

The Role of Categories and Spatial Cuing in Global-Scale Location Estimates

Alinda Friedman
University of Alberta

Seven independent groups estimated the location of North American cities using both spatial and numeric response modes and a variety of perceptual and memory supports. These supports included having location markers for each city color coded by nation and identified by name, giving participants the opportunity to see and update all their estimates throughout the task, and allowing them to respond directly on a map. No manipulation mitigated the influence of categories on the judgments, but some manipulations improved within-region ordinal accuracy. The data provide evidence that the city and regional levels are independent, spatial and numeric response modalities affect accuracy differently at the different levels, biases at the regional level have multiple sources, and accurate spatial cues improve estimates primarily by limiting the use of global landmarks to partition the response space. Results support J. Huttenlocher, L. V. Hedges, and S. Duncan's (1991) theory of spatial location estimates and extend it to the domain of real-world geography.

Keywords: spatial cognition, spatial representation, location estimation, cognitive mapping, categorization

When people estimate the locations of North American cities, they have a tendency to do four things: They divide the cities into regions that do not overlap and that are not necessarily coincident to political boundaries; they do not distinguish among the north–south locations of cities within most of the regions; they leave relatively large gaps between regions; and they produce increasingly biased estimates as the cities being estimated are located farther south. Each of these effects has been observed with Canadian university students (Friedman & Brown, 2000a, 2000b) and, more recently, with American and Mexican university students (Friedman, Kerkman, & Brown, 2002; Friedman, Kerkman, Brown, Stea, & Cappello, 2005; Friedman & Montello, 2006) and children as young as 11 years of age (Kerkman, Friedman, Brown, Stea, & Carmichael, 2003).

All four characteristics observed in global-scale location judgments are consistent with the assumption (Friedman & Brown, 2000a, 2000b) that although people may generally have inexact item-level knowledge (e.g., they do not know the specific latitudes of most cities), they have a wealth of other specific and general beliefs—correct or incorrect—about the cities and the geographical regions and continents to which the cities belong. Thus, in the absence of specific knowledge to the contrary, decisions about where to locate the continents and the regions within them may be

arrived at by a process of plausible reasoning from these beliefs (Collins & Michaelski, 1989; Friedman & Brown, 2000a, 2000b). For example, people may assume that all of North America is north of the equator and all of South America is south of the equator or that Mexican and southern European cities are near the equator because they are often destinations for warm vacations. Indeed, we have assumed that the qualitative nature of the biases observed in our prior data was due to the existence of geographical regions (i.e., categories) and people's beliefs about where these regions were located with respect to global landmarks, such as the equator, the oceans, and the poles (Friedman & Brown, 2000a, 2000b; Friedman & Montello, 2006).

We adapted process assumptions from a metrics and mapping approach to making numeric estimates in real-world domains (Brown & Siegler, 1993) to explain some of the other characteristics in the data (e.g., the mutual exclusivity between regions). In this approach, when people make numeric latitude estimates, we assume that they use an ordinal conversion strategy (Brown, 2002b; Brown & Siegler, 1993) to divide the numeric range of latitudes (which are given by instruction) among the continents and whatever regions they believe lie within. Further, after deciding where North America lies within the 90° to –90° range, participants divide that portion of the range into mutually exclusive parts associated with each of the geographical regions. Finally, they generate estimates for each city by identifying its region and then selecting a value within that region's range. Plausible reasoning may enter the process at this point too. For example, if little is known about a city's location, the estimate might be based on the region's prototype location; this would truncate the within-region response range. On this view, when an ordinal conversion strategy is combined with beliefs about the location of regions relative to global landmarks, such as the equator, the resulting representation of North America is stretched southward. Arguably,

This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada. I am grateful to Bernd Kohler for programming the experiments and data collection, Katrina Wiebe and Magda Piwowarczyk for data collection and analysis, Todd Williams for help with the mass-testing questionnaire, and Norman Brown and Don Heth for commenting on an earlier draft of this article.

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

however, geographic (and other) location estimates have a heavily spatial component inherent in the knowledge representation underlying the task; we have used primarily numeric judgments in our location estimate tasks (Friedman & Brown, 2000a, 2000b; Friedman et al., 2005; Friedman & Montello, 2006; but see Friedman, Brown, & McGaffey, 2002). Thus, one focus in the present study was to extend the plausible reasoning and metrics and mapping frameworks to a location estimate task that requires a spatial, rather than a numeric, response.

Of course, the four characteristics observed in global-scale numeric location judgments are also consistent with a long tradition of research on the effects of categorization and hierarchical knowledge on spatial location estimates (e.g., Hirtle & Jonides, 1985; Huttenlocher, Hedges, Corrigan, & Crawford, 2004; Huttenlocher, Hedges, & Duncan, 1991; McNamara, 1986; McNamara & Diwadkar, 1997; Stevens & Coupe, 1978). It is common for categories to influence judgments in many domains, and real-world geography is one of them (Friedman & Brown, 2000a, 2000b; Friedman & Montello, 2006; Maki, 1981; Montello, 2003; Stevens & Coupe, 1978).

Understanding how people use categories to reason about real-world geography is important for understanding spatial cognition in general and the processes underlying location estimates in particular. This is because geography is a complex domain that is learned about over the life span from a variety of eclectic sources and experiences, so reasoning about geographic entities can potentially be affected by perceptual (Huttenlocher et al., 1991), cognitive (Friedman & Brown, 2000b), and affective (Carbon & Leder, 2005; Kerkman, Stea, Norris, & Rice, 2004) factors. Thus, understanding how people reason about geographic categories should inform theories of spatial cognition about the interplay among these factors and how they contribute to spatial judgments and biases. However, the most elaborated current theory of the role of categories in spatial location judgments (Huttenlocher et al., 1991, 2004) has thus far primarily been tested in a more abstract domain than that of real-world geography: remembering the locations of single dots presented in circular frames.

In Huttenlocher et al.'s (1991) category-adjustment model, location information is represented at two independent levels: a fine-grained level and a category level. Both levels are assumed to be free of bias. An induced category is a bounded region that covers a range of fine-grained stimulus values; its prototype is, for example, the mean or median location of the observed instances. Categories are assumed to affect spatial judgments in at least two ways. First, a remembered location may be weighted by the prototype for its category: The prototype is given more weight in the judgment to the extent that the item values are uncertain and inexact. This weighting results in a truncation of the range of estimates, because they move away from the category boundaries and toward the prototype. Second, people are constrained to respond within the category boundaries, so the accuracy and precision of the boundaries' locations also affect the accuracy of the judgments. Thus, the amount of bias in the estimates is predictable from the amount of inexactness in the representations of the items, the locations assigned to their category boundaries and prototypes, and, ultimately, from the manner in which the item- and category-level information is combined. Further, the observed biases in the estimates are construed as a constructive process in this model, because "memory storage increases uncertainty about the stimulus,

causing category-level information to be given more weight in the estimates" (Crawford, Huttenlocher, & Engebretson, 2000, p. 281). Finally, Duffy, Huttenlocher, and Crawford (2006) extended the category adjustment model to children and concluded that, like adults, children use categories for increasing average accuracy in estimating inexact stimuli.

Huttenlocher et al. (1991; see also Huttenlocher et al., 2004) tested their model by having participants estimate the locations of single dots within a circle they had seen just moments before; the task thus tested relatively short-term visual memory. Participants' responses were found to be clustered as if horizontal and vertical boundaries had been imposed that divided the circles into quadrants; the estimates were misplaced toward the centers of the quadrants (i.e., toward the diagonals, which were the category prototypes). Huttenlocher et al. (2004) further showed that the locations of category boundaries used in judgments of dot locations were robust against various manipulations: People continued to use imagined horizontal and vertical axes as boundaries, even when the distribution of stimuli was clustered at the diagonals and they had to categorize the stimuli before each trial. Thus, functionally, none of their manipulations caused people to alter the locations of the category boundaries they used to estimate dot locations.

Most of our previous data from real-world geographic judgments are consistent with the basic tenets of the category adjustment model. For example, when participants estimate the location of a city, they tend to place it toward the locational center (i.e., the prototype) of a psychologically defined region and as a result, the physical extent of the regions is relatively truncated compared with the actual range of distances between cities (Friedman & Brown, 2000a, 2000b; Friedman et al., 2005). Similarly, we have evidence that the category (regional) and item (city) levels are independent (Friedman & Montello, 2006). However, the effect of categories on the accuracy of geographic judgments differs from that on short-term memory for dot locations. In particular, on average, category-level information improves estimation accuracy for dot locations; it is primarily for this reason that its use is rationalized in Huttenlocher et al.'s model (1991, 2004). However, category-level information generally does not improve estimation accuracy in geography; indeed, it often makes the estimates quite inaccurate (e.g., Friedman & Brown, 2000a; Friedman et al., 2005). Thus, the main focus in the present study was to examine some of the possible causes of this difference in estimation accuracy between domains and to provide evidence for its locus.

The various sources of information that constitute geographic knowledge imply that geographic categories and their boundaries are induced from a mix of information that includes maps and political borders, world history and geopolitics, climate, geographic landmarks, family background, etc. Thus, with respect to global-scale spatial location estimates, geographic categories are generally not isomorphic to countries or other a priori political entities (Friedman & Brown, 2000a, 2000b). On the basis of their latitude estimates, for example, Canadian participants treat some countries as if they were single regions (e.g., Mexico), some countries as if they contained more than one region (e.g., the northern and southern United States), and some regions as if they contained more than one country (e.g., northern and southern Europe). Furthermore, it is entirely possible that the factors used to induce geographic categories change with scale (Montello, 2003).

At the urban level, for example, neighborhoods may function as categories, but these categories (and the factors from which they were induced) are not likely to be retrieved when participants estimate global-scale locations. Finally, because of the many factors from which geographic boundaries are induced, geographic boundaries are likely to be less precise than the horizontal and vertical axes of a circle. Once a category's boundaries have been established by whatever means, however, the locational prototype of the category should be approximately at its physical center, as predicted by Huttenlocher et al. (1991).

Like Crawford et al. (2000), Duffy et al. (2006), and Huttenlocher et al. (1991), we hypothesize that category-level beliefs generally carry more weight in the estimates than do item-level beliefs and that this weighting contributes to the observed biases. However, in contrast to these authors we believe that, in real-world geography, the memory representations of the category boundaries are probably not free of bias, due to the many sources from which geographic information is learned and the years taken to learn it. The biases inherent in the geographic categories that participants bring to the task are inherited by the items during the estimation process; we thus observe a dissociation in accuracy between item- and category-level knowledge (Friedman & Montello, 2006): People can have accurate ordinal information at the item level (e.g., they place cities within a given region in the correct north-to-south order) while, at the same time, the amount of absolute error in their placement of these cities is off by hundreds of miles, presumably because of their beliefs about the location of the category boundaries.

It is important to note that the use of geographic landmarks to establish continental and/or category boundaries does not imply that these landmarks are used in making the location estimates themselves (e.g., on a trial-by-trial basis). Rather, the tenets of the category adjustment model dictate that once the category boundaries have been established, the observed biases in judgment are formed from overweighting them. When these boundaries are horizontal and vertical lines in circles, the item estimates become more accurate on average; when the boundaries are those for geographical regions, the estimates can become very inaccurate indeed.

Nevertheless, it is important to note that the metrics and mapping (Brown & Siegler, 1993) and category-adjustment approaches (Huttenlocher et al., 1991) are not necessarily in opposition. Indeed, both models should apply to estimation processes in a variety of domains. However, it is fair to say that, thus far, the metrics and mapping framework, even when applied to geography (e.g., Friedman & Brown, 2000a, 2000b), has been formulated and applied principally to numerically based estimates (e.g., of populations, dates, locations; see, e.g., Brown, Cui, & Gordon, 2002; Brown & Siegler, 1993), whereas the category adjustment model has been formulated and applied principally to spatially and visually based estimates (e.g., of locations, sizes, shades of gray; see, e.g., Huttenlocher et al., 1991; Huttenlocher, Hedges, & Vevea, 2000).

Thus, as noted above, one possible reason for the difference in the role of categories in estimation accuracy across domains is that the dot location task requires an explicit spatial response (participants indicate where in a blank circle they think the dot they just saw was located), whereas the geographic estimate task has been implemented primarily by asking people to make numeric estimates (e.g., of latitudes). A second possibility is that participants

do not need to remember their responses on previous trials of the dot location task to be accurate on the current trial; however, to the extent that at least some geographic knowledge in long-term memory is relational (e.g., the ordinal north-south relations among cities within a region), memory for responses on previous trials can be useful. The third possibility is that the horizontal and vertical lines that serve as category boundaries in the dot location task are relatively precise and accurate; in contrast, those in geography may be relatively vague and highly inaccurate. Each possibility is discussed in turn below. They were addressed in the present study by using a spatially based response in all three experiments and by allowing participants to track (Experiment 1) and correct (Experiments 2 and 3) their prior estimates throughout the task.

A priori, a spatially based response has more face validity for making location judgments than does a numeric estimate task. However, on the one previous occasion in which we used a spatially based response to obtain global-scale location estimates (Kerkman et al., 2003), there were several serious constraints on the grid task used that made it virtually impossible to compare its results with previously obtained numeric estimates. For example, participants made their responses by "dragging and dropping" an X displayed on a computer screen to a location on a latitude-by-longitude grid; however, they were instructed only about latitudes, and this made interpretation of the longitude data ambiguous. Several other constraints of the original implementation of the grid task are discussed below and ameliorated in Experiment 1. In Experiment 2, the new version of the grid task was used to compare numerically and spatially based responses, and in Experiment 3 participants responded directly on an outline map. Experiment 3 thus also provided an explicit test of the relative importance of accurate spatial cues and category information in global-scale location estimates.

The second possible reason for the difference in the role of categories in estimation accuracy across domains is whether it is useful (or even possible) to induce or make use of distributional information about the items across trials. As noted, for the dot task, the location of the dots on previous trials is irrelevant to performance on the current trial, and one need only rely on relatively short-term memory to be accurate. Indeed, Huttenlocher et al. (2004) showed that people did not induce the spatial relations across trials even when the distributions of dot locations were extreme; one reason for this might be that there is no point to (or perhaps no possibility of) accurately remembering the relations between the individual dots from trial to trial. In contrast, much of the item- and category-level information about geography is learned over the life span, and at least some of it is relational (e.g., one may know that Edmonton is to the north of Calgary and to the east of Vancouver, which in turn is to the north of Seattle). Consequently, remembering prior responses is relevant if one is to make use of relational knowledge in a series of geographic location estimates. Thus, distributional information in long-term memory, in the form of interitem ordinal and metric spatial relations for cities within and between regions, is potentially relevant to overall estimation accuracy in the real-world domain. In the task itself, one can take advantage of distributional information only by remembering responses from prior trials.

Faulty memory for prior responses might also contribute to the observed lack of overlap and the gaps between regions. For instance, via both instructions and a prior knowledge rating task

(Friedman & Brown, 2000a, 2000b), participants knew what states/provinces and countries the cities belonged to, as well as the complete set of cities themselves, before they made their location estimates. Thus, participants might settle on a range of numbers to use for each region, not because they are parsing the response space per se but because this response rule mitigates the effect of having to remember prior responses. Similarly, participants might know that the national borders have overlapping latitudes but deliberately keep the regions mutually exclusive, so that their estimates remain within the categories' boundaries and are thus consistent in this sense. This sort of consideration is not relevant to the dot location task, because keeping track of prior estimates really has no bearing on the current estimate.

The methods used in the present study were an attempt to mitigate some of the effects on judgments that may be due to faulty trial-to-trial memory by eliminating any reliance on it. I used a strategy analogous in its logic to that of Crawford et al. (2000). They tested the reconstructive predictions of the category adjustment model (Huttenlocher et al., 1991) by comparing performance on line-length estimations, either while the lines were in view or after a short interval. Estimation biases were present in the memory condition but not in the perceptual condition. Crawford et al. interpreted these findings to mean that introducing a delay increased the uncertainty within the memory representation and that bias in the estimates was introduced during retrieval. Similarly, Choplin, Huttenlocher, and Kellman (2001) found that the tendency for newly acquired categories to affect judgments was present only in a sequential discrimination condition and not when the stimuli were displayed simultaneously. Thus, it appears that perceptual supports can eliminate the biasing effects of category boundaries in some circumstances.

In the present experiments, a perceptual support methodology was used to eliminate the effects of memory for prior responses and to allow participants to display both their item- and category-level knowledge as accurately as possible. For example, in some conditions, identity cues for the cities and the spatial locations of prior judgments were provided throughout the task. When prior estimates are continuously available, participants can compare the on-screen configuration of the estimated locations with their long-term memory representations of those configurations and, in some conditions, could change their estimates until they were satisfied that these estimates were accurate. Notably, these conditions are analogous to Crawford et al.'s (2000) conditions, in which they did not obtain evidence for bias in line-length estimates.

If biases in geographic estimates remain when participants can track and correct their prior responses because the responses are available to perception, the biases must persist in long-term memory and are not due to limitations imposed by the task. This would mean that the biases in the real-world spatial domain cannot be entirely due to the process of combining category- and item-level representations that are unbiased. Rather, such data would support the third reason proposed for the difference in accuracy between the dot location and real-world geography tasks: that biases in the representations of geographic category boundaries exist a priori and are inherited by the items. In other words, as predicted by the category adjustment model, short-term memory for dot locations and long-term memory for city locations are both influenced by the spatial relations among the items and by their category boundaries. However, there are likely to be differences

in the relative accuracy of category boundaries across domains and, thus, in the role that categories play in estimation accuracy overall. This third reason is the one we believe is primarily responsible for the inaccuracies and specific biases observed in geographic estimates (Friedman & Montello, 2006).

The same manipulations that should mitigate the influence of trial-to-trial short-term memory factors on performance by providing participants with a way to monitor the spatial relations among the items also allows them to monitor the relations between those items and their beliefs about the locations of the category boundaries relative to those items. Thus, if these manipulations do not significantly alter the main characteristics of the data (e.g., the regions remain mutually exclusive; the continent appears to be stretched southward), the data would provide the strongest evidence thus far that these characteristics are fundamental to the long-term representation of geographic information.

To summarize, the first goal in the present study was to provide an explicit comparison of spatially and numerically based estimates; the second was to provide perceptual and memory supports that allowed participants to take advantage of and express explicitly their item-level relational (i.e., distributional) knowledge, to the extent that it existed both within and between geographical regions; and the third was to determine whether the mutual exclusivity between the regions and the biases in their absolute locations would be ameliorated with use of these various perceptual and memorial aids. To anticipate our results: The data provide evidence that Huttenlocher et al.'s (1991) model applies to global-scale geography, but accuracy is poor in the real-world context because of participants' beliefs about the location of geographic boundaries. Further, the mutual exclusivity between regions remains strong even when people respond directly on a map; the data thus have important implications for the role of spatial cuing in estimation accuracy.

Experiment 1

In Experiment 1, participants performed a location estimate task in which the response space was a latitude-by-longitude grid displayed on a computer screen. The control group dragged and dropped an *X* to its estimated location, and it disappeared when they pushed the *Enter* key. In one experimental group, all of the *X*s stayed on the screen as they were placed throughout the task. In the second experimental group, the *X*s stayed on the screen and were color coded by nation (Canada, the United States, and Mexico). These features effectively provided participants with an external "scratch pad" and the means to delineate among national borders. Indeed, the color-coded condition might be the first to reveal that participants knew that the national borders are intermingled. Note, however, that the color coding used throughout the present study did not divide the United States into the two regions (northern and southern) that we have previously observed in the estimates of both Canadian and American university students (Friedman & Brown, 2000a; Friedman, Kerkman, & Brown, 2002). Thus, if the biased estimates and gaps between regions obtained previously were mainly due to an inability to remember responses from trial to trial, the bias and the gaps should be lessened in the black marker condition and possibly eliminated in the colored marker condition, at least for the three nations.

The grid task used previously (Kerkman et al., 2003) was implemented in a way that precluded a direct comparison between its data and the data from numeric estimates. First, a disproportionate number of the cities were located within the participants' home region to avoid having the youngest children overwhelmed by too many unfamiliar place-names. However, it is unclear how weighting the stimuli in this manner affected the data. For example, Haubensak (1992) proposed that the range of a response scale is centered on the most frequent stimuli. Thus, throughout the present study, we used an equal number of cities in each of the four regions.

Second, though both latitude and longitude lines were labeled on the grid every 10°, the instructions in Kerkman et al. (2003) referred only to latitudes so the children would consider only a single dimension. Thus, we could not unambiguously assess the longitude data, even for the adult participants. In the present study, we gave participants instructions about both dimensions and told them to take both into account for their estimates.

Third, though Kerkman et al. (2003) instructed participants that latitudes range between +90° and -90°, the grid itself showed latitude lines between +90° and -10°. This display effectively set a lower bound for responses. Though accurate for the cities in that study, it was arbitrarily different than the endpoint of the dimension, and this fact made comparisons with previous numeric estimates moot. In the present study, I used the entire range of latitudes on the grid.

Method

Participants. Ninety-eight volunteers (36 men, 62 women) from the Department of Psychology's undergraduate pool participated to fulfill a course requirement. They were randomly assigned to one of the three experimental groups: no memory aids, black markers, or color-coded markers. Data from 8 participants (3 men, 5 women) were not used. Of these participants, 2 were omitted because of a program malfunction, 1 was omitted because the participant had been in a similar study, and 5 were omitted because they did not follow the instructions and appeared to be responding randomly. This left 30 participants in each group; 71 of these had been born in Canada. On average, the 90 included participants had lived in the Edmonton metropolitan area for 10.7 years ($SE = 0.9$).

Stimuli, design, and procedure. The stimuli were nine cities from each of the four regions found in prior studies with both Canadian and American participants (e.g., Friedman, Kerkman, & Brown, 2002). From north to south, with their latitudes given in parentheses, the Canadian cities were Edmonton (54°), Saskatoon (52°), Calgary (51°), Regina (50°), Winnipeg (50°), Vancouver (49°), Montreal (45°), Ottawa (45°), and Halifax (45°). The northern U.S. cities were Seattle (48°), Minneapolis (45°), Buffalo (43°), Chicago (42°), Detroit (42°), Boston (42°), New York (41°), Pittsburgh (40°), and Cincinnati (39°), and the southern U.S. cities were Las Vegas (36°), Los Angeles (34°), Atlanta (34°), Phoenix (33°), Dallas (33°), Houston (30°), New Orleans (30°), Tampa (28°), and Miami (26°). The Mexican cities were Tijuana (33°), Cabo San Lucas (23°), Mazatlán (23°), Puerto Vallarta (21°), Cancún (21°), Mexico City (19°), Veracruz (19°), Acapulco (17°), and Oaxaca (17°).

Each group performed two tasks. In each task, the names of the cities were presented, one at a time, on a computer screen together with their state or province and country. The cities were presented in a different random order for each task and participant.

Participants first rated their knowledge of each city on a scale from 0 (*no knowledge*) to 9 (*a lot of knowledge*). The stimuli appeared along with the labeled endpoints of the rating scale, and participants responded using the numeric keypad on the keyboard in front of them. They next estimated the locations of the same 36 cities. They were told, "Latitude ranges from 0° at the equator to +90°N at the North Pole and -90°S at the South Pole." They were then instructed that longitude lines run vertically, meet at the poles, and measure how far east or west a place is. They were further told that the longitude line that runs through Greenwich, England, is the 0° longitude line and that moving westward from Greenwich one eventually arrives at the place on the opposite side of the world from England, Fiji, which has a longitude of approximately 180°. Finally, they were informed that cities on the North American continent are located somewhere between 160°W (Pacific Ocean) and 40°W (Atlantic Ocean).

The display was a 17-in. color monitor with a 1,024-by-768 pixel resolution. The grid was 704 pixels wide by 728 pixels high and had latitude and longitude lines that were numbered every 10° on each of its sides. The latitude lines ranged from +90N to -90S, and the longitude lines ranged from 40W to 160W. On each trial, the name of the city and its state or province and country appeared in bold font to the left of the grid in a box, with an X below the names. The X remained there until the participant clicked on a location within the grid or dragged the X from its start location to a location within the grid. The X continued to move to new places until the *Enter* key was pressed for that trial. If the participant left the X on a location outside the grid, a reminder appeared to click within the limits of the grid.

For the group that did not receive memory aids, the X disappeared after every trial. For the black marker group, each black X stayed where it had been placed until all 36 locations had been estimated. For the color-coded group, the Xs also remained where they had been placed and were color-coded by country: green for Canadian cities, blue for U.S. cities, and red for Mexican cities. It was possible to overlap the markers to within 1 pixel of each other.

Results

A $p < .05$ criterion for significance was adopted for all of the analyses throughout the article, and η_p^2 was used, as appropriate, as the measure of effect size.

Knowledge ratings. Each participant's knowledge ratings were averaged for each region and subjected to a Group \times Region analysis of variance (ANOVA). The ratings differed markedly as a function of region, $F(3, 261) = 423.59$, $MSE = 0.73$, $\eta_p^2 = .830$, with the averages for Canadian, northern U.S., southern U.S., and Mexican cities being 6.4, 3.8, 4.3, and 1.9, respectively. This pattern replicates the pattern found for Canadian participants in previous studies (e.g., Friedman & Brown, 2000a). Neither the main effect of group nor the Group \times Region interaction was significant ($F < 1$ in both cases).

Location estimates. Figure 1 shows scatter plots of the average estimates for each city and group. An equirectangular map is superimposed on each plot to indicate where the correct location of the North American continent is on this type of projection; the choice of projection is appropriate because the response grid used by the participants was equirectangular. Note that participants did not see this map at any time during the experiment.

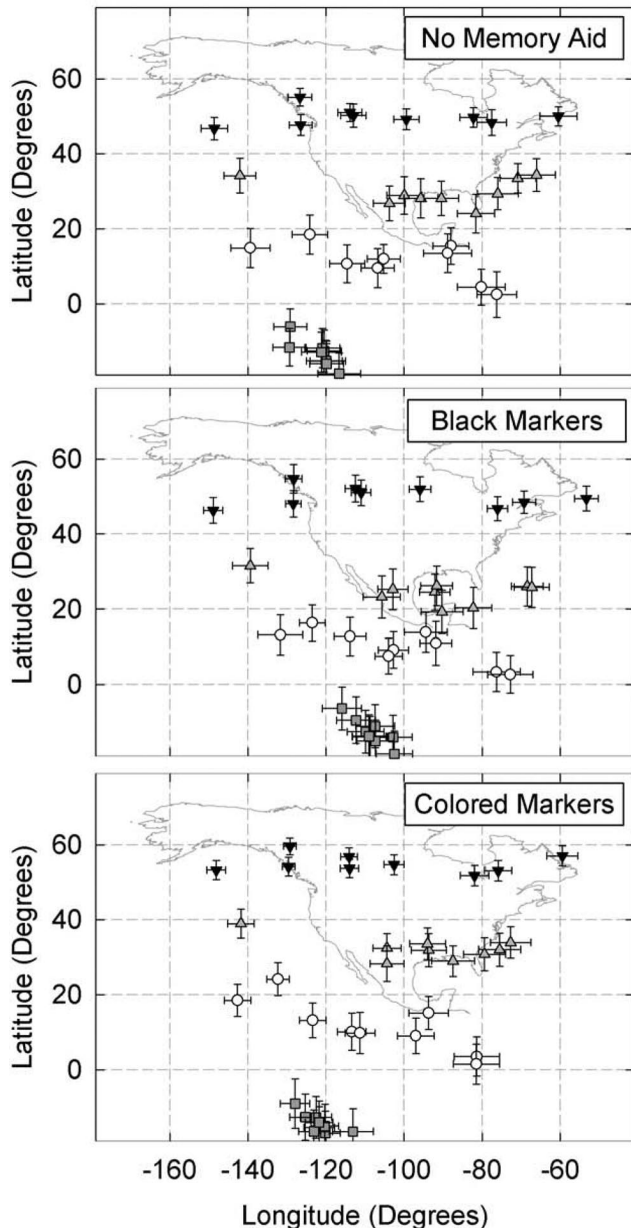


Figure 1. Mean estimate for each city computed across participants for each group in Experiment 1. Error bars are standard errors of the mean across participants for each city computed independently for each dimension. Inverted triangles are Canadian cities, triangles are northern U.S. cities, circles are southern U.S. cities, and squares are Mexican cities. Participants did not have an outline of the continent on their response displays; it is provided here for reference.

It is obvious that the groups' estimates differed little, if at all. In particular, all three groups displayed the same four characteristics in their estimates: the division of North American cities into four nonoverlapping regions, the gaps between regions, the north-south truncation of cities within regions, and the overall southward bias, particularly in the two southernmost regions. In this experiment, participants placed Mexican cities slightly below the equator; we have seen this pattern once before with Texan participants who made numeric estimates (Friedman, Kerkman, & Brown, 2002, Figure 1).¹

The configurational similarity among the estimates for the three groups was examined by converting each participant's estimates from xy pixel coordinates to latitudes and longitudes. A bidimensional regression analysis (Friedman & Kohler, 2003) was conducted, in which the actual latitudes and longitudes were used as the independent variables and their estimates as the dependent variables. The bidimensional regression coefficient quantifies how much variance in the estimates is accounted for by the Euclidean configuration of actual latitudes and longitudes of the cities.

The Fisher-transformed regression coefficients were submitted to a one-way ANOVA with group as the between-subjects factor. The back-transformed means for the no memory aid, black marker, and color-coded groups were 0.813, 0.795, and 0.829, respectively. Though these are quite respectable and are all significantly above zero, $t(29) = 23.61, 16.60,$ and 25.12 , respectively, they did not differ among the three groups.

As a further means of assessing group differences, the 36 latitude estimates from each participant were submitted to a Group \times City ANOVA. The large main effect of city was not surprising, $F(35, 70) = 279.44, MSE = 192.33, \eta_p^2 = .763$, but there was no hint of an effect of group, $F(2, 87) < 1$, or of a Group \times City interaction, $F(70, 3045) = 1.00, MSE = 192.33, p = .47$, despite a very large number of dfs in both numerator and denominator. There was also no interaction in an analysis in which the participants' mean latitude estimates for each region were submitted to a Group \times Region ANOVA, $F(6, 26) < 1$. The latter analysis is most instructive, because it eliminates individual differences in the placement of cities within regions and examines only the differences between groups in the means for the regions themselves.

Finally, all three groups had significant differences between cities estimated to be at the national borders (see Table 1), so there was no evidence that participants intermingled the cities at the category boundaries. Further, all the cities that had been assigned a priori to the northern U. S. region were placed to the north of all

¹ To be sure, there is noise in the data across samples. In some of our original studies with Canadian participants (e.g., Friedman & Brown, 2000b), the average for the Mexican cities in two experiments was almost exactly at the equator (e.g., Friedman & Brown, 2000b, Figures 1 and 4), as it was for our sample of Mexican participants (Friedman et al., 2005, Figure 1). Our Californian sample (Friedman & Montello, 2006) estimated Mexican cities to be, on average, at about 10°N, and our Texan sample estimated them to be about 10°S. Thus, we are certainly not claiming that all participants in all cases use the equator to partition the response space in exactly the same way. Clearly, some people will have more precise (or imprecise) beliefs than others, though it is notable that the Californian sample comprised university students who were taking a final examination on North American regional geography.

Table 1
Mean Latitude Estimates and t Tests for the Difference Between Means for Cities Estimated to Be at the National and Psychological Borders for the Three Groups in Experiment 1

Marker type	Canada–U.S.			Northern U.S.–Southern U.S.			Southern U.S.–Mexico		
	Southernmost	Northernmost	<i>t</i> (29)	Southernmost	Northernmost	<i>t</i> (29)	Southernmost	Northernmost	<i>t</i> (29)
No memory aid	Vancouver (46.8°)	New York (34.4°)	4.44*	Pittsburgh (24.1°)	Las Vegas (18.5°)	2.09*	Tampa (2.5°)	Tijuana (–6.1°)	2.07*
Black	Vancouver (46.3°)	Seattle (31.6°)	6.05*	Pittsburgh (19.3°)	Las Vegas (16.3°)	1.11	Miami (2.7°)	Tijuana (–6.4°)	2.21*
Color-coded	Ottawa (51.8°)	Seattle (39.0°)	5.46*	Cincinnati (28.3°)	Las Vegas (24.2°)	1.57	Miami (1.5°)	Tijuana (–9.0°)	2.43*

Note. U.S. = United States.

* $p < .05$.

the cities that had been assigned to the southern U.S. region in all three groups.

Experiment 2

When participants were allowed to see their prior estimates, their estimation accuracy did not improve in comparison with that of participants who had no memory aids. Because this is a null result, the first purpose in Experiment 2 was to provide a statistically more powerful demonstration of the pattern of results obtained in Experiment 1. To this end, participants in Experiment 2 made two sets of estimates and thus provided their own baseline. The first estimates were made without any perceptual or memory aids. The second estimates were made with a variety of aids, including a continuous display of the cities' identities and estimated locations and the ability of participants to change any of their estimates at any time throughout this phase. It is important to note in this context that mere provision of the opportunity to make a second (or third, or fourth) set of estimates does not by itself produce higher estimation accuracy (Brown & Siegler, 1993).

Allowing people to change their estimates throughout the second phase accomplished a second goal, because it allowed people to take advantage of their knowledge, insofar as such knowledge existed, about both within- and between-region spatial relations among cities. For example, if people knew that Tijuana, Baja California, is north of Miami, Florida, they would be able to show this directly by placing those cities in that relation, irrespective of the order in which the cities appeared on the screen and in which the estimates were made. Similarly, if people had ordinal or interval metric knowledge about the spatial and configural relations within familiar regions (Friedman & Montello, 2006), they could rearrange the markers until they were satisfied that the configuration on the screen reflected this knowledge accurately.

If the regions remain mutually exclusive under these circumstances, it is difficult to argue that this is because of flaws in the retrieval of long-term memory representations. Rather, if the regions remain mutually exclusive when participants can see all their estimates at once as well as change them at will, it is more likely that the estimates accurately reflect participants' beliefs about the actual locations of the regional boundaries as well as the relations among the cities within each region.

The third purpose in Experiment 2 was to make the first direct comparison between spatially and numerically based methods of making global-scale estimates. To this end, we compared estimates from three groups of participants: The first used the grid task from

Experiment 1, the second made numeric latitude estimates, and the third made estimates by placing an *X* on a vertical line marked with latitudes every 10°, like the grid.

The comparison between numeric and spatial methods of responding is theoretically important for several reasons. We have argued, first, that biases in both location and distance estimates are attributable to the use of global landmarks and other beliefs (e.g., about climate) to partition the response range and, second, that both location and distance estimates are based on the same item- and category-level representations of the locations of the cities and their superordinate regions (Friedman & Montello, 2006). However, though we have strong evidence for both assertions, that evidence is based solely on numeric estimates. Third, there is increasing acceptance of the idea that conceptual representations of the sort that we believe underlie global location estimates are grounded in modality-specific systems (Barsalou, Simmons, Barbey, & Wilson, 2003); indeed, early interpretations of the locus of biases in memory for maps claimed that their basis was in perceptual representations distorted through gestalt processes (Tversky, 1981, but see Friedman, Brown, & McGaffey, 2002, for an alternative interpretation). If modality-specific representations underlie global-scale estimates, providing a direct comparison of different response modes should help reveal this relation.

The grid and numeric estimate tasks differ in more than one way. Thus, we included the line task because, like the grid task, it is spatially based. In particular, the line task's display was identical in length and position to the middle vertical line of the grid task's display, and participants responded in the same drag and drop manner. But, as with the numeric task, the line task requires participants to consider only one dimension. Thus, if estimates from the line and grid tasks are more similar to each other than to numeric estimates, is likely that the spatial versus numeric response mode is responsible (i.e., the ability to see the ordinal and metric spatial relations among one's responses on a line or grid vs. responding with a numeric estimate of latitude). On the other hand, if the latitude estimates for the line and numeric tasks are more similar to each other than they are to the latitude estimates from the grid task, the difference is more likely attributable to the need to consider one versus two dimensions in the estimates. In contrast to these possibilities, if our representational and process assumptions are correct, the observed north–south estimates should be similar in all three groups. These assumptions include the notions that the charac-

teristics of the data reflect a mental representation of global-scale geography that is regionalized and that the southward bias reflects how the regions are referenced to global landmarks during the initial parsing of the response space, rather than any particular response mode or the number of dimensions being considered.

Notably, the identity and location cues provided during the second estimates in this experiment gave participants no information about the appropriate scale to use for North America or its regions. Thus, they were free to partition the response space as they wished, and each of the spatial tasks was similar in kind to the numeric estimate task in this sense. The groups thus provided the first test of whether the metrics and mapping framework (Brown & Siegler, 1993) applies equally to numeric and spatial estimates.

Method

Participants. 109 undergraduates (42 men, 67 women) from the Department of Psychology’s pool volunteered to be in the study and were randomly assigned to the grid, line, or numeric estimate group. Four participants (2 men, 2 women) were excluded from the data analysis, 1 because of an equipment malfunction; 1 because his session was interrupted unexpectedly with some loud construction noise; and 2 because the correlation between their estimated and actual latitudes was negative, which indicated that they had reversed the scale. The remaining 105 participants (35 per group) were all born in Canada and had lived in the Edmonton metropolitan area for an average of 11.2 years (*SE* = 0.85).

Stimuli, design, and procedure. The stimuli were 36 North American cities (9 from each of the four regions). They differed slightly from those used in Experiment 1. Names of cities from Experiment 2 are shown in Figure 2.

Each participant performed three tasks. The first was the knowledge rating task; the instructions and procedure were identical to

those used in Experiment 1. During the first location estimates for all groups, the city names appeared one at a time and disappeared after the participant had made his or her estimate, but during the second estimates all three groups received multiple perceptual and memory aids. First, the city names as well as the Xs themselves were color coded as a function of nation (Canada, the United States, and Mexico). Second, the name of each city remained on the screen; it was either directly next to or connected to its estimated location by a color-coded line. Third, after each placement, the city names in the latitude-only groups were sorted in the participant’s current north-to-south order (they remained beside their estimated locations for the grid group). Finally, participants could change their minds about their estimates for any of the cities and relocate the estimates at any time throughout this phase.

All three groups received instructions that were identical to those used in Experiment 1 regarding how latitudes and longitudes work, except that we omitted the information about the range of longitudes for North America so that participants could use the entire scale if they wished. The instructions then diverged according to how participants in each group were to implement their estimates. For each participant, the cities appeared in the same random order for each estimate, but a different random order was used for each participant.

During the first estimates, the grid group was given instructions and a procedure that was identical to that for the no memory aid group in Experiment 1: The Xs to be moved were displayed in black font; they could be placed by clicking on or dragging and dropping to a location, and they disappeared when the *Enter* key was pressed. The procedure for the line group was identical, except that the display consisted of a vertical line in the center of the screen with latitudes marked every 10°. It was identical in length to the center latitude line of the grid.

For the first numeric estimates, the instruction “Please enter the latitude (90° S to 90° N) of *city, state/province, country*” appeared at the top of the screen in a box, with an underscored blank space to its right. Participants entered their estimates on the numeric keypad, followed by an *n*, an *s*, or nothing, depending on whether they thought the city was located in the Northern Hemisphere, in the Southern Hemisphere, or on the equator, respectively. As in the other two groups, the city name and estimate disappeared once the *Enter* key was pressed. Thus, this group provided an exact replication of all of our previous numeric latitude estimate tasks.

For the second estimate task, all participants were informed that they would be estimating the locations of the same cities a second time and that “we will help you by changing the task so that you can see all the cities and your estimates all the time.” For all groups, when the task began, the 36 city names appeared at the left of the screen with an X to the left of each name. The city names and the Xs were color coded by country, as in Experiment 1. Further, in all groups, whenever a participant clicked on an X, the city’s name and its state or province and country appeared in a distinctive yellow box at the top of the screen. For the grid and line groups, that information remained on the screen as long as the participant was still dragging and dropping the X associated with that city; in the numeric group, it appeared on the screen whenever the participant clicked the X, either during the initial estimate or during later changes. Participants were told they could make their estimates in whatever order they wished, that they could change any of the estimates at any time they wished, and that all their

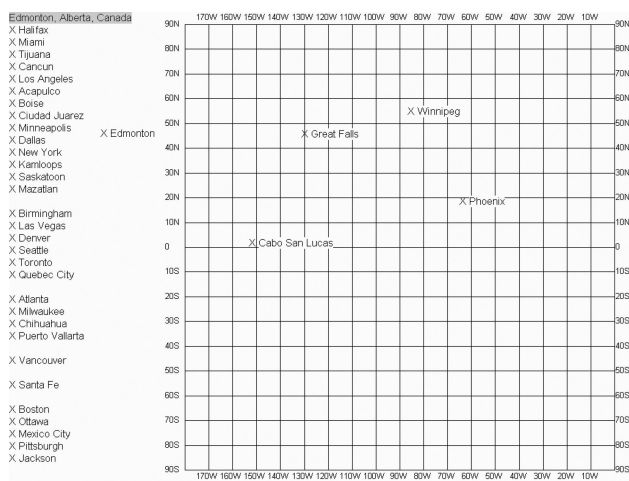


Figure 2. Example of the stimulus display for the grid group of Experiment 2 after four estimates had been made and as the fifth was being made. In the actual display the city names and their Xs were color coded by nation, but cities from the two regions of the United States were coded in the same color.

estimates would remain on the screen until they were satisfied that everything was as accurate as they could make it. Participants were told to press the *Enter* key only when they had finished estimating the locations of all 36 cities and were satisfied with the accuracy of their estimates.

For the grid group, the names of the cities moved with their location markers (see Figure 2). The procedure for the line group was similar, except that once an X had been placed on the line, it was connected to its color-coded name at the right of the screen with a color-coded line. Further, as each city was placed or if a city's location was changed, the city names to the right of the latitude line were resorted to reflect the participant's current north-to-south order. Thus, the markers on the line represented participants' beliefs about the interval north-south distances between cities, and the names attached to them represented the cities' identities.

For the numeric group, when a participant clicked on the X for a city it became "live" and the city's name, state or province, and country appeared at the top of the screen in the yellow box. The participant could then enter his or her numeric response and *n*, *s*, or no letter. Once an estimate had been made, the color-coded city name and its estimate moved to the right side of the screen and the names were sorted in the participant's current north-to-south order. Though the cities and estimates on the right side of the screen correctly represented a given participant's beliefs about their ordinal positions, there was no interval information about distances between the cities that was represented spatially in the display. Moreover, if the ordinal information represented in the display was particularly helpful to participants in the numeric group, this would be evident in the correlations between actual and estimated latitudes.

At the end of the second estimate task, participants filled out a 7-item questionnaire presented on the computer. The first two questions were demographic. The next three questions asked about the strategies participants had used. The response alternatives appeared in a different random order for each participant. The alternatives were selected on the basis of responses given to similar questions during interviews with previous participants. The final two questions asked where participants thought the northern and southern boundaries of North America were located. The response choices were all displayed at the same time and could be selected by clicking on a box next to them. For expository and comparative purposes, a discussion of the questionnaire results will be deferred until the results of Experiment 3 have been presented.

Results

Knowledge ratings. The knowledge ratings replicated the pattern seen in Experiment 1. The means for Canadian, northern U.S., southern U.S., and Mexican cities were 6.36, 2.99, 3.37, and 1.63, respectively, $F(3, 306) = 612.32$, $MSE = 0.68$, $\eta_p^2 = .857$.

Location estimates. Figure 3 shows the data for the grid group. There are striking similarities to the data from Experiment 1: The regions are distinct, there was little north-south discriminability among cities within regions, and the north-south biases in the estimates became increasingly large as the cities were actually located farther to the south. Although participants had the city names available at all times and were able to rearrange their

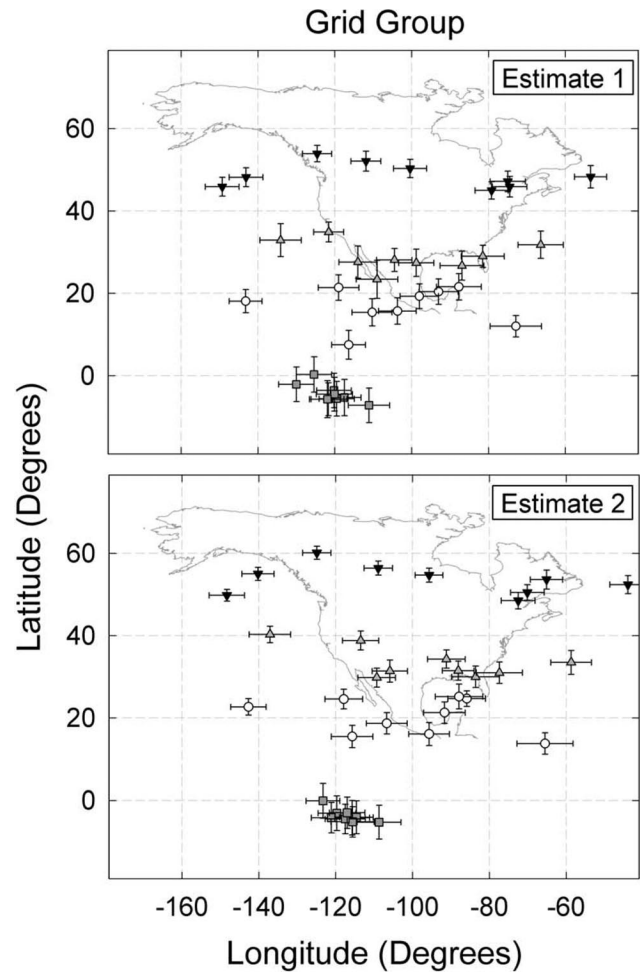


Figure 3. Mean first and second estimates for the grid group of Experiment 2. Error bars are standard errors of the mean across participants for each city computed independently for each dimension. Inverted triangles are Canadian cities, triangles are northern U.S. cities, circles are southern U.S. cities, and squares are Mexican cities. Participants did not have an outline of the continent on their response displays; it is provided here for reference.

locations during the second estimates, these fundamental characteristics did not change.

However, the availability of the cities' identities during the second estimates did affect configurational accuracy in Experiment 2. The bidimensional regression coefficients for the first and second estimates were .791 and .830, respectively, $F(1, 34) = 12.16$, $MSE = .02$, $\eta_p^2 = .263$. However, the gaps between regions still were present in the data and the placement of the regions themselves—especially the two southernmost regions—was still biased, so the improvement in accuracy was not achieved by "fixing" these aspects of the data. Rather, the improvement in accuracy was likely to have been achieved because the availability of prior responses and the ability to change the location markers throughout the second phase allowed participants to display more precisely the within-region configurational knowledge they had about at least some cities; this is what the bidimensional regression coefficient measures.

Figure 4 shows the mean latitude estimates for each city across participants, sorted in subjective north-to-south order independently for the first and second estimates in each group. All three groups had significant differences between the individual cities estimated to be at the national borders (see Table 2), and all the cities assigned a priori to the northern U.S. region were estimated to be north of all the cities assigned to the southern U.S. region.

To compare across groups statistically, we examined three different measures of accuracy in the latitudes via Group \times Region \times Estimate Number ANOVAs. The Fisher-transformed correlations between actual and estimated latitudes within regions reflect within-region ordinal accuracy at the item level and thus reflect differences in item knowledge within regions without regard to inaccuracies in the placement of the regions themselves. In con-

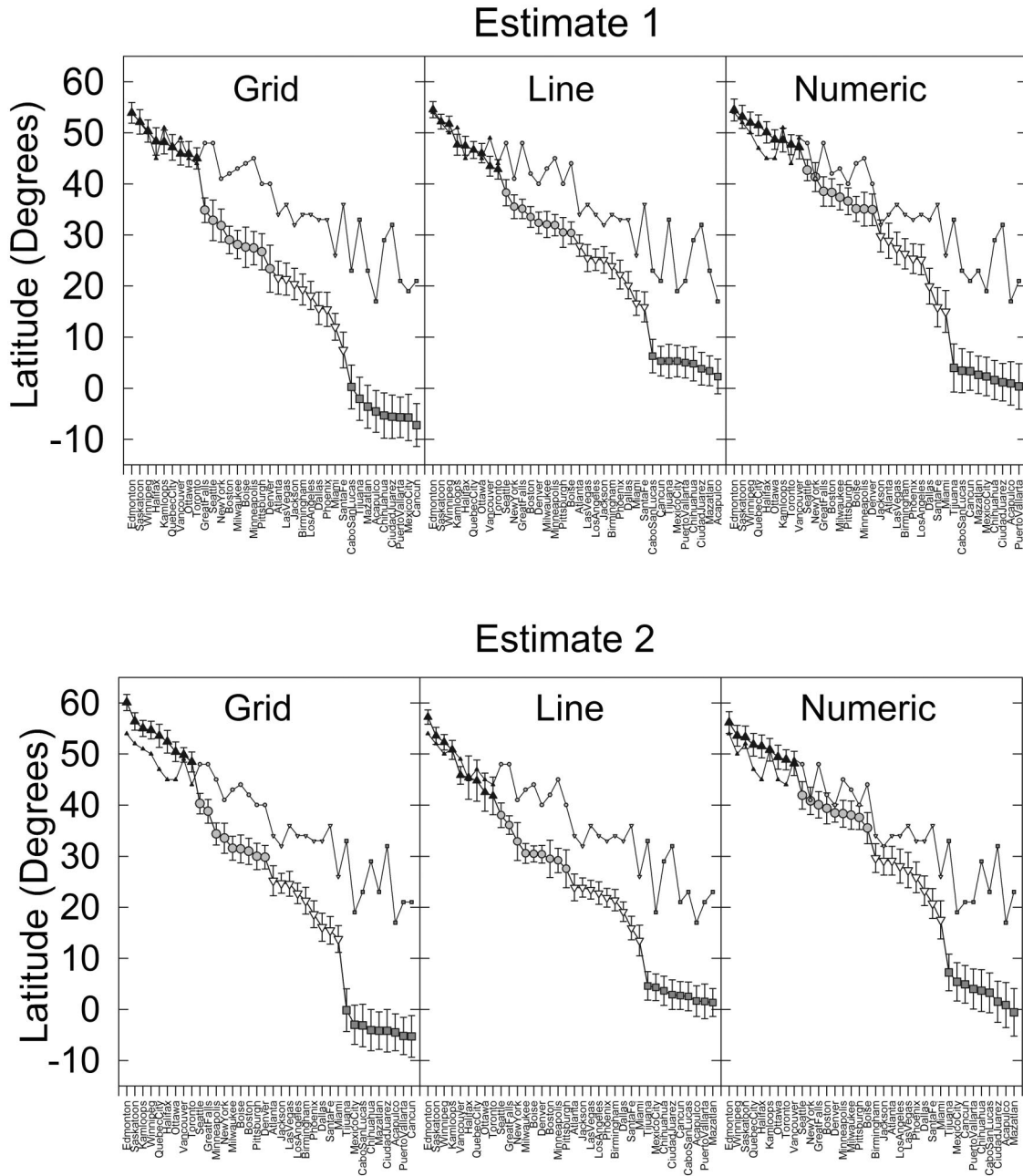


Figure 4. Mean first and second latitude estimates for each city, computed across participants in each group in Experiment 2. The estimates are ordered from the subjectively northernmost to southernmost city. Error bars are standard errors of the mean for each city and estimate number, computed across participants. Triangles are Canadian cities, circles are northern U.S. cities, inverted triangles are southern U.S. cities, and squares are Mexican cities.

Table 2

Mean Latitude Estimates and *t* Tests for the Difference Between Means for Cities Estimated to Be at the National and Psychological Borders for Experiments 2 and 3

Group, Est	Canada–Northern U.S.			Northern U.S.–Southern U.S.			Southern U.S.–Mexico		
	Southernmost	Northernmost	<i>t</i> (34)	Southernmost	Northernmost	<i>t</i> (34)	Southernmost	Northernmost	<i>t</i> (34)
Num, E1	Vancouver 47.2°	Seattle 42.7°	3.85*	Denver 34.9°	Jackson 29.8°	1.25†	Miami 15.0°	Tijuana 4.0°	3.31*
Num, E2	Vancouver 48.2°	Seattle 41.9°	5.15*	Boise 35.5°	Birmingham 29.6°	1.79*	Miami 17.5°	Tijuana 7.2°	4.35*
Line, E1	Toronto 42.9°	Seattle 38.3°	2.53*	Boise 30.4°	Atlanta 27.9°	1.33**	Santa Fe 15.9°	Cabo San Lucas 6.3°	5.47*
Line, E2	Toronto 41.8°	Seattle 38.0°	<1	Pittsburgh 27.5°	Atlanta 23.9°	1.73*	Miami 13.5°	Tijuana 4.6°	4.97*
Grid, E1	Toronto 45.0°	Great Falls 34.9°	4.43*	Denver 23.4°	Atlanta 21.6°	<1	Santa Fe 7.5°	Cabo San Lucas 0.3°	2.42*
Grid, E2	Toronto 48.5°	Seattle 40.3°	5.73*	Denver 29.8°	Atlanta 25.2°	1.98*	Miami 13.8°	Tijuana -0.1°	4.70*
Map, E1	Toronto 46.0°	Seattle 44.9°	1.26†	Pittsburgh 40.6°	Las Vegas 39.5°	1.35**	Miami 28.5°	Tijuana 25.2°	2.80*
Map, E2	Toronto 46.7°	Seattle 45.1°	2.26*	Pittsburgh 39.9°	Las Vegas 39.1°	<1	Miami 28.0°	Tijuana 25.7°	2.47*

Note. Est = estimate; Num = numeric; E1 = Estimate 1; E2 = Estimate 2.

* $p < .05$. ** $p < .10$. † $p = ns$.

trast, signed errors reflect how inaccurately the regions themselves have been placed (i.e., they reflect inaccuracies at the category level that have been inherited by the items), because they are computed by subtracting the estimated from the actual latitude for each city before averaging over cities within regions. Absolute errors are the absolute values of the signed errors, taken before averaging. They thus also reflect uncertainty at the regional level but without regard to direction.

The comparison between item- and category-level measures is important because of the theoretical assumption about the independence of the levels (Huttenlocher et al., 1991), and our assumption that so much of the variance (and bias) in latitude estimates is accounted for by biases at the regional level (Friedman et al., 2005; Friedman & Montello, 2006). For example, it is possible for the ordinal accuracy of items within a region to be relatively good even when the absolute values of the estimates are inaccurate because of the placement of the region.

Not surprisingly, there were main effects of region for all measures: for the correlations, $F(3, 306) = 34.27$, $MSE = 0.237$, $\eta_p^2 = .251$; for signed errors, $F(3, 306) = 114.03$, $MSE = 120.21$, $\eta_p^2 = .528$; and for absolute errors, $F(3, 106) = 58.18$, $MSE = 105.04$, $\eta_p^2 = .363$. The back-transformed mean correlations between actual and estimated locations across groups for cities in Canada, the northern United States, the southern United States, and Mexico were .520, .304, .265, and .101; the mean signed errors were 1.0°, -9.5°, -13.0°, and -23.4°, and the mean absolute errors were 9.4°, 12.8°, 16.0°, and 24.8°. Thus, there were differences in item-level ordinal accuracy within regions and parallel differences between regions in the accuracy of the placement of their category boundaries.

There were also main effects of estimate number for each measure. This result reflects an improvement in accuracy from the first to the second estimates: for the correlations, $F(1, 102) = 30.83$, $MSE = .12$, $\eta_p^2 = .232$; for signed error, $F(1, 102) = 241.75$, $MSE = 95.43$, $\eta_p^2 = .703$; and for absolute error, $F(1, 102) = 84.96$, $MSE = 91.45$, $\eta_p^2 = .187$. For the correlations between actual and estimated latitudes, the back-transformed means for the first and second estimates were .245 and .363, respectively; for the signed errors the means were 11.7° and 10.7°, and for the absolute errors they were 16.0° and 15.5°. Although these effects are small in absolute terms, this is one of the only

times an improvement in accuracy has been observed in this paradigm in the absence of participants' being given accurate metric information (Brown & Siegler, 1993; Friedman & Brown, 2000a, 2000b; but see Brown & Siegler, Experiment 4, and LaVoie, Healy, & Bourne, 2002). That is, the perceptual supports and memory aids improved both item- and category-level accuracy without giving participants specific metric information.

The most theoretically interesting differences between the measures were revealed by the effect of group and its interactions with the other factors. For the correlations between actual and estimated latitudes, the effect of group was significant, $F(2, 102) = 3.50$, $MSE = .41$, $\eta_p^2 = .064$, and reflects higher average within-region (i.e., item-level) correlations for the grid (.360) and line (.318) groups compared with the numeric group (.234). Further, the interaction between group and estimate number approached significance, $F(2, 102) = 2.61$, $MSE = .115$, $\eta_p^2 = .049$, $p = .08$. Analyses of the simple interactions between group and region as a function of estimate number showed that, during the first estimates, there was relatively little difference between groups on this measure; the back-transformed means for the grid, line, and numeric groups were .277, .253, and .205, respectively, $F(2, 102) < 1$. However, during the second estimates the effect of group was significant, $F(2, 102) = 5.82$, $MSE = .261$, $\eta_p^2 = .103$. Ordinal accuracy for both the grid (.442) and line (.378) groups was better than it was for the numeric group (.262) during the second estimates. This is an important finding because it implies that for item-level spatial information, being able to respond spatially and having perceptual and memorial supports produced more accurate responses than did responding numerically, even though the numeric group also had these supports.

In the omnibus ANOVA, the interaction between group and region also approached significance for the correlations between actual and estimated latitudes, $F(6, 306) = 1.92$, $MSE = .24$, $\eta_p^2 = .036$, $p = .08$, and the Wilks's lambda for the interaction (.879) was significant, $F(6, 200) = 2.43$, $p = .03$. For Canadian cities, the back-transformed means were .553, .582, and .415, respectively, for the grid, line, and numeric groups. For northern U.S. cities the means were .433, .302, and .165, respectively, and for southern U.S. cities the means were .255, .265, and .276. For Mexican cities, the means were .170, .067, and .063. In sum, the correlations between actual and estimated

latitudes indicate that both spatial response modes facilitated within-region ordinal accuracy more than did the numeric response mode, especially in the two northern regions (Canada and the northern United States).

However, a somewhat different picture emerged across groups for signed and absolute errors. Note again that whereas the within-region correlations reported above assess item-level accuracy, signed and absolute errors assess the accuracy with which the regions were placed (signed errors indicate the direction of error and absolute errors indicate accuracy without regard to direction). For signed errors, there was a two-way interaction between region and group, $F(6, 306) = 3.90$, $MSE = 120.21$, $\eta_p^2 = .071$, and a three-way interaction between region, estimate number, and group, $F(6, 306) = 7.15$, $MSE = 21.3$, $\eta_p^2 = .123$. The three-way interaction is shown in the top panel of Figure 5.

For absolute errors, there were two-way interactions between region and group, $F(6, 306) = 2.30$, $MSE = 105.04$, $\eta_p^2 = .043$, and region and estimate number, $F(3, 306) = 23.42$, $MSE = 21.20$, $\eta_p^2 = .187$. The three-way interaction between region, estimate number, and group was also significant, $F(6, 306) = 2.22$, $MSE = 21.20$, $\eta_p^2 = .042$, and is shown in the bottom panel of Figure 5. Overall, for both category-level measures, the Canadian cities for all three groups were estimated accurately whether or not there

were perceptual and memory supports. However, with the exclusion of these cities, the numeric group was relatively accurate compared with the line and grid groups, particularly for cities in the United States. For Mexican cities, the numeric and line groups were roughly equivalent, and the grid group still had the poorest accuracy.

Thus, the item- and category-level measures tell interestingly different stories. The correlations between actual and estimated latitudes within regions indicate that during the second estimates, participants who responded spatially displayed higher within-region ordinal accuracy among the items in familiar regions than did participants who made numeric estimates. In contrast, at the regional level the differences between tasks did not matter much in the placement of the items in the most familiar region (Canada), but the grid task produced the least accurate placement overall for the two southernmost regions.

In sum, the spatial and numeric tasks differentially affected information hypothesized to be retrieved from different levels in the spatial representation of global-scale geography. This dissociation provides further evidence for the independence of item- and category-level spatial information and replicates and generalizes to many more cities the pattern of estimates obtained by Friedman and Montello (2006). In that study, which used participants from Southern California, the ordinal north-south placement of cities within the two U.S. regions was very accurate, compared with the placement of cities in Canada and Mexico, but there was still a large bias in the placement of cities in the southern United States. Thus, bias at the regional level was not due to lack of familiarity per se but rather to metric beliefs about the absolute locations of the regions and their boundaries, which were then inherited by the cities.

Experiment 3

When participants were provided with a means to identify the cities and correct their responses, both configural accuracy and some aspects of north-south, within-region ordinal accuracy for latitudes improved; however, this information did not eliminate the mutual exclusivity, diminish the gaps, or improve accuracy in the overall placement of the regions for any of the groups. Some of the improvement in configural accuracy for the grid group may have been achieved because, in the absence of any trial-to-trial memory requirements, participants could reveal their item-level knowledge of familiar regions more precisely. Nevertheless, though there were interesting differences between spatial and numeric response modalities, the modalities did not change the fundamental nature of the pattern of data observed.

Although both the grid and line tasks allowed for spatially based responding, they did not provide specific metric information beyond the initial instructions about how latitudes and longitudes work. Further, the numeric reference points on the grid and line might have influenced responses in some subtle or not so subtle way (e.g., by reminding participants that the equator is at 0°). In any case, there were still numeric reference points in the stimulus displays of the two spatial conditions. In Experiment 3, participants were asked to respond directly on an outline map of North America that had no numeric or grid information, to eliminate the possibility that any numeric display characteristics could be used to facilitate responding. Direct response on a map provided an

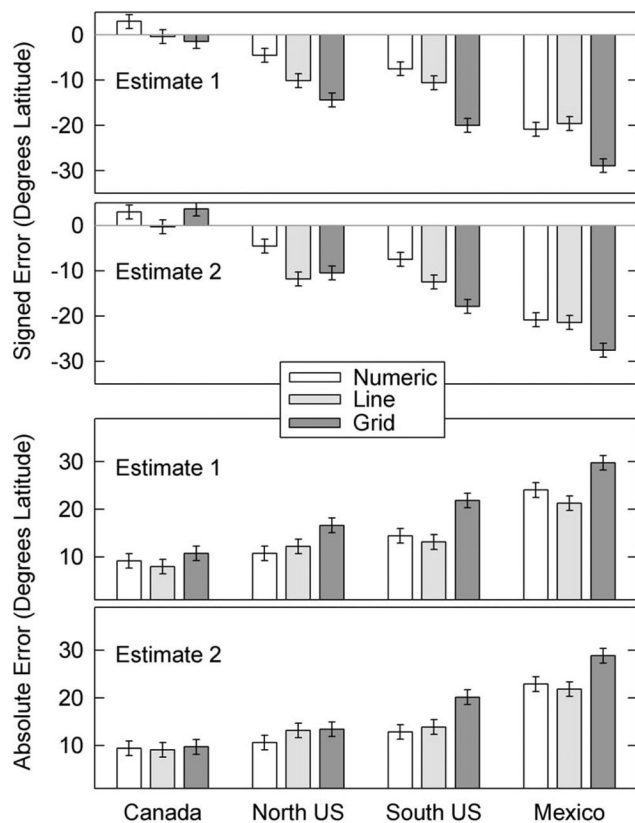


Figure 5. Mean signed and absolute errors in degrees latitude as a function of condition, estimate number, and region in Experiment 2. Error bars are the 95% confidence intervals computed with the error term for the Group \times Estimate Number \times Region interaction (Loftus & Masson, 1994).

explicit test of the relative importance of accurate spatial cues and prior category information in the accuracy of global-scale estimates.

Spatial cues, such as maps, have been shown to improve the accuracy of geographic location estimates at an urban scale. For example, Kitchin and Fotheringham (1998) had participants estimate the locations of 25 places in the Swansea region of the United Kingdom on a piece of paper that either had two locations indicated on it or had an outline of the coast of the region in addition to the two reference locations. Participants who had the map outline were approximately 30% more accurate than were participants who had only the reference locations. In general, giving participants an outline map is functionally equivalent to giving them the correct metric "space" in which to respond; in the present case, that meant they could not use the equator as a lower bound for the response range. Indeed, because participants were responding on a map, there should be relatively little or no signed error in the placement of southern U.S. and Mexican cities, because the map's boundaries constrain the responses.

However, absolute error measures accuracy without regard to the direction of error. On one view, there should be less absolute error made to Mexican cities than to cities in the United States when participants are responding on a map because there is actually less physical space on the map for Mexico than there is for the United States, given the map projection being used. On the other hand, if participants are uncertain of the exact locations of most Mexican cities and are thus heavily weighting a prototype to make their estimates, most of the Mexican estimates would be relatively inaccurate. In this case, although overall accuracy would necessarily improve relative to the accuracy of the previous grid groups, the absolute error to Mexican cities would be larger than for any other region, as we have found in all previous research. In addition, if participants are still using a regionalized long-term memory representation to respond and the regional borders do not overlap in their representations, provision of a map should not change the overall configuration of the cities relative to that seen in the grid group of Experiment 2.

Method

Participants. Thirty-eight individuals (23 women, 15 men) from the same pool as Experiments 1 and 2 participated in Experiment 3. Of the women, 3 had their data dropped, 2 because they failed to follow the instructions and 1 because her questionnaire data were missing. The remaining 35 individuals were all born in Canada and had lived in the Edmonton metropolitan area for an average of 12.8 years ($SE = 1.4$).

Stimuli and procedure. The procedure and instructions were identical to those used for the grid group in Experiment 2, including the information about latitudes and longitudes (which was actually irrelevant for the map task). The exceptions were the display itself and some slight variations that were necessary for reference to the map.

For both estimate tasks, an equirectangular outline projection of North America was displayed on the computer screen without a grid or any other geographic features (e.g., the Great Lakes). Thus, this task was a "pure" spatial location estimate task, insofar as there was no latitude or longitude lines or any other numeric reference points. The rectangle that enclosed the map was 850

pixels across and 425 pixels high, which preserved the approximately 2:1 longitude to latitude ratio for North America that results from this kind of projection. The northwestern and southeastern coordinates of the corners of the rectangle corresponded to 73°N, 171°W, and 14°N, 53°W, respectively.

Results

Knowledge ratings. The knowledge ratings replicated the pattern seen in Experiments 1 and 2. The means for Canadian, northern U.S., southern U.S., and Mexican cities were 6.4, 3.1, 3.4, and 1.8, $F(3, 102) = 160.51$, $MSE = .83$, $\eta_p^2 = .825$.

Location estimates. Figure 6 shows the results for both estimates. Most important, and somewhat surprising, all of the features seen in previous research are evident here, in particular the lack of overlap and the gaps between regions. There was no evidence that the configuration as a whole became more accurate during the second set of estimates. The bidimensional regression coefficients for the first and second estimates (.877 and .879, respectively) were both high, but there was certainly room to move higher.

A Group \times Estimate Number ANOVA was conducted on the Fisher-transformed bidimensional coefficients to compare the present results with those of the grid group of Experiment 2. There was an overall improvement in configural accuracy with the map, $F(1, 68) = 10.90$, $MSE = 0.18$, $\eta_p^2 = .138$; the back-transformed average bidimensional correlations between actual and estimated locations were .812 for the grid group and .878 for the map group. There was also an overall main effect of estimate number, $F(1, 68) = 8.80$, $MSE = 0.015$, $\eta_p^2 = .115$, and a Group \times Estimate Number interaction, $F(1, 68) = 6.58$, $MSE = 0.015$, $\eta_p^2 = .088$. The interaction was attributable to the improvement between Estimates 1 and 2 for the grid group (.791 and .830, respectively) but not for the map group (.877 and .879 respectively), as well as to the overall higher accuracy of the map group's estimates.

Examination of the data in Rows 5 through 8 of Table 2 shows that, except for estimates for the Canadian cities, the map group's estimates were generally farther north than were the grid group's estimates. In addition, the difference between groups increased as regions were farther south. For example, for second estimates of the cities at the regional borders shown in Table 2, the difference between the estimates of the map and grid groups (left to right) was 1.8°, 4.8°, 10.1°, 13.9°, 14.2°, and 25.8°. These differences follow the increase in bias observed for the grid group and all the other groups we have tested, both with numeric and spatial formats. The implication is that having the map prevented participants from stretching North America southward; nevertheless, having the map did not ameliorate the lack of overlap between borders or the gaps between regions. Also notable is the fact that, exactly like the Californian participants in Friedman and Montello (2006), participants were correct that Miami was the southernmost U.S. city and Tijuana was the northernmost Mexican city in the set. However, they would have been more accurate if they had estimated Tijuana to be north of Miami and closer to Los Angeles. That they did not do so implies that, like the Californians, the present participants subscribe to the general belief that all cities in the United States are north of all cities in Mexico.

On average, for the grid group, the second estimates for Canadian, northern U.S., southern U.S., and Mexican cities were 52.2°,

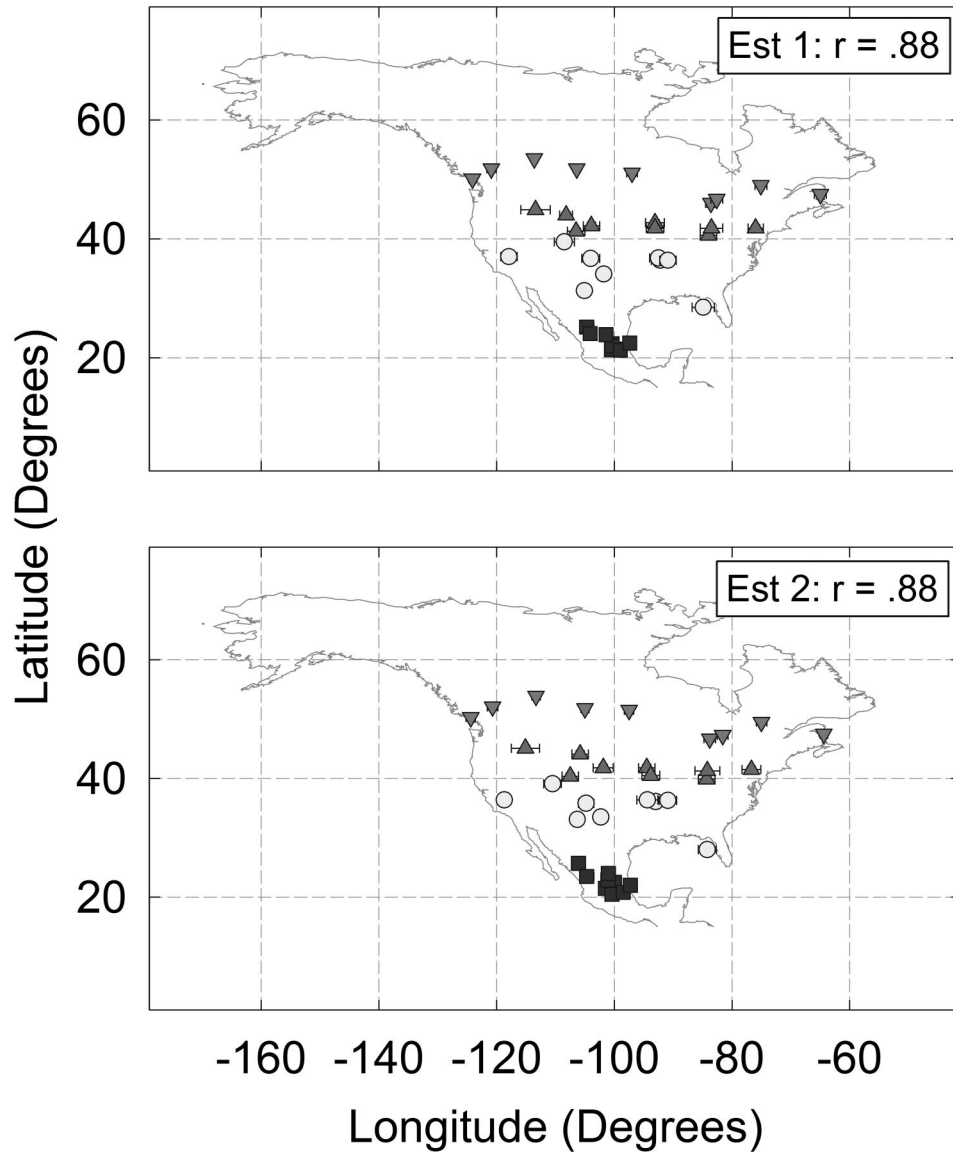


Figure 6. Mean first and second estimates for the group that responded on the map in Experiment 3. Error bars are standard errors of the mean across participants for each city, computed independently for each dimension. Inverted triangles are Canadian cities, triangles are northern U.S. cities, circles are southern U.S. cities, and squares are Mexican cities. Participants responded on the outline map of the North American continent that is displayed in the figure, but they did not have either a grid or any other information on the display.

30.9°, 12.9°, and -4.3°, respectively; for the map group, they were 50.0°, 41.8°, 35.0°, and 22.6°. Thus, relative to the map group, the grid group had indeed stretched the continent southward, and the amount of “stretch” or bias increased with each region. Yet, the absolute errors for the map group, though small, showed the identical pattern across regions as did those for the grid group: For the first estimates in the map group, the average absolute errors to Canadian, northern U.S., southern U.S., and Mexican cities were 2.3°, 2.8°, 3.8°, and 5.2°, respectively; for the second estimates, they were 2.3°, 3.0°, 3.3°, and 5.4°. The effect of region was reliable, $F(3, 102) = 45.33$, $MSE = 2.62$, $\eta_p^2 = .057$, as was the Region \times Estimate Number interaction, $F(3, 102) = 6.26$, $MSE =$

0.295, $\eta_p^2 = .155$. A planned contrast on just the second set of estimates showed that the difference in absolute errors between southern U.S. and Mexican cities was significant, $F(1, 34) = 46.14$, $MSE = 3.23$, $\eta_p^2 = .576$. Thus, even when participants responded on a map and were able to change their responses at will, the largest absolute error in the estimates was for the Mexican cities.

The pattern of within-region correlations between the actual and estimated latitudes for the map group was also remarkably similar to that for the grid group. In particular, from north to south, the back-transformed means for the four regions were .815, .541, .554, and .131, respectively, for the map group’s first set of estimates and were .846, .560, .606, and .070, respectively, for the second

estimates. Only the effect of region was significant, $F(3, 102) = 49.00$, $MSE = .284$, $\eta_p^2 = .590$. Thus, the within-region estimates were again the most accurate for cities in the home region (Canada) and the least accurate for the Mexican cities.

Questionnaire Responses: Experiments 2 and 3

After the second estimate task in Experiments 2 and 3, participants filled out the same questionnaire. For this analysis, we treated the data from the numeric, line, grid, and map groups as if the experiments were conducted simultaneously and participants had been randomly assigned to the four conditions.

Strategy questions. There were three strategy questions. For the first, the strategy that participants across the four groups claimed to use "most of the time" was to visualize a map (30%), or to position the cities relative to the location of their province/state (22.9%), other cities that had already been placed (22.1%), the national borders (17.1%), the equator (5.7%), Edmonton (0.7%), or some other familiar city (1.4%).

Only two strategies distinguished among the groups. First, placing each city in relation to the borders of its province or state was favored primarily by the map group, $\chi^2(3, N = 32) = 20.25$, $p < .05$. (The percentages of participants who favored this strategy in the numeric, line, grid, and map groups were 11.4%, 14.3%, 11.4%, and 54.3%, respectively.) Presumably, it was easier for the map group to visualize provincial and state borders than it was for the other groups. This conjecture is supported by Simmering and Spencer (2007), who showed that people cannot impose arbitrary spatial category boundaries without perceptual support. Second, placing each city in relation to others that had already been placed was favored by a substantial portion of the participants in all groups except the map group, although this difference only approached reliability. The percentages for the numeric, line, grid, and map groups were 25.7%, 31.4%, 25.7%, and 5.7%, respectively, $\chi^2(3, N = 31) = 6.03$, $p < .11$. Here, it is clear that participants who did not have a map were aware that they were using interitem spatial relations to make their responses, though the within-region correlations show that the map group used these relations equally well.

The location estimates for all the groups imply that participants believed that the southern boundary of North America is at the equator or below, but few participants said they used the equator most of the time as a reference point to place individual cities. This result provides important support for the theoretical claim that people first partition the response space (numeric or spatial) among the categories (continents and their regions) and during this initial stage, they use beliefs about the location of the equator and other reference points to set the upper and lower bounds for North America. The bias in the individual estimates is therefore due to the bias in the initial partitioning of the response space, which in turn likely reflects participants' a priori beliefs.

The second strategy question was designed to provide evidence that use of multiple strategies (Brown, 1995 2002a) and plausible reasoning (Brown & Siegler, 1993; Collins & Michaelski, 1989; Friedman & Brown, 2000a, 2000b) is a normal feature of estimation processes, whether the estimates are numerically or spatially based; participants selected as many strategies as needed to indicate how they performed the task overall. When the alternatives that involved using Edmonton (7.9%) or another city (8.1%) were

excluded, there was a relatively even distribution of choices selected among strategies and groups: visualizing a map (21.3%), positioning a city relative to other cities (19.5%) or to state/provincial (17.9%) or national (15.7%) borders. The only strategy selected in different proportions across groups was that of trying to position each city in relation to the equator, $\chi^2(3, N = 47) = 16.06$, $p < .05$. Participants in the map group never selected that strategy, whereas participants in the numeric, line, and grid groups said they used it 11.1%, 12.3%, and 13.2% of the time, respectively. This result, too, suggests that the equator is used relatively infrequently when placing individual items. In addition, the fact that no participants in the map group said they used the equator as a reference point provides an important manipulation check: Provision of a map precluded their use of the equator to partition the response space; the space that was partitioned was the space within the map itself.

The third question asked what single piece of information could be provided that would help people make more accurate location estimates. The choices selected were the latitude of a national border (47.1%) or another (familiar) city in the same country as the city being estimated (20.7%), the distance between Edmonton and the city to be estimated (12.9%), the average annual temperature of the city (10%), the latitude of a geographic landmark in that city's country (6.4%), and the number of degrees of latitude there are in a kilometer (2.9%), $\chi^2(5, N = 140) = 109.17$, $p < .05$. These data indicate that participants were sensitive to the notion that several types of reference information and proxy cues (Lee, 2005), such as climate, could be differentially useful for those making accurate location estimates.

The only type of information that was mentioned differentially among the groups was the distance between Edmonton and the city to be estimated: A substantial percentage of the participants in the map group (31.4%) selected that as their choice, whereas only 2.9%, 8.6%, and 8.6% of participants in the numeric, line, and grid groups selected it, $\chi^2(3, N = 18) = 13.11$, $p < .05$.

Boundary questions. The final two questions asked about the northern and southern boundaries of North America. The alternatives for both questions appeared as numeric answers, every 10°, from 90°N to 90°S, with the 0° alternative annotated as the equator.

Participants were relatively accurate with respect to the northern boundary. When the Arctic islands (e.g., Baffin Island; Ellesmere Island) are included, the approximately correct answer is 80°N. Across groups, 40% of the participants selected this response and a full 30% more selected the North Pole, which is reasonable given current international news about Canadian and other governments' views on Canadian sovereignty over the Northwest Passage (Carnaghan & Goody, 2006). Across groups, the percentages of participants who selected the alternatives 90°, 80°, 70°, 60°, 50°, and any category below 50° were 30.0%, 40.0%, 17.9%, 5%, 2.9%, and 4.3%, respectively, $\chi^2(3, N = 140) = 101.11$, $p < .05$. However, there was no difference between groups in the percentage of participants who chose each category.

In contrast to the accuracy for the northern boundary, a substantial 53.5% of the participants across groups selected either the equator (15.0%) or a category that was below the equator (38.5%) as the southern boundary for North America. Indeed, the percentages of participants who selected a category that was at the equator or below for the lower bound of the continent were 40.0%, 48.5%,

62.8%, and 42.8%, respectively, for the numeric, line, grid, and map groups. This result suggests that the estimation accuracy of the map group was by default. Participants in the map group explicitly agreed with those in the other groups regarding the loci of the continent's boundaries; however, because we scored the data according to the latitudes on the map projection, the map group's estimates could not reflect that belief. Across groups, the proportion of participants who selected alternatives that either were above 20° north or were 20°N, 10°N, 0°, 10°S, and below 10°S was 9.3%, 12.1%, 25.0%, 15.0%, 7.1%, and 31.4%, respectively, $\chi^2(3, N = 140) = 38.29, p < .05$.

General Discussion

Across three experiments, seven independent groups had more similarities than differences among them in the pattern of data observed. For all the groups, the data demonstrated that participants' assumptions about the locations of geographical regions and their boundaries were highly influential in determining the observed biases in the estimates. In a sense, the data across groups stand as a statistically powerful replication of the same effect; it is exactly this kind of "null" replication, observed in their 2004 study, that led Huttenlocher et al. to conclude that when estimating dot locations, it is "better to leave boundaries where precision is greatest, at the vertical and horizontal axes, rather than to use inductive categories to estimate location even though distributions are uneven" (p. 93). Finding that various manipulations do not change the pattern of results in each domain strengthens conclusions about the important role played by categories in both of them, despite some otherwise significant differences between tasks and stimuli.

The present data support the idea that, unlike the category boundaries imposed by participants upon remembered circles, the long-term representations of real-world geographic category boundaries are biased and these biases are inherited by the items during the estimation process. There are, of course, other important differences between use of category boundaries to estimate dot locations and use of category information to estimate the locations of real-world cities. For example, despite the apparent immutability of category boundaries in the dot location task (Huttenlocher et al., 2004), participants who are provided with accurate information about the location of one or two cities can shift the location of geographic boundaries and improve metric accuracy (Friedman & Brown, 2000a, 2000b). However, giving people accurate information about the locations of cities is different from giving people information about the locational distribution of the cities relative to their superordinate category boundaries. Changing locational distributions (i.e., the ordinal and metric relations among items) is impossible in the real world but quite possible with arbitrary stimuli.

Further, the influence of categories on estimation accuracy may change with the domain, because it is unlikely that real-world category boundaries are as precise or well specified as the vertical and horizontal axes of a circle. Thus, in the dot location task, use of the horizontal and vertical axes as category boundaries generally contributes to more accurate estimations overall (Huttenlocher et al., 2004). In contrast, as noted earlier, geographical representations are influenced by a multiplicity of sources; these can lead people locate a given category's boundaries quite inaccurately (Friedman & Brown, 2000a, 2000b; Friedman & Montello, 2006).

Geographic category boundaries may be referenced to global landmarks (Friedman et al., 2005; Friedman & Montello, 2006), or they may derive from beliefs about regional climates (Friedman & Brown, 2000b; the present data) or the locations of noncontiguous but "conceptually coordinated" regions. For example, the southern United States and Mediterranean Europe are regions whose cities are both estimated to be too far south by approximately the same amount; it is possible to eliminate the bias in both regions by giving participants correct information about the latitude of a particular city or two—a "seed city"—in only one of them (see Friedman & Brown, 2000b). Additionally, category boundaries can be influenced by geopolitical beliefs about the regions (Carbon & Leder, 2005) or feelings about the people who live there (Kerkman et al., 2004). Thus, for geographical categories, inaccuracies in the long-term memory representations of the location of the category boundaries and, thus, their prototypes are often quite large. This leads to large biases in the estimates.

Beliefs about the location of real-world boundaries are not necessarily ill-informed; rather, both the estimates and the questionnaire data provide evidence that the beliefs result from plausible, but incorrect, assumptions. In the present experiments, all the groups except the map group, which was constrained by the map's borders, stretched North America to the equator or beyond. Indeed, even in the map group, 42.8% of the participants explicitly stated that they believed that the southern boundary of North America was at or below the equator; even our Mexican participants (Friedman et al., 2005) appeared to subscribe to this belief. If a person holds this belief, he or she will necessarily have biased estimates for cities in the more southern regions of the continent.²

Other real-world beliefs about the cities, if deemed relevant to the dimension being estimated, can affect real-world location estimates (e.g., climate, the number of time zones). These dimensions stand as better known proxies for the dimension that is actually being estimated (Lee, 2005). The questionnaire responses in the present study indicate that it is more likely that beliefs about the location of regions and their boundaries affected the initial

² Because the strategy reports of participants in Experiments 2 and 3 were necessarily retrospective, with a consequent potential bias as a function of memory for their estimates, an independent group of 120 Canadian-born participants from the same population as the participants in Experiments 1–3 (39 men, 53 women, 28 of unknown gender because of missing names) volunteered to participate in a brief questionnaire study conducted in a large classroom setting. The instructions that were given about how latitudes work were identical to those used in the present experiments; the questions were in multiple-choice format. Participants were asked to imagine making latitude estimates for cities in North America and to indicate "the single piece of information we could give you, along with each city, that would help you make a more accurate latitude estimate than if you did not have this information." In general, participants anticipated that the latitude of either a national border or a familiar city, followed by climate information and information about the latitudes of global geographical landmarks, would help their accuracy most. More important, participants were next asked to indicate the latitudes of the northern and southern borders of North America. They were relatively accurate about the location of the northern border of North America, yet over 30% of them thought that the southern border of the continent was at or below the equator, which they knew to be at 0° from the instructions. Thus, a priori, a substantial portion of participants believed that North America is bounded at the equator or below it.

parsing of the response space into continents and the regions within them, rather than the placement of the cities within the regions once the category boundaries had been established. The persistence of the gaps between regions provides further evidence for the use of plausible (though erroneous) reasoning. For example, although Miami, Florida, is about 500 miles south of Tijuana, Baja California, participants responded as if they held the belief that all Mexican cities are to the south of all cities in the United States (see Table 2). Thus, the data in the present study provide important new evidence that plausible reasoning may often underlie biases in both spatially and numerically based location estimates.

The present study also extends the notion of ordinal conversion in the general metrics and mapping framework (Brown & Siegler, 1993) to support the theoretical position that even when the response is nonnumeric, the response space is divided among the regions that are believed to exist in that space. Accordingly, partitioning the total response range—numeric or spatial—among preexisting psychological regions is an inherent part of the process of making spatial location estimates and is largely responsible for their mutual exclusivity. Thus, when they estimate the location of dots in circles, people divide the response space into four quadrants based on the horizontal and vertical axes (Huttenlocher et al., 1991, 2004); when they estimate the location of cities in continents, people divide the response space according to their beliefs about the locations of preexisting geographical regions relative to global landmarks and other types of beliefs. However, it is certainly possible for people to select response ranges that are mutually exclusive and yet not leave gaps between the regions. Thus, it is likely that the gaps have a different cause. In particular, and in accordance with Huttenlocher et al.'s (1991) model, the gaps are likely to result from the overweighting of regional prototypes, which also tends to truncate the range of estimates within a region.

Notably, seed cities do not affect the mutual exclusivity between regions, but they do improve accuracy in the placement of the regions themselves (e.g., Friedman & Brown, 2000b). Thus, seed information is effective, when identity information is not, because seeds ameliorate incorrect beliefs about metric properties that can be propagated upward to the parent categories; all the items within those categories and within any coordinated categories will be affected and will become more accurate as a result. In contrast, in the present study we showed that access to item identities and prior estimates improves ordinal accuracy within regions but not metric accuracy in the placement of the regional boundaries. This is an important constraint on theories of both numeric and spatial estimation, and it provides further support for the independence of the levels of representation.

In addition to the implications for theories of spatial location estimates, the present study had several new empirical findings. For example, the perceptual and memory aids and the ability to respond spatially did facilitate some aspects of performance: For the grid group in Experiment 2, these manipulations led to an improvement in configural accuracy; for the grid and line groups, there was an improvement in the north–south ordinal accuracy for items within familiar regions. This finding is in contrast to the data in Crawford et al. (2000) and Choplin et al. (2001). Further, the improvements in both item and configural accuracy occurred without participants being given any specific metric information or relevant general information about the metric (e.g., as was done by Brown & Siegler, 1993, and LaVoie, Healy, & Bourne, 2002). The

improvements in accuracy obtained by allowing participants to see and change their former responses indicate that memory for prior responses does influence performance in this task. Therefore, it may be an important consideration in other kinds of estimate tasks, both numeric and spatial, if those tasks involve information about relations among items in long-term memory.

A second new finding was that the ability to respond spatially rather than numerically improved the ordinal accuracy of items within regions but was often disadvantageous at the regional level. This finding implies that numeric judgments may underestimate some forms of accuracy in spatial domains of expertise (e.g., within categories), whereas spatial judgments may underestimate others (e.g., between categories); thus, a complete theory of estimation processes must account for performance in tasks and domains that differ in kind.

Third, the responses to the strategy questions in Experiments 2 and 3 provide several new insights. Participants used a variety of strategies to make their estimates. They were aware of the relation between latitude and climate and on occasion used climate as a proxy (Lee, 2005) for locations in the north–south dimension. Although we have always assumed this to be true, this study is the first in which people were asked to indicate explicitly how they made judgments in this domain. In addition, the questionnaire responses converge on the idea that a substantial portion of participants believe that North America stretches to at least the equator or beyond. Yet, they do not generally claim to use the equator as a reference point in making individual estimates. Thus, global landmarks may be used to set the response range and perhaps to place the regions within that range in an absolute sense. But after that, individual cities are placed using information that is primarily based on within-region knowledge, including the inter-item relations between cities and the prototype for the region. Because prior estimates were available to be adjusted as needed, participants could display this knowledge more fully.

The most surprising new results are those from Experiment 3. Though the configuration of the second set of estimates for the map group was more accurate overall than was the configuration for the grid group of Experiment 2, the map group's estimates still exhibited mutual exclusivity between regions, gaps between the national borders, little north–south discrimination among cities within regions, and an increase in absolute error as the cities being estimated were actually located farther south. As noted earlier, Kitchin and Fotheringham's (1998) participants improved their accuracy substantially when they responded on an outline map. The present participants did too; however, that was principally because the map constrained the absolute amount of error they could display for the two southernmost regions. Yet, their overall pattern of responding, including the pattern of absolute error, was the same as in all the other groups. Thus, the role of spatial cuing in general, and of maps in particular, in improving accuracy for real-world location estimates may be more limited than previously thought: Maps may improve estimation accuracy merely because they limit the use of global landmarks to partition the response space.

The data from the present study converge on the following conclusions. First, the division of North American cities into mutually exclusive regions is a general phenomenon of global-scale location estimates. It has been observed among all groups of North American participants tested thus far, with both numeric and

spatially based estimates, whether they involved one dimension (e.g., the line and numeric groups) or two (e.g., the grid and map groups). We have seen similar features in North Americans' estimates of cities in Europe and Africa (Friedman & Brown, 2000a, 2000b). Thus, the use of regional categories in global-scale geographic location estimates appears to be ubiquitous.

Second, the hallmarks of regionalization cannot be attributed to perceptual errors or biases that derive from perception per se—even the perception of maps—because the map group's data were also regionalized. Third, whereas perceptual supports do eliminate the bias in line-length and size estimates (Choplin et al., 2001; Crawford et al., 2000), they do not eliminate the bias in real-world geographic estimates. Fourth, the regions and biases are unlikely to be due to faulty retrieval processes either, because participants had a constant means of checking and updating their responses. And finally, because the map group's estimates had characteristics similar to those observed in the other groups' estimates, it is unlikely that the biases observed are a function of making numeric estimates or of having numbers available as referents (as on the grid). It is also unlikely that the conceptual regionalization of the continents in both the Old and New Worlds is something participants learned during the task; rather, whatever people learn from maps (and from general geography, politics, history, and so on), the result is a regionalized representation of the world. Indeed, a recent study (Battersby, 2006) showed that the effect of different map projections on spatial cognition is minimal, even though different map projections vary widely in the way that they depict the areas and positions of the continents and major bodies of water. Our present and past data imply strongly that the conceptual regionalization of the continents must be part of the representations that participants bring to the task, and the observed biases in the data are produced by processes that combine a remembered but inexact item (city) value with biased category (regional) information.

The idea that biases are produced by combining information from two levels was originally hypothesized by Huttenlocher et al. (1991; see also Newcombe & Huttenlocher, 2000, for discussion). The present study supports Huttenlocher et al.'s (1991) theory and extends it to global-scale, real-world estimates. Further, the present study identifies the main source of the difference in accuracy between estimating dots and cities as being inherent in errors in the beliefs about, and precision of, the locations of the category boundaries for the real-world stimuli. The present study also extends the metrics and mapping framework (Brown & Siegler, 1993) to situations in which the knowledge base and response modalities are principally spatial.

The way that various factors contribute to the categorization of cities in the first place, as well as to the biased placement of real-world geographic categories, remains an interesting issue for further research. People have many sources and kinds of information about geography, so there are many factors that could possibly contribute to the mental partitioning of the world to provide a challenge for future research.

References

- Barsalou, L. W., Simmons, W. K., Barbey, A. K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*, 7, 84–91.
- Battersby, S. E. (2006). Distortion of area in the global scale cognitive map: A geographic perspective. Retrieved from ProQuest Digital Dissertations. (AAT 3226251).
- Brown, N. R. (1995). Estimation strategies and the judgment of event frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 1539–1553.
- Brown, N. R. (2002a). Encoding, representing, and estimating event frequencies: A multiple strategy perspective. In P. Sedlmeier & T. Betsch (Eds.), *Frequency processing and cognition* (pp. 37–53). Oxford, England: Oxford University Press.
- Brown, N. R. (2002b). Real-world estimation: Estimation modes and seeding effects. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory*. Vol. 41 (pp. 321–359). San Diego, CA: Academic Press.
- Brown, N. R., Cui, X., & Gordon, R. D. (2002). Estimating national populations: Cross-cultural differences and availability effects. *Applied Cognitive Psychology*, 16, 811–827.
- Brown, N. R., & Siegler, R. S. (1993). Metrics and mappings: A framework for understanding real-world quantitative estimation. *Psychological Review*, 100, 511–534.
- Carbon, C. C., & Leder, H. (2005). The Wall inside the Brain: Overestimation of distances crossing the former iron curtain. *Psychonomic Bulletin & Review*, 12, 746–750.
- Carnaghan, M., & Goody, A. (2006, January). *Canadian Arctic sovereignty* (Publication No. PRB 05–61e). Retrieved from <http://www.parl.gc.ca/information/library/PRBpubs/prb0561-e.htm>
- Choplin, J., Huttenlocher, J., & Kellman, P. (2001). Perceptual discrimination and memory. *Journal of Vision*, 1, 472.
- Collins, A., & Michalewski, R. (1989). The logic of plausible reasoning: A core theory. *Cognitive Science*, 13, 1–49.
- Crawford, L. E., Huttenlocher, J., & Engebretson, P. H. (2000). Category effects on estimates of stimuli: Perception or reconstruction? *Psychological Science*, 11, 280–284.
- Duffy, S., Huttenlocher, J., & Crawford, L. E. (2006). Children use categories to maximize accuracy in estimation. *Developmental Science*, 9, 597–603.
- Friedman, A., & Brown, N. R. (2000a). Reasoning about geography. *Journal of Experimental Psychology: General*, 129, 193–219.
- 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.
- Friedman, A., Brown, N. R., & McGaffey, A. (2002). A basis for bias in geographical judgments. *Psychonomic Bulletin & Review*, 9, 151–159.
- Friedman, A., Kerkman, D. D., & Brown, N. R. (2002). Spatial location judgments: A cross-national comparison of estimation bias in subjective North American geography. *Psychonomic Bulletin & Review*, 9, 615–623.
- 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.
- 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.
- 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: General*, 32, 333–346.
- Haubensak, G. (1992). The consistency model: A process model for absolute judgments. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 303–309.
- Hirtle, S. C., & Jonides, J. (1985). Evidence of hierarchies in cognitive maps. *Memory & Cognition*, 13, 208–217.

- 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.
- Huttenlocher, J., Hedges, L. V., & Vevea, J. L. (2000). Why do categories affect stimulus judgment? *Journal of Experimental Psychology: General*, *129*, 220–241.
- Kerkman, D. D., Friedman, A., Brown, N. R., Stea, D., & Carmichael, A. (2003). The development of geographic categories and biases. *Journal of Experimental Child Psychology*, *84*, 265–285.
- Kerkman, D. D., Stea, D., Norris, K., & Rice, J. L. (2004). Social attitudes predict biases in geographic knowledge. *Professional Geographer*, *56*, 258–269.
- Kitchin, R. M., & Fotheringham, A. S. (1998). The effects of spatial and locational cuing on the analysis of aggregate cognitive mapping data. *Environment and Planning A*, *30*, 2245–2253.
- LaVoie, N. N., Healy, A. F., & Bourne, L. E., Jr. (2002). Memory seeding: Representations underlying quantitative representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 1137–1153.
- Lee, P. J. (2005). The role of ordinal information and the effects of cue validity in numerical estimation. Retrieved from ProQuest Digital Dissertations. (AAT NR08671).
- Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, *1*, 476–490.
- Maki, R. H. (1981). Categorization and distance effects with spatial linear orders. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *7*, 15–32.
- McNamara, T. P. (1986). Mental representations of spatial relations. *Cognitive Psychology*, *18*, 87–121.
- McNamara, T. P., & Diwadkar, V. (1997). Symmetry and asymmetry in human spatial memory. *Cognitive Psychology*, *34*, 160–190.
- 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: Taylor & Francis.
- Newcombe, N. S., & Huttenlocher, J. (2000). *Making space: The development of spatial representation and reasoning*. Cambridge, MA: MIT Press.
- 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.
- Stevens, A., & Coupe, P. (1978). Distortions in judged spatial relations. *Cognitive Psychology*, *10*, 422–437.
- Tversky, B. (1981). Distortions in memory for maps. *Cognitive Psychology*, *13*, 407–433.

Received October 9, 2006

Revision received July 11, 2008

Accepted July 24, 2008 ■

Low Publication Prices for APA Members and Affiliates

Keeping you up-to-date. All APA Fellows, Members, Associates, and Student Affiliates receive—as part of their annual dues—subscriptions to the *American Psychologist* and *APA Monitor*. High School Teacher and International Affiliates receive subscriptions to the *APA Monitor*, and they may subscribe to the *American Psychologist* at a significantly reduced rate. In addition, all Members and Student Affiliates are eligible for savings of up to 60% (plus a journal credit) on all other APA journals, as well as significant discounts on subscriptions from cooperating societies and publishers (e.g., the American Association for Counseling and Development, Academic Press, and Human Sciences Press).

Essential resources. APA members and affiliates receive special rates for purchases of APA books, including the *Publication Manual of the American Psychological Association*, and on dozens of new topical books each year.

Other benefits of membership. Membership in APA also provides eligibility for competitive insurance plans, continuing education programs, reduced APA convention fees, and specialty divisions.

More information. Write to American Psychological Association, Membership Services, 750 First Street, NE, Washington, DC 20002-4242.