 7345–7350.

Anderson, D. E., Vogel, E. K., & Awh, E. (2011). Precision in visual working memory reaches a stable plateau when individual item limits are exceeded. *The Journal of Neuroscience, 31*(3), 1128–1138.

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science, 12*(2), 157–162.

Bays, P. M., Gorgoraptis, N., Wee, N., Marshall, L., & Husain, M. (2011). Temporal dynamics of encoding, storage, and reallocation of visual working memory. *Journal of Vision, 11*(10):6, 1–15, doi:10.1167/11.10.6. [PubMed] [Article]

Bays, P. M., & Husain, M. (2008, August 8). Dynamic shifts of limited working memory resources in human vision. *Science, 321*, 851–854.

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.

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

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences, USA, 105*(38), 14325–14329.

Downloaded from jov.arvojournals.org on 07/01/2019

Brady, T. F., Konkle, T., Gill, J., Oliva, A., & Alvarez, G. A. (2013). Visual long-term memory has the same limit on fidelity as visual working memory. *Psychological Science, 24*(6), 981–990.

Brady, T. F., & Tenenbaum, J. B. (2013). A probabilistic model of visual working memory: Incorporating higher order regularities into working memory capacity estimates. *Psychological Review, 120*(1), 85–109.

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.

Endress, A. D., & Potter, M. C. (2014). Large capacity temporary visual memory. *Journal of Experimental Psychology: General, 143*(2), 548–565.

Ferguson, T. S. (1983). Bayesian density estimation by mixtures of normal distributions. *Recent Advances in Statistics, 24,* 287–302.

Geman, S., & Geman, D. (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 6,* 721–741.

Gershman, S. J., Tenenbaum, J. B., & Jäkel, F. (in press). Discovering hierarchical motion structure. *Vision Research.*

Hopper, W. J., Finklea, K. M., Winkielman, P., & Huber, D. E. (2015). Measuring sexual dimorphism with a race—gender face space. *Journal of Experimental Psychology: Human Perception and Performance, 40*(5), 1779–1788.

Hollingworth, A. (2007). Object-position binding in visual memory for natural scenes and object arrays. *Journal of Experimental Psychology: Human Perception and Performance, 33*(1), 31–47.

Huttenlocher, J., Hedges, L. V., & Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review, 98*(3), 352.

Im, H. Y., & Chong, S. C. (2014). Mean size as a unit of visual working memory. *Perception, 43*(7), 663–676.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics, 14*(2), 201–211.

Lawrence, B., Myerson, J., & Abrams, R. (2004). Interference with spatial working memory: An eye movement is more than a shift of attention. *Psychonomic Bulletin & Review, 11*(3), 488–494.

Lew, T. F., Pashler, P. E., & Vul, E. (in press). Fragile associations coexist with robust memories for precise details in long-term memory. *Journal of Experimental Psychology: Learning, Memory and Cognition.*

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

Miller, G. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*(2), 81–97.

Mutluturk, A., & Boduroglu, A. (2014). Effects of spatial configurations on the resolution of spatial working memory. *Attention, Perception, & Psychophysics, 76*(8), 2276–2285.

Orhan, A. E., & Jacobs, R. A. (2013). A probabilistic clustering theory of the organization of visual short-term memory. *Psychological Review, 120*(2), 297–328.

Orhan, A. E., Sims, C. R., Jacobs, R. A., & Knill, D. C. (2014). The adaptive nature of visual working memory. *Current Directions in Psychological Science, 23*(3), 164–170.

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

Tudusciuc, O., & Nieder, A. (2010). Comparison of length judgments and the Müller-Lyer illusion in monkeys and humans. *Experimental Brain Research, 207*(3–4), 221–231.

Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology, 24*(3), 295–340.

Zhang, W., & Luck, S. J. (2008, May 8). Discrete fixed-resolution representations in visual working memory. *Nature, 453*, 233–235.

Appendix A: Mechanical Turk replication

To test whether our results generalized when screen size was uncontrolled and in an online sample, we replicated our in-lab experiment using Amazon Mechanical Turk for 10 new environments that contained two clusters each composed of four objects (2C4). Fifty-nine subjects participated, receiving a monetary bonus based on their performance.

The stimuli were identical to our main experiment except we decreased the size of the environments to 600 $\times$ 1100 pixels due to smaller space in Mechanical Turk's interface.

We again used our error similarity measure (q) to measure whether subjects recalled clustered objects

Downloaded from jov.arvojournals.org on 07/01/2019

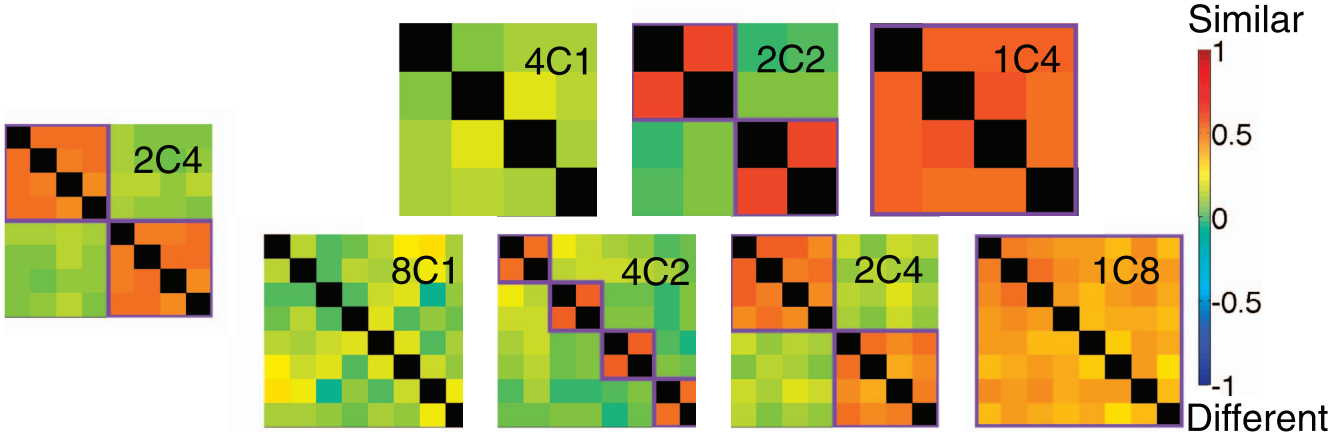

Figure A1. Error similarity heat maps for the Mechanical Turk replication (single 2C4 heat map on the left) and the main experiment (seven heat maps on the right). The format of the figure is identical to Figure 2. Subjects recalled objects in the same cluster with similar errors.

with more similar errors. The error similarity of objects in the same cluster was consistently greater than 0, $t(58) = 23.83$, $p < 0.001$ (Figure A1), indicating that memory errors did not accumulate homogeneously for all objects. Instead, subjects' responses respected the clustering structure of the objects.

Appendix B: Did subjects encode objects based on their positions?

Subjects may have remembered objects using salient positions or landmarks. For example, subjects may have used the center or the axes of the environments or visible landmarks, such as corners and edges (Hollingworth, 2007; Huttenlocher, 1991) to help them recall objects. We expected that if subjects used salient positions or landmarks they would recall objects near such locations more accurately (given Weber noise on relative positions).

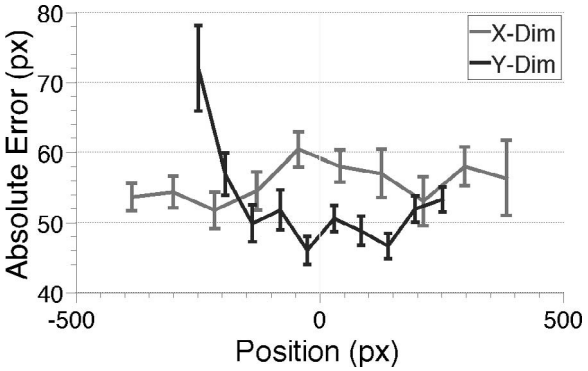

Figure A2. Absolute error in the X- and Y-dimensions based on X- and Y-positions. Lines indicate the binned results; (0,0) indicates the center of the environment, (−500, −350) indicates the bottom left corner of the environment. Error bars indicate *SEM*. The locations of objects had little effect on subjects' errors except when the object was located toward the bottom.

To evaluate these strategies, we examined the magnitude of errors in the X-dimension given the X-position and the magnitude of errors in the Y-dimension given the Y-position and binned the positions (Figure A2). There was no significant effect of position on the magnitude of errors in the X-dimension, $t(438) = 1.56$, $p = 0.12$ for the linear effect of the X-position bin in a model including the fixed effect of the X-position bin). However, the Y-dimension of an object's position did affect the magnitude of errors, $t(438) = 2.90$, $p = 0.003$ for the linear effect of the Y-position bin in a model including the fixed effect of the Y-position bin), such that errors in the Y-dimension increased toward the bottom of the environment. Given that the environments were symmetrical, this most likely reflects subjects initially dragging objects from below the environment to place them rather than subjects using salient positions or landmarks.

Because objects were arranged in clusters, encoding objects in a relative position tree may have been more effective than landmark-based strategies. Objects were typically very close to their cluster centers, making positions relative to clusters easy to remember. If our stimuli were reliably near salient position or landmarks, we expect subjects would have used those alongside the clustering structure of objects.

Appendix C: Markov chain Monte Carlo error model fit

We used a Markov chain Monte Carlo algorithm to fit the parameters of our error model. Let $\Theta^{(i)}$ be the

Downloaded from jov.arvojournals.org on 07/01/2019

	4C1	2C2	1C4	8C1	4C2	2C4	1C8
p_M	.082 (.010)	.076 (.017)	.071 (.008)	.14 (.017)	.11 (.014)	.096 (.028)	.13 (.033)
β_c	.17 (.014)	.14 (.015)	NA	.21 (.012)	.19 (.015)	.11 (.016)	NA
β_o	NA	.12 (.018)	.024 (.006)	NA	.41 (.031)	.25 (.040)	.14 (.039)
σ_g	29.5 (2.0)	22.2 (3.3)	35.7 (2.5)	33.6 (2.1)	31.5 (3.0)	27.5 (4.0)	45.8 (2.0)
σ_c	44.0 (3.4)	29.8 (2.8)	NA	77.7 (2.3)	45.8 (3.3)	40.0 (4.9)	NA
σ_o	NA	28.9 (1.9)	22.9 (1.9)	NA	55.5 (2.6)	47.8 (2.4)	46.0 (4.8)

Table A1. Error model parameter fits for each clustering structure. *Notes*: Each cell indicates the mean parameter value, and the values in parentheses indicate *SEM*. Cells containing "NA" indicate cases in which the parameter and clustering condition are not compatible (e.g., because objects are not clustered in 4C1 and 8C1, the model cannot measure objects' bias toward their cluster [β_o]).

set of parameters $\{p_M^{(i)}, \beta_c^{(i)}, \beta_o^{(i)}, \sigma_g^{(i)}, \sigma_c^{(i)}, \sigma_o^{(i)}\}$ at iteration i and $f^{(i)}$ be the mapping of true locations to response locations at iteration i. In each iteration, the algorithm samples the values of the parameters that compose Θ conditional on the current mappings of f and then samples the mappings of f conditional on the previously sampled value of Θ. The exact algorithm is

1. Choose random starting values for the parameters $f^{(0)}$ and $\Theta^{(0)}$.
2. At iteration i, draw a candidate Θ^* from its proposal distribution $P(\Theta^*|\Theta^{(i-1)})$
3. Compute an acceptance ratio (probability):

$$a = \frac{LIK\left(s|t, f^{(i-1)}, \Theta^*\right)}{LIK\left(s|t, f^{(i-1)}, \Theta^{(i-1)}\right)}$$

4. Accept Θ^* as $\Theta^{(i)}$ with probability $min(a,1)$. If Θ^* is not accepted, then $\Theta^{(i)} = \Theta^{(i-1)}$.
5. Draw a candidate f^* from its proposal distribution $Q(f^*|f^{(i-1)},\Theta^{(i)})$.
6. Compute an acceptance ratio (probability):

$$a = \frac{LIK\left(s|t, f^*, \Theta^{(i)}\right)}{LIK\left(s|t, f^{(i-1)}, \Theta^{(i)}\right)}$$

7. Accept f^* as $f^{(i)}$ with probability $min(a,1)$. If f^* is not accepted, then $f^{(i)} = f^{(i-1)}$.
8. Repeat steps 2–7 N times to get N samples of f and Θ.

For the proposal function $P(\Theta^*|\Theta^{(i-1)})$, we used truncated normal distributions for each parameter's proposal distribution (the truncation enforced the constraints that the noise parameters must be greater than zero and the bias and misassociation probabilities must be between zero and one). Noise proposal distributions had a standard deviation of 2.5 and bias and probability proposal distributions had a standard deviation of .1.

For the proposal function $Q(f^*|f^{(i-1)},\Theta^{(i)})$, we sampled two unique objects based on the inverse likelihood that they came from their currently assigned locations. Intuitively, this selects the two objects that are currently least likely to be assigned to the correct locations. We then swapped the assignments of the sampled objects to create a new mapping proposal assignment.

We set N to 3200 and treated the first 800 samples as burn-in.

Appendix D: Error model parameter estimates

To distinguish different forms of structured representations in visual working memory, our primary analyses focused on the extent to which subjects remembered objects biased toward their clusters and noisily remembered the centers of clusters. In addition, our error model allowed us to examine how the structure of objects influenced other types of errors in visual memories (Table A1). We used fixed effects models that included the fixed effects of the number of objects and the number of clusters to examine how different conditions affected the types of errors subjects made (Table A2).

Subjects may have used the hierarchical structure of objects to help remember associations between objects and their locations. We found that although the rate of misassociations (p_M)[5] increased with the number of objects, it was unaffected by the clustering structure of objects. This suggests that subjects did not use the clustering structure of objects to minimize binding errors.

As objects were arranged in fewer clusters, subjects recalled the locations of clusters with less bias toward the global center (β_c). The decreasing bias of clusters toward the global center may suggest that subjects relied on a representation of objects' hierarchical generative model when remembering the locations of clusters, relying less on the location of the global center

Downloaded from jov.arvojournals.org on 07/01/2019

		Number of objects		Number of clusters	
	DF	*t*	*p*	*t*	*p*
p_M	67	2.0	0.042	.71	0.48
β_c	47	5.3	<0.001	5.2	<0.001
β_o	47	9.4	<0.001	7.1	<0.001
σ_g	67	1.2	0.23	2.7	0.009
σ_c	47	9.2	<0.001	7.7	<0.001
σ_o	47	6.2	<0.001	2.8	0.007

Table A2. The linear effects of a model, including the fixed effects of the number of objects and the number of clusters. *Notes*: DF indicates the degrees of freedom. Under "Number of objects" and "Number of clusters," values in the left and right columns indicate *t* and *p* values, respectively.

as the number of clusters decreased. However, it is unclear why this pattern did not extend to objects' bias toward their clusters.

Subjects also recalled the locations of clusters (σ_c) and objects (σ_o) more accurately. The decreasing noise of cluster and object memories is consistent with the relative position model—organizing objects into fewer clusters should decrease the magnitude of the relative positions needed to represent the objects' and clusters' locations. The clustering structure of objects had an unclear effect on the noise of the global center (σ_g), i.e., the error that is shared among all objects in a display. Subjects appeared to remember the global center more accurately as the number of clusters decreased, but this benefit went away when objects were arranged in a single cluster. The sudden increase in the noise of the global center may reflect subjects focusing on encoding the locations of the individual objects at the cost of the global center when they do not need to remember the clustering structure of objects. Consequently, it is difficult to determine exactly how the objects' clustering structure influenced memories of the global center.

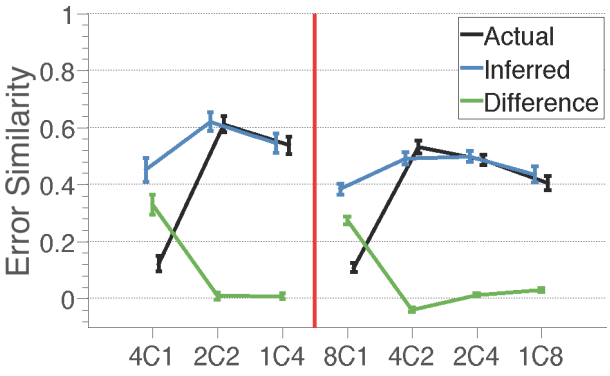

Figure A3. The mean error similarity (*q*) of objects in the same cluster. The black line indicates the error similarity of objects that were actually generated from the same cluster. The blue line indicates the error similarity of objects that the Dirichlet process inferred were generated from the same cluster. The green line indicates the difference between the actual and inferred clusters. Error bars indicate *SEM*. The error similarity of objects was indistinguishable or higher for objects using the inferred groupings compared to the actual groupings used to generate the objects.

Appendix E: Did the nonparametric clustering process predict subjects' errors?

Our cognitive models used a nonparametric Dirichlet process to infer the clustering structure of objects. To determine whether subjects grouped objects like our cognitive models, we examined how well the groupings inferred by the Dirichlet process predicted the error similarities (q) of objects compared to the actual clustering structures used to generate the locations of the objects. For each condition, we found the average error similarity of objects in the same cluster (Figure A3). If no objects were in the same cluster, we calculated the average error similarity over all objects.

The groupings inferred by the Dirichlet process were either comparable to or better than the actual groupings at predicting the similarities of subjects' errors. The Dirichlet process was notably better than the actual clustering structures in unstructured conditions 4C1, $t(34) = 8.20$, $p < 0.001$, and 8C1, $t(34) = 14.20$, $p < 0.001$. This demonstrates that the Dirichlet process grouped objects like subjects did even when there was no intended clustering structure. In the other conditions, the error similarity of objects that were actually from the same cluster versus those that the Dirichlet process inferred were from the same cluster were similar, suggesting that both subjects and the Dirichlet process recovered the intended clustering structures.

Downloaded from jov.arvojournals.org on 07/01/2019