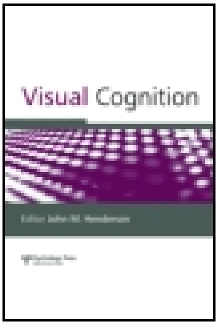

Visual Cognition
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pvis20

Environment sensitivity in hierarchical representations

Timothy F. Lew[a] & Edward Vul[a]
[a] Department of Psychology, University of California, San Diego, CA, USA
Published online: 14 Oct 2013.

To cite this article: Timothy F. Lew & Edward Vul , Visual Cognition (2013): Environment sensitivity in hierarchical representations, Visual Cognition, DOI: 10.1080/13506285.2013.844963

To link to this article: http://dx.doi.org/10.1080/13506285.2013.844963

Visual Cognition, 2013
http://dx.doi.org/10.1080/13506285.2013.844963

Routledge
Taylor & Francis Group

Environment sensitivity in hierarchical representations

Timothy F. Lew and Edward Vul

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

Real-world objects often have hierarchical structure—feathers comprise the wings of birds that are members of flocks, leaves hang from the branches of trees in forests, etc. People are capable of extracting the aggregate ("ensemble") statistics of these groups (Alvarez & Oliva, 2008; Ariely, 2001). This ensemble information is used by visual short-term memory (VSTM) to inform the representation of individual objects (Brady & Alvarez, 2011; Brady & Tenenbaum, 2013; Orhan & Jacobs, 2013). This process is similar to chunking (Cowan, 2001): encoding multiple objects as parts of a larger entity. Chunking assumes that these chunks completely determine their subparts (e.g., encoding "FBI" fully determines its constituent letters). In hierarchical encoding, however, the overarching structure provides only a soft constraint on the components.

Imposing different structures upon objects influences the integration of object and ensemble information in these hierarchical representations. The chosen structure determines what ensembles are used, how levels of information are combined, and, ultimately, the accuracy of object representations. Here we ask: Does the hierarchical structure used by VSTM affect the encoding of individual objects? In the present study, we examined the composition, structure, and use of hierarchical encoding in VSTM as evident in how subjects recalled the locations of objects in different spatial arrangements. We predicted that organizing objects into fewer, larger clusters (here called denser clustering) would enable subjects to use cluster information more effectively and consequently recall the locations of objects with greater accuracy.

Please address all correspondence to Timothy F. Lew, Department of Psychology, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA, USA. E-mail address: tflew@ucsd.edu

METHODS

Thirty-five students from University of California, San Diego studied the same set of 70 environments containing images of real-world objects (from Brady, Konkle, Alvarez, & Oliva, 2008). Each environment had one of seven clustering structures: four clusters of one object (4C1), 2C2, 1C4, 8C1, 4C2, 2C4, 1C8; and there were 10 unique environments of each clustering structure (Figure 1a). Presentation time was 4 s for the four-object environments (4C1, 2C2, and 1C4) and 8 s for the eight-object environments (8C1, 4C2, 2C4, and 1C8) and was followed by a 1 s mask. We then presented an empty environment with the objects located at the bottom of the screen. Subjects were given unlimited time to place all the objects in their locations.

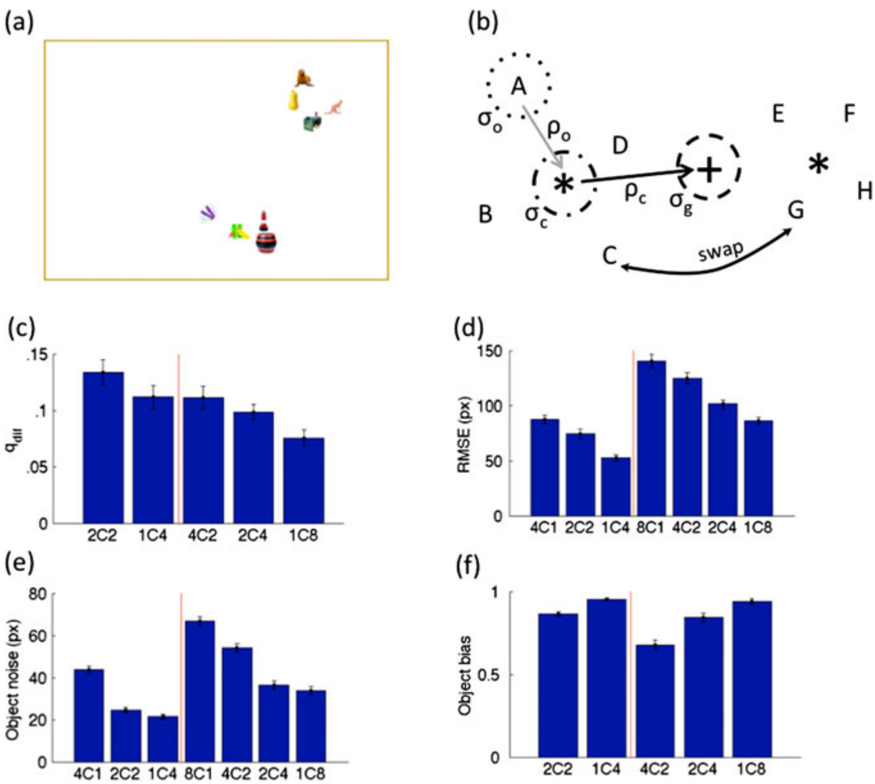

Figure 1. (A) Environment containing two clusters of four objects (2C4). (B) Error model schematic. + is the overall centre, *'s are cluster centres and letters are objects. (C) Error similarity (q) difference of clustered versus nonclustered objects. (D) Subject RMSE. (E) Estimation of object noise. (F) Estimation of object bias. To view this figure in colour, please see the online issue of the Journal.

RESULTS

Did subjects encode objects according to their clustering structure? If subjects utilized clustering structure, we would expect errors to be more similar for objects in the same cluster than for objects that are not organized into clusters. We defined the similarity of the errors (q) in reporting the locations of two objects as:

$$q_{ij} = ||obj_i + obj_j|| \;/\; (||obj_i|| + ||obj_j||)$$

Where obj_i and obj_j are vectors containing the spatial translational error of the two objects' reported locations. If the recalled locations of two objects were both shifted in the exact same direction q would be 1 and if they shifted in the exact opposite directions q would be 0. We found the difference between the average q in the clustered conditions and the average q in the nonclustered condition with an equal number of objects (4C1 or 8C1) (Figure 1c). The q of objects in the same cluster was consistently greater than that of nonclustered objects for all conditions, smallest $t(34) = 10.64$, indicating that memory errors did not accumulate homogeneously for all objects, but rather respected their cluster structure.

To determine the effect of denser clustering upon the recall of object location, we calculated the root mean square error (RMSE) of subjects' responses. We found that RMSE decreased as cluster density increased for both the four-object, $F(2, 34) = 39.69$, $p < .001$, and eight-object, $F(3, 34) = 62.7$, $p < .001$, conditions (Figure 1d).

Although RMSE changed as a function of clustering, this could reflect several different types of errors. We considered noise, bias, and misassociations as potential sources of error and constructed a hierarchical error model to disentangle them (Figure 1b). People may remember locations with some noise resulting from perception or memory. The model includes three levels of correlated spatial noise: that which is shared across all objects (σ_a), for objects of the same cluster (σ_c), and individual objects (σ_o).

Furthermore, when recalling locations people may use memories of ensemble properties rather than memories of the locations themselves; this would result in locations being biased towards higher-level ensembles. We accounted for two types of bias: the degree to which cluster centres are remembered (or are drawn towards the centre of all the objects) (ρ_c) and the degree to which objects are remembered (or are drawn towards their cluster centre) (ρ_o). We jointly estimated these error parameters as well as the extent of misassociations, or swapping, of locations between objects.

The three levels of noise allowed us to determine whether denser clustering actually improved the recall of individual object locations. Object noise decreased as cluster density increased: four-object, $F(2, 34) = 87.98$, $p < .001$; eight-object,

$F(3, 34) = 69.85$, $p < .001$ (Figure 1e), confirming that, as performance improved, individual objects were recalled with greater fidelity.

Denser clustering could have improved performance either by decreasing the error of random guesses around the clusters or by allowing people to better remember the relative locations of objects. If the latter improved accuracy, object bias should have increased with denser clustering. Object bias did increase with cluster density: four-object, $F(2, 34) = 117.14$, $p < .001$; eight-object, $F(3, 34) = 35.08$, $p < .001$, and overall remained quite high (Figure 1f). These results suggest that denser clustering resulted in higher fidelity representations of objects rather than improved random guessing.

DISCUSSION

We found that environmental structure influences how people encoded object information in hierarchical representations. Not only was performance superior for all clustered conditions compared to nonclustered conditions, but it also improved as objects were more densely clustered. We attributed this trend to the increased retention of object information. Our results are consistent with a model that assumes people encode the locations of objects relative to their clusters and clusters relative to the centre of all objects (similar to the relative encoding of objects in scenes from Hollingworth, 2007) and have difficulty recalling larger relative distances. These findings suggest that the arrangement of objects can be manipulated to influence how people encode those objects and to potentially improve the fidelity of recall given the constraints of VSTM.

REFERENCES

Alvarez, G. A., & Oliva, A. (2008). The representation of simple ensemble features outside the focus of attention. *Psychological Science*, *19*(4), 392–398. doi:10.1111/j.1467-9280.2008.02098.x

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science*, *12*(2), 157–162. doi:10.1111/1467-9280.00327

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. doi:10.1177/0956797610397956

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. doi:10.1073/pnas.0803390105

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. doi:10.1037/a0030779

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

Hollingworth, A. (2007). Object-position binding in visual memory for natural scenes and object arrays. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 31–47. doi:10.1037/0096-1523.33.1.31

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. doi:10.1037/a0031541