Journal of Experimental Psychology:
Learning, Memory, and Cognition
2010, Vol. 36, No. 3, 590–604

© 2010 American Psychological Association
0278-7393/10/$12.00 DOI: 10.1037/a0019293

A Category Adjustment Approach to Memory for Spatial Location in Natural Scenes

Mark P. Holden, Kim M. Curby, Nora S. Newcombe, and Thomas F. Shipley
Temple University

Memories for spatial locations often show systematic errors toward the central value of the surrounding region. This bias has been explained using a Bayesian model in which fine-grained and categorical information are combined (Huttenlocher, Hedges, & Duncan, 1991). However, experiments testing this model have largely used locations contained in simple geometric shapes. Use of this paradigm raises 2 issues. First, do results generalize to the complex natural world? Second, what types of information might be used to segment complex spaces into constituent categories? Experiment 1 addressed the 1st question by showing a bias toward prototypical values in memory for spatial locations in complex natural scenes. Experiment 2 addressed the 2nd question by manipulating the availability of basic visual cues (using color negatives) or of semantic information about the scene (using inverted images). Error patterns suggest that both perceptual and conceptual information are involved in segmentation. The possible neurological foundations of location memory of this kind are discussed.

Keywords: category adjustment, scenes, inversion, contrast reversal, categorization

The spatial structure of the world is hierarchically organized (Newcombe & Huttenlocher, 2000; Stevens & Coupe, 1978). For example, cities are divided into districts, each of which consists of many buildings. In turn, these buildings are each made up of several distinct floors, which can be further split into individual offices. The offices themselves may even be subdivided in terms of areas defined by furnishings or functionality, and so on.

This hierarchical structure of space may affect representations of location and capacity to make spatial judgments. Stevens and Coupe (1978), for example, demonstrated that most people erroneously believe that Reno, Nevada, is east of San Diego, California, because they know that much of Nevada is east of much of California. Similarly, Nelson and Chaiklin (1980) noted that participants produced a systematic pattern of error in recalling the location of a point placed on the midline within a circle. Some researchers have suggested that such errors simply show the schematized, distorted quality of spatial memory representations (e.g., Tversky, 1981), whereas others have proposed that such cases of bias reflect the use of inferences based on information from different levels of a spatial hierarchy (e.g., Stevens & Coupe, 1978). Building on the latter position is the idea that systematic errors reflect neither distorted representations nor erroneous inferences but rather are simply the result of an adaptive process of optimal combination across multiple sources of information—a process that minimizes overall error (e.g., Alais & Burr, 2004; Ernst & Banks, 2002).

One class of such adaptive systems is Bayesian models. The general Bayesian framework involves the combination of multiple sources of information about a quantity (e.g., a spatial location) in an optimal fashion. A useful hypothetical example, outlined by Deneve and Pouget (2004), is estimating the location of an object for which there is both visual and auditory information about its location. They ask the reader to imagine that the two sensory modalities offer slightly different estimates of the location and that the variance of the visual estimate is smaller than that of the auditory estimate. That is, the reliability—which is the inverse of variance—of the visual information is greater than that of the auditory information. The two estimates of location can then be imagined as two overlapping probability distributions—one wider and one thinner—each centered about its mean. Under a Bayesian framework, the cross-modal estimate of location is a linear combination of the two unimodal estimates, weighted by their respective reliabilities. Therefore, the final estimate of position will be closer to—but likely not equal to—the visual mean estimate (for a review of cross-modality Bayesian combination, see Deneve & Pouget, 2004). Note that although the combination of the two unimodal estimates leads to bias away from the visual estimate, it also produces a combined distribution of estimates that is less variable than either of the two unimodal estimates, reducing the average error.

Similarly, Ernst and Banks (2002) have demonstrated this effect of optimal combination across sensory modalities by exploring the integration of visual and haptic information. They systematically varied the amount of visual noise, making this information less reliable. The results suggest that the human nervous system inte-

Mark P. Holden, Kim M. Curby, Nora S. Newcombe, and Thomas F. Shipley, Department of Psychology, Temple University.

This research was supported by National Science Foundation Grant SBE-0541957. Portions of this research were presented at the seventh annual meeting of the Vision Sciences Society, May 2007; the 48th annual meeting of the Psychonomic Society, November, 2007; and the 49th annual meeting of the Psychonomic Society, November 2008. We would also like to thank Alinda Friedman for her helpful comments and insight. A sample of scenes used is available from Mark P. Holden upon request.

Correspondence concerning this article should be addressed to Mark P. Holden, Department of Psychology, Temple University, Philadelphia, PA 19122. E-mail: mholden@temple.edu

grates haptic and visual information by weighting them according to the reciprocal of their variances; that is, visual information dominated haptic only for tasks in which the variance associated with visual estimation was lower than that associated with haptic estimation (Ernst & Banks, 2002).

Optimal or Bayesian combination may also occur when the sources of information arise from the same continuous, hierarchically structured dimension, such as time or color (e.g., Huttenlocher, Hedges, & Duncan, 1991). In such cases, fine-grained and coarser estimates are combined. For example, one could remember a color as being in the broader family of blues (represented by a prototypical value) but also remember it as being of a particular shade. Evidence for combination across such values has been found for estimates of the time at which an event occurred (Huttenlocher, Hedges, & Bradburn, 1990; Huttenlocher, Hedges, & Prohaska, 1988; Lee & Brown, 2003), memory for the fatness of fish-like stimuli (Duffy, Huttenlocher, & Crawford, 2006; Huttenlocher, Hedges, & Vevea, 2000), and shades of grey (Huttenlocher et al., 2000). It has even been found in social dimensions, such as judgments of gender and ethnicity for photomorphed faces (Corneille, Huart, Becquart, & Bredart, 2004; Huart, Corneille, & Becquart, 2005).

This combination of hierarchically organized values is central to Huttenlocher et al.'s (1991) category adjustment (CA) model. In terms of spatial location, the CA model posits that individuals remembering a given location combine information from different levels of the hierarchical structure of space. The final estimate of location is hypothesized to be an optimal combination of a fine-grained metric estimate and coarser categorical information, represented by a central or prototypical value. For example, if the task is to remember where one's keys are, one might combine the remembered value of approximately 4 in. from the edges of a table with the categorical estimate of somewhere on the top left quadrant of the surface. According to the CA model, this combination proceeds in a Bayesian manner, with each type of information weighted by its reliability. For example, the reliability of the 4-in. estimate might be higher immediately after dropping the keys than if they were dropped hours ago. Thus, in the latter case, the metric estimate would be relatively less strongly weighted than the central value of the quadrant. The combination across these two types of information introduces bias away from the metric estimate but improves accuracy across multiple instances by reducing the variability of the final estimates.

Empirical evidence for this model has been derived from experiments in which people are asked to recall locations presented within simple geometric spaces, such as a blank circle (e.g., Huttenlocher et al., 1991). Participants used both fine-grained and categorical information (corresponding to the quadrants of the circle) in remembering the locations. Bias was seen toward the category prototype (center of mass) of the surrounding quadrant. Subsequent research has demonstrated similar biases in location memory using a number of different paradigms, including reproduction of the location (e.g., Fitting, Wedell, & Allen, 2007; Wedell, Fitting, & Allen, 2007), pointing tasks (e.g., Schutte & Spencer, 2002; Spencer & Hund, 2002), and verbal responses (e.g., Spencer, Simmering, & Schutte, 2006). Furthermore, as predicted by the CA model, the degree of bias increases with greater uncertainty about fine-grained information (e.g., due to longer delays between studying and recalling a location; Hund & Plumert, 2002; Huttenlocher et al., 1991; Huttenlocher et al., 1994; Plumert & Hund, 2001; Recker, Plumert, Hund, & Reimer, 2007). It is important, however, that in this model each type of information—metric and categorical—is assumed to be unbiased; the resultant bias toward the category prototype does not occur from misrepresentations or distortions of stimuli but from the use of categorical information to adjust inexactly remembered metric values.

It is interesting to compare the CA model to Kosslyn's (1987) distinction between categorical and coordinate information. Kosslyn was not focused on how these types of information interact or can be combined but did propose neurological bases for the two kinds of information, which appear closely related and perhaps identical, to the kinds of information posited by the CA model. Results of several studies demonstrate the role of the left frontal and parietal cortices in categorical spatial relations, whereas the corresponding regions of the right hemisphere are more involved in fine-grained metric relations (e.g., Kosslyn, Thompson, Gitelman, & Alpert, 1998; Slotnick & Moo, 2006). Such lateralization effects have been demonstrated in the posterior parietal cortex (Laeng, 1994; Trojano, Conson, Maffei, & Grossi, 2006), the angular gyri (Baciu et al., 1999), and the prefrontal cortex (Slotnick & Moo, 2006). Kosslyn et al. (1998) have also reported evidence of a specific role of the left frontal cortex in categorical processing, as well as roles for the right superior parietal lobe and right precuneus in metric spatial relations. Postma, Kessels, and van Asselen (2008) further suggested that the hippocampus may play a role in long-term retention or when a task demands viewpoint independence, whereas the parietal structures may play a larger role in shorter term tasks, like those in the traditional location memory experiments (e.g., Huttenlocher et al., 1991).

However, at least two uncertainties about the CA model need exploration. First, it is unclear whether the location memory bias predicted by the CA model applies to memory for locations in natural spaces. The stimuli used in many studies of location memory have used highly simplified spaces, such as the interior of a single geometric shape. These spaces seem likely to invite the use of horizontal and vertical axes defined by gravitation or lateral symmetry of the human body to define the categorical frame of reference. It is unclear whether the CA model applies to memory for locations in natural spaces, which contain irregular categories, as it should if it is to provide an overarching framework for thinking about spatial memory.

Thus, it is of critical importance to address the question of whether the CA predictions hold for more complex, naturalistic environments. Some of the key issues that might differentiate the coding of spatial location in natural space from the simplified stimulus space of past studies were outlined by Newcombe and Huttenlocher (2000). One issue is the possibility of segmentation into spatial categories by higher order processing, such as functional categorization (e.g., the picnic area of a park). Second, naturally occurring spatial categories are rarely regular geometric shapes, which may affect the use of Cartesian or polar coordinates in estimating fine-grained locations. Third, because real-world space is rarely uniform, the presence of landmarks or distinctive features within a region is highly probable and may change the location of the prototypical value away from the geometric mean of the area.

Given the adaptive value of a system that reduces the average error of estimates, it seems likely that the same basic principles of optimal combination across levels of information should apply to location memory in more complex spaces. However, we cannot assume that this is the case. This article, therefore, examines memory for locations in photographs of complex natural scenes. Specifically, in Experiment 1 we use complex scenes rather than simple geometric shapes in a location memory experiment. To preview, our results support the CA hypothesis for location memory in these scenes.

Complex natural images are imbued with semantic content as well as perceptual information. Spatial categories in complex images might be defined by either or both of these types of information. For example, a mountain may be perceptually distinct from the plateau beneath it and the sky above it, but *mountain* is also associated with specific semantic knowledge that may help to define the spatial category; for example, a mountain may be subdivided into the *peak* and *base* areas, even without strong local perceptual boundaries to distinguish these areas. It is possible that either or both of these types of information are used in forming spatial categories when individuals remember locations.

Dynamic field theory (DFT), as outlined by Spencer and colleagues, suggests that only perceptual information is used in segmenting space into categories (e.g., Johnson, Spencer, & Schöner, 2008; Simmering & Spencer, 2007; Spencer, Simmering, Schutte, & Schöner, 2007). We argue that although the rigidity of categories defined purely on a perceptual basis may be appealing—higher precision of the boundaries would lead to fewer miscategorization errors—a more traditional Bayesian approach would treat the categories as Bayesian priors, learned or derived from experience; this suggests a potential role for conceptual knowledge in segmenting scenes into constituent spatial categories as well. Indeed, Hund and Plumert (2003, 2005) have shown some evidence that both knowledge about object identity and spatiotemporal experience can aid in forming spatial categories in relatively simple spaces. Furthermore, if we assume that the goal of combining categorical and metric information is to minimize errors, then having multiple categories from which to choose could furnish the individual with options that might further minimize errors (by reducing category size, for example). Through the use of our more complex scene stimuli, which include both perceptual and semantic information, we can, for the first time, address the question of whether perceptual and/or conceptual information is used in defining spatial categories that can be used to help recall locations in a scene. In Experiment 2, we manipulated the images to reduce the influence of conceptual structure. We then compared the patterns of errors between image types to investigate whether these manipulations had altered category formation.

Experiment 1

In Experiment 1, participants were asked to recall locations within color photographs depicting natural scenes. Specifically, we examined the distribution of errors about the correct value for bias toward a spatial category prototype (center of mass), as predicted by Huttenlocher et al.'s (1991) CA model. Such bias would be evidence in favor of the combinatory process suggested by the model. The spatial categories within the pictures were initially defined using a computer algorithm based on perceptual cues (see Analysis 1A). To ensure the validity of these categories, a separate group of participants was also asked to identify the categories containing tested locations (see Analysis 1B). In each case, the prototypical value of the category was defined as the center of mass of the two-dimensional region of the image occupied by the category. We predicted that the direction of errors would be biased toward the center of mass of the surrounding category. To preview, the results of the two analyses do not differ from each other, and both indicate that individuals' memory for spatial locations in natural scenes are biased toward central (prototypical) locations.

Method

Participants. The participants were 19 undergraduate students at Temple University who participated for course credit. Two additional students participated; however, these data were not included in the final analyses because the participants did not follow the instructions. Specifically, subjects were discarded if they (a) identified a location that was impossible (outside the image boundaries) on more than 10% of the trials, or (b) badly misremembered a location, erring by more than 25% of the image length (225 pixels, or 6.5 cm) on more than 15% of the trials.

Materials. The presentation of stimuli and collection of data were controlled by a program (PsyScope X B46; Cohen, MacWhinney, Flatt, & Provost, 1993) that ran on a Macintosh computer connected to a large (37 cm wide × 28 cm tall) flat-screen CRT monitor set to a resolution of 1280 × 960 pixels. The computer directly recorded responses with pixel precision, allowing even subtle response patterns to be detected.

Stimuli. The stimuli were free downloaded images of natural landscape vistas. Thirty-five such images were used, with four possible locations per image. Landscapes were evenly distributed between seven types: desert scenes (e.g., sand dunes), plains (e.g., grasslands), rolling hills, coastal scenes, rocky regions (e.g., hoodoos), lakeside scenes, or mountain scenes. For example, Figure 1 is considered a lakeside scene, whereas Figure 2 depicts a sand dune and is therefore considered a desert scene.

Locations in Experiment 1 were small yellow dots, which stood out against the natural background colors. The dots were made to be elliptical, corresponding to the slope that they appeared on, and were blurred slightly at the edges. They were designed to support the impression that the location depicted was a location within the scene, rather than a point on the photographic surface (Figure 1A). Locations themselves were chosen pseudorandomly, so that the relative direction and distance between the correct locations and category centers (as well as between correct locations and nearby category edges) varied across locations. The only constraints placed on potential locations were that they could not appear in the sky or within 2 cm of any border of the image. The latter constraint was meant to reduce use of the location's distance from a border, rather than scene content, to recall metric information.

All images were 900 pixels (26 cm) in length but varied in height. Study images (those that portrayed the dot) were cropped by removing 10% of the image from a nearby border, with the constraint that it could not eliminate any part of the surrounding category (see Figure 1). This cropping further prevented subjects from relying on cues such as distance to image edges and encouraged them to rely on the image content. For similar reasons, the study and test images were each randomly assigned to one of nine locations on the screen. Participants were informed of both of these

(a) Study Image

(b) Test Image

Figure 1. Example of natural scene stimuli used in Experiment 1. The study image (a) was made smaller in area by 10% than the test image (b) by removing information from the top border of the image. The location for recall is found on the island of the study image (a).

operations beforehand and were given two sample trials to demonstrate them.

Procedure. The experiment consisted of four blocks, with each of the 35 landscape images appearing in a random order once per block. There were four possible dot locations per landscape, each randomly assigned to one of the four blocks for each subject.

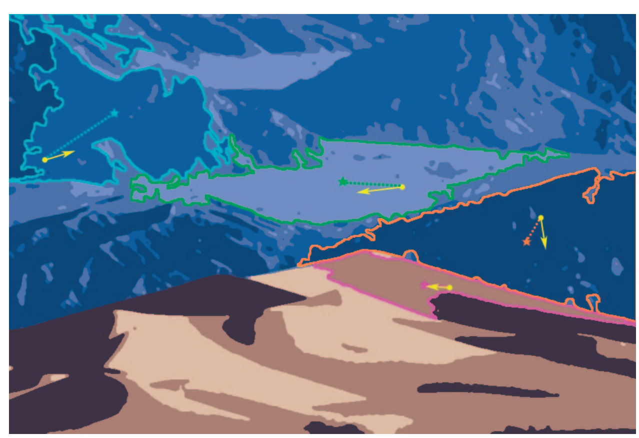
On every trial of the experiment, two landscapes were presented serially, each containing one location. These were followed by the full-sized images of the landscapes in the same order, without locations. The participants' task was to indicate the location of the dots using the mouse. The primary advantage of this procedure is that it doubles the data collected over the course of the experiment compared with a procedure that employs a simple delay between study and test images, while still retaining the characteristic error patterns of bias toward category prototypes. That is, Huttenlocher et al. (1991) demonstrated that taxing spatial memory by including an interference task yielded similar and even increased biases in location memory estimates when compared with trials with a delay between study and test; using two landscapes per trial, then, taxes spatial memory and provides an interference task while increasing the amount of data collected per participant. As we shall see, analyses confirmed that no difference in error patterns was found for locations shown first or second in a trial.

Trials began with a 500-ms fixation cross in the center of the screen. The locations were then shown for 2,500 ms each, separated by a 750-ms blank screen. Following a 1,000-ms delay, the recall images were then presented. These remained on screen until subjects made a response. If participants were unsure of the location, they were encouraged to guess. Trials continued until all 140 image locations had been seen.

Analysis 1A

Category identification. As a preliminary method to identify a priori the spatial categories surrounding each location, the clus-

(a) Original Landscape Image



(b) Clustered Imaged with Categories and Average Response Vectors

Figure 2. An example landscape (a) and the categories indentified by the *k*-means clustering algorithm (b). The test image (a) was determined to contain seven color-based clusters (b). Categories for Analysis 1A were determined on the basis of these clusters and are outlined in (b). The locations to be remembered are shown as yellow dots, whereas the center (or prototype) values for each category are depicted by stars. The mean response vector is shown as a yellow arrow. The angle of difference between this arrow and the dotted line (connecting the correct location and the category center) was measured.

tering plug-in of a widely available software package—ImageJ—was used (Abramoff, Magelhaes, & Ram, 2004). This plug-in employs a pixel-based *k*-means clustering method. That is, pixels in the image may be defined by their red, green, and blue (RGB) color values; to create clusters, the plug-in randomly chooses *k* such values, called cluster values, and calculates the difference between each pixel's RGB profile and these values. Each pixel within the image is assigned to the nearest cluster value. The cluster values are then iteratively adjusted, with each pixel again being assigned to the nearest cluster value, to minimize the average distance between them. The value of *k* for each image was determined by the number of distinct peaks present on the image's RGB profile histogram, determined by two independent observers.

The result of this plug-in is an image consisting of *k* color values, which was then despeckled to remove noise (see Figure 2). The category was then identified by using the software's wand tool at the correct location; this tool creates a selection by seeking the nearest edge and following it until it loops back on itself. The center of mass of the two-dimensional shape surrounding each dot location served as the central or prototype value for our experiments. An example image with all four dot locations, spatial categories, prototypes, and average errors is shown in Figure 2B.

Scoring. Again, the participants' task was to indicate the location of the dots using the mouse. On some trials, however, participants did not make a true response; the button was accidentally pressed before the mouse had moved from the previous response location. Other large errors occurred because we had encouraged participants to guess if they were unsure of the location. To prevent such cases from affecting the mean error, we deleted those responses whose error was so great that it was likely that the participant had accidentally responded or momentarily failed to attend to the task. Specifically, those responses greater than 25% of the image length (225 pixels, or 6.5 cm) from the correct location were removed. These responses accounted for 3.8% of the total data. After this culling, the standard deviation of the distance of responses from the correct location was calculated for each location (Levine, 2004); responses greater than three standard deviations were deleted, eliminating an additional 0.17% of the data.

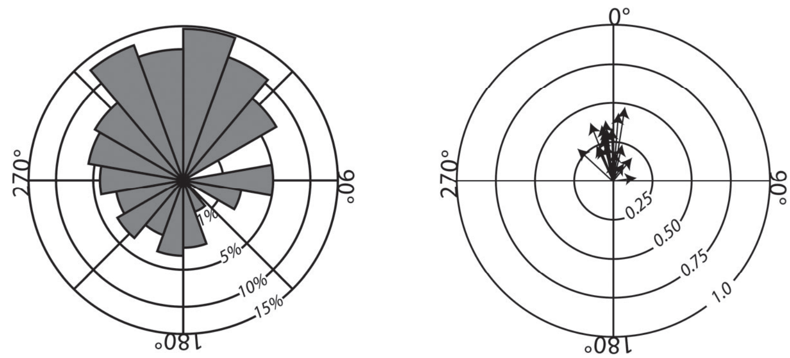
The remaining responses were converted to vectors, each originating at the associated correct location. These error vectors were then added and the length of the summed vector divided by the number of responses to obtain an average error for each location (e.g., see Figure 2B). Note that averaging error vectors as described automatically weights the angle of an individual response by the magnitude of error. For each location, the weighted average error vector was then compared to the error direction predicted by the CA model (the angle between the correct location and the category prototype; e.g., Figure 2B). Therefore, each location was associated with a specific difference angle representing the difference between the predicted and observed error. Note that this method differs from that of previous studies (e.g., Huttenlocher et al., 1991). Rather than examining responses with respect to a fixed point across all locations—the circle center, in the past—we examined the direction and magnitude of error vectors from the correct location itself. The primary advantage of this approach over those used in previous studies (e.g., Huttenlocher et al., 1991) is that whereas previous studies used separate analyses for the direction and the distance of errors, this method directly assesses both types of information simultaneously. This has additional advantages in that it allows us to easily visualize errors with respect to the predicted value, even across irregularly shaped categories. It also allows for comparison of error distributions (e.g., we can compare the error distributions of images seen first or second in a trial pair).

Results. Response vectors for each location were used to plot bias (the mean direction of recall errors versus the predicted direction). That is, the average error vector was calculated for each location, as described above, and the direction of this vector was compared to the direction of error predicted by the CA model (e.g., Figure 2B). Therefore, each location was associated with a specific difference angle representing the difference between the predicted and observed error. The resultant distribution of differences across all locations is shown in Figure 3A. Note that in this figure, an overall pattern of bias consistent with the CA model would lead to errors of approximately 0°. This figure implies that, for the majority of locations shown, participants appear to have erred in the direction of the category prototype.

However, it would be inappropriate to use these difference angles to determine the overall mean error or to perform statistical analyses (Zar, 1998). One reason is simply that each participant contributed multiple data points, which we cannot treat as independent. Furthermore, using the difference angles would imply that each mean difference for a given location had an angular deviation of zero—that all participants erred in precisely the same direction. An example helps illustrate this latter point: Consider two locations, each remembered by two individuals. On the first location, both individuals err in precisely the direction of 90° by precisely 1 unit of length. The average vector is therefore directed toward 90° and is 1 unit long, implying zero angular variance. For the second location, both individuals make the same 1-unit error, but one errs in the direction of 1°, whereas the other misremembers the location in the direction of 179°. For this location, the angular variance of responses is much higher. Even though the average angle is still 90°, the length of the average vector is very small (approximately 0.02 units), implying that the errors are more or less randomly distributed. So, returning to our analysis, it would be inappropriate to use difference angles alone without taking into account the angular variance associated with each value.

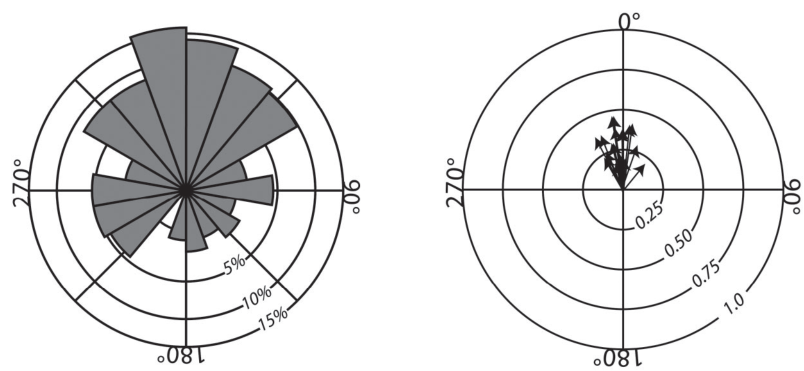
The proper analysis should also take into consideration the magnitude of location memory errors. That is, it is likely that for some locations—those immediately adjacent to prominent textural elements, for example—errors would be small and randomly distributed about the correct location. In such cases, the only source of error might be motor response and not location memory per se. Therefore, cases in which the average error magnitude is small are less informative about location memory processes. Indeed, for a distribution in which errors are small and randomly distributed, shifting a single response by 1 or 2 mm could dramatically shift the direction of the average error vector for that location.

Therefore, to address these issues and to more directly assess the question of whether individuals tend to recall locations as being closer to the category prototype (center of mass) than they were, we performed Hotelling's one-sample second-order analysis (Zar, 1998). Specifically, rather than calculating the difference angle between the average and predicted errors for each location, we calculated the difference between each individual response and the predicted error. The average of these difference angles across all



(a) Difference Angles for All Locations (using algorithm-defined categories)

(b) Weighted Error Vectors for All Subjects (using algorithm-defined categories)



(c) Difference Angles for All Locations (using drawn categories)

(d) Weighted Error Vectors for All Subjects (using drawn categories)

Figure 3. Difference of location memory errors from the predicted direction (toward category center). (a) and (b) defined categories using the *k*-means clustering algorithm, whereas (c) and (d) defined categories using the category identification task. (a) and (c): Equal-area rose diagram of the angular difference between the mean error vector for a location and the predicted direction of error for that location. Angular difference values for all locations are represented. (b) and (d): Mean error vectors per subject. The difference between each response and the predicted direction of error is recorded, and a mean error vector (with respect to the predicted direction) is calculated. Each vector therefore represents a single subject's tendency to err in the predicted direction or not. Vector length varies inversely with angular dispersion (i.e., longer vectors represent more consistent errors in a given direction, such as toward the prototype). Note that the predicted direction of error in all figures is 0°.

140 locations was then calculated for each participant. Thus, each participant is associated with one mean difference angle. Furthermore, because we calculated each subject's average error vector in the same manner as above (adding the vectors, then dividing the summed vector length by the number of responses), we automatically took into account the issue of error magnitude. The second-order (also called the mean of means) test is appropriate, as each average error vector represents the mean difference angle for a single participant. Finally, Hotelling's analysis tests the distribution of these average error vectors, which include both direction and vector length (angular variance) information, against the null hypothesis that there is no mean direction of bias (i.e., random error).

The second-order distribution—the average error vectors for each participant—is shown in Figure 3B. With an alpha of .05, subjects erred in a significantly directed fashion, $M = 355.20°$, 95% CI [342.98°, 8.05°], $F(2, 17) = 30.62$, $p < .01$, $d = 2.68$. We also tested whether individuals erred toward the centers of the quadrants of the image, treating it as a simple rectangular shape, and found no evidence of this, $F(2, 17) = 1.44$, *ns*. Using Hotelling's second-order test for paired data, we found no difference in the errors made between images seen first or second within a trial, $F(2, 17) < 1$, *ns*. Thus, participants' memory for locations was largely biased toward the predicted category centers.

Analysis 1B

Category identification. To ensure the validity of the categories identified by the clustering algorithm, a separate group of participants was asked to identify the spatial categories. Participants were 35 undergraduate students at Temple University who participated for course credit.

The categories were identified by participants on physical copies of the images used in the location memory task. Physical versions were identical in size to those seen on the computer screen. The images were presented serially inside transparent plastic sheet protectors, upon which participants were asked to trace the "region[s] surrounding or enclosing the dot[s]" using felt markers. As with the location memory task, the tracing task consisted of four blocks, with each of the 35 landscape images appearing in a random order once per block. The four possible dot locations per landscape were again each randomly assigned to one of the four blocks for each subject. Participants were allowed to take as much time as needed to complete the task.

To combine the multiple drawn categories into composite ones for analysis, each outlined category was transformed into a digital copy. The categories for each dot location were then overlaid using photo editing software. Portions identified by only a single subject as belonging to the category were discarded. The remaining areas were used to calculate the category prototypes. As with Analysis 1A, the two-dimensional center of mass served as the category prototype; however, in this case the pixels were weighted by the degree to which they were agreed upon. Thus, regions that all subjects identified as part of a category were weighted twice as heavily as areas identified by only half of the subjects.

Scoring. Scoring was identical to that of Analysis 1A, except that the values of the category prototype were altered to reflect the results of the category identification task.

Results. As in Analysis 1A, the average error vector was calculated for each location and compared to the predicted error direction (the angle between the correct location and the category center, derived from the category identification task). The resultant distribution of differences between these angles is shown in Figure 3C. Again, note that an overall pattern of bias consistent with the CA model would lead to errors of approximately 0°. This figure implies that, for the majority of locations shown, participants appear to have erred in the direction of the category prototype.

However, as above, it would be inappropriate to use these difference angles to determine the overall mean error (Zar, 1998). We therefore again performed Hotelling's one-sample second-order analysis, using the category prototype (center of mass) values from the category identification task. The second-order distribution is shown in Figure 3D. With an alpha of .05, subjects erred in a significantly directed fashion, $M = 353.46°$, 95% CI [343.82°, 3.23°], $F(2, 17) = 58.94$, $p < .01$, $d = 3.72$. Again, using Hotelling's second-order test for paired data, we found no difference in the errors made between images seen first or second within a trial, $F(2, 17) < 1$, *ns*. Therefore, as in Analysis 1A, participants' memory for locations was largely biased toward the predicted category centers.

Furthermore, to test the similarity between the categories identified by the clustering algorithm (Analysis 1A) and those identified by a separate group of participants (Analysis 1B), the second-order error distributions from the two analyses (Figures 3B and 3D) were compared. Again we used Hotelling's second-order test for paired data. This analysis subtracts the paired vectors from each other and tests the distribution of differences against the null hypothesis of a random distribution. If the category centers used in the two analyses differed significantly from one another, then the distribution of error vectors with respect to these predicted values would also differ; that is, if the distributions differ significantly either in terms of variance or direction of average error, then we can say that the category centers differed between the two analyses. However, there was no significant difference, $F(2, 17) = 2.55$, *ns*, implying that the categories identified by the clustering algorithm were similar to those identified by a separate group of participants.[1]

Discussion

This experiment supports the validity of the CA model in complex environments. The results clearly demonstrate that memory for locations within visually rich images is biased and that this bias tends to be in the direction of the category center, as predicted by the CA model. We take this as the first evidence of the combinatory process outlined by Huttenlocher et al. (1991) in visually complex, natural scenes. Furthermore, the categories used as a frame of reference for remembering locations were based on information captured by the image and were not related to the geometric information (e.g., boundaries or axes of symmetry) of the rectangular image space, because errors were not biased toward the quadrant prototypes.

However, these results leave open the question of the specific kind(s) of information used by participants to identify categories. That is, because images such as those of Experiment 1 are rich with semantic information as well as perceptual data, and because these two types of information are highly correlated in such images, the spatial categories in complex scenes might be defined by either (or both) of these types of information. In Experiment 2, we address this question.

Experiment 2

Experiment 1 showed that memory for locations within complex natural scenes was biased toward category prototypes. However, we do not know what sort of information is being used by participants to parse these scenes into their constituent categories. This information might include basic perceptual cues (the brownish region adjacent to blue), functionality (the sandy beach), or semantic knowledge (a river delta), among others. The software we used to define categories in Experiment 1 (Analysis 1A) used only perceptual information, so that may have been the dominant factor for humans as well. However, because the perceptual and conceptual information in these images are highly correlated, it is difficult to know whether one type of information or both were used. In Experiment 2, we therefore explore the role of conceptual information in segmenting scenes into spatial categories.

It seems likely that conceptual information plays a role in identifying spatial categories in everyday situations. The advantage of any Bayesian system lies in its ability to adapt estimates as

[1] The difference in F values of the two sets of analyses (30.62 and 58.94, respectively) might seem to imply that a difference might be found, but visual inspection of Figures 3B and 3D suggests that the difference in F values is driven essentially by one or two subjects. That is, in Figure 3B, the overall variability of the distribution is increased by one data point biased toward 90° and one toward 310°, whereas the data points in Figure 3D are slightly more clustered. However, the difference in variance between the two distributions is still quite small, so no significant difference was found.

more information becomes available. In terms of such input, a system in which categories are rigidly defined without regard to semantic information would be severely limited at best. Knowledge of the interactions of light and shadow, for example, would be critical to choosing optimal spatial categories to encode. Consider remembering the location of one's keys in a grassy park: recalling that they were within the dark shadow of a tree (a region defined by a change in luminance) may be optimal in a task of short-term recall but will not be of much use after a few hours, or past sunset. Even in the short term, conceptually defined categories may have the potential to reduce location memory errors to a greater degree than some perceptually defined categories. For example, in our task a novice may remember a location as being "on the rock face"; an expert geologist, however, may encode the location with respect to some conceptually identified structure, such as the "Mesozoic portion of rock." Although such categories do admittedly have perceptual correlates—different profiles due to erosion, for example—expertise effects as outlined above would still certainly demonstrate a role for top-down processing in category identification. Furthermore, as this example demonstrates, conceptually defined categories need not always incur the cost of having less precise boundaries than those of perceptually defined categories.

The purpose of Experiment 2, therefore, was to explore the nature of segmentation into spatial categories within complex scenes in a situation in which perceptual and conceptual influences could be distinguished. We again asked participants to recall locations within scenes, but we manipulated the images to reduce the influence of conceptual structure. Specifically, some images were presented upside down, whereas others were presented as color negatives of the original, canonical versions.

Inverted (upside down) versions of scenes retain the low-level visual cues (e.g., color, texture, luminance cues) that are available in the canonical images. Inversion, however, interferes with the extraction of meaning from scenes (Brockmole & Henderson, 2005; Shore & Klein, 2000; Velisavljevic & Elder, 2008). For example, inverted depictions of scenes are more difficult to recognize (Rock, 1974) and produce less conceptual masking (Intraub, 1984). Thus, if spatial categorization in Experiment 1 depended on conceptual structure, it might be impeded by using inverted images. This is not to say that conceptual information is completely inaccessible in these images but simply that it should be less readily available, thereby affecting the categories that individuals form. For example, the location depicted in Figure 1A is clearly located on the island. This much is obvious for both canonical and inverted orientations. However, the island itself may be subdivided into smaller categories, such as the left and right sides of the island, the peak and base areas, or the portions more thickly and more sparsely covered by trees. Because processing of semantic information of complex scenes becomes somewhat impaired with inversion, individuals may be less inclined or less able to refine categories into such subdivisions. Similarly, a location depicted on the tree in Figure 1A might be remembered as being located on the upper or lower portions of the tree or on a particular branch in an upright image, versus the tree as a whole for an inverted image. In any case, a difference in the error patterns between canonical and inverted images, because they contain the same low-level visual information, would suggest that inverting the images alters the manner in which individuals identify spatial categories.

Color negatives, on the other hand, do not retain all low-level visual information that is available in the corresponding canonical image. This manipulation may affect the use of cues such as surface texture or shading cues (Vuong, Peissig, Harrison, & Tarr, 2005) that may have been used in identifying spatial categories in Experiment 1. Photographic negation may or may not also affect identification of spatial categories on the basis of semantic information. For example, photographic negation has been shown to affect judgments of facial identity (e.g., White, 2001), but the impairment of item-level recognition does not necessarily impact identification of semantic categories (i.e., whether the object is a face or not). On the other hand, it has been suggested that the extraction of semantic information may also be disrupted when objects are colored in an unfamiliar way (Goffaux et al., 2005). Nevertheless, it is certainly the case that low-level visual information is disrupted in color negatives.

The effects of the image manipulations are schematically shown in Figure 4. Because the two manipulations alter different aspects of the images, pairwise differences that arise in the error patterns of our location memory task between image types may suggest the mechanism by which individuals perform this segmentation. Because color negatives may affect both semantic and visual processing, some findings might be considered ambiguous when taken alone (such as a difference between color negatives and canonical images). However, three patterns of pairwise differences would lead to unambiguous conclusions. First, a difference between the error patterns of the upside down and the upright images would suggest a role for conceptual information in the segmentation process, because basic visual cues are retained in this manipulation. Second, a difference in the error distributions of the color negative and canonical images (but no difference between inverted and canonical) would suggest that low-level visual cues are likely involved in determining categories; this conclusion would be bolstered if there were also a difference in errors between the color negative and inverted images, because the local low-level visual cues are identical for canonical and inverted images. Finally, finding significant pairwise differences between the errors of all

Image Type	**Semantic Information**	**Low-Level Visual Cues**
Canonical	✓	✓
Negative	○	●
Inverted	●	✓

✓ Cue is unaffected by image manipulation
○ Effect on cue is uncertain
● Cue is affected

Figure 4. Schematic of the information retained and altered in the three image types of Experiment 2. Pairwise differences in error patterns between the conditions may indicate which cue(s) are important to forming spatial categories. A pairwise difference between lines 1 and 3 would indicate that semantic content is important in forming spatial categories. A difference between lines 1 and 2, especially with an additional difference between 2 and 3, would indicate that low-level visual cues are important. Pairwise differences between all three lines would indicate that both visual cues and semantic content are important to forming spatial categories.

three image types would imply that both conceptual and perceptual information were used to form spatial categories. It is possible that such a pattern would arise if the same process were disrupted by each manipulation but to different degrees. However, if the differences between color negatives and canonical images and between upside-down and canonical images are different in kind, then this would strongly suggest that different processes were disrupted by each manipulation.

Method

Participants. The participants were 25 undergraduate students at Temple University who participated for course credit. Three additional students participated in the location memory task, but these data were not included in the final analyses because these participants failed to attend to the task, using the criteria described in Experiment 1.

Materials. The materials for the location memory task were the same as in Experiment 1.

Stimuli. The stimuli were a subset of those used in Experiment 1. Twenty-five images with two locations per image were selected for use in Experiment 2. The locations that showed the greatest consistency in errors across all subjects were selected. These locations were evenly distributed in terms of their direction relative to the prototype. The images were then converted into photographic negatives, using a common photographic software package, or were rotated 180 degrees to create upside-down versions. Thus, there were three versions of each image location: canonical, color negative, and upside down. A sample image that has undergone these manipulations is shown in Figure 5.

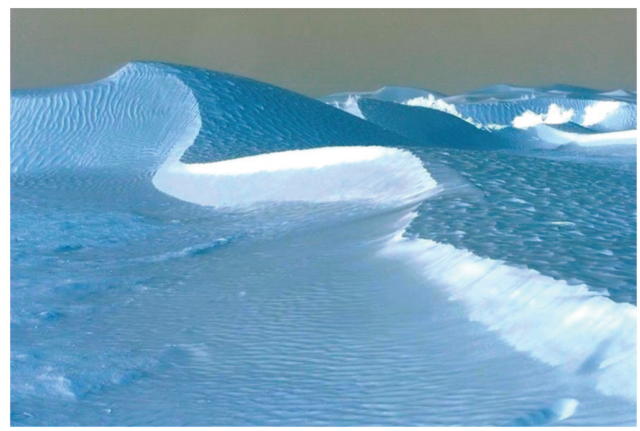
Locations were depicted in the same manner as in Experiment 1, except in the case of the color negative images. The color of the dots marking the to-be-recalled locations was not clearly visible in this condition, so the dots were recolored in these images to be a bright green. All other aspects of the dots (e.g., elliptical shape, fuzziness of boundaries) were kept identical.

Furthermore, the size of study and test images was the same as in Experiment 1. Also, the study and test images were again each randomly assigned to one of nine locations on the screen to prevent the use of cues such as distance to the edge of the computer screen. Again, participants were informed that there would be differences in cropping and position beforehand and were given two sample trials to demonstrate.

Procedure. The location memory experiment consisted of six blocks, with each of the 25 landscape images appearing in a random order once per block. There were two possible dot locations per landscape and three image types for each location. Each image type–dot location pair was randomly assigned to one of the six blocks for each subject. Trials were of the same format as in Experiment 1 and continued until all 150 image locations had been seen. If subjects were unsure of the location, they were instructed to press a key on the keyboard rather than make a response.

Scoring. As in Experiment 1, on some trials of the location memory task participants accidentally pressed before the mouse had moved from the previous response location. Again, to prevent such cases from affecting the mean error, we deleted responses whose error was so great that it seemed likely that the participant must have accidentally responded or momentarily failed to attend to the task. Specifically, responses more than 25% of the image length (225

(a) Canonical



(b) Color Negative

(c) Inverted

Figure 5. Sample images used in Experiment 2. The same locations were tested using (a) canonical, (b) color negative, and (c) inverted versions of the image. Inverting the images is assumed to reduce the influence of conceptual structure in forming spatial categories while retaining basic, low-level visual cues (e.g., color, luminance), whereas photographic negation affects perceptual cues (and potentially conceptual structure as well).

pixels, or 6.5 cm) from the correct location were removed. These responses accounted for 4.06% of the total data. In addition, after this culling the standard deviation of the distance of responses from the correct location was calculated for each location (Levine, 2004); responses greater than three standard deviations were deleted, eliminating an additional 0.08% of the data. The remaining responses were converted to vectors, as in Experiment 1.

Results

The goal of Experiment 2 was to compare the responses for each location across the image types to determine if the image manipulations affected participants' segmentation of scenes. As in Experiment 1, error vectors for each location were used to plot bias and to determine the average vector for each location. To determine if there is a difference in the error distributions across the image types, a common reference point for the category center is required. The clustering algorithm of Experiment 1 was found to identify the same categories regardless of the image manipulations, because it works on the basis of minimizing the difference between RGB values of pixels and cluster values. Therefore, we used these categories as the reference for our analyses. Thus, the difference between the angle of the average vector for each location and the angle of the vector predicted by the CA model was then calculated across all image types. The resultant distributions of differences between these angles can be examined visually (Figure 6, A–C). This figure implies that for the majority of locations shown and across the image types, participants appear to have erred in the direction of the category prototype. However, as in Experiment 1, it would be inappropriate to consider each of these angles, which represent a mean error for a given location, as a single value by which we can determine the overall mean error (Zar, 1998).

Therefore, as in Experiment 1, we performed Hotelling's one-sample second-order analysis for each image type (Zar, 1998). Figure 6 (D–F) shows the resultant second-order distributions for each image type. We found that subjects did err in a significantly directed fashion for the canonical images, $M = 359.93°$, 95% CI [351.60°, 9.33°], $F(2, 23) = 79.94$, $p < .01$, $d = 3.73$; for the color negatives, $M = 359.02°$, 95% CI [344.11°, 15.76°], $F(2, 23) = 23.03$, $p < .01$, $d = 2.00$; and for the upside-down images, $M = 8.59°$, 95% CI [354.41°, 26.06°], $F(2, 23) = 42.12$, $p < .01$, $d = 2.71$.

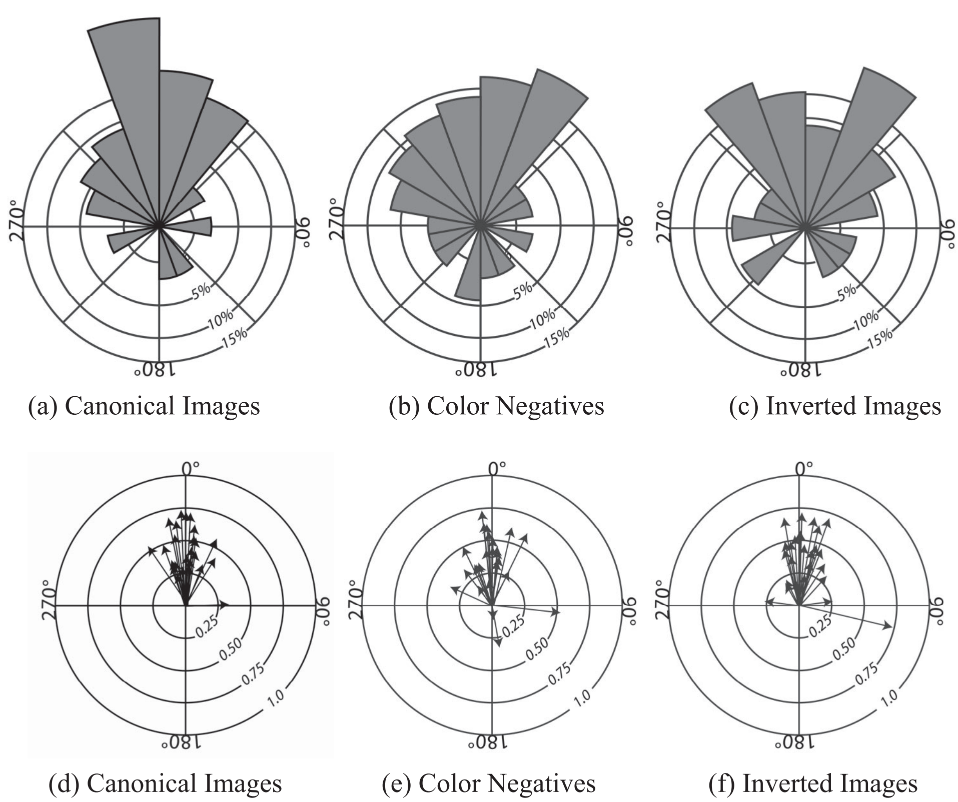


Figure 6. Difference of location memory errors from the predicted direction (toward category center) for Experiment 2. (a), (b), (c): Equal-area rose diagrams of the angular difference between the mean error vector for a location and the predicted direction of error for that location in (a) canonical, (b) color negative, and (c) inverted images. Angular difference values for all locations within an image type are represented. (d), (e), (f): Mean error vectors per subject in (d) canonical, (e) color negative, and (f) inverted images. The difference between each response and the predicted direction of error is recorded, and a mean error vector (with respect to the predicted direction) is calculated for each image type. Each vector therefore represents a single subject's tendency to err in the predicted direction or not for the given image type. Vector length varies inversely with angular dispersion. The predicted direction of error for all figures is 0°.

Of greatest interest, however, was whether there existed differences in the distributions of mean error vectors between the image conditions. If participants used different categories between the conditions, then the pattern of error vectors for one type of image (e.g., canonical) would significantly differ from that for another type of image (e.g., color negative). Hotelling's second-order test for paired data was used to examine this possibility. Again, this analysis subtracts paired vectors from each other and tests the resultant distribution of difference vectors against the null hypothesis of a random distribution. The angle of the mean of these difference vectors can also inform us whether the distributions differed in terms of overall direction, angular variance, or both; a difference in angular variance (i.e., the lengths of the vectors in the distributions) but not direction would result in a mean difference vector with an angle near 0° or 180° (depending on which vector set is subtracted from the other). Similarly, a mean difference vector with an angle of approximately 90° or 270° would indicate a difference in the overall direction of errors but no difference in variance. Finally, a mean angle between these values would suggest that both direction and variance differed.

Using Hotelling's second-order test for paired data, we found that a significant difference exists between the error distributions of the canonical images and the color negatives, $M = 182.22°$, 95% CI [136.07°, 227.81°], $F(2, 23) = 5.88$, $p < .01$, $d = 1.01$; and between the canonical images and the upside-down images, $M = 143.50°$, 95% CI [56.42°, 206.11°], $F(2, 23) = 3.78$, $p = .04$, $d = 0.81$. Finally, there also existed a significant difference in the distributions of errors between the color negative and inverted images, $M = 230.49°$, 95% CI [171.98°, 321.07°], $F(2, 23) = 4.45$, $p = .02$, $d = 0.88$.

The significant difference between the color negatives and the canonical images appears to be primarily due to a difference in the length of vectors (which is related to angular variance) in the two distributions, because the angle of the mean difference vector is near 180°. Recall from the results of Experiment 1 that a greater length of these second-order vectors does not imply that individuals made larger errors but rather that the errors were more consistently in a given direction—that of the predicted category prototypes, in this case. Therefore, there was greater within-subjects angular variance in participants' recall in color negatives compared with canonical images. That is to say, for color negative images, subjects frequently erred in the direction of the category prototype but with less reliability than for canonical images; in some cases, errors were directed in another direction suggesting the use of alternate categories.

On the other hand, the difference between the canonical and upside-down images appears to be related to both the length as well as the direction of the vectors of the two distributions, because the angle of the mean difference vector is between 90° and 180°. This suggests that individuals again used different categories for canonical and inverted images. It is important to note that the difference in mean error direction is not due to a simple leftward or rightward response bias (associated with using a mouse to make responses, for example), which would be expected to flip with inverting the image (e.g., Wedell et al., 2007), or to an artifact of the coding process, which makes certain assumptions about the category prototypes (i.e., that they are the center of mass of the region). As coded in our analyses, a positive mean error (e.g., 8.59° for inverted images) simply shows that errors were in a clockwise direction. Because the direction of correct locations to category prototypes was evenly distributed, this rules out a simple response bias, as a leftward bias would produce both clockwise and counterclockwise errors, depending on whether the correct location was above or below the category prototype, for example. Neither is the difference in mean error directions between conditions due to an artifact of defining the category prototype. For example, it has been pointed out that the center of mass of a triangle is below the midway point and that when this triangle is inverted, the center of mass is above the midway point; if individuals used the midpoint as the prototype (for both the upright and inverted versions), might this explain the difference in error patterns without assuming different category prototypes? Again, because our data are scored as clockwise and counterclockwise with respect to the predicted angle and because the inverted images are 180° rotations (not flipped versions) of the canonical images, such an explanation cannot account for our data; even if individuals used the midpoint in each case, the clockwise and counterclockwise errors would be scored as identical. After ruling out these alternative explanations, the significant difference in error patterns between canonical and inverted images suggests that individuals again used different categories across the two image types.

Finally, the significant difference in the distributions of errors between the color negative and inverted images also appears to be due to differences in both the length and the direction of the vectors of the two distributions, because the mean difference vector is between 180° and 270°. The previous paragraphs explain that the category prototypes used in inverted and in color negative images are different from those used in canonical images. The fact that the overall directions of the error distributions of the two altered image conditions differ is important because it implies that the category prototypes used in each of these conditions also differ; because the significant result is partially due to a difference in the overall directions of the distributions and not solely vector length (angular variance), this implies that the difference between manipulated images is not simply one of degree but also of kind. That is, the two image manipulations did not result in the use of the same category prototypes but in two different ones.

Discussion

The purpose of Experiment 2 was to explore the nature of segmentation into categories within complex scenes in a situation in which perceptual and conceptual influences could be distinguished. Subjects were asked to recall locations in canonical, inverted, and color negative images of landscapes. Previous work has suggested that inverting scenes interferes with the ability to extract meaning (e.g., Brockmole & Henderson, 2005; Intraub, 1984; Rock, 1974; Shore & Klein, 2000; Velisavljevic & Elder, 2008), whereas alterations involving unfamiliar color schemes disrupt certain basic visual cues (Vuong et al., 2005) and may also affect the ability to extract semantic information (Goffaux et al., 2005). Because the stimulus locations were identical across all versions of an image (canonical, inverted, or color negative), differences in location memory error patterns would indicate the information used in the process of category identification.

The results of Experiment 2 suggest that categorization uses both perceptual and conceptual information. Because inverted images alter the ease with which semantic or conceptual information is extracted, yet retain basic visual information, the finding that participants' errors significantly differed between the upright and inverted versions suggests that inversion disrupted some aspect of the process of category identification. In addition, the error patterns between color negative and canonical versions differed significantly; however, as alteration of familiar color schemes may affect the ability to extract semantic information in addition to altering perceptual features of the image, this finding taken alone is ambiguous. Nevertheless, because we also found a significant difference between the two altered image types—inverted and color negative—and because this difference was one of kind (direction of errors) not degree (variance), it seems that the two image types disrupt different aspects of the categorization process. For example, inversion may have led to categories that were more defined by perceptual boundaries that might not have been as rigidly used in the canonical images. On the other hand, color negation may have impaired the use of some of these perceptual cues, such as texture (Vuong et al., 2005), causing the increased variability of location memory errors. Explanations for our findings are considered further in the general discussion.

Some may interpret these findings as suggesting that the two image manipulations simply interfered with the same aspect of category identification, but to different degrees. However, if the two image manipulations disrupted the same process to different degrees, one would expect that the errors of both image types would differ from canonical in the same manner and that the two altered image types would differ from each other in terms of degree. But our results show that the two error distributions differ from the canonical error distribution in dissimilar manners (i.e., the inverted image errors differ from canonical in terms of variance and of direction, whereas the color negative errors differ from canonical only in terms of variance), suggesting that different aspects of category identification are disrupted by each process. That the error distributions of the altered images differ from each other in terms of direction furthers this argument, because this is a difference of kind, not degree. This strongly suggests that the two image manipulations in Experiment 2 affected different aspects of category identification rather than one aspect to different degrees. Therefore, category identification in our tasks makes use of at least two types of information: conceptual information (disrupted in the inverted images and possibly in color negative images) and perceptual information (disrupted in color negative images).

General Discussion

Numerous experiments have provided support for the CA model, suggesting that it accurately characterizes recall of spatial location in simplified spaces and tasks. However, such stimuli do not approach the complexity of the environments that individuals find themselves in on a daily basis. Experiment 1 represents an important step for the CA model in showing its usefulness in much more complex situations. The results of Experiment 1 show that individuals carve up the world into its constituent components and remember a location using the broad category that surrounds it as a frame of reference, as predicted by the CA model.

Experiment 2 provides further information about the nature of category identification in complex environments. Because the stimuli in Experiment 1 were imbued with both semantic and perceptual information that could define categories, it left unclear which type of information, or both, was important to segmenting the images into constituent categories. Therefore, in Experiment 2 we manipulated the images to reduce the influence of semantic information. Because upside-down images interfere with the extraction of semantic information while leaving basic visual cues (color, luminance, texture) intact, a difference in the error patterns between the upside-down and the canonical images suggested that semantic information is involved in the segmentation process for visually complex scenes. In addition, because the color negative manipulation affects basic visual cues and may also affect extraction of semantic information, and because upside-down images only affect the extraction of semantic information, a difference in the error patterns between the color negative and upside-down versions—given that both differed from the canonical images—suggests that perceptual information is also involved in the segmentation process for visually complex scenes. Taken together, the results of Experiment 2 argue that perceptual and conceptual information are used in the segmentation of visually complex scenes into categories.

Alternatively, one might suggest that the effect of inversion on category formation may stem from a decrement in the ability to process spatial relations within the scenes. There is an extensive body of literature linking the inversion cost to face recognition performance with a disruption to the processing of the relations between features within a face (e.g., Leder & Bruce, 1998, 2000; Leder, Candrian, Huber, & Bruce, 2001; Murray, Yong, & Rhodes, 2000; Rhodes, Brake, & Atkinson, 1993; Searcy & Bartlett, 1996; Tanaka & Sengco, 1997; Thompson, 1980). However, this classic effect (widely termed the "face inversion effect") is disproportionately larger for faces compared with other nonface objects and scenes (Diamond & Carey, 1986; Wright & Roberts, 1996; Yin, 1969). Notably, face perception is distinct from scene perception with respect to the importance of subtle differences in the relations between features. Specifically, Diamond and Carey (1986) noted that because all faces share the same first-order relations between features (e.g., two eyes aligned horizontally over a centrally aligned nose, which is over a mouth), this information cannot support the recognition of individual faces. Instead, face perception is sensitive to subtle differences in the spatial relations between the features forming the shared first-order configuration. In contrast, scenes do not share a common first-order relation between features (e.g., a boulder can be to the left or to the right of a tree, and a creek could run between them or around them). Thus, Diamond and Carey (1986) argued that because of the relatively unconstrained nature of the first-order relations within scenes, this information is typically sufficient for the perception and recognition of such stimuli. Because inversion has been shown to specifically disrupt the processing of second-order rather than first-order relations, it seems unlikely that the effect of inversion in Experiment 2 arises through a disruption to the processing of spatial relations.

Although the inversion manipulation may not have disrupted the processing of basic spatial relations within the scenes, there is evidence that it likely changed the manner in which the scenes were processed. Specifically, a recent study found reduced activation for inverted, compared to upright, scenes in the classically scene-selective brain region, commonly referred to as the parahippocampal place area (PPA; Epstein & Kanwisher, 1998; Epstein, Higgins, Parker, Aguirre, & Cooperman, 2006). Notably, this reduction in activation in the PPA with inversion was accompanied by a corresponding increase in activation in the lateral occipital complex, a region typically associated with more generic object processing mechanisms (Grill-Spector, Kourtzi, & Kanwisher, 2001; Malach et al., 1995). Thus, inversion appears to induce a shift from the recruitment of specialized scene processing mechanisms to more generic object processing mechanisms.

As described earlier, behavioral evidence suggests that one consequence of this apparent shift in processing—from specialized mechanisms to more generic ones—is a loss in the ability to access semantic information embedded in scenes (e.g., Brockmole & Henderson, 2005; Intraub, 1984; Rock, 1974; Shore & Klein, 2000; Velisavljevic & Elder, 2008). It is possible that the recruitment of generic perceptual processing mechanisms, rather than those specifically tailored for representing scenes, may weaken the representations of scenes such that it is more difficult to activate relevant stored semantic knowledge about scenes. This account is in accordance with distributed models of semantic memory that suggest that the same areas responsible for perceiving a stimulus, and thus active when the semantic information is acquired, are also involved in storing the semantic knowledge associated with that stimulus (Allport, 1985; Martin, 1998, 2007). Thus, because our semantic knowledge of scenes is likely accrued through our extensive everyday experience (with upright scenes), this information may be linked in brain networks to the specialized PPA-based representations of upright scenes. Thus, the somewhat unfamiliar generic representations of inverted scenes created by the lateral occipital complex may be less efficient at accessing semantic information, because they were unlikely involved to the same extent during the acquisition of this knowledge.

One could argue that it is not surprising and possibly even expected that our spatial memory system relies, at least in part, on semantic information and other information accrued through our experience with the world. The advantage of any Bayesian system lies in its ability to adapt estimates as more information becomes available. In terms of such input, a system in which categories are rigidly defined without regard to semantic information would be severely limited at best. However, DFT, as outlined by Spencer et al. (2007), does not explicitly account for such information. That is, Spencer et al. state that geometric category bias—involving the use of edges and symmetry axes to subdivide space into what Huttenlocher et al. (1991) call categories—are always tied to perceived reference frames. Because the upside-down images in Experiment 2 were visually identical to the canonical images, the fact that the error patterns differed between these two types of images is hard to account for with perceived reference frames alone. Rather, the results of Experiment 2 indicate that individuals remember locations by using an allocentric frame of reference grounded in both the perceptual cues and the semantic content of the task space. Thus, it seems that Spencer et al.'s DFT account of spatial memory biases needs to address the use of conceptual information. Further contrasts between the DFT and CA model will need further empirical work to adjudicate.

It is also worth noting that, as with the canonical images, these frames of reference—perceptually and conceptually defined—frequently converge on the same category boundaries. However, as outlined previously, there are instances in which they might disagree (e.g., the base of a mountain, or the difference between a crack in a rock and a fault). In these cases, segmenting the world into categorical frames of reference by using either perceptually or conceptually defined cues might lead to very different results. However, it is unclear as to how these different patterns would be used. One possibility is that individuals may effectively choose to use one or the other sets of categories; if we assume that the goal of combining categorical and metric information is to minimize errors, then having multiple potential categories from which to choose could furnish the individual with options that might further minimize errors (e.g., by reducing category size or by having more precise boundaries). In cases of disagreement, then, one might predict that the frame of reference that reduces error to the largest degree would be used to define the category. A second possibility is that the multiple segmentation patterns are combined so that the size of the frame of reference is minimized. This case, with categories akin to the overlapping region of a Venn diagram, would make use of all available information to minimize errors. In such instances, it might also be predicted that the relative weighting applied to each segmentation pattern would be proportional to the reliability of that pattern. Future work will begin to explore these and other issues by comparing the location memory errors of novice and expert geoscientists, who likely differ in their conceptual understanding of geological formations.

In summary, Experiment 1 addressed a significant gap in the literature by demonstrating a bias in location memory errors within visually complex scenes. This is the first evidence to date that the same combinatory process that leads to bias in location memory errors within blank geometric shapes is used in more visually complex scenes. Using this new paradigm, in Experiment 2 we demonstrate that the information that people use to segment images into categories includes both low-level visual cues and semantic information. That semantic information is used strongly suggests that the process of categorization depends, in part, on the conceptual information brought to bear on the task by the individual.

References

Abramoff, M. D., Magelhaes, P. J., & Ram, S. J. (2004). Image processing with ImageJ. *Biophotonics International, 11,* 36–42.

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology, 14,* 257–262.

Allport, D. A. (1985). Distributed memory, modular subsystems and dysphasia. In S. K. Newman & R. Epstein (Eds.), *Current perspectives in dysphasia* (pp. 32–60). Edinburgh, Scotland: Churchill Livingstone.

Baciu, M., Koenig, O., Vernier, M. P., Bedoin, N., Rubin, C., & Segebarth, C. (1999). Categorical and coordinate spatial relations: fMRI evidence for hemispheric specialization. *NeuroReport, 10,* 1373–1378.

Brockmole, J. R., & Henderson, J. M. (2005). Using real-world scenes as contextual cues for search. *Visual Cognition, 13,* 99–108.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers, 25,* 257–271.

Corneille, O., Huart, J., Becquart, E., & Bredart, S. (2004). When memory shifts toward more typical category exemplars: Accentuation effects in the recollection of ethnically ambiguous faces. *Journal of Personality and Social Psychology, 86,* 236–250.

Deneve, S., & Pouget, A. (2004). Bayesian multisensory integration and cross-modal spatial links. *Journal of Physiology–Paris, 98,* 249–258.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General, 115,* 107–117.

Duffy, S., Huttenlocher, J., & Crawford, L. E. (2006). Children use categories to maximize accuracy in estimation. *Developmental Science, 9,* 597–603.

Epstein, R., & Kanwisher, N. (1998, April 9). A cortical representation of the local visual environment. *Nature, 392,* 598–601.

Epstein, R. A., Higgins, J. S., Parker, W., Aguirre, G. K., & Cooperman, S. (2006). Cortical correlates of face and scene inversion: A comparison. *Neuropsychologia, 44,* 1145–1158.

Ernst, M. O., & Banks, M. S. (2002, January 24). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415,* 429–433.

Fitting, S., Wedell, D. H., & Allen, G. L. (2007). Memory for spatial location: Cue effects as a function of field rotation. *Memory & Cognition, 35,* 1641–1658.

Goffaux, V., Jacques, C., Mouraux, A., Oliva, A., Schyns, P. G., & Rossion, B. (2005). Diagnostic colours contribute to the early stages of scene categorization: Behavioural and neuropsychological evidence. *Visual Cognition, 12,* 878–892.

Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research, 41,* 1409–1422.

Huart, J., Corneille, O., & Becquart, E. (2005). Face-based categorization, context-based categorization, and distortions in the recollections of gender ambiguous faces. *Journal of Experimental Social Psychology, 41,* 598–608.

Hund, A. M., & Plumert, J. M. (2002). Delay-induced bias in children's memory for location. *Child Development, 73,* 829–840.

Hund, A. M., & Plumert, J. M. (2003). Does information about what things are influence children's memory for where things are? *Developmental Psychology, 39,* 939–948.

Hund, A. M., & Plumert, J. M. (2005). The stability and flexibility of spatial categories. *Cognitive Psychology, 50,* 1–44.

Huttenlocher, J., Hedges, L. V., & Bradburn, N. M. (1990). Reports of elapsed time: Bounding and rounding processes in estimation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 196–213.

Huttenlocher, J., Hedges, L. V., & Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review, 98,* 352–376.

Huttenlocher, J., Hedges, L. V., & Prohaska, V. (1988). Hierarchical organization in ordered domains: Estimating the dates of events. *Psychological Review, 95,* 471–484.

Huttenlocher, J., Hedges, L. V., & Vevea, J. L. (2000). Why do categories affect stimulus judgment? *Journal of Experimental Psychology: General, 129,* 220–241.

Huttenlocher, J., Newcombe, N., & Sandberg, E. H. (1994). The coding of spatial location in young children. *Cognitive Psychology, 27,* 115–147.

Intraub, H. (1984). Conceptual masking: The effects of subsequent visual events on memory for pictures. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 115–125.

Johnson, J. S., Spencer, J. P., & Schöner, G. (2008). Moving to higher ground: The dynamic field theory and the dynamics of visual cognition. *New Ideas in Psychology, 26,* 227–251.

Kosslyn, S. M. (1987). Seeing and imagining in the cerebral hemispheres: A computational approach. *Psychological Review, 94,* 148–175.

Kosslyn, S. M., Thompson, W. L., Gitelman, D. R., & Alpert, N. M. (1998). Neural systems that encode categorical versus coordinate spatial relations: PET investigations. *Psychobiology, 26,* 333–347.

Laeng, B. (1994). Lateralization of categorical and coordinate spatial functions: A study of unilateral stroke patients. *Journal of Cognitive Neuroscience, 6,* 189–203.

Leder, H., & Bruce, V. (1998). Local and relational aspects of face distinctiveness. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 51*(A), 449–473.

Leder, H., & Bruce, V. (2000). When inverted faces are recognized: The role of configural information in face recognition. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 53*(A), 513–536.

Leder, H., Candrian, G., Huber, O., & Bruce, V. (2001). Configural features in the context of upright and inverted faces. *Perception, 30,* 73–83.

Lee, P. J., & Brown, N. R. (2003). Delay related changes in personal memories for September 11, 2001. *Applied Cognitive Psychology, 17,* 1007–1015.

Levine, N. (2004). Crimestat III: A spatial statistics program for the analysis of crime incident locations (Version 3.0) [Computer software]. Houston, TX: Levine & Associates.

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., . . . Tootell, R. B. (1995). Object-related activity revealed by functional magnetic-resonance-imaging in human occipital cortex. *Proceedings of the National Academy of Sciences, USA, 92,* 8135–8139.

Martin, A. (1998). Organization of semantic knowledge and the origin of words in the brain. In N. G. Jablonski & L. C. Aiello (Eds.), *The origin and diversification of language* (pp. 69–98). San Francisco, CA: California Academy of Sciences.

Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology, 58,* 25–45.

Murray, J. E., Yong, E., & Rhodes, G. (2000). Revisiting the perception of upside-down faces. *Psychological Science, 11,* 492–496.

Nelson, T. O., & Chaiklin, S. (1980). Immediate memory for spatial location. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 529–545.

Newcombe, N. S., & Huttenlocher, J. (2000). *Making space: The development of spatial representation and spatial reasoning.* Cambridge, MA: MIT Press.

Plumert, J. M., & Hund, A. M. (2001). The development of memory for location: What role do spatial prototypes play? *Child Development, 72,* 370–384.

Postma, A., Kessels, R. P. C., & van Asselen, M. (2008). How the brain remembers and forgets where things are: The neurocognition of object-location memory. *Neuroscience and Biobehavioral Reviews, 32,* 1339–1345.

Recker, K. M., Plumert, J. M., Hund, A. M., & Reimer, R. (2007). How do biases in spatial memory change as children and adults are learning locations? *Journal of Experimental Child Psychology, 98,* 217–232.

Rhodes, G., Brake, S., & Atkinson, A. P. (1993). What's lost in inverted faces? *Cognition, 47,* 25–57.

Rock, I. (1974). The perception of disoriented figures. *Scientific American, 230,* 78–85.

Schutte, A. R., & Spencer, J. P. (2002). Generalizing the dynamic field theory of the A-not-B error beyond infancy: Three-year-olds' delay- and experience-dependent location memory biases. *Child Development, 73,* 377–404.

Searcy, J. H., & Bartlett, J. C. (1996). Inversion and processing of component and spatial-relational information in faces. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 904–915.

Shore, D. I., & Klein, R. M. (2000). The effects of scene inversion on change blindness. *Journal of General Psychology, 127,* 27–43.

Simmering, V. R., & Spencer, J. P. (2007). Carving up space at imaginary joints: Can people mentally impose arbitrary spatial category boundaries? *Journal of Experimental Psychology: Human Perception and Performance, 33,* 871–894.

Slotnick, S. D., & Moo, L. R. (2006). Prefrontal cortex hemispheric specialization for categorical and coordinate visual spatial memory. *Neuropsychologia, 44,* 1560–1568.

Spencer, J. P., & Hund, A. M. (2002). Prototypes and particulars: Geometric and experience-dependent spatial categories. *Journal of Experimental Psychology: General, 131,* 16–37.

Spencer, J. P., Simmering, V. R., & Schutte, A. R. (2006). Toward a formal theory of flexible behavior: Geometric category biases generalize across pointing and verbal response types. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 473–490.

Spencer, J. P., Simmering, V. R., Schutte, A. R., & Schöner, G. (2007). What does theoretical neuroscience have to offer the study of behavioral development? Insights from a dynamic field theory of spatial cognition. In J. Plumert & J. P. Spencer (Eds.), *The emerging spatial mind* (pp. 320–361). New York, NY: Oxford University Press.

Stevens, A., & Coupe, P. (1978). Distortions in judged spatial relations. *Cognitive Psychology, 10,* 422–437.

Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory & Cognition, 25,* 583–592.

Thompson, P. (1980). Margaret Thatcher: A new illusion. (1980). *Perception, 9,* 483–484.

Trojano, L., Conson, M., Maffei, R., & Grossi, D. (2006). Categorical and coordinate spatial processing in the imagery domain investigated by rTMS. *Neuropsychologia, 44,* 1569–1574.

Tversky, B. (1981). Distortions in memory for maps. *Cognitive Psychology, 13,* 407–433.

Velisavljevic, L., & Elder, J. H. (2008). Visual short-term memory of local information in briefly viewed natural scenes: Configural and non-configural factors. *Journal of Vision, 8,* 1–17.

Vuong, Q. C., Peissig, J. J., Harrison, M. C., & Tarr, M. J. (2005). The role of surface pigmentation for recognition revealed by contrast reversal in faces and greebles. *Vision Research, 45,* 1213–1223.

Wedell, D. H., Fitting, S., & Allen, G. L. (2007). Shape effects on memory for location. *Psychonomic Bulletin & Review, 14,* 681–686.

White, M. (2001). Effect of photographic negation on matching the expressions and identities of faces. *Perception, 30,* 969–981.

Wright, A. A., & Roberts, W. A. (1996). Monkey and human face perception: Inversion effects for human faces but not for monkey faces or scenes. *Journal of Cognitive Neuroscience, 8,* 278–290.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81,* 141–145.

Zar, J. H. (1998). *Biostatistical analysis* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Received June 19, 2009
Revision received February 17, 2010
Accepted February 19, 2010 ■