

Location Memory for Dots in Polygons Versus Cities in Regions: Evaluating the Category Adjustment Model

Alinda Friedman
University of Alberta

Daniel R. Montello and Heather Burte
University of California at Santa Barbara

We conducted 3 experiments to examine the *category adjustment model* (Huttenlocher, Hedges, & Duncan, 1991) in circumstances in which the category boundaries were irregular schematized polygons made from outlines of maps. For the first time, accuracy was tested when only perceptual and/or existing long-term memory information about identical locations was cued. Participants from Alberta, Canada and California received 1 of 3 conditions: *dots-only*, in which a dot appeared within the polygon, and after a 4-s dynamic mask the empty polygon appeared and the participant indicated where the dot had been; *dots-and-names*, in which participants were told that the first polygon represented Alberta/California and that each dot was in the correct location for the city whose name appeared outside the polygon; and *names-only*, in which there was no first polygon, and participants clicked on the city locations from extant memory alone. Location recall in the dots-only and dots-and-names conditions did not differ from each other and had small but significant directional errors that pointed *away* from the centroids of the polygons. In contrast, the names-only condition had large and significant directional errors that pointed toward the centroids. Experiments 2 and 3 eliminated the distribution of stimuli and overall screen position as causal factors. The data suggest that in the “classic” category adjustment paradigm, it is difficult to determine a priori when Bayesian cue combination is applicable, making Bayesian analysis less useful as a theoretical approach to location estimation.

Keywords: spatial memory, location estimation, cognitive maps

Memory for spatial locations (where food and home are; where we are) is crucial in many contexts for all mobile organisms. Location memory has been characterized as multileveled and thus hierarchical in nature (e.g., Friedman, 2009; Huttenlocher, Hedges, & Duncan, 1991; McNamara, 1986; Newcombe & Huttenlocher, 2000; Stevens & Coupe, 1978). As with many kinds of categories, properties believed to be true of one level of the hierarchy are believed to be true of the levels below it. Thus, for example, Stevens and Coupe (1978) famously demonstrated that most people incorrectly believe that San Diego, California, is west of Reno, Nevada, presumably because they believe that most or indeed all of California is west of all of Nevada. In this case the categories are states, and the items are cities within them; the incorrect inference at the item level is constrained by incorrect beliefs held

at the category level. In the present study we examined estimation accuracy for the locations of dots in polygons that represented map outlines of the state or province in which the participants lived. The dots represented actual city locations. Some participants knew this (as well as what the particular location was), and others did not. To our knowledge, this is the first time that perceptual and extant long-term conceptual knowledge about the identical locations within the same perceptual frame have been directly compared.

Comparing perceptual and long-term conceptual knowledge about locations is theoretically important because at least one prominent theory of location memory (i.e., the *category adjustment model*, Huttenlocher et al., 1991) has relied on empirical paradigms that do not clarify this distinction. For example, there are occasions when the model seems to apply to principally perceptual (recent) memory (e.g., the original Huttenlocher et al., 1991, article and many others) and other occasions when, at least in discussion, the model should apply to location memory more generally, including that which is derived from extant knowledge (e.g., Holden, Curby, Newcombe, & Shipley, 2010).

It is also valuable to distinguish location memory based on short-term perceptual memory stores derived from a single presentation to location memories in long-term memory stores that are derived from a variety of sources, including perception in various modalities, language, school, maps, schematic expectations, etc. It may be, for example, that the two different conditions require two theories because different retrieval and other mechanisms are involved. To our knowledge, and including the data from Holden et al. (2010), this issue has never been directly tested. Thus, this study is an attempt to compare memory for dot locations that had

This article was published Online First April 30, 2012.

Alinda Friedman, Department of Psychology, University of Alberta, Edmonton, Alberta, Canada; Daniel R. Montello, Department of Geography, University of California at Santa Barbara; Heather Burte, Department of Psychology, University of California at Santa Barbara.

This research was funded by a grant to Alinda Friedman from the Natural Sciences and Engineering Research Council of Canada. We would like to thank Bernd Kohler for programming the experiment and help with the data analysis and Tyler Astring, Kimberly Botts, Michael Feldman, and Arian Monem for help with data collection. We also appreciate Mark Holden's thorough critique of an earlier version of the article and insights about the analyses.

Correspondence concerning this article should be addressed to Alinda Friedman, Department of Psychology, University of Alberta, Edmonton, Alberta, Canada T6G 2E9. E-mail: alinda@ualberta.ca

no semantic reference and were recently learned visually to memory for dot locations that had deep semantic reference and were learned over long time periods from a variety of sources.

Finally, to test some of the Bayesian cue combination predictions of the category adjustment model, we included a condition in which the two types of cues—recent visual short-term memory (STM) and extant long-term memory (LTM)—were combined. This is of relevance to any theory of location memory because it speaks to the major explanatory concepts of such models (e.g., centroid prototype locations, cognitive regions, cue combination, movement toward or away from a prototype) and evaluates the generality of these concepts. Thus, the “combined” condition in the present study allowed us to potentially disentangle the contributing effect of the semantic content of the stimuli and the immediate perceptual conditions under which they were learned.

Huttenlocher et al. (1991) first proposed a theory about how spatial information from two different sources—a category level and an item level—is combined adaptively to form a location estimate (see also Huttenlocher, Hedges, Corrigan, & Crawford, 2004; Huttenlocher, Hedges, & Vevea, 2000; Newcombe & Huttenlocher, 2000). Their category adjustment (CA) model takes a Bayesian approach to combining information from multiple sources. The main idea is that errors in judgments occur not because the information being combined is itself biased but because when correct but uncertain information from different levels is combined, less reliable information is given less weight than more reliable information. To the extent that there is uncertainty at the item level, for instance, this type of Bayesian combination implies that judgments on a given dimension will shift away from the value that belongs to the item and toward some value that is determined from the item’s category. For location estimates, that value is usually determined by the category’s borders and its locational prototype (assumed to be the mean item location or the center of mass of the category’s borders). The borders and prototype are thus paramount in determining the direction of the bias.

Huttenlocher et al. (1991) wrote that “When the uncertainty of the memory is very large (compared with that of the prototype) so that memory provides essentially no information, we assume that participants give essentially total weight to the prototype and essentially no weight to the memory” (p. 373). In this model, then, judgments are typically biased toward the prototype. Bayesian combination thus creates optimal judgments overall in the sense that the average error across judgments is less than it would be if the prototype was not used. This type of optimality is the main justification for proposing that Bayesian combination is the psychological process used in combining cues to make location judgments (Cheng, Shettleworth, Huttenlocher, & Rieser, 2007). Notably, in the psychological literature, there are many other ways to combine information from multiple cues and dimensions to make an estimate (e.g., stimulus generalization [Shepard, 1987], viewpoint interpolation [Edelman, 1999; Friedman & Waller, 2008; Thrash, Waller, & Friedman, 2011], metrics and mapping [Brown & Siegler, 1993; Friedman & Brown, 2000a]), but these are often overlooked or downplayed by Bayesian theorists. In the present article we take a neutral stance with respect to these other theories, because we suspect strongly that each (as well as Bayesian theories) will have strengths and weaknesses depending on the domain of application.

Huttenlocher et al. (1991) tested their model by presenting a single dot randomly located within in a circular frame on each trial, removing the stimulus, and then having participants reproduce the location of the dot from memory in a blank paper circle after a 2-s interval. The seminal finding was that the remembered locations of the randomly distributed dots were clustered toward the oblique diagonals of the circle as well as toward the center of these radii. The categories were thus assumed to be quarter sections bounded by imagined horizontal and vertical lines, with the diagonals as the prototypes. Evidence that estimates shift toward the category’s prototype has been found many times in many perceptual domains (e.g., fatness of fish-like stimuli [Duffy, Huttenlocher, & Crawford, 2006], shades of gray [Huttenlocher et al., 2000], location to the left or right of the body midline [Spencer & Hund, 2002], and other regular polygons [Wedell, Fitting, & Allen, 2007]). In large-scale geographic domains, however, the bulk of the evidence shows that estimates shift away from the actual subregional prototype (e.g., the southern United States, or southern Europe) but toward the prototype/centroid of the region in which a given subcategory is believed to be located (in the two cases mentioned, too far to the south; Friedman & Brown, 2000a, 2000b; Friedman & Kohler, 2003; Friedman & Montello, 2006).

Holden et al. (2010) found that the bias toward the center of mass of a category predicted by the category adjustment model generalized to natural scenes (e.g., landscapes) in which the category boundaries were very irregular (e.g., parts of sand dunes that were a uniform color). In a second experiment, they examined what kind of information was used to segment the scenes into categories by using inverted scenes and colored negatives; inversion interferes with the extraction of meaning but leaves low-level visual information intact (Brockmole & Henderson, 2006). Negatives do not contain all the low-level information that exists in the corresponding canonical image (Vuong, Peissig, Harrison, & Tarr, 2005) and may also alter the ability to extract semantic information (Goffaux et al., 2005). In both cases, the errors were still biased toward the prototype locations, but they differed significantly between the canonical upright images and the two types of non-canonical images. Thus, Holden et al. (2010) concluded that both conceptual and perceptual information were used in the categorization of the scenes. However, because the scenes were novel, “conceptual” must be at the level of object/scene recognition, as might be required in a naming task, for example. In this sense, recognition required long-term memory information but at the type, rather than the token, level (i.e., categorizing an exemplar). Thus, one can recognize that an “object” in a scene is part of a sand dune without ever having seen that particular instance of a sand dune. In addition, the placement of the dots within the category boundaries did not signify anything about the identity of the category and more important, the category boundary did not signify anything per se about where a dot should be placed. Thus, the locations of the dots were arbitrary and deliberately bore no relation to each other or to the category boundary.

Cheng et al. (2007) pointed out circumstances in which the Bayesian combination of cues for *navigation* does not make sense (e.g., cue competition). When two cues indicate conflicting information, it makes little sense to combine them. However, they concluded their review by noting that

the notion of subjective discrepancy is crucial. Discrepancy is found when two cues point to different values on some metric parameter (e.g., direction). Subjective discrepancy is measured not in physical terms but in psychological terms: in units of standard deviation in the subjective measure of a parameter, a measure of subjective uncertainty about the measure . . . small discrepancies led to integration, whereas large discrepancies lead to using one cue or the other, not an averaging of cues. *Both sides of this principal may be compatible with Bayesian principles.*" (p. 633; emphasis added)

What is difficult to understand from this analysis of Bayesian cue combination is how, in a given circumstance, one can *predict* what is a large versus small discrepancy. Without this capability, a Bayesian model can be made compatible with almost any pattern of behavior. For example, Cheng et al. (2007) go on to state that ". . . rats rely on a beacon displaced by 45° over path integration . . . *This pattern may not be contrary to Bayesian principles.* The landmark cue might be that much better and receive a strong weighting" (p. 633; emphasis added). Thus, by definition or default, if optimal combination of cues is not found, then it must be because one cue was not weighted highly (or at all). Cheng et al. assumed that the alternative hypothesis to Bayesian combination of cues for navigation—path integration—is used as a backup system when a landmark cannot be viewed as the same landmark that was learned previously, but they did not indicate what circumstances would cause that to happen. This situation is untenable because it is virtually unfalsifiable.

There is growing evidence that the CA model does not generalize to estimates of locations of known or remembered places (e.g., Friedman, 2009) because, as noted above, even under situations of uncertainty, such estimates are often biased *away* from category prototypes and the data indicate that both category- and item-level information are inaccurate. That is, although there are certainly spatial categories in large-scale geography (cognitive geographic regions; see Montello, 2003), the estimated locations of both the category boundaries and the cities within some of the regions are often quite biased (Friedman, Kerkman, Brown, Stea, & Cappello, 2005; Friedman & Montello, 2006) and often directionally *away* from the actual categories' prototypes. This implies that judgments about locations of cities may involve a different set of processes (e.g., metrics and mapping; Brown & Siegler, 1993) than Bayesian combination. We do not wish to go into these alternative model(s) in depth because they have been well described elsewhere (e.g., Brown & Siegler, 1993; Friedman & Brown, 2000a, 2000b) and because the present experiments were set up simply to test performance for the same locations when participants knew they represented an actual place versus when they did not. However, we will point out that metrics and mapping, for instance, can predict that there are occasions when estimates will move away from a prototype, does not require information at category and item levels to be independent and allows extant long-term memory information to be inaccurate (not only uncertain). However, it is a model of retrieval from only LTM.

Although location estimates of cities are often biased in an absolute sense, location memories about known domains can be quite accurate in a relative sense. For example, if one looks at the estimates within familiar regions, the north-south positioning of the cities *relative to each other* is often quite accurate (Friedman & Montello, 2006); so (sometimes) is configurational information estimated on a grid (Friedman, 2009). Thus, on an absolute level,

the item locations can be quite biased, whereas in a relative or configurational sense, they might not be. This final point makes it unlikely that the *representations* are unbiased (or even independent), which are important assumptions of the CA model. Thus, Bayesian combination may not always be used in location estimates of this sort. Rather, geographic category-level knowledge may be biased because of other beliefs about the world that are used in plausible reasoning processes (Collins & Michaelski, 1989), which combine those beliefs with whatever metric information is available (see Brown & Siegler, 1993). The overall bias within a given geographic region (however ill-placed absolutely) may still be toward its centroid, however.

Experiment 1

A priori, it is likely that remembering dot locations relies more on some type of visiospatial short-term memory, whereas remembering city locations relies more on what is stored in long-term memory. In Experiment 1, we compared three conditions to try to disentangle the respective roles of visual short-term versus longer-term memories about locations. We used participants from two universities—the University of Alberta, in Edmonton, Canada and the University of California in Santa Barbara. The outline of both the province of Alberta and the state of California can be made into schematized polygons by drawing their borders as straight lines (see Figures 1 and 2).

Albertan and Californian participants were presented with the Albertan and Californian polygons, respectively, in one of three conditions, which we named according to the stimuli that were given to be reproduced; a tentative task analysis using Bayesian principles follows. In Experiment 1, the dots' locations in all three conditions were identical and corresponded to the real locations of cities in each place. In the *dots-only* condition, we used a paradigm similar to that used by Huttenlocher et al. (1991): The polygon appeared on a computer screen with an unidentified dot, then a dynamic masking stimulus appeared (which Huttenlocher et al., 1991, did not use), then a blank polygon appeared. The task was to click a mouse cursor on the location where participants remembered the dot to be located. No mention was made of the "identity" of the polygon or the dots. The mask was added to the paradigm to prevent responding on the basis of iconic memory (Phillips, 1974). On the other hand, the overall time frame between stimulus offset and presentation of the response frame was well within the limits of visual short-term memory (Matsukura & Hollingworth, 2011).

In the remaining two conditions, participants were told the identities of the polygons (either Alberta or California) at the outset of the experiment. In the *dots-and-names* condition, when the first polygon appeared with its dot, the name of the city whose location the dot represented appeared at the top of the polygon; participants were truthfully told that this was the correct location of the city. Then the mask appeared and then the blank polygon appeared. Participants responded by clicking the mouse at the location within the polygon where they thought the dot had been; except for the instructions and city names, the dots-only and dots-and-names procedures were identical. In the *names-only* condition there was no "first polygon" or mask. That is, all trials consisted of blank polygons with a city name at the top; the participants' task was to click the mouse on the location within

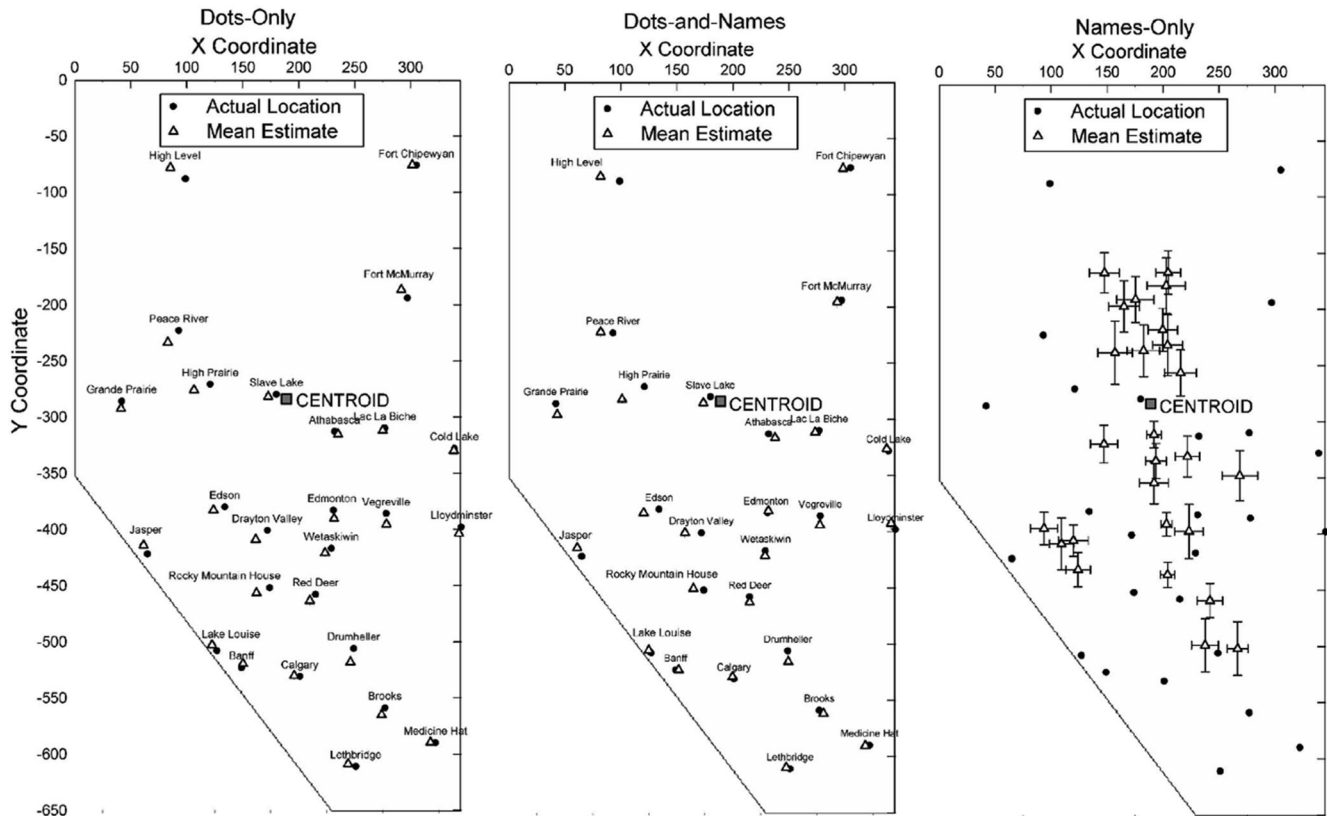


Figure 1. Albertan actual and estimated locations in Euclidean space, superimposed on the actual (equirectangular) map polygon that was used in Experiment 1. In the experiment, there were no city names shown on the polygon in any condition; in the *dots-and-names* and *names-only* conditions the city names appeared above the polygon on each trial.

the polygon where they thought each city belonged. Participants in the names-only condition were never shown the dot locations during the experiment but estimated their locations from any extant knowledge of the cities they had.

A task analysis of the three conditions using the CA model would be something like the following: For all three conditions in both countries, the category boundary information provided by the schematic polygon outline was perceptually available (and correct) during both the initial presentation (for the dots-only and dots-and-names groups) and test (for all three groups); thus, participants should give a high weight to the category boundary information. In addition, whatever the centroid is should be computable “online” at test; we assume participants err toward the centroid of the given category boundaries because the boundaries form irregular polygons (cf. Holden et al., 2010; Simmering & Spencer, 2007).

In the dots-only condition, the mask should make the item-level information unavailable to iconic memory and relatively uncertain (compared with when it is directly present or there is no mask). However, the information should be available in visual short-term memory (VSTM); we remain uncommitted to any particular model of visual/spatial short-term store but do assume we are well within its time frame, particularly with only one item to remember (Matsukura & Hollingworth, 2011). Thus, a Bayesian analysis of the dots-only condition predicts that the item-level information should

be weighted somewhat less than the category-level information, and we should see some movement toward the prototype.

From prior research, city location information in the names-only condition should be less biased than when participants respond on a grid and must infer the category boundaries (e.g., Friedman, 2009, Experiments 1 and 2; Friedman, Mohr, & Brugger, 2011). This is because the category-level information given at test may be assumed to be the correct border information. Thus, in the names-only condition, the category-level information may have a much higher weighting than the item-level information compared with the dots-only condition. This prediction derives from the assumptions that participants know the border is correct and that the item-level information will activate the city names. At this point, they may then become aware of their uncertainty about this item-level information, relative to the certainty of the category. Thus, there should be larger errors toward the prototype, compared with the dots-only condition. However, there may be subregions that are not seen in the dots-only condition, as there often are with retrieval from long-term memory of items in large-scale geographic spaces, even within map boundaries (e.g., Friedman, 2009, Experiment 3).

In the dots-and-names condition, there is the possibility that participants can use the category-level perceptual frame available at test (with a high weighting), the correctly given item location information at learning (with a lower weighting due to some

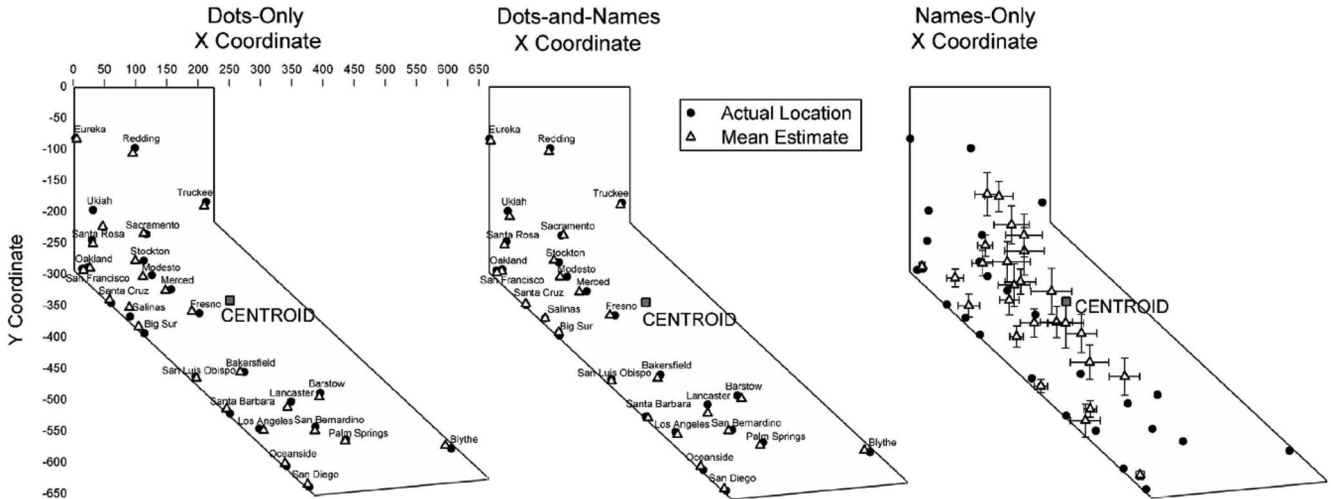


Figure 2. Californian actual and estimated locations in Euclidean space, superimposed on the actual (equirectangular) map polygon that was used in Experiment 1. In the experiment, there were no city names shown on the polygon in any condition; in the *dots-and-names* and *names-only* conditions, the city names appeared above the polygon on each trial.

uncertainty, as in the dots-only condition) and possibly whatever location information is activated in LTM by the cities' names. Thus, this condition should plausibly be midway in accuracy between dots-only and names-only, because there is potentially a third source of information to use (i.e., the presented locations, now associated with real places). Thus, although there might be some movement toward the centroid in the dots-and-names condition, it should be less than in the names-only condition and possibly more than in the dots-only condition because of activation of the items in long-term memory. Of course, the Bayesian prediction can be modified to say that information in LTM is given a weight of zero, in which case the dots-only and dots-and-names conditions will appear to be the same as each other. However, to be fair, this should only happen if the names-only condition produces such error-prone data that it amounts to nearly random performance. Otherwise, as noted above, the Bayesian model can predict any outcome.

The Bayesian predictions above should be true when we average the data across all 26 cities. However, a priori it is likely that participants know more about the locations (and other things) of some cities than others. Thus, we took knowledge ratings of all the cities after the estimate task and also analyzed the data for the "top five" most familiar cities for each participant. We chose five because more than that resulted in cities that were too idiosyncratic across participants. In this case, the Bayesian approach should predict that the item-level information should have a relatively higher weighting than it did across all 26 cities for both the names-only and dots-and names conditions. Movement toward the centroid should thus be lessened, particularly in the names-only condition.

In contrast to the Bayesian predictions, if the present names-only condition is at all similar in its demands to the many other geographical scales we have data for (e.g. Friedman & Brown, 2000a, 2000b; Friedman & Montello, 2006; Uttal, Friedman, Liu, & Warren, 2010) then we should see evidence for subregions (e.g.,

for Alberta, the Rocky Mountains) in addition to a bias to make errors toward the prototype. Furthermore, any bias in any condition that is in a direction away from the prototype goes against any plausible reading of a Bayesian combination of information.

To summarize the Bayesian predications, all three conditions should show responses that are biased toward the center of mass of the polygons. In addition, the names-only condition should be less accurate than the dots-only condition because, presumably, information retrieved from long-term memory should be less certain than information retrieved from visual short-term memory. However, the long-term memory information is not negligible, as has been shown in our previous research (e.g., Friedman, 2009; Friedman & Montello, 2006). Consequently, the dots-and-names condition should show accuracy somewhere in between the other two conditions.

On a metrics-and-mapping account, which is described fully elsewhere (Brown & Siegler, 1993; Friedman & Brown, 2000a, 2000b), the names-only condition should show further regionalization of the estimates. This subregionalization might be such that for some of the subregions, estimates move toward the map polygon's centroid and for others it does not. Because this is a model of retrieval from long-term memory (and of numeric information; but see Friedman, 2009), it does not directly address the dots-only condition but may address the dots-and-names condition. Whether it does is an empirical question.

Method

Participants and design. The Albertans were 96 volunteers (36 men, 60 women) from the University of Alberta's Psychology Department participant pool, who received partial credit toward their grade for participating. They were randomly assigned to one of the three experimental conditions. The data from six volunteers (three men, three women) were not analyzed because they had not been raised in Alberta. This left 30 participants in each of the

experimental groups. The Californians were 68 volunteers (25 men, 43 women) from the participant pool in the Department of Geography at the University of California at Santa Barbara who received extra credit for participating. Twelve of these individuals (three men and nine women) were either from a different state than California or a different country than the United States and had not lived in California for 10 years or more; the data from these participants were eliminated from further consideration. This left 18, 19, and 19 participants, respectively, in the dots-only, dots-and-names, and names-only conditions. The order in which the stimuli appeared was randomized separately for each participant.

Stimuli, apparatus, and procedure. The same computer program was used to collect the data in both places, differing only in the bitmap that was used for the background polygon and, of course, the cities. There were 26 stimulus cities in Alberta and 26 in California (see Figures 1 and 2). City names in the dots-and-names and names-only conditions appeared outside the boundaries of the province or state, in the margin above the polygon image.

The trials were presented on a PC computer in each place, and participants were run individually. After signing a consent form, the participants were seated comfortably in front of the computer screen, which in Alberta was a 19-in. (48.26-cm) Samsung LCD monitor with a .294 dot pitch and in California was a 17-in. (43.18-cm) NEC LCD Monitor with a .263 dot pitch. The pixel bitmaps used for the stimuli were the same pixel height in each location (652 pixels); the Alberta bitmap was 345 pixels wide, and the California bitmap was 663 pixels wide.

Participants next read a screen of instructions. In the dots-only condition they were told that their task was to remember the location of dots that would appear in a “polygon shape.” In the dots-and-names condition, they were told the same thing and in addition, that the polygon was an abstract representation of either Alberta or California and that each dot would correspond to the actual location of the city that was named at the top of the screen. In the names-only condition, participants were told the same information as in the dots-and-names condition, except they were told their job was to click the mouse cursor on the place within the Alberta or California polygon that corresponded to each city’s location, as best as they could remember. All participants were told there was no time constraint on their responses and to try to be as accurate as possible.

For each trial in the dots-only and dots-and-names conditions, the first stimulus appeared after a 1-s warning beep and stayed on the screen for 3 s. When it disappeared, it was replaced with an animated mask of random black line segments for 4 s, and then the blank polygon appeared and stayed on the screen until the participant responded. There was a 1-s intertrial interval. For each trial in the names-only condition, a blank polygon appeared with a city name and stayed on the screen until the participant responded. There was also a 1-s intertrial interval in this condition.

When the estimation task was finished, participants in all three conditions were asked to rate their familiarity with each of the cities on a 0–9 scale, where 0 was to be used if they had “no knowledge of a city” and 9 was to be used if they knew “a great deal about a city,” e.g., they had visited it. They were to rate their knowledge with respect to their knowledge of other Albertan/Californian cities. Each city name appeared centered on the screen one at a time with an abbreviated scale beneath it (0 = *none*, 9 = *a lot*). Participants responded by using the numeric keypad on the

keyboard and then pressing the enter key. There were again no time constraints, and the cities were randomized separately for each participant.

Results

Scoring. A mean and standard deviation across error angles was obtained (as described below) and any response that was more than three standard deviations from the mean (calculated separately within each condition) was eliminated. This procedure eliminated 9.7% of the Alberta data and 7.6% of the California data; the outliers were spread roughly equally across the three groups in each country.

Data analyses. We analyzed the Alberta and California data separately, as we regard them as replications. However, we report the analyses together, in part because the outcomes were so similar. We used $p < .05$ as the alpha level and η_p^2 as the measure of effect size, where appropriate. For data reported as vectors, we report the 95% confidence limits and use d as the measure of effect size. We did not assume that participants used horizontal and vertical lines as category boundaries (as Holden et al., 2010, did not) because Spencer and Hund (2002) showed that it is quite difficult to impose categories on irregular shapes.

There are several issues to take account of when dealing with angular data; these are elegantly described by Holden et al. (2010), so we do so only briefly here. The main goal was to assess whether individuals tend to recall locations as being closer to the category prototype than the locations actually were. Following Holden et al. (2010), rather than examining all the responses with respect to a fixed point across all locations (e.g., the centroid, or center of mass, of the polygons), we examined the magnitude and direction of the error vectors from the correct locations themselves. That is, we converted each participant’s response for each dot/city to a vector originating at the associated correct location and ending at the estimated location. This is analogous to putting all the locations in the same “space” by computing vectors from the actual location to the centroid, from the actual location to the estimated location, and then computing the angle of the directional error between them. This method thus directly assesses both the magnitude and direction of errors simultaneously and allows one to examine error distributions across irregularly shaped categories, as we used here. Each error vector for each participant and location was thus associated with its own angular “difference error.”

To assess differences between experimental groups, each participant’s error vectors were averaged across all 26 locations (all of the vectors were added and then the summed vector length was divided by the number of responses); this automatically takes into account the issue of error magnitude in that if all the errors are toward the same direction, the average length of the resultant vector will be longer. Further, the resultant vector represents the mean difference angle for a single participant (see Holden et al., 2010, for further details). Because each participant’s data represent a mean error angle across cities, it was appropriate to analyze them using Hotelling’s one-sample second-order analysis of angles (Zar, 1996; *second-order* refers to the averaging across the items). This analysis indicates whether the participants within a group erred in a significantly directional fashion. We analyzed differences between groups using Hotelling’s two-sample analysis of angles. We also first transformed the data so that the direction of a vector was

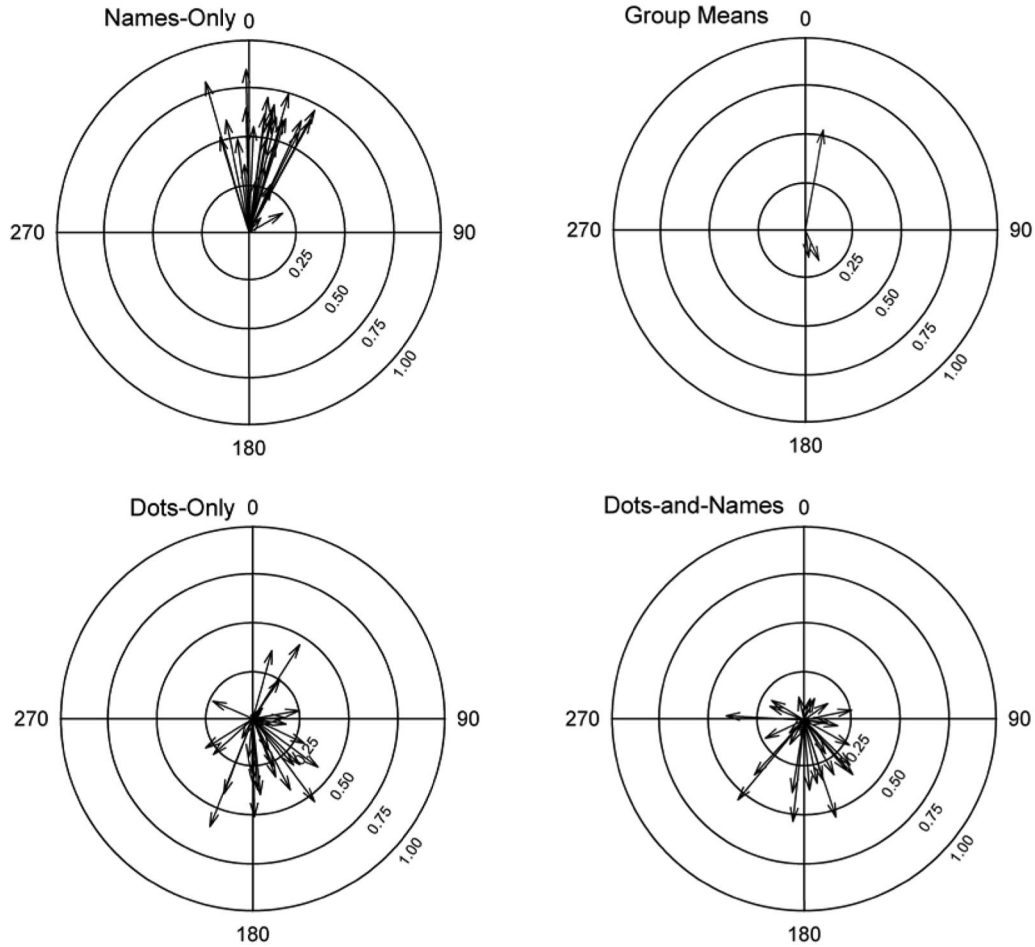


Figure 3. Vector graphs of the Alberta data in Experiment 1. Each vector is the average for one participant; the direction of the vector is that participant's average direction of error across the 26 dots/cities; the magnitude (length) of the vector is a measure of the consistency of angular error toward the centroid of the polygon (0°). More participants in the names-only condition erred consistently toward the centroid than in the other two groups.

0° when it pointed directly to the centroid, 180° when it pointed directly away from the centroid, and between 0° and 360° when it pointed elsewhere.

Visual analyses. We first examine the data visually, because the results are so striking. Figure 1 shows the data for Alberta and Figure 2 for California, plotted in Euclidean pixel coordinates; the 0,0 point is at the upper left because that is where it was on the computer screen. It is clear that the estimates from the dots-only and dots-and-names conditions were nearly identical and very accurate, whereas estimates from the names-only conditions were inaccurate but *not random*. Furthermore, in the dots-only and dots-and-names conditions, the estimates moved away from the center of mass of the polygons, whereas in the names-only condition, they moved toward the centroids. Finally, in the names-only conditions, it is possible to discern within-province/state subregions from the clustered pattern of distorted locations, whereas these are not observed in the other two conditions. In Alberta there appeared to be four regions: northern, middle, southern, and the Rocky Mountains (in the lower left). In California, there appeared

to be at least two regions: coastal and central and perhaps northern and southern.

Figures 3 and 4 corroborate the conclusions about the directionality of the errors; they are the average second-order error vectors of each individual participant in the three conditions and then the mean vector across all participants. Only the names-only condition has second-order error vectors that point on average toward the centroid. Together, these findings imply that different mechanisms were being used to remember locations across the two types of conditions, one type including the two conditions potentially based on only short-term visual memory of the dots and the second type including the condition based only on extant beliefs in long-term memory. These conclusions are supported by the analyses below.

Directional errors in estimating locations. All three Alberta groups had directional errors that were significant (see Table 1 and Figure 3). However, it is clear from Figure 3 that only the names-only group had estimates whose direction was approximately toward the centroid and that the errors in the other two groups, although significant, were generally small. Importantly, Hotell-

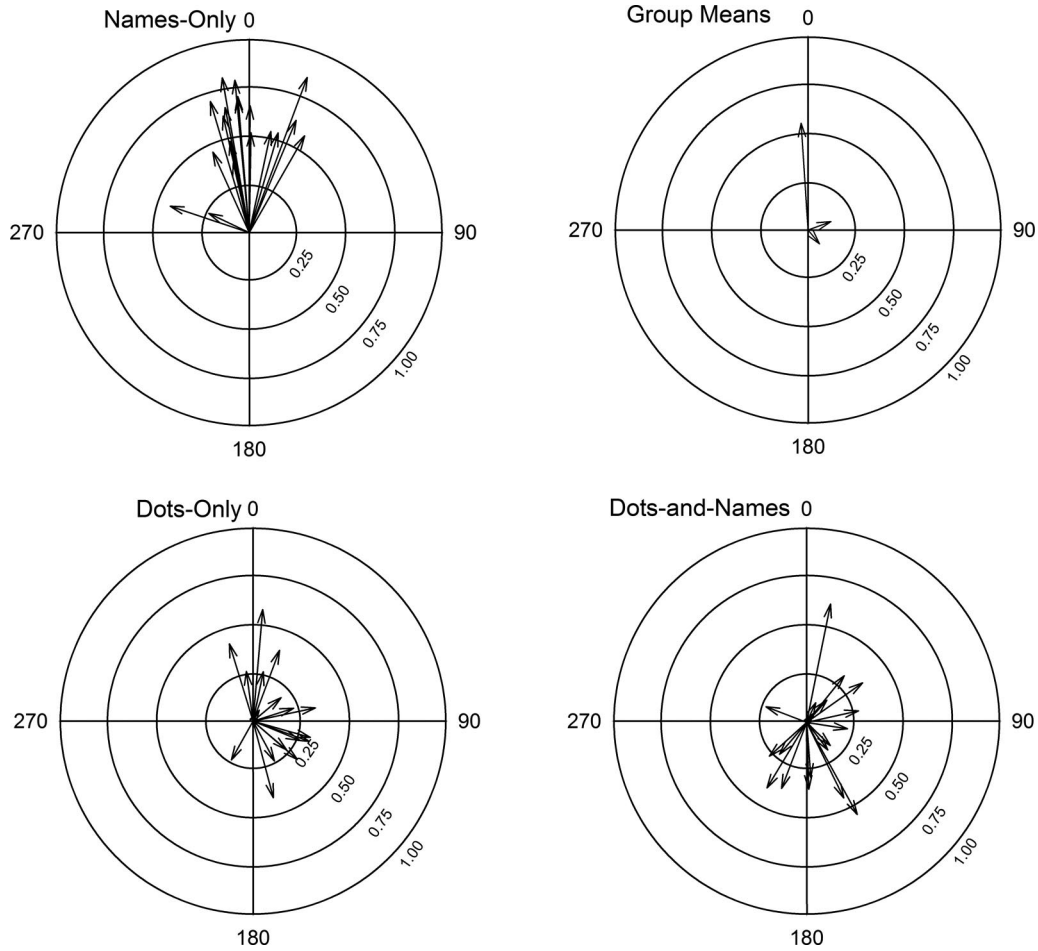


Figure 4. Vector graphs of the California data in Experiment 1. Each vector is the average for one participant; the direction of the vector is that participant's average direction of error across the 26 dots/cities; the magnitude (length) of the vector is a measure of the consistency of angular error toward the centroid of the polygon (0°). More participants in the names-only condition erred consistently toward the centroid than in the other two groups.

ing's two-sample second-order analysis of angles showed that there was no directional difference between the dots-only and dots-and-names groups ($F < 1.0$), but there were significant differences between both the dots-only group and the names-only group, $F(3, 57) = 75.32$, $M = 181.92^\circ$, confidence interval (CI)

[177.86°, 185.93°], $d = 28.47$, and between the dots-and-names and names-only groups, $F(3, 57) = 92.22$, $M = 186.50^\circ$, CI [182.17°, 190.91°], $d = 34.85$. Because both mean differences were close to 180°, the interpretation is that, whereas the names-only group erred in the direction of the centroid, the dots-only and

Table 1
Means, 95% Confidence Limits, F-ratios, and Effect Sizes for All 26 Stimuli in Experiment 1

Groups	M - CI	M	M + CI	F	d
Alberta					
Dots-Only	115.3°	156.0°	182.5°	$F(2, 28) = 11.3$	4.26
Dots-and-Names	136.4°	170.0°	211.5°	$F(2, 28) = 7.86$	2.97
Names-Only	3.30°	10.15°	17.40°	$F(2, 28) = 121.32$	45.86
California					
Dots-Only	13.5°	69.9°	134.19°	$F(2, 16) = 8.71$	4.11
Dots-and-Names	218.29°	140.5°	175.8°	$F(2, 17) = 2.30, p = .13$	1.05
Names-Only	342.7°	356.3°	7.8°	$F(2, 17) = 87.47$	40.13

Note. CI = confidence interval.

dots-and-names groups tended to err in approximately the direction *opposite* to that of the centroid.

For the California data, the dots-only group and the names-only groups had significant directional errors, but the dots-and-names group did not (see Table 1 and Figure 4). Once again, Hotelling’s two-sample second-order analysis of angles showed that there was no directional difference between the dots-only and dots-and-names groups, $F(3, 37) = 1.55, p = .25$, but there were significant directional differences between the dots-only group and the names-only group—the mean difference was $163.13^\circ, F(3, 37) = 27.82, CI [154.99^\circ, 170.67^\circ], d = 9.15$ —and between the dots-and-names and names-only groups—the mean difference was $171.40^\circ, F(3, 38) = 41.12, CI [163.89^\circ, 178.12^\circ], d = 13.33$. Once again it is clear that the names-only group was the only group that erred in the direction of the centroid of the polygon; the other two groups erred in directions that were closer to 180° .

Errors in estimating locations of best known cities. The three groups from Alberta did not differ from each other in their self-rated knowledge of all 26 cities, $F(2, 87) < 1.00$. The means for the dots-only, dots-and-names, and names-only groups were 4.0, 4.0, and 3.7, respectively. For the Californians the means were 3.1, 4.0, and 3.3, respectively, $F(2, 53) = 3.22, \eta_p^2 = .108$, and in this case post hoc analyses showed that the mean rated knowledge for the dots-only and names-only groups were less than for the dots-and-names group. However, for both Albertans and Californians, the mean ratings in all groups were roughly at the midpoint of the scale. Because the knowledge ratings were taken after the estimates, they could not have influenced them directly. Nevertheless, we correlated the post hoc knowledge ratings and the Euclidean distance error between actual and estimated locations separately for each participant and then averaged across participants within each group (we used the distance errors because the vectors are already averaged over cities). Not surprisingly, for both countries, there was virtually no correlation for the dots-only and dots-and-names groups (for Alberta, the correlations were .03 and .06, respectively, and for California they were $-.12$ and $-.13$, respectively). This is not surprising because there was hardly any error variance in these groups. In contrast, the correlations for the names-only groups were $-.36$ for Alberta ($p < .10$) and $-.49$ for California ($p < .05$). The direction of the correlations for the names-only groups, who could only use their extant knowledge in both tasks, is what one would expect: Those cities for which it was claimed to have high knowledge had distance errors between

actual and estimated locations that were smaller than those cities for which the knowledge claimed was low.

Nevertheless, because the overall mean self-rated knowledge across cities was at the midpoint of the scale, we redid the vector analyses reported above using the five cities that both Albertans and Californians self-rated as the most well-known on the grounds that these cities would have much less uncertainty associated with them. As noted, we used five cities because these happened to be the same cities for many of the participants; more than five “well-known” cities became more idiosyncratic across participants. This analysis thus equalizes across groups any effects that familiarity per se might have implicitly had on the estimates.

We first sorted the data on knowledge ratings, and then computed the average error vectors, as above, for the five most well-known cities for each participant in each group. We did not eliminate outliers because the data were now relatively sparse (i.e., five observations per participant). The mean knowledge ratings for the Albertans in the dots-only, dots-and-names, and names-only groups were now 7.9, 7.9, and 7.6, respectively. Each of these means were significantly above the midpoint of the scale, $t(29) = 31.31, 40.57, \text{ and } 26.77$, respectively. The differences between groups were not significant ($F < 1.0$). For the Californians, the means for the dots-only, dots-and-names, and names-only groups were 7.3, 8.3, and 7.7, respectively, and these three means were also significantly above the scale’s midpoint, $t(17) = 22.46, t(18) = 51.07, \text{ and } t(18) = 29.82$, respectively. Here, the differences between groups were significant, $F(2, 53) = 3.70, \eta_p^2 = .12$, showing the same pattern as the overall knowledge ratings.

With respect to the estimated locations, for the Albertans, the mean directional error for the names-only group for the five most well-known cities was still tending toward the centroid, and the mean directional error for the other two groups was not (see Table 2). The dots-and-names group did not have a significant directional error, indicating that when the cities were well known, that group was quite accurate but still erring in a direction away from the centroid in many cases. This makes sense because most of the well-known cities in Alberta are located south of its centroid. Hotelling’s second-order difference between samples showed that the dots-only group now differed from the dots-and-names group; the mean vector difference was $106.13^\circ, F(2, 28) = 3.72, CI [67.70^\circ, 151.72^\circ], d = 1.19$; the names-only group differed from both the dots-only group and the dots-and-names group. For the dots-only group the mean difference was $172.16^\circ, CI [163.37^\circ,$

Table 2
Means, 95% Confidence Limits, F-ratios, and Effect Sizes for the Five Self-Rated Most Familiar Cities in Experiment 1

Groups	M - CI	M	M + CI	F	d
Alberta					
Dots-Only	60.2°	110.5°	172.2°	$F(2, 28) = 8.74$	3.31
Dots-and-Names	224.6°	173.9°	150.6°	$F < 1.00$	0.01
Names-Only	356.6°	7.8°	18.36°	$F(2, 28) = 97.23$	36.75
California					
Dots-Only	350.0°	37.0°	37.2°	$F < 1.00$	0.05
Dots-and-Names	230.2°	200.0°	157.2°	$F < 1.00$	0.03
Names-Only	336.6°	2.4°	23.0°	$F(2, 17) = 20.18$	9.26

Note. CI = confidence interval.

179.89°, $F(2, 28) = 41.90$, $d = 15.84$; and for the dots-and-names group the mean difference was 187.44°, CI [178.83°, 195.52°], $F(2, 28) = 26.10$, $d = 9.86$. Notably, the mean vector difference between the names-only group and the other two groups was again approximately 180°.

For the Californians, neither the dots-only nor the dots-and-names group had significant directional errors, but the names-only group did (see Table 2); their mean error again pointed toward the centroid. In addition, whereas the dots-only and dots-and-names groups did not differ from each other ($F < 1.00$), the average vector difference between the names-only group and the dots-only group was significant; the mean was 178.89°, $F(2, 28) = 6.46$, CI [159.74°, 192.71°], $d = 2.12$. The names-only group also differed from the dots-and-names group; the mean difference was 183.63°, $F(2, 28) = 10.60$, CI [167.50°, 196.81°], $d = 3.44$. In sum, and surprisingly, using the most well-known cities in each state/province did not change the pattern of the data.

Configural knowledge. We conducted bidimensional regressions on the xy estimates for each participant using the actual values as the independent variable and the estimated values as the dependent variable (Friedman & Kohler, 2003). When computed in this manner, the bidimensional regression (BDR) coefficient is interpreted as the amount of variance in the estimated locations that can be accounted for by the Euclidean configuration of actual locations. Thus, even if, for example, the names-only groups had relatively inaccurate estimates on an individual city or dot level, they might show some accuracy in terms of the overall configuration of the cities.

We analyzed the Fisher-transformed regression coefficients in a one-way analysis of variance. For both Alberta and California, the effect of group was significant, $F(2, 87) = 540.504$, $\eta_p^2 = .926$, and $F(2, 53) = 289.47$, $\eta_p^2 = .916$, respectively. The back-transformed means (i.e., “undoing” the Fisher transformation and yielding the means of the original correlation coefficients) for Alberta for the dots-only, dots-and-names, and names-only conditions were .993, .994, and .657, respectively, and for California they were .997, .996, and .529, respectively. Clearly, the actual configuration of locations that were presented in the dots-only and dots-and-names group accounted for nearly all of the variance in the responses; whereas in the names-only group, the configuration accounted for between 28.0% and 43.2% of the variance. The estimates in the names-only group were clearly reflecting a sizable influence of the prototype location. However, responses were also clearly not random and at some level, claiming that they were given a weighting of zero in the dots-and-names group seems to be an heroic effort to save a Bayesian account of the data. Further, the fact that the dots-only and dots-and-names groups erred in a direction opposite to that of the prototype argues against the Bayesian expectation, and for the dots-only group is the first time, to our knowledge, that this has occurred for purely perceptually based stimuli.

Discussion

To the extent that Bayesian combination implies that directional errors should be in the direction of the center of mass of a polygonal shape, the data from Experiment 1 showed that this occurred only in the names-only condition. The dots-only and dots-and-names conditions were both very accurate in terms of the actual magnitude of error per dot and in terms of configurational

error. Further, to the extent that directional errors did occur, they were in a direction *away* from the centroid. Thus, these two conditions do not appear to have behaved in a Bayesian manner. Further, the dot-and-names condition did not appear to have been influenced by information from extant long-term memory, because its errors were not in between those of the dots-only and names-only groups. We consider the theoretical implications of this finding further in the General Discussion.

Experiment 2

In Experiment 1, because we used real-world stimuli, the actual locations were not uniformly distributed around the map polygons, as they were in Huttenlocher et al. (1991) and, roughly, in Holden et al. (2010). Huttenlocher et al. (1991) originally suggested that erring toward the prototype might depend on the distribution of exemplars; Huttenlocher et al. (2004) suggested that it did not matter; but more recent studies have shown that the distribution of dots can matter for accuracy (e.g., Hund & Spencer, 2003; Spencer & Hund, 2002; Spetch, Friedman, Bialowas, & Verbeek, 2010).

As it happens, most Albertan cities are in the southern half of Alberta; most Californian cities are in the western half of California. To address the possible confound(s) that might be caused by uneven distributions of dots in irregular polygons, we conducted an experiment in Alberta in which we added 26 more dots and distributed them quasirandomly in the “empty” spaces on the map polygon. We reran the dots-only group with all 52 dots and a second group of participants in which the 26 dots that were cities were labeled with their names, as in the previous dots-and-names group; indeed, they were the same cities. The remaining 26 dots were not cities and had no label. We refer to this condition as the *mixed condition* to distinguish it from the former dots-and-names condition. We did not run a names-only condition because half the locations tested had no actual city to name.

Method

Participants and design. The participants were 64 new volunteers (29 men, 34 women) from the University of Alberta Department of Psychology participant pool; all were from Alberta. They were randomly assigned to either the dots-only condition or the mixed condition. The data from three women and one man were replaced due to experimenter or computer error, leaving 30 participants per group. Half the dots in each condition were located where actual cities were located, and the other half were distributed quasirandomly within spaces that were otherwise empty.

Stimuli and procedure. The dots-only condition was run with the identical procedure as in Experiment 1, but with 26 additional dots added to the stimuli. In the mixed condition, participants saw the same 52 dots. Participants in this condition were told, truthfully, that when a dot appeared with a city’s name, then that was the location of the city and that when the dots did not have a name that appeared, there was no city there. They were further asked to pay equal attention to dots with and without city names; that they were equally important. Otherwise the procedures were identical to those of Experiment 1, including the mask and the timing.

Results

Directional errors in estimating locations. We first tested for outlier responses in the same way as in Experiment 1, separately for dots representing cities and those representing noncities; these accounted for 1.96% of the possible responses. We next converted each participant's remaining responses to error vectors, as before, and summed these separately for locations that corresponded to cities versus those that corresponded to noncities. Note that for the group instructed that they would be trying to remember dot locations, the within-participant factor was maintained to keep the analyses parallel; it does not represent a meaningful manipulation to the participants, in that all dots simply appeared to be dots.

We analyzed the vectors for the dots that represented cities separately from those that did not for each of the between-participants groups (dots only and mixed; see Figure 5). For the 26 dots that represented cities, there were significant directional errors for both groups: For the dots-only group, the mean was 175.12°, $F(2, 28) = 5.78$, CI [125.10°, 212.43°], $d = 2.19$, and for the mixed group, the mean was 156.01°, $F(2, 28) = 5.40$, CI [76.62°, 188.68°], $d = 2.04$. However, Hotelling's two-sample second-

order analysis of angles showed that there was no directional difference between the groups ($F < 1.0$). This pattern of means and significance is similar to that found for the same cities in Experiment 1 and shows that the means tended to point away from the centroid.

For the 26 dots that did not represent cities there was a similar pattern of means and significance (see Figure 5): For the dots-only group, the mean directional error was 216.41°, $F(2, 28) = 3.86$, CI [161.05°, 301.390°], $d = 1.46$, and for the mixed group the mean was 205.82°, $F(2, 28) = 10.92$, CI [170.94°, 249.73°], $d = 4.13$. In addition, there was no difference between the two groups in Hotelling's two-sample analysis. These data again replicate the data for the dots only and dots-and-names groups of Experiment 1 rather well.

Knowledge ratings. The groups did not differ from each other in their mean rated knowledge about the 26 cities, $F(1, 58 = 2.51)$, $p = .12$. The mean rating for the STM group was 4.0 for the dots-only group and for the mixed group it was 3.4.

Configurational knowledge. Each participant's data were submitted to two BDR analyses, separately for the dots that rep-

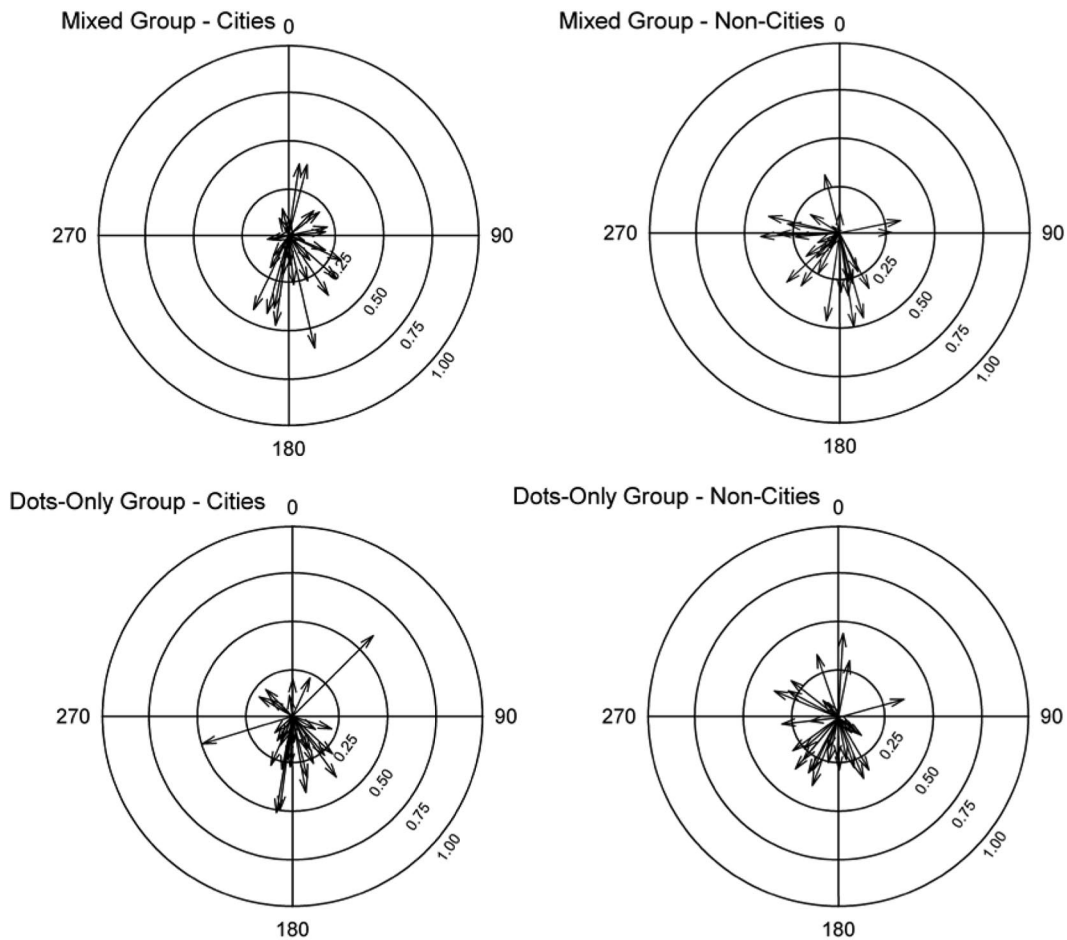


Figure 5. Vector graphs of the Alberta data for Experiment 2. Each vector is the average for one participant; the direction of the vector is that participant's average direction of error across the 26 dots that were either cities or not cities; the magnitude (length) of the vector is a measure of the consistency of angular error toward the centroid (0°).

resented cities and those that did not. The back-transformed mean regression coefficients for dots that were cities versus those that were not were functionally identical (.994 and .996) and similar to those obtained in Experiment 1. The coefficients indicate that virtually all the variance in the estimated locations was due to the actual locations.

Discussion

It does not appear that the distribution of dots was responsible for the absence of a difference between the dots-only and dots-and-names groups of Experiment 1. Nor does it seem that the distribution prevented generally accurate performance and errors directed away from the centroid of the polygon. Thus, this experiment serves as a replication, both within and between-participants, of the data from the dots-only and dots-and-names groups of Experiment 1.

Experiment 3

Although both Experiments 1 and 2 had 4 s of dynamic masking stimuli between the stimulus and response polygons, the polygons remained in the same place on the screen. To eliminate the possibility that participants could perform the task by simply focusing on each stimulus location, Experiment 3 was a replication of Experiment 2, in which the second (response) polygon was moved around the screen. If we replicate the results of Experiment 2 for the dots-only and mixed groups, then we will have eliminated this potential explanation of the results.

Method

Participants. We tested 60 new volunteers (32 men, 28 women) from the University of Alberta participant pool; all were from Alberta. They were randomly assigned to either the dots-only or the mixed condition (30 per condition).

Design, stimuli, and procedure. The design and procedure were identical to those used in Experiment 2. The only difference between experiments was that the first (stimulus) polygon appeared in the center of the screen, as before, but the second (blank) polygon could now appear at random in one of six places on the computer screen. None of the six was in the center of the screen so participants could not rely on screen position to facilitate responding.

Results

Neither of the groups had significant directional errors, whether we tested just the dots that represented cities or just the dots that did not represent cities, as can be seen in Figure 6. For all 52 dots, the mean for the dots-only group was 181.16° , $F(2, 28) = 2.41$, $p = .11$, and for the mixed group it was 143.17° ($F < 1.00$).

There were generally no differences between groups in the BDR analysis either, $F(1, 58) = 1.91$, $p = .17$; the back-transformed means for the dots-only group for city and noncity dots (again, a distinction of no meaning to participants) were .981 for each, and for the mixed group they were .985 each. Thus, participants were very accurate at remembering the configurations of these dots (presented one at a time), and their accuracy could not be based on a specific position on the screen. Moreover, even though mean

directional errors were not significantly consistent, there was no indication that they tended either toward or away from the centroid of the polygon. That is, there is no indication of bias in any direction.

Discussion

Moving the polygon around the screen at test for the dots-only and mixed groups did not affect overall accuracy; actual locations were still accounting for approximately 96% of the variance in the estimated locations, but there was no consistency to the directional errors. This result is also inconsistent with expectations of both the CA model and Bayesian combination.

General Discussion

In three experiments participants made location estimates of dots inside irregular schematized polygonal frames. For eight independent groups, performance in the dots-only and dots-and-names conditions was extremely similar in kind. This finding was true both between-participants (Experiment 1) and within-participant (Experiments 2 and 3). Performance in these two conditions was very accurate at a high level of precision, and the directional errors that existed were generally away from the centroid if they were significant. It was only in the two names-only conditions that participants consistently made consistent errors in the direction of the centroid of the polygons. Furthermore, Experiment 2 ruled out the uneven distribution of the dots (cities) as a possible explanation for the similarity between the dots-only and dots-and-names conditions, and Experiment 3 ruled out absolute screen position as a possible explanation.

We already have evidence that estimates can be biased away from a category's prototype when location information is retrieved from long-term memory (Friedman, 2009; Friedman & Montello, 2006) and, in the present case, when the information is retrieved from visual short-term memory (in dots-only and perhaps also in dots-and-names). The results obtained from the dots-only and dots-and-names conditions in Experiment 1 are not expected on a Bayesian analysis and Experiments 2 and 3 rule out two obvious possible confounds. The names-only conditions fit a metrics-and-mapping analysis as well as a Bayesian model, but metrics and mapping does not deal with estimates from short-term memory. We are unaware of any particular model that can explain the data from all three conditions in Experiment 1 without further assumptions.

One very important aspect of the data is that in none of the dots-only or dots-and-names conditions did the direction of error point toward the centroid of the polygon, even when we "evened-out" the distribution of dots in Experiments 2 and 3. In addition, the data from both these conditions were quite accurate. It was only in the two names-only conditions in Experiment 1 that the errors pointed toward the centroid. Thus, we think that Bayesian combination occurred only in the latter condition, if "moving toward the centroid" is a hallmark of such combination.

It would thus seem that only the data from only the names-only condition support the prediction of the category adjustment model and the use of Bayesian combination for location estimates, even when we examined just the five most familiar places in each region. The overall directional errors for the names-only groups in

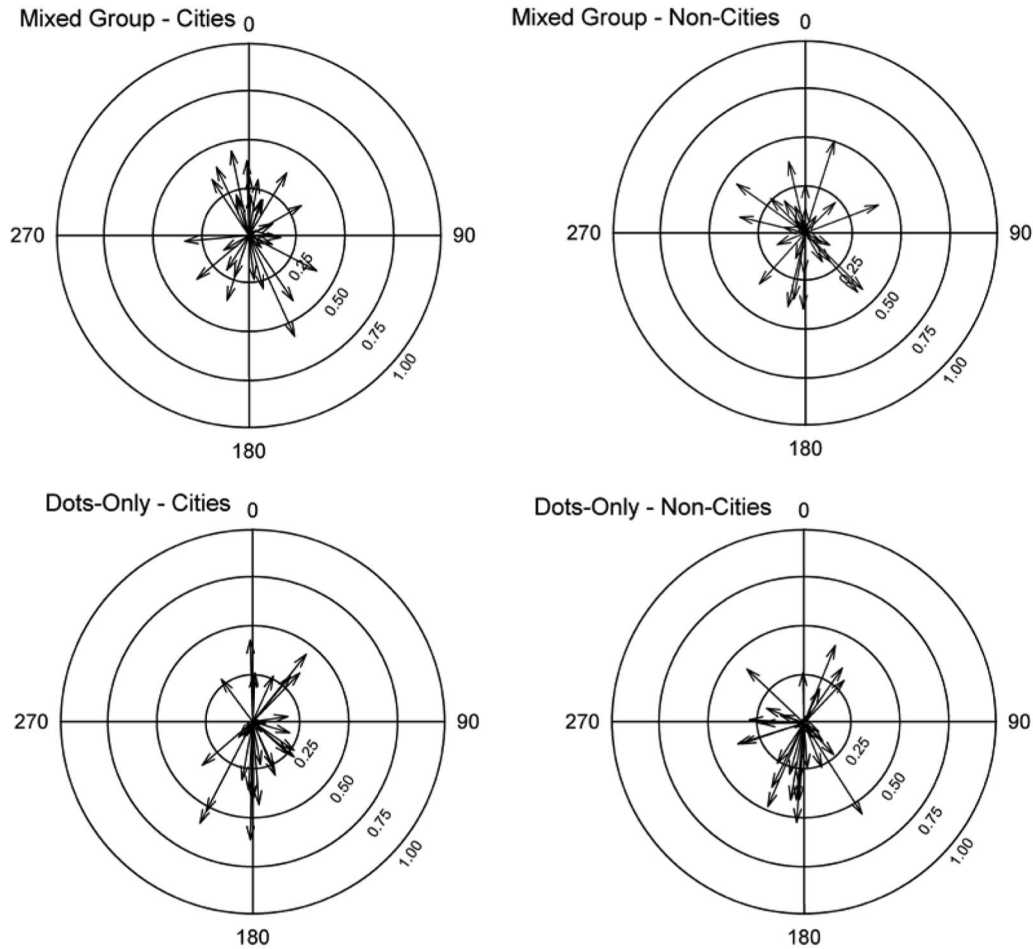


Figure 6. Vector graphs of the Alberta data for Experiment 3. Each vector is the average for one participant; the direction of the vector is that participant's average direction of error across the 26 dots that were either cities or not cities; the magnitude (length) of the vector is a measure of the consistency of angular error toward the centroid (0°).

Experiment 1 were definitely biased toward the center of mass of both the Albertan and Californian polygons. Presumably, the a priori knowledge about city locations was relatively vague, compared with the accurate outline of the polygon participants were to respond in. Given the outline, the centroid could be computed either “online” or also known a priori, and estimates therefore tended to move toward it. The names-only conditions in Experiment 1 thus produced estimates that were, indeed, in error in the direction of the centroid of the regularized map outlines. This type of “erring toward the prototype” (or learning it more quickly than other instances) is found in many perceptual and semantic domains in which the instances can be presumed to exist in long-term memory (cf. Rosch, 1973; Rosch & Mervis, 1975), even when the prototype is not central (cf. Rosch, 1973).

In previous research on location estimates of cities in which an accurate category boundary was not given (e.g., Friedman & Brown, 2000a, 2000b; Friedman & Montello, 2006), the category-adjustment model also did not do so well at predicting performance. In these studies, it could be inferred from the participants' responses that the category/regional boundaries

were quite incorrect, rendering estimates of the cities equally incorrect. The existence in long-term memory of incorrect category boundary information is itself contrary to the assumptions of the category adjustment model; however, the “inheritance” of the incorrect absolute locations by the items within each category is what one would normally expect to happen in a category-item hierarchical relationship (cf. Rosch & Mervis, 1975; Stevens & Coupe, 1978).

Thus, it would appear that one factor that plays a role in whether long-term memory information about items is used and whether errors tend toward the direction of the prototype is whether the category boundaries are given and are accurate. In the present context, since both attributes were true, errors tended toward the prototype. When the boundaries must be inferred (for real-world categories), the errors tend toward *regional* prototypes (as they have in our previous study; Friedman, 2009, Experiment 3). There is a hint that this tendency even occurred in the present case, but as we do not know for sure what the regional boundaries are (see Montello & Friedman, 2012), we cannot pursue this question with the current data.

In contrast to the names-only conditions, responses in the dots-only and dots-and-names conditions were extremely similar to each other and quite accurate, despite the presence of a relatively long and distracting masking stimulus and/or movement of the response frame around the screen. These results imply that prior knowledge *about city locations* was not being used to make the estimates in the dots-and-names conditions of the present experiments, despite the availability of this knowledge (as shown in the names-only condition).

If prior long-term memory knowledge was being used in the dots-and-names condition, the estimates for this group should have been less accurate than they were, because the names-only groups' estimates were inaccurate and directed toward the prototype. So we did not obtain what would appear to be a Bayesian combination of category-level and long-term item-level knowledge in the dots-and-names condition; rather, it appears that category-level information, if it was used at all, was combined with information in short term-visual memory traces from the dots themselves in these conditions. Thus, in this sense the CA model does not do well with the present data: First, there is no evidence of a Bayesian combination of knowledge *from long-term memory* in the dots-and-names condition, even though the data from the names-only condition was clearly not random; on a Bayesian account this outcome would indicate that the long-term memory knowledge had a weight of zero, which at least seems odd (cf. Figures 1 and 2). It seems odd because the information available in LTM is rich and learned over the lifespan from a variety of sources. It is also odd in the sense that, as noted in the introduction, it is not readily predictable when a cue will be weighted high and when it will not. Further, this study was designed to be as close to a "perfect" cue combination situation as possible, insofar as we had two single-cue conditions (dots-only and names-only) and a combined condition (dots-and-names), and there was no ambiguity in the latter condition regarding the fact that the cues were in fact from the same source. In addition, it is clear that there was relatively decent knowledge in the names-only condition, so this condition does not meet the criteria for giving it a weight of zero. Thus, information from the dots-only and names-only conditions should have shown evidence of having been combined. It would appear that these facts, taken together, make Bayesian analysis *as a theoretical approach to location estimation* not particularly useful. Perhaps even more important, in both the dots-only and dots-and-names conditions, there was no indication at all that participants erred in the direction of the centroid, which goes against the prediction of the CA model. And that finding was replicated three times with eight independent groups of participants.

So can the category adjustment model predict this outcome a priori? Well, it might be conjectured that the participants knew (subconsciously or otherwise) that the representations of the response frame and the location of the presented dots/cities were more accurate than what they had stored in long-term memory (for the dots-and-names group), so the long-term memory information was given a small or zero weighting in the dots-and-names group (even though the information per se was not random), and it was only the presented information (frame as category and dot as item) that was combined. Indeed, the response frame certainly was optimal (as a guide to dot location). And perhaps the reason the errors did not point toward the centroid had something to do with the "irregularity" of the polygonal shapes and the fact that all the

edges were straight (unlike Holden et al., 2010); this might have allowed for better memory of the relation between a given dot and one or more edges. But this post hoc explanation has the flavor of allowing a Bayesian model to predict almost anything; we prefer to conjecture that Bayesian combination was simply not used in the dots-only and dots-and-names conditions. Instead, for reasons we do not yet understand, the memory for locations per se was very accurate. The frames might have been used as a cue to location, but we do not think they were combined with locations because we did not see movement toward the prototype. However, compared with a circle or a very irregular polygon, these particular irregular but relatively simple and straight-lined polygons may have been helpful in this way.

Because of our observed data, we believe that Holden et al.'s (2010) conclusion that both conceptual and perceptual knowledge was being used to categorize their upright scenes is true at a very different level than our data are. Theirs was a categorization task; this is a different "level" of conceptual information than what we are talking about when we discuss knowledge about locations and other aspects of cities extant in long-term memory in the present case.

Another finding in our study indicating that the names-only condition was qualitatively different than the other two conditions was that in that condition there were apparently four regions in the Alberta data and at least two (or more) in the California data. This is theoretically interesting because these are the "real" regions (categories) that participants think of when they make their estimates from long-term memory, and it is prior beliefs about the location of the regions that contribute to bias in the estimates made from long-term memory. No regions of this sort were observed in the dots-only or dots-and-names groups. Indeed, for a different purpose than the present study, we conducted another experiment (Montello & Friedman, 2012) in both California and Alberta, in which participants were given actual maps of each state/province and asked to draw regions on the map (the instructions defined regions as being "pieces of the earth's surface that enclose fairly similar or homogenous areas. The similarity may be based on natural or cultural variables, or any combination."). Fifty out of 67 (74.6%) of the Albertans drew three to four regions labeled, for example, "north, central, south, and Rocky Mountains" (see Figure 1). The Californian data were more complex: Out of 141 participants 59 divided the state from north to south with between two and four regions, and 39 divided the state from east to west with between two and four coastal/inland regions, for a total of 63.8% using one of these schemes. The remaining participants identified particular cities as regions. The north-south versus east-west divisions can be seen to a certain extent in Figure 2. Some participants in both countries added further details (e.g., they put a small square around the location of Edmonton or labeled the Bay area in California), but by and large these obtained representations were similar in kind to the regions obtained (by eye) in the names-only conditions of Experiment 1 (see Figures 1 and 2). We could not precisely analyze the subregions in the present data or determine their centroid because we do not know their precise borders and, indeed, have evidence that the borders are likely to be imprecise (Montello & Friedman, 2012).

It is also interesting that we found at the state and provincial scale in the names-only groups the same kind of gaps between regions that we and others have previously observed at the level of

the continent (e.g., Friedman, 2009; Friedman & Montello, 2006), the city (Hirtle & Jonides, 1985), and the college campus (Uttal et al., 2010). Thus, for real places the data from several geographic scales suggest that participants form regions in their long-term memory representations and once formed, those representations may be biased in some way; the bias is inherited by the items believed to belong within the regions (cf. Stevens & Coupe, 1978). Clearly, in the present case, although participants claimed to be quite familiar with some of these cities, information from long-term memory appeared not to have been given any weight in making location estimates in the dots-and-names conditions.

Studies, such as the original work by Huttenlocher et al. (1991) and Wedell et al. (2007), found biases in location estimates in a direction toward the spatial centroid (or centroids). However, currently the category adjustment model lacks an a priori theoretical basis for identifying the spatial centroid as the prototype, as opposed to other possible measures of spatial central tendency, such as the center of the maximum enclosed circle, the centroid of the convex hull, and so on (Fotheringham, Brunson, & Charlton, 2000). It is unclear what recall patterns would occur with extremely concave polygons, the center of mass of which can even be outside the boundaries. Furthermore, in both the Alberta and the California data, we found that a single prototype point attracted biased LTM recall. This contrasts with previous research with regular, convex polygons that finds multiple points, often four, serving as prototype attractors for biased recall (e.g., Wedell et al., 2007). Considerations such as these suggest that the CA model would benefit from the development of a stronger theoretical basis for identifying a priori how many points will serve as prototype locations in a given region of a given shape, where they are likely to be located, and why. It does not seem obvious why the center of mass would be the attractor for some polygons but not others.

We do not know why circles, regular polygons, and irregular regions in natural scenes produced such markedly different patterns of bias than that observed in the irregular but schematic polygons used here. However, it is clear that the data from the dots-only condition are easily replicated and are very similar to the data from the dots-and-names condition (also easily replicated), in which participants could have used long-term item information in their responses if they so chose. We believe, therefore, that the replication across two sites, two polygons, and two participant populations is a strong indicator that the data seriously constrain the CA model.

As noted, with respect to the present data, this study is not the first time that estimates have tended away from the centroid of a region (e.g., Friedman & Brown, 2000a, 2000b; Friedman & Montello, 2006). However, prior studies have used only extant LTM knowledge for making location estimates. To our knowledge, the present data represent the first time that estimates from perception alone (i.e., dots-only) have exhibited this behavior. This is a very serious constraint on the CA model and, indeed, is contrary to the model's predictions.

We do not wish, at this point, to construe a model to explain the data from the dots-only and dots-and-names conditions; it would clearly be ad hoc. Nor do we wish to claim that one of the models we mentioned previously (e.g., viewpoint interpolation; generalization) would necessarily do better. We do not know enough about how prototype locations are chosen in irregular polygons for a short-term visual memory task. It is plausible that participants in

the dots-only and dots-and-names conditions used asymmetries in the boundaries of the polygons, rather than their centroids, as locations from which to recall dot locations. One obvious approach to the issue of when to expect movement toward a centroid and when to expect movement away from it for perceptually based estimates would be to test location memory systematically using a continuum of irregular polygonal frames with specific features (e.g., concavity; amount of asymmetry). Further, as noted above, any theory of spatial location estimates needs a basis for predicting, given a certain shape, when a spatial centroid is likely to be the center of its mass versus when it is likely to be defined in some other manner, particularly for concave polygons.

It is also obvious that more research is needed using geographic categories in which there are different strengths of long-term memory location knowledge (cf. Uttal et al., 2010) to sort out some of these theoretical issues. It is also plausible that reliance on long-term knowledge will vary as a function of the relative ease of the task, such as reproducing or merely recognizing the presented location (e.g., Sampaio & Wang, 2009), and other contextual aspects of the task, such as whether characteristics of places besides location (e.g., linguistic traditions) need to be considered. For example, Uttal et al.'s (2010) participants took approximately a year of navigation (in addition to being exposed to maps, verbal directions, etc.) to learn the (linguistically based) categories of a campus and have those categories affect location judgments of the buildings within them. Further, the point that Bayesian models may be too flexible should not be lost: It is too facile to say, post hoc, that a given dimension was weighted or not. That is, we could claim that in the dots-only and dots-and-names conditions the long-term memory location information was simply not given any weight after the data were available, but it would presumably be better to be able to do that a priori. Equally, we could claim that in the names-only condition, long-term memory information had to be given weight (and the regression coefficients reflecting configurational knowledge were certainly not zero), but again, it would presumably be better to be able to do this a priori. Thus, why extant knowledge was weighted zero in the dots-and-names condition is not obvious. It seems clear that more research using real-world preexisting spatial categories is essential to resolve the issue of when Bayesian combination is used (and how) and when it is not.

References

- Brockmole, J. R., & Henderson, J. M. (2006). Using real-world scenes as contextual cues for search. *Visual Cognition, 13*, 99–108. doi:10.1080/13506280500165188
- Brown, N. R., & Siegler, R. S. (1993). Metrics and mappings: A framework for understanding real-world quantitative estimation. *Psychological Review, 100*, 511–534. doi:10.1037/0033-295X.100.3.511
- Cheng, K., Shettleworth, S. J., Huttenlocher, J., & Rieser, J. J. (2007). Bayesian integration of spatial information. *Psychological Bulletin, 133*, 625–637. doi:10.1037/0033-2909.133.4.625
- Collins, A., & Michaelski, R. (1989). The logic of plausible reasoning: A core theory. *Cognitive Science, 13*, 1–49. doi:10.1207/s15516709cog1301_1
- Duffy, S., Huttenlocher, J., & Crawford, L. E. (2006). Children use categories to maximize accuracy in estimation. *Developmental Science, 9*, 597–603. doi:10.1111/j.1467-7687.2006.00538.x
- Edelman, S. (1999). *Representation and recognition in vision*. Cambridge, MA: MIT Press.

- Fotheringham, A. S., Brunson, C., & Charlton, M. (2000). *Quantitative geography*. London, England: Sage.
- Friedman, A. (2009). The role of categories and spatial cuing in global-scale location estimates. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 94–112. doi:10.1037/a0013590
- Friedman, A., & Brown, N. R. (2000a). Reasoning about geography. *Journal of Experimental Psychology: General*, *129*, 193–219. doi:10.1037/0096-3445.129.2.193
- Friedman, A., & Brown, N. R. (2000b). Updating geographical knowledge: Principles of coherence and inertia. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 900–914. doi:10.1037/0278-7393.26.4.900
- Friedman, A., Kerkman, D. D., Brown, N. R., Stea, D., & Cappello, H. H. (2005). Cross-cultural similarities and differences in North Americans' geographic location judgments. *Psychonomic Bulletin & Review*, *12*, 1054–1060. doi:10.3758/BF03206443
- Friedman, A., & Kohler, B. (2003). Bidimensional regression: Assessing the configural similarity and accuracy of cognitive maps and other two-dimensional data sets. *Psychological Methods*, *8*, 468–491. doi:10.1037/1082-989X.8.4.468
- Friedman, A., Mohr, C., & Brugger, P. (2011). Representational pseudoneglect and reference points both influence geographic location estimates. *Psychonomic Bulletin & Review*, DOI 10.3758/s13423-011-0202-x
- Friedman, A., & Montello, D. R. (2006). Global-scale location and distance estimates: Common representations and strategies in absolute and relative judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 333–346.
- Friedman, A., & Waller, D. A. (2008). View combination in scene recognition. *Memory & Cognition*, *36*, 467–478. doi:10.3758/MC.36.3.467
- Goffaux, V., Jacques, C., Mouraux, A., Oliva, A., Schyns, P. G., & Rossion, B. (2005). Diagnostic colours contribute to the early stages of scene categorization: Behavioral and neuropsychological evidence. *Visual Cognition*, *12*, 878–892. doi:10.1080/13506280444000562
- Hirtle, S. C., & Jonides, J. (1985). Evidence of hierarchies in cognitive maps. *Memory & Cognition*, *13*, 208–217. doi:10.3758/BF03197683
- 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. doi:10.1037/a0019293
- Hund, A. M., & Spencer, J. P. (2003). Developmental changes in the relative weight of geometric and experience-dependent location cues. *Journal of Cognition and Development*, *4*, 3–38.
- Huttenlocher, J., Hedges, L. V., Corrigan, B., & Crawford, L. E. (2004). Spatial categories and the estimation of location. *Cognition*, *93*, 75–97. doi:10.1016/j.cognition.2003.10.006
- Huttenlocher, J., Hedges, L. V., & Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review*, *98*, 352–376. doi:10.1037/0033-295X.98.3.352
- Huttenlocher, J., Hedges, L. V., & Vevea, J. L. (2000). Why do categories affect stimulus judgment? *Journal of Experimental Psychology: General*, *129*, 220–241. doi:10.1037/0096-3445.129.2.220
- Matsukura, M., & Hollingworth, A. (2011). Does visual short-term memory have a high-capacity stage? *Psychonomic Bulletin & Review*, *18*, 1098–1104. doi:10.3758/s13423-011-0153-2
- McNamara, T. P. (1986). Mental representations of spatial relations. *Cognitive Psychology*, *18*, 87–121. doi:10.1016/0010-0285(86)90016-2
- Montello, D. R. (2003). Regions in geography: Process and content. In M. Duckham, M. F. Goodchild, & M. F. Worboys (Eds.), *Foundations of geographic information science* (pp. 173–189). London, England: Taylor & Francis. doi:10.1201/9780203009543.ch9
- Montello, D. R., & Friedman, A. (2012). *Using a high-resolution rating task to demarcate vague cognitive regions in geographic spaces*. Manuscript in preparation.
- Newcombe, N. S., & Huttenlocher, J. (2000). *Making space: The development of spatial representation and reasoning*. Cambridge, MA: MIT Press.
- Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics*, *16*, 283–290. doi:10.3758/BF03203943
- Rosch, E. (1973). Natural categories. *Cognitive Psychology*, *4*, 328–350. doi:10.1016/0010-0285(73)90017-0
- Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*, 573–605. doi:10.1016/0010-0285(75)90024-9
- Sampaio, C., & Wang, R. F. (2009). Category-based errors and the accessibility of unbiased spatial memories: A retrieval model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 1331–1337. doi:10.1037/a0016377
- Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, *237*, 1317–1323. doi:10.1126/science.3629243
- 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. doi:10.1037/0096-1523.33.4.871
- Spencer, J. P., & Hund, A. M. (2002). Prototypes and particulars: Geometric and experience-dependent spatial categories. *Journal of Experimental Psychology: General*, *131*, 16–37. doi:10.1037/0096-3445.131.1.16
- Spetch, M. L., Friedman, A., Bialowas, J., & Verbeek, E. (2010). Contributions of category and fine-grained information to location memory: When categories don't weigh in. *Memory & Cognition*, *38*, 154–162. doi:10.3758/MC.38.2.154
- Stevens, A., & Coupe, P. (1978). Distortions in judged spatial relations. *Cognitive Psychology*, *10*, 422–437. doi:10.1016/0010-0285(78)90006-3
- Thrash, T., Waller, D., & Friedman, A. (2011). *Enhanced prototypes do not necessarily lie in the extremes of psychological space*. Manuscript in preparation.
- Uttal, D. H., Friedman, A., Liu, L. L., & Warren, C. (2010). Learning fine-grained and category information in navigable real-world space. *Memory & Cognition*, *38*, 1026–1040. doi:10.3758/MC.38.8.1026
- 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. doi:10.1016/j.visres.2004.11.015
- Wedell, D. H., Fitting, S., & Allen, G. L. (2007). Shape effects on memory for location. *Psychonomic Bulletin & Review*, *14*, 681–686. doi:10.3758/BF03196821
- Zar, J. H. (1996). *Biostatistical analysis* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Received July 11, 2011

Revision received January 29, 2012

Accepted February 16, 2012 ■