CrossMark

Bayesian average or truncation at boundaries? The mechanisms underlying categorical bias in spatial memory

Cristina Sampaio[1] · Ranxiao Frances Wang[2]

Abstract

Spatial memory is often biased by various factors, such as the region a target belongs to, which can be defined based on physical, perceptual, or implicit boundaries. In the typical dot-localization task first introduced by Huttenlocher, Hedges, and Duncan (Psychological Review 98: 352-376, 1991), individuals normally divide the task space into four quadrants delineated at the Cartesian axes (forming "default categories") and show systematic bias in target localization toward the center of the category. At least two mechanisms have been proposed to account for these categorical biases, namely (a) weighted-average of a metric representation and the category prototype representation and (b) truncation of an un-biased metric representation at the category boundary. Both models can account for these findings and cannot be differentiated by existing research methods. Using a new distribution analysis, the current study sought to differentiate between these two models. Participants viewed a dot inside a circle and recalled its location after a delay either with the same blank circle (i.e., the standard dot-in-circle paradigm) or when an alternative V-shaped category boundary was visually presented at retrieval. The data from three experiments showed symmetrical distribution of the errors that shifted toward the category center when people primarily used the default category, supporting the weighted-average model. In contrast, when people primarily used the alternative category, the errors showed a highly skewed distribution, more consistent with the truncation model. Overall, these results provided the first experimental evidence for both mechanisms separately.

Keywords Spatial memory · Category-adjustment model · Category bias · Spatial memory bias · Spatial cognition

Introduction

Spatial memory is an important cognitive capability for most animal species, and a vast literature has been devoted to studying the encoding and processing of spatial information across many areas, including animal behavior, human development, neural disorders, and brain functions. Within the spatial cognition research, the format of spatial coding is one of the central theoretical issues for understanding spatial memory in both humans and other animals (e.g., Shelton & McNamara, 2001; Wang, 2012), and many research paradigms have been developed. For example, categorical bias is a well-known phenomenon that reveals the underlying structure of spatial representations and spatial coding (Huttenlocher, Hedges, & Duncan, 1991; McNamara, 1986). Previous research has shown that when people remember a target location, they tend to show a systematic bias according to the target's region (Huttenlocher, Hedges, & Duncan, 1991; McNamara, Hardy, & Hirtle, 1989; Steven & Coupe, 1978). For example, people often mistakenly judge Los Angeles to be to the west of Reno, supposedly because they use the relative location of California and Nevada in their assessment of the spatial relationship between the cities within the respective states. These findings suggest that region membership is an integral component of the location memory representation, which can be defined based on the physical, perceptual, or implicit boundaries of the region.

A common paradigm to study this type of category bias in spatial memory is the dot localization task, in which individuals see a dot inside a simple geometric space (e.g., a circle) and replace the dot in that space after a brief delay. In this standard task, people implicitly impose Cartesian axes as the boundaries of the regions to categorize the circular space. That

✉ Cristina Sampaio
   cristina.sampaio@wwu.edu

[1] Department of Psychology, Western Washington University, Bellingham, WA, USA

[2] Department of Psychology, University of Illinois at Urbana-Champaign, Urbana-Champaign, IL, USA

is, the data indicate that individuals partition these simple spaces into default categories bounded at the x- and y-axis and misplace the dot in the direction of the center of the quadrant in which the dot had appeared (i.e., its L-shaped region or spatial category). The phenomenon of spatial bias is prevalent across a variety of other paradigms and also occurs with orientation (Stevens & Coupe, 1978) and distance judgments (McNamara, 1986) as well as many types of spaces, such as world maps (Friedman & Brown, 2000), pictures of nature (Holden, Curby, Newcome, & Shipley, 2010), natural environments (Sampaio & Cardwell, 2012), and virtual environments (Sampaio et al., 2017). Together, the literature suggests that metric information is calibrated within a frame of reference, and this frame can bias spatial judgments from memory.

The well-known pattern of misjudging a target's location in relation to its spatial region is frequently viewed as evidence for the metric and categorical sources of information blending in a Bayesian manner. However, the mechanism that leads to the bias is not completely understood. The reliability of cues is key in determining the relative weight of each cue in estimates of location from memory. As certainty of a cue increases, the influence of the others in location judgments decreases (Engebretson & Huttenlocher, 1996; Huttenlocher et al., 1991; Newcombe et al., 1999). Engebretson and Hunttenlocher (1996), for instance, manipulated the certainty of the metric information and found that as it degraded, people relied more on the categorical information. Huttenlocher et al. (1991) decreased the precision of the metric information by introducing a distractor task and also found that, in that case, categorical information (in the form of a central value or prototype) was weighed more heavily in people's estimates of location compared to the standard condition. Moreover, the amount of bias depends on the degree of inexactness of category values (i.e., prototypes and boundaries). Generally, there is little bias in remembering targets that appear near the category prototypes and/or boundaries (Huttenlocher et al., 1991). The imposed axes of symmetry in circular spaces are invisible and inexact, which influences the probability of correctly assigning items to their respective categories. The size of the category effect also increases with longer delays (Hund & Plumert, 2002; Huttenlocher et al., 1991; Spencer & Hund, 2002, 2003). Sampaio and Wang (2012), for example, asked participants to reproduce the location of target dots within a circle after a short (300 ms) or a long (5,000 ms) delay, and showed larger biases in target localization in the long delay condition. The hypothesis is that the metric information decays more rapidly over time compared with the categorical representation. Thus, there is a greater reliance on the categorical coding than on the metric coding with the passage of time.

In this paper, we investigated the mechanisms underlying the bias. The dominant account for the use of categories in spatial memory is that a weighting process based on Bayesian principles occurs, yielding a response that combines the multiple sources of information available at the time of estimation. Specifically, the location of a target is coded at two levels of detail: a coarse level (the category representation) and a fine level (the metric representation). Huttenlocher et al. (1991) proposed that both codings are represented as a distribution of values, with the category coding centered at the category mean and the metric coding centered at the actual location of the target. These representations are associated with some level of uncertainty and may be modeled as probability distributions centered at the corresponding mean with the corresponding variance. The retrieval of the target location during the responding stage involves obtaining a random sample from each distribution, and the two samples are averaged with weights determined by the relative variance of the two representations.

Huttenlocher et al. also proposed that recalled locations may reflect a bias toward the center of the target's region because the distribution is truncated by the category boundaries. According to this hypothesis, a target's location is also represented at two levels of detail: a coarse level (category representation) and a fine level (metric representation). The metric representation is the same as that in the weighted-average model. However, unlike the weighted-average model, the category representation consists of the boundaries of the category, instead of the prototype. At the time of retrieval, a random sample is obtained from the metric representation, but the selection is restricted to the range that resides within the boundary of the category representation. In other words, the distribution of the metric representation is "truncated" at the category boundaries, which results in systematic biases in the retrieved target locations toward the center of the category.

Both hypotheses make the same predictions on the traditional measure of mean bias, and there has been no experimental test that could differentiate them in the literature. In this paper we sought to differentiate among these accounts for category-biased responses based on a new distribution analysis extended from the analysis recently developed by Sampaio and Wang (2017). Although the mechanisms in the two proposed accounts predict the same mean bias, each mechanism produces a different *shape* in error distributions and thus an analysis of distribution *shape* can potentially discriminate these mechanisms. Specifically, a weighted average of the category and the metric location information should produce an error distribution with its center at a category-biased value, and the distribution's shape should remain symmetrical. On the other hand, truncation at boundaries such that the borders of a particular category restrict the placement of recalled locations should produce a distorted distribution with a skewed tail away from the boundaries and toward the mean of the category used. Moreover, the peak of the distribution should remain at the true location of the target and unbiased (Sampaio & Wang, 2017). Therefore, in the current work, we focused on the *shape/skewness* of error distributions along with the position of the peak.

Our aim for this research was to first examine the mechanism of category use in a standard paradigm with a blank circle by re-analyzing the data of an experiment by Sampaio and Wang (2017), which addressed the cause of the category bias using the classic dot-in-circle paradigm. In this case, data indicate that people impose invisible axes of symmetry onto the circular space and carve up categories bounded at these axes (upper-left, upper-right, lower-left, and lower-right). We applied our new analysis of distribution *shape*, and then we used the same analysis to investigate the mechanisms underlying the bias when a set of alternative category boundaries was presented to participants during retrieval (Experiments 2 and 3).

Experiment 1

In this study, we re-analyzed the data in Sampaio and Wang (2017) using a new extension of our distribution analysis to examine the shape of the error distribution using the classic dot localization paradigm (with a blank circle). If participants primarily used the weighted-average mechanism, then their error distribution should be symmetrical with a biased peak position. In contrast, if they primarily used the truncation mechanism, then their error distribution should show significant skew but little or no shift in the peak position.

Method

Materials and procedure

In their experiment, 32 participants from Western Washington University took part in a version of the basic dot-in-circle task developed by Huttenlocher et al. (1991). There was a total of 160 trials, and in each trial participants saw a circular space for 2,000 ms in the center of a computer screen and then with a small target dot for 250 ms in the space. After a blank white screen appeared for 3,000 ms, the circle reappeared and participants were asked to locate the dot from memory (Fig. 1a shows a schematic procedure) by clicking the mouse on the position they thought that the dot had been shown. To prevent individuals from utilizing the cursor as a placeholder for the target, participants had to move the cursor to the top of the screen, above the circle, before responding. Targets randomly appeared within the 5–15° range from the default category boundaries (the Cartesian axes). The decision to use angles at least 5° away from the boundaries was due to the fact that targets presented at boundaries usually exhibit no bias. The radii were also randomly selected from 1.0 to 8.2 cm from the center of the circle.
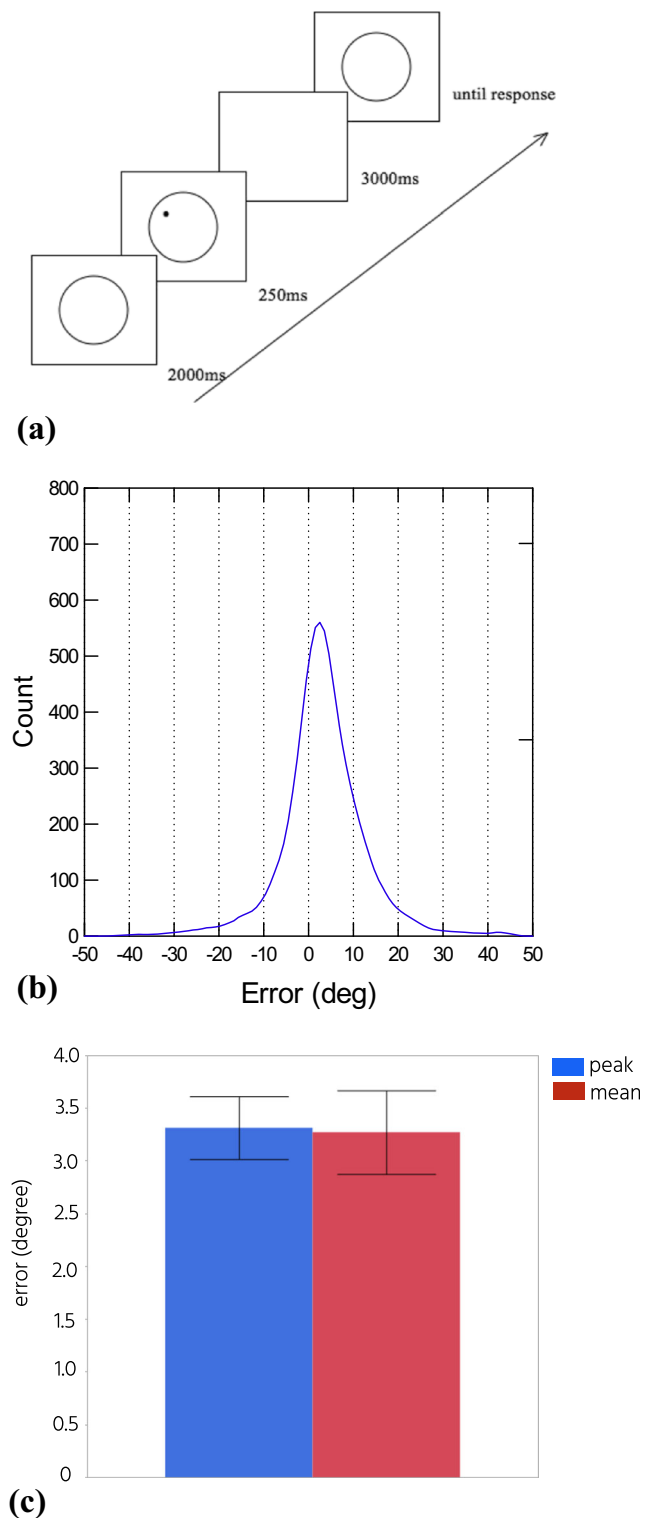

Fig. 1 Procedure and results for Experiment 1 (reanalysis of Sampaio & Wang, 2017). Panel (**a**) shows the schematic procedure of one trial. Panel (**b**) shows the kernel curve for the overall combined error distribution. Panel (**c**) shows the comparison of the mean error and peak of the error distribution

Results and discussion

Sampaio and Wang (2017) used standard exclusion criteria, which resulted in 6.3% of the trials being removed from the analysis; the placements/recalled locations in these trials were more than 45° away from the true target location. Moreover, six individuals did not follow instructions and their data were excluded from the analysis. In the first step, angular errors were calculated as the difference between the recalled target angle and the actual angle. Signs were then recoded in order to make all errors toward the categorical mean (i.e., the diagonal axes) positive for all targets and all errors away from the categorical mean (i.e., toward the vertical or horizontal axes) negative for all targets. The analysis of the mean errors revealed the typical categorical bias effect, with systematic bias toward the center of the category (mean bias = 3.27°, one-sample, two-tailed t-test against zero, $t(25) = 8.24$, $p < .001$).

To test whether people used a weighted-average mechanism or a truncation mechanism to estimate the target locations, we further examined the data in Sampaio and Wang (2017) on the *shape* of the error distribution by extending the distribution analysis used in the original study. First, a distribution analysis was used to obtain the peak value of the distribution by fitting the angular errors with a Kernel curve for each participant individually. As Sampaio and Wang (2017) explained, "Kernel density estimation is a nonparametric method to estimate the probability density function of a random variable with no a priori assumption about the nature of the density function such as normality or symmetry. The density estimation uses a kernel function to weight the data at each point, which decreases as the distance from the point increases." (p. 1990). Figure 1b shows the Kernel curve for all participants combined.

Next, we examined the skew of the distribution by comparing its mean and peak. We reasoned that if the distribution were symmetrical, then the mean would be the same as the peak; in contrast, if the distribution were skewed toward one side, then the mean would be larger (positive skew) or smaller (negative skew) than the peak value. We found that the mean ($M = 3.3$, $SD = 2.0$) and peak ($M = 3.4$, $SD = 2.3$) are not statistically different (paired $t(25) = .39$, $p = .70$). These results are consistent with the hypothesis that weighted-average of the codings is the mechanism responsible for the bias, at least in the standard paradigm with a blank circle (Fig. 1c).

The analysis of the shape of the distribution thus suggests that the weighted average mechanism underlies the categorical bias in the standard paradigm. In Experiment 2, we performed the same analysis of the shape of the error distribution to investigate which mechanism is responsible for the bias when a salient alternative category containing two boundaries of a V shape was provided during retrieval. Both the weighted average and the truncation mechanism are plausible. People may use the same mechanism regardless of which or how many category sources are available; that is, people may use a Bayesian combination of all the categorical and metric codings to generate an average that reflects the reliability of the various sources. Alternatively, people may instead use different mechanisms depending on the nature of the categories being used (e.g., default or alternative categories). Our aim was to use this new analysis of distribution shape to differentiate possible mechanisms underlying the category bias under different conditions.

We note that these possibilities should not be viewed as necessarily competing hypotheses, and different individuals may use different mechanisms. For example, Crawford, Landy, and Salthouse (2016) reanalyzed a previously published data set and found variability on the strategies that individuals use to remember locations within a space. Although they did not use the dot localization task first reported by Huttenlocher et al. (1991), their point that data aggregation may be mis-informative is highly relevant to the present study. Specifically, when multiple categories are available (e.g., the implicit default category bounded on the Cartesian axes and the visible category bounded on the diagonals of the circle), different individuals may use different categories in their judgments. For example, some people may predominantly use the default category and therefore should show a systematic bias toward the default category mean, while other people may primarily use the alternative category and therefore show a systematic bias toward the alternative category mean. Because these groups use different categories, they may exhibit different interaction mechanisms, therefore in Experiment 2 we also examined the mechanism according to the primary strategy people used.[1]

Experiment 2

In this experiment, we tested the prevailing view to explain category biases when markings of the boundaries of an alternative category are presented. The markings reflect the carving of an alternative categorization to the default axes of symmetry typically found with circular spaces. In fact, in the dot-in-circle paradigm used in our experiments, the robust default Cartesian categorization has resisted various manipulations. In a multi-experiment paper, for example, Huttenlocher, Hedges, and Crawford (2004) attempted to induce an alternative categorization scheme in the circular space by presenting uneven distributions of targets. In their four experiments, targets were clustered along the vertical and horizontal axes, thus forming categories bounded at the diagonals of the circle with central values at the vertical and horizontal axes. Their prediction was

---

[1] We define "primary strategy" as the strategy used to any degree of preference over the other. Thus, unless the weighting is absolutely even, people will be grouped according to the category they prefer (even if only slightly).

that if people categorized the stimuli according to their distribution in the circle, biases locating the stimuli towards the vertical and horizontal axes should result. However, if people continued to use the default geometric categories bounded at the horizontal and vertical axes, then the same pattern of bias should result regardless of the distribution (that is, biases towards the diagonals of the circle). They found that people used the same spatial categories regardless of the distribution of locations, and they argued that the spatial organization used maximizes accuracy of estimates because the exact category boundaries minimize misclassification of stimuli. Even when participants made explicit classification judgments according to the alternative "X" categorization scheme (thus, a dot would be classified as belonging to one of four V-shaped regions rather than the L-shaped regions), they still continued to use the default L-shaped regions in their estimates of location.

A few years later, Sampaio and Wang (2010, 2012) and Crawford and Jones (2011) reported that in specific circumstances, individuals can flexibly use an alternative categorization scheme in a circle: (a) when clear, reliable visual borders of an alternative category are provided during retrieval, and (b) when identifying membership information based on unique target features (i.e., semantic information) is available during encoding. These findings are consistent with research in the developmental literature showing that visible boundary cues can be used to form spatial categories in space (e.g., Hund, Plumert, & Benney, 2002). The data are moreover consistent with a more recent view that spatial categories can indeed be flexible and created by the individual to perform a given task (Hund & Plumert, 2005). Hund and Plumert suggest that people combine remembered information in memory, perceptually available information, and task goals to carve up the space.

In Experiment 2, we used the shape of distribution analysis to differentiate among possible mechanisms by which these sources of categorical information (in this case, the default and alternative) are integrated with metric information to form biased responses. To account for potential multiple mechanisms, we conducted a distribution analyses for each participant individually. Next, we divided the participants into two groups based on whether they showed systematic bias toward the default L-category or toward the alternative V-category and examined the shape (skewness) of the distribution for the two groups separately to assess the underlying mechanism in each case.

Method

Participants

The participants were 30 undergraduate students, 80% female, from Western Washington University, who carried out the experiment to partially fulfill a course requirement.

Materials and procedure

We adopted the stimuli from Sampaio and Wang's (2010) Experiment 2 (see Fig. 2). The procedure was similar to the typical dot localization task, in which subjects briefly see a target within a circular space and reproduce its location from memory. The only variation from the standard task was that a "V" marking the boundaries of an alternative category for each target was presented at retrieval. That is, the "V" was 90° in size and rotated 45° from the Cartesian axes.

There were 160 trials in the experiment. The circle in this experiment measured 18.5 cm in diameter. The targets were blue squares randomly placed inside the circle, between the range of 5–15° away from the default category boundaries (i.e., the vertical and horizontal axes) or the range of 5–15° away from the alternative category boundaries (i.e., the diagonals). The targets were placed at a random radius, varying between 1.0 and 8.2 cm.

In each trial, subjects viewed a blank circle at the center of the computer screen for 1,000 ms, and then they viewed a small square target within the circle for 250 ms. Both the circle and the square disappeared, and the screen was blank for 3 s. The subjects were then presented with a circle containing a "V" inside (two diagonal radius lines) to mark the boundaries of an alternative spatial category where the square had appeared. In each trial, the V frame appeared in one of the four segments of the canonical axes probed according to the location of the target, thus including sideways and inverted Vs in addition to the upside V shown in Fig. 2. The alternative boundary cue was 100% valid, and subjects were told that the region in which the target had appeared would be marked to aid their localization of the target. Subjects responded by moving the mouse cursor to the desired location and clicking there.

Results and discussion

Responses that were more than 45° away from the actual target location were excluded from the analysis (2.9% of responses). Three participants were excluded from the analysis due to failure to follow instructions. We defined angular errors as the difference between the reported and the actual target angle, following the standard convention. Bias caused by the use of the default L-shaped categories is in the reversed direction of those caused by the use of the alternative V-shaped categories. For simplicity, we changed the sign of the bias for all angles between the ranges 5–15° and 30–40° in quadrant I, and the corresponding ranges in all quadrants. With this coding system, the sign of the bias indicated which category was applied in estimation: default bias appeared positive and alternative bias appeared negative for all targets.

We also combined target angles into two bins: targets near the orthogonal axes were placed in bin 1 (between the ranges 5–15° and 75–85° in quadrant I, and corresponding ranges in
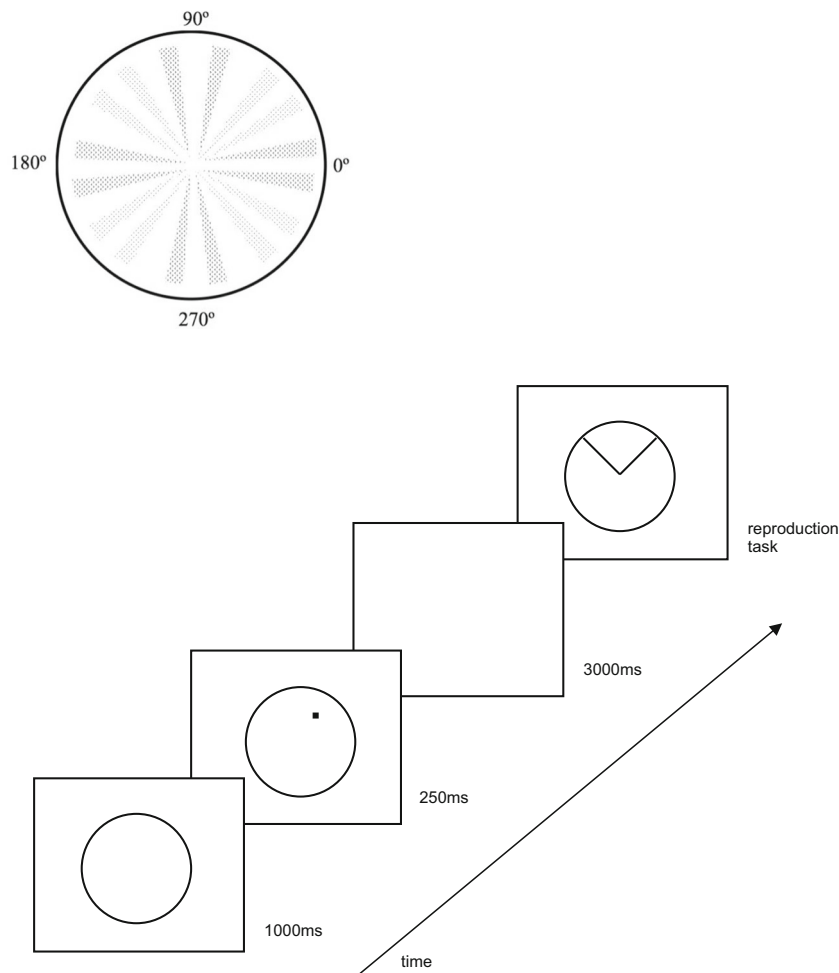
Fig. 2 Schematic distribution of targets in bin 1 (darker shading) and bin 2 (lighter shading) and experimental procedure for Experiment 2

the other quadrants) and targets near the diagonal axes were placed in bin 2 (between the ranges 30–40° and 50–60°, and corresponding ranges in the other quadrants) (Fig. 1a). The reason for the bins is that the effect of a category is most strongly observed away from its center, hence the usage of alternative categories will mostly noticeably surface in bin 2 (near-diagonal targets) while the usage of the default category will most noticeably surface in bin 1 (near-orthogonal targets). Because our primary interest was to observe the effects of alternative categories, 90% of the trials had near-diagonal targets; the 10% filler trials were in the near-orthogonal bin and were excluded from the analyses.

Our results showed an overall bias toward the center of the alternative category, with a negative mean bias (M = -1.15, SD = 2.5, one-sample, two-tailed t-test against zero, $t(26)$ =2.3, p = .028). These results showed that overall people used the alternative category to determine target positions. Moreover, we hypothesized that if the distribution is symmetrical, then the mean should be the same as the peak. If the distribution is skewed toward one side, then the mean should be larger (positive skew) or smaller (negative skew) than the peak value. Overall, we found an unbiased peak (M

= .51, SD = 2.4, one-sample, two-tailed t-test against zero, $t(26)$ = 1.1, p = .28), with a significant difference between the mean and the peak (paired two-tailed t-test, $t(26)$ = 7.47, p < .001), suggesting that the distribution is significantly skewed toward the alternative category (Fig. 3) but the peak position remained at the true target location. These results are consistent with the truncation hypothesis but not the weighted-average hypothesis.

There were apparent variations in the mean bias across participants, however, with some showing negative mean bias and some showing positive bias. To further examine whether there are individual differences in the use of categories, participants were divided into two distinctive groups (Fig. 4) based on their mean bias, the conventional criterion for determining which category individuals use. The first group consists of participants with a positive mean bias toward the default category (nine participants, mean signed error = 1.76°,SD = 1.5, one-sample, two-tailed t-test against zero, $t(8)$ = 3.6, p < .01), while the second group consists of participants with a negative mean bias toward the alternative category (18 participants, mean signed error = -2.60°, SD = 1.5, one-sample, two-tailed t-test against zero, $t(17)$ = 7.35, p <.001).
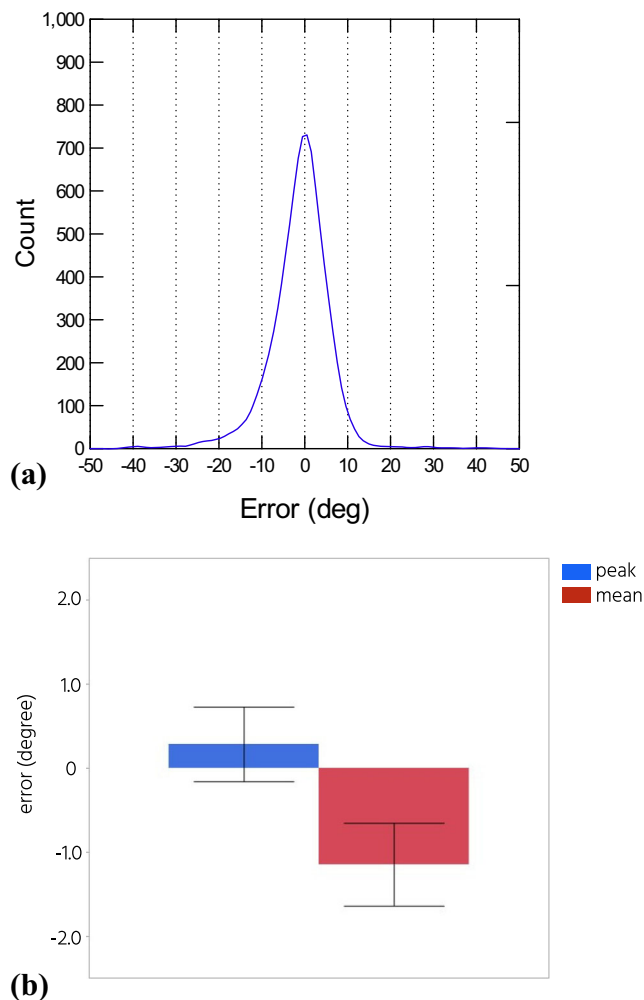
Fig. 3 Results for Experiment 2. Panel (**a**) shows the kernel curve for the overall combined distribution of errors. The "V" boundary line (and the center of the default mean) was located between 5~15° from the true target location (0° in the graph) on the positive side, with the mean position at 10°. The center of the alternative category was located between -40 ~ -30°, with a mean position at -35°. Panel (**b**) shows the mean error and the mean peak position of the error distributions

To examine how participants in these groups used the categorical information, we fitted the angular errors with a Kernel curve (as in Experiment 1) for each individual participant and estimated the main peak of the distribution. Next, we compared the mean bias to the peak position to examine the skew of the distribution, which provides information on the mechanism of the bias. When we examined the data by group, the distribution of errors for the default group showed a large shift of peak towards the diagonal (mean peak = 2.8°, SD =2.2, one-sample, two-tailed t-test against zero, $t(8) = 3.8$, $p <.01$), and a small but significant difference between the mean and the peak position (paired two-tailed t-test, $t(8) = 3.7$, $p < .01$). Figure 4 shows the combined distribution for the two groups (panels a and b) and representative participants in the default (panels c and d) and the alternative group (panels e and f). This

pattern of error distribution is similar to that of Experiment 1 and is most consistent with the weighted-average mechanism.

In contrast, the distribution of errors for the alternative group did not show a significant shift of peak (mean peak = -0.65°, SD=1.5, one-sample, two-tailed t-test against zero, $t(17) = 1.84$, $p = 0.08$), but showed a large difference between the mean and the peak position (paired, two-tailed t-test, $t(17) = 7.0$, $p < .001$), suggesting a significant skew away from the boundary (Fig. 4). This pattern of error distribution cannot be explained by the weighted-average hypothesis, but instead is consistent with predictions of the truncation hypothesis. These results suggest that these two groups relied on different mechanisms in their use of the categorical information.

In summary, our new distribution shape analysis allowed us to provide the first empirical evidence for the weighted-average and the truncation mechanisms separately. Our goal in Experiment 3 was to further examine the robustness of these findings under some variant conditions.

Experiment 3

Method

The participants were 15 undergraduate students, 28% male, from the University of Illinois at Urbana-Champaign, who carried out the experiment to partially fulfill a course requirement. We conducted a variant of Experiment 2 to test the reliability of our findings. The general procedure was similar to that used in Experiment 2 except that: (a) the "V" frame to mark the alternative category was only presented in half of the trials determined at random; (b) there were 320 trials in total to provide sufficient data in both the default condition (without "V" boundary) and the alternative condition (with the "V" boundary); (c) the delay was reduced to 1 s so that we can accommodate more trials within a 1-hour experimental session; and (d) the targets were randomly distributed around all directions, instead of clustered in certain bins as in Experiments 1 and 2, to test whether the distribution of targets affected the results. We reasoned that the modified procedure would further establish our paradigm to differentiate the mechanisms.

Results and discussion

Responses that were more than 45° away from the actual target location were excluded from the analysis (1.0% of responses). One participant was excluded from the analysis due to failure to complete the experiment. We calculated angular bias, as we did in Experiment 2, as the difference between the reported and the actual target angle. We also followed the same coding scheme, so that the sign of the bias indicated which category was applied in estimation: default bias appeared positive and alternative bias appeared negative for all
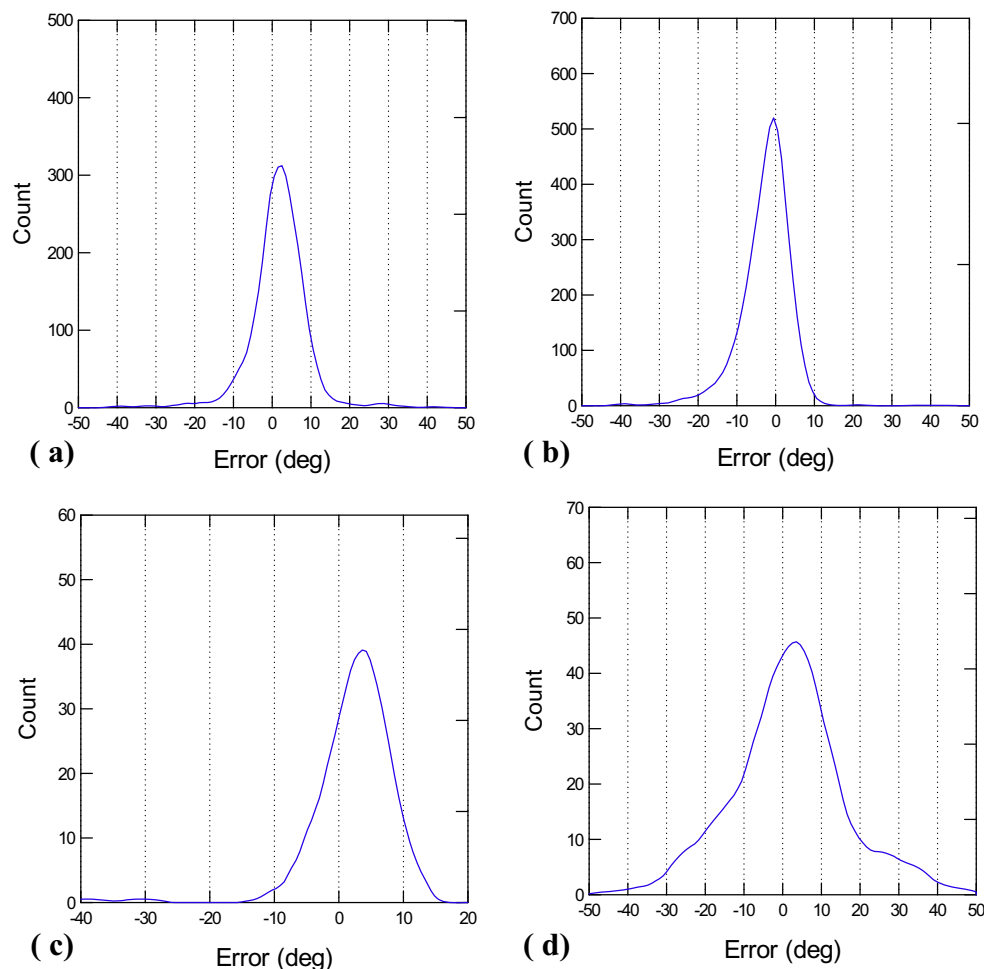
Fig. 4 Error distribution for the two groups of participants in Experiment 2. (**a**) The combined distribution for participants in the default category group; (**b**) the combined distribution for those in the alternative category group; (**c**) and (**d**) two representative participants in the default category group; (**e**) and (**f**) two representative participants in the alternative category group. In all these plots, the "V" boundary line (and the center of the default category) was located between 5~15°, with the mean position at 10°. The center of the alternative category was located between -40 ~ -30°, with a mean position at -35°. (**g**) Mean error and peak of the error distributions for the two groups separately

targets. Moreover, we conducted the distribution analysis and fitted the kernel curve for each individual participant and for the default condition and the alternative condition separately.

Our results showed a positive bias toward the center of the default category when the V frame was absent at retrieval (i.e., a condition comparable to that of Experiment 1). The mean bias was marginally significant, $M = 0.73$, $SD = 1.4$, one-sample, two-tailed t-test against zero, $t(13) = 2.0$, $p = .066$, and the peak position of the distribution of errors was positive, $M = 0.54$, $SD = 1.1$, one-sample, two-tailed t-test against zero, $t(13) = 2.6$, $p = 0.01$. We then compared the mean bias to the peak position to examine the skew of the distribution, which provides information on the mechanism of the bias, and found no significant difference between the mean and the peak position (paired, two-tailed t-test, $t(13) = 0.4$, $p = 0.67$). These data indicate a symmetrical distribution of errors for the trials without the V frame and generally replicated the pattern in Experiment 1.

For trials in which the V frame was presented at retrieval, the mean bias toward the alternative category did not reach significance (one-sample, two-tailed t-test against zero, $M = -1.15$, $SD = 2.6$, $t(13) = 2.0$, $p = 0.13$). As for the distribution, the peak was centered at the true target location ($M = 0.44$, $SD = 1.6$, one-sample, two-tailed t-test against zero, $t(13) = 1.1$, $p = 0.31$). Moreover, the peak and mean bias were significantly different (paired, two-tailed t-test, $t(13) = -2.8$, $p = 0.01$), as shown in Fig. 5, revealing a skewed distribution. The error distribution data in the trials with a visible V frame are thus more consistent with the predictions of the truncation hypothesis while the data in the trials without a V frame are more consistent with the predictions of the weighted average hypothesis. The pattern of data when the V frame was absent replicates that of Experiment 1 and the default group of Experiment 2, while the pattern of data when the V frame was presented generally
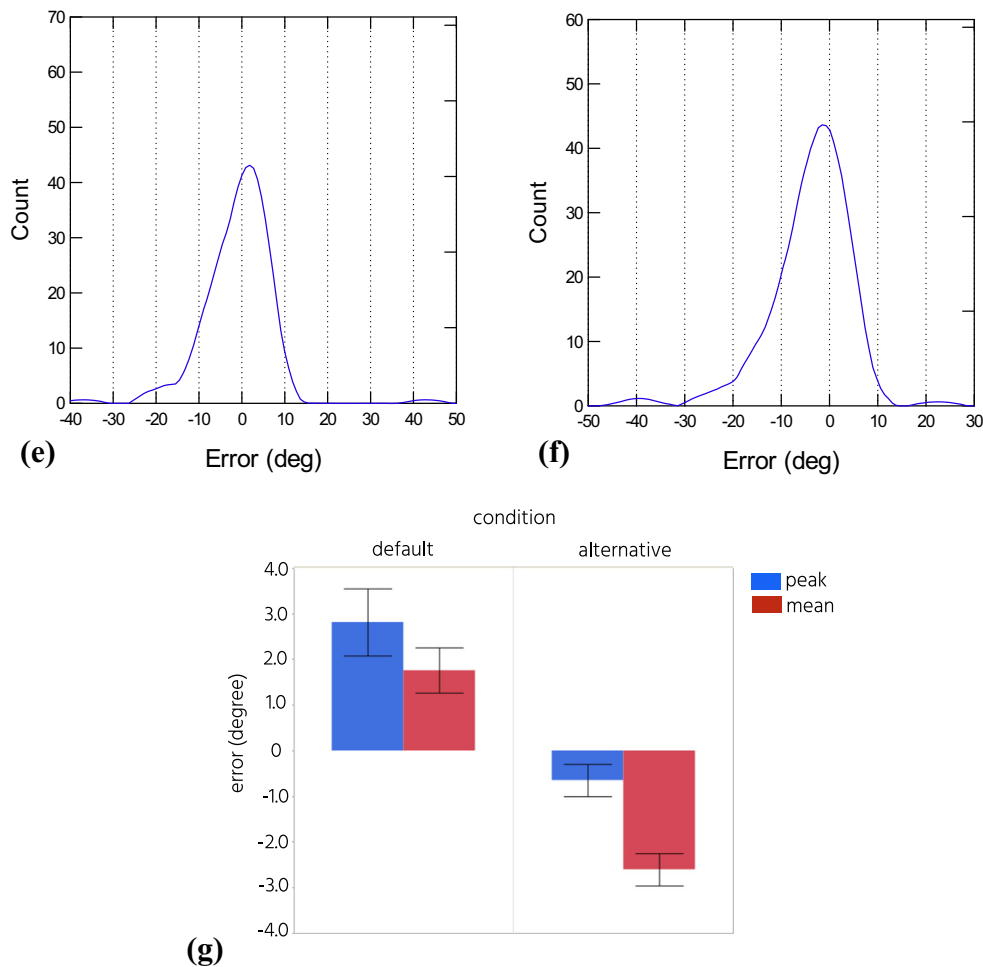
(e)

(f)

(g)

Fig. 4 continued.

replicates the overall pattern and that of the alternative group in Experiment 2, providing further evidence of the two mechanisms separately under experimental conditions of different target distribution and memory delay.

General discussion

Spatial representation is hierarchical in nature. The evidence comes from phenomena such as the effect of region membership on location memory (Huttenlocher et al., 1991), orientation judgments (Maki, 1981), and proximity judgments (Allen, 1981), as well as the effect of barriers on estimations of distance (Newcombe & Liben, 1982). In this paper, we examined the mechanisms underlying the effect of spatial category on estimates of location. Specifically, using a distribution shape analysis, our aim was to develop a new paradigm to differentiate possible mechanisms underlying the category bias in spatial memory. We distinguished between the main hypotheses using

a standard paradigm with a blank circle and when the boundaries of an alternative category to the default symmetrical axe were visible during retrieval. We found evidence for multiple mechanisms underlying the category bias in spatial memory.

First, we reanalyzed the data from Sampaio and Wang (2017) and showed that the mechanism causing the bias with the standard dot-in-circle task (no alternative category to the default category bounded at the Cartesian axes is visually presented) is consistent with a weighted average between the coordinate information and the category (a symmetrical distribution of errors). In Experiment 2, when the boundaries of an alternative category were visible at retrieval (participants recalled dot locations in a circle with a V marked), our results showed two patterns of skew dependent on which category individuals used. Specifically, some people primarily used the default categories bounded at the horizontal and vertical axes and displayed a large shift of the peak towards the boundary with a slight skew in the distribution of errors. Other people primarily used the alternative "V" categories bounded at the
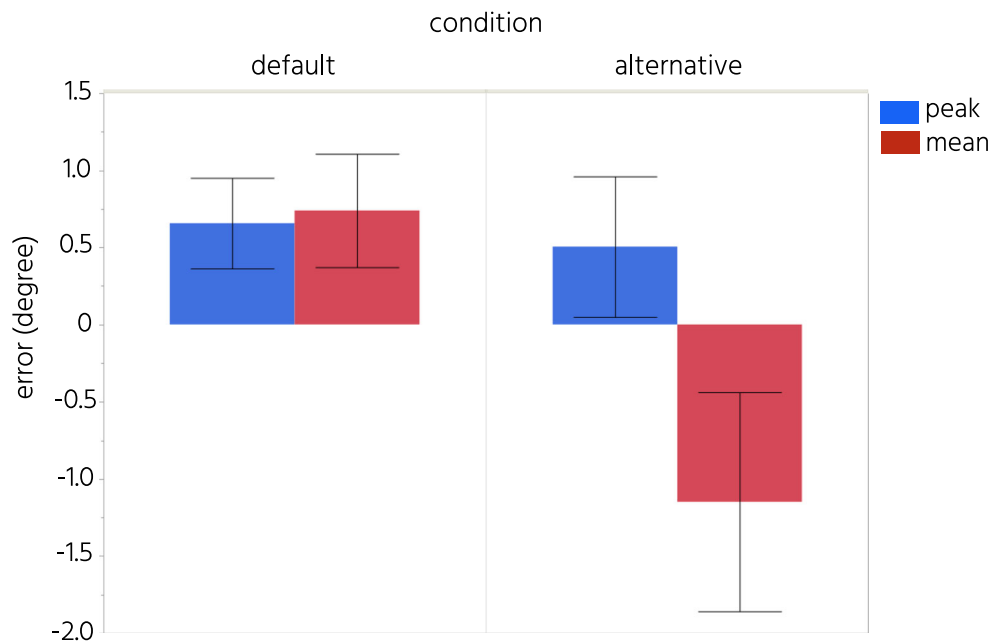
Fig. 5 Mean error and peak of the error distributions for the two conditions in Experiment 3

diagonals and displayed little shift of peak but a large skew away from the boundary. Experiment 3 replicated the overall pattern of data: a symmetrical distribution of errors shifted toward the center of the default category in trials when a V frame was absent during retrieval, and a skewed distribution of errors away from the boundaries in trials when the V frame was present.

Overall, our data suggest that the default category information mainly shifted the distribution toward the category mean and only slightly changed the shape of the distribution. This pattern is indeed most consistent with a weighted average mechanism (Huttenlocher et al., 1991). In contrast, the alternative category information mostly changed the shape of the distribution and only minimally the peak. This pattern is most consistent with a truncation effect at the category boundaries. In addition, the data indicate that when an additional source of categorical information is available in the form of visible boundary lines of an alternative category, the bias is primarily caused by truncation in some cases and by weighted average in others. These findings provided the first experimental evidence for these two mechanisms separately using the new distribution analysis.

We acknowledge that when the V frame was presented, the lack of peak shift alone (i.e., a distribution centered at the true target location) could be a result of the weighted average of two opposite categories instead of a true truncation effect. However, note that we defined the evidence of truncation primarily on the skew of distribution, with the absence of peak shift only as a secondary feature. Because the skew cannot be produced by the weighted average mechanism, we believe the truncation effect we found was not due to multiple category averages.

When we made multiple categories experimentally available to participants (e.g., in Experiments 2 and 3), the distribution pattern showed some slight deviation from the predictions of a pure weighted-average or a pure truncation mechanism. These deviations likely reflect occasional usage of a secondary strategy. For example, the slight skew in the distribution for the default category group could be due to participants using the alternative category boundaries by truncation in a small subset of trials, which could affect the shape of the distribution slightly. Similarly, when the conditions were mixed and unpredictable in Experiment 3, the effect of the default category also seemed weaker, possibility due to using the default category less consistently.

The finding that some individuals still used the default categorization scheme despite the alternative categorization cues being perceptually visible supports previous research that found that default categorization is challenging to overcome, even when people are instructed to use an alternative category (Huttenlocher, Hedges, & Crawford, 2004). We do not know what led individuals to use one or another category. Crawford et al. (2016) found that working memory capacity is related to how people structure a space and remember targets within it. They found that greater bias is associated with lower working memory capacity. We speculate that it is possible that individuals with lower working memory capacity relied more on the visible alternative category, while individuals with higher working memory capacity relied more on the invisible default category. Other individual differences such as spatial ability may also play a role in the variations in category use. How different individuals use and weight multiple categories remains a question for future research.

The evidence of both models raised an important theoretical question, i.e., when and/or why each mechanism is used. Speculatively, it is possible, for example, that people are more inclined to use a truncation mechanism for categories with salient boundaries (they may represent the boundaries for categories with salient boundaries). On the other hand, when the boundaries are not directly presented, people may be more likely to represent the prototype of the categories and thus the weighted-average mechanism may be more prominent. Although effects of boundary salience have been examined in the literature (e.g., Simmering & Spencer, 2007; Spencer & Hund, 2002), the topic has not yet been linked to the *cause* of the bias. Future research can manipulate boundary salience (e.g., dim/thin vs. bright/thick lines), for example, as a path to start answering the question on when and why each mechanism is used.

Another issue that future research needs to examine is with regard to whether categorical and metric cues actually combine when people estimate locations from memory. Sampaio and Wang (2017) found that, in the standard dot-in-circle paradigm, individuals integrate the metric and categorical sources to respond in each trial, as revealed in their distribution of errors (unimodal and centered at a biased location), rather than alternating the use of a single cue in each trial/judgment. Similar to the case of a standard dot-in-circle paradigm, when multiple categories are available at the time of retrieval, there are different possibilities for how different sources of information may be blended. For example, it is possible that the multiple categorical codings merged with the metric information in a given trial to form a single, integrated representation (full combination). Alternatively, each category may compete to combine with the metric information individually and succeed with a certain probability (independent combination). That is, in any given trial, only one of the categories is combined with the metric information. Finally, it is also possible that all sources of information compete individually and no combination occurs in a given trial (full alternation). Thus, every response is made exclusively based on the metric, the default category, or the alternative category, each with a certain probability. Although these scenarios are clearly different from each other, the traditional analysis of mean bias cannot distinguish among them, as it only reflects overall effects. How to differentiate these potential scenarios remains a challenge for future research.

In summary, our studies provided the first experimental evidence for the weighted-average hypothesis and the truncation hypothesis separately using a new distribution shape analysis. Moreover, we found individual differences in estimating locations when multiple categories are available. Some people preferred to combine the metric information with the default and others with the alternative category. Importantly, individuals used different mechanisms of integration for the default and alternative categories; those using the default category primarily result from a weighted average, while those using the alternative category result from a truncation effect at the visible category boundaries.

References

Allen, G. L. (1981). A developmental perspective on the effects of "subdividing" macrospatial experience. *Journal of Experimental Psychology: Human Learning and Memory*, 7, 120–132.

Crawford, L. E., & Jones, E. L. (2011). The flexible use of inductive and geometric spatial categories. *Memory & Cognition, 39,* 1055–1067

Crawford, L.E., Landy, D., Salthouse, T.A. (2016). Spatial working memory capacity predicts bias in estimates of location. *Journal of Experimental Psychology: Learning, Memory and Cognition,42,* 1434–1437.

Engebretson, P. H., & Huttenlocher, J. (1996). Bias in spatial location due to categorization: Comment on Tversky and Schiano. *Journal of Experimental Psychology: General, 125,* 96–108.

Friedman, A., & Brown, N. R. (2000). Reasoning about geography. *Journal of Experimental Psychology: Learning, Memory and Cognition,129,* 193–219.

Holden, M.P., Curby, K.M., Newcombe, N.S., & Shipley, T.F. (2010). A category adjustment approach to memory for spatial location in natural scenes. *Journal of Experimental Psychology: Learning, Memory, and Cognition 36,* 590–604.

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. (2005). The stability and flexibility of spatial categories. *Cognitive Psychology, 50,* 1–44.

Hund, A. M., Plumert, J. M., & Benney, C. (2002). Experiencing nearby locations together in time: The role of spatial-temporal contiguity in children's memory for location. *Journal of Experimental Child Psychology, 82,* 200–225.

Huttenlocher, J., Hedges, L.V., Corrigan, B, & Crawford, L. E. (2004). Spatial categories and the estimation of location. *Cognition, 93,* 75-97.

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

Maki, R. H. (1981). Categorization and distance effects with spatial linear orders. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 15–32.

McNamara, T.P. (1986). Mental representations of spatial relations. *Cognitive Psychology 18,* 87–121.

McNamara, T.P., Hardy, J.K., & Hirtle, S.C. (1989). Subjective hierarchies in spatial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 211–227.

Newcombe, N., & Liben, L. S. (1982). Barrier effects in the cognitive maps of children and adults. *Journal of Experimental Child Psychology, 34,* 46–58.

Newcombe, N.S., Huttenlocher, J., Sandberg, E., Lie, E., & Johnson, S. (1999). What do misestimations and asymmetries in spatial

judgement indicate about spatial representation? *Journal of Experimental Psychology, 25,* 986–996.

Sampaio, C. & Cardwell, B.A. (2012) Bias in long-term location memory in the real world. *Quarterly Journal of Experimental Psychology,* 65, 1865–71.

Sampaio, C., & Wang, R.F. (2010). Overcoming default categorical bias in spatial memory. *Memory & Cognition, 38,* 1041–1048.

Sampaio, C., & Wang, R.F. (2012). The temporal locus of the categorical bias in spatial memories. *Journal of Cognitive Psychology, 24,* 781–788.

Sampaio, C., & Wang, R.F. (2017). The cause of category-based distortions in spatial memory: A distribution analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition,43,* 1988–1992.

Shelton, A. L. & McNamara, T. P. (2001). Systems of spatial reference in human memory. *Cognitive Psychology, 43,* 274–310.

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 & Performance, 33,* 871–894.

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., & Hund, A. M. (2003). Developmental changes in the relative weighting of geometric and experience-dependent location cues. *Journal of Cognition and Development, 4,* 3–38.

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

Wang, R. F. (2012). Theories of spatial representations and reference frames: What can configuration errors tell us? *Psychonomic Bulletin & Review, 19(4),* 575–587.