International Journal of Geographical Information Science, 2014
http://dx.doi.org/10.1080/13658816.2014.900178

Taylor & Francis
Taylor & Francis Group

Vague cognitive regions in geography and geographic information science

Daniel R. Montello[a]*, Alinda Friedman[b] and Daniel W. Phillips[a]

[a]*Department of Geography, University of California–Santa Barbara, Santa Barbara, CA, USA;*
[b]*Department of Psychology, University of Alberta, Edmonton, Canada*

(*Received 5 November 2013; accepted 23 February 2014*)

Cognitive regions are regions in the mind, reflecting informal ways individuals and cultural groups organize their understanding of earth landscapes. Cognitive region boundaries are typically substantially vague and their membership functions are substantially variable – the transition from outside to inside the region is imprecise or vague, and different places within the region are not equally strong or clear as exemplars of the region. Methods for assessing and cartographically depicting cognitive regions, as with other vague geographic regions, have traditionally implied an inappropriate level of boundary sharpness and membership uniformity, such as when boundaries are mapped as precise lines. Research in recent decades has explored methods for assessing and depicting boundary vagueness and membership variability, either within or across individuals, but has still assumed homogeneity and regularity in the vagueness and variability. In this article, we present two studies that assess the cognitive regions of 'Northern' and 'Southern' California, and, for comparison, 'Northern' and 'Southern' Alberta. The first study uses a standard boundary-drawing task; the second uses a novel task in which participants rate cells of a high-resolution grid laid over an outline map. This technique allows us to assess and depict vagueness and nonuniformity that is heterogeneous and irregular across different areas. Differences in the conceptualization of 'Northern' and 'Southern' regions in California, as compared to those in Alberta, point to thematic influences on cognitive regions in California but not in Alberta. As is often true with cognitive regions, Northern and Southern California are about attitude, not just latitude.

Keywords: regions; boundary vagueness; cognition

1. Introduction

One of the most venerable concepts in geography is the *region* (Semple 1907, Minshull 1967, Martin 2005). Geographic regions are (approximately) two-dimensional features that geographers and lay people use to understand the (near) earth surface, as well as reason and communicate about it. A region[1] encompasses places that are internally similar to each other and externally dissimilar to places outside the region. Thus, regions are spatial categories. We recognize that categories and the concepts upon which they are based are neither naively realistic entities that exist independent of cognitive and social acts, nor arbitrarily created fantasies that have no particular correspondence to anything in reality independent of semiotic acts. In other words, we advocate a deeply interactionist (or transactionist) view of the metaphysics of regions: They often do correspond non-arbitrarily to real entities, properties, and processes but are always created as intellectual

*Corresponding author. Email: montello@geog.ucsb.edu

or cognitive acts (Montello 2003). A virtually unlimited number of regions and systems of regionalization can be identified, but some are easier to identify and are more useful than others because they correspond more readily to the reality of a heterogeneous earth surface or they more readily serve the needs of a particular problem.

All regions have boundaries, which are the transitions between places inside the region and places outside. One of the most important properties of boundaries is that they vary in their precision or sharpness or, conversely, their *vagueness*. Some boundaries are very sharp, essentially linear features with little or no width, like the formal geometric concept of a line. In reality, geographic boundaries are never strictly geometric lines, except potentially. Quite commonly, however, geographic boundaries are not sharp at all but are really two-dimensional features – regions – themselves. Vague boundaries are transition zones rather than lines between neighboring regions (as in frontier zones – Prescott 1965), but they are just as real as sharp boundaries. Montello (2003) and several authors in Burrough and Frank (1996) discuss different reasons for boundary vagueness. Other work in geography has focused on quantifying and representing vague boundaries in computational systems (e.g., Mark and Csillag 1989, Burrough and Frank 1996, Wang and Hall 1996) or focused on their cartographic depiction (Montello *et al.* 2003, Kronenfeld 2005). Several researchers have referred to vague boundaries as *fuzzy* and explored fuzzy logic as a formalization of vague boundaries (Leung 1987, Didelon *et al.* 2011), but fuzzy logic is just one possible formalization to deal with boundary vagueness.

Taking traditional taxonomies from geography as a starting point, Montello (2003) updated a taxonomy of geographic regions, identifying the four types of *administrative*, *thematic*, *functional*, and *cognitive* regions, the latter being our focus in this article. Cognitive regions (traditionally called 'perceptual') are regions in the mind, reflecting informal ways that people organize places. Cognitive regions can be idiosyncratic to a single person but are often shared among members of cultural groups. Such informal cognitive region concepts held by groups of lay people are known in geography as *vernacular regions*. As used informally, concepts like 'Midwest,' 'the Bible Belt,' and 'downtown' are vernacular cognitive regions.[2] Cognitive regions reflect the type of categorical thinking (spatially categorical in this case) that so highly characterizes human thought and communication.

Like thematic and functional regions, cognitive regions usually have substantially vague boundaries and variable membership functions. The transition from inside to outside the cognitive region is usually a probabilistically graded zone of significant width. As a corollary to boundary vagueness, their membership functions are variable or probabilistically graded so that all places within the region are not equally strong or clear as members or exemplars of the region.

1.1. Research on cognitive regions in psychology, geography, and geographic information science

The implications of regionalization for human cognition have been studied outside of geography, especially in psychology and cognitive science. Much of this work has focused not on cognitive regions per se, but on the cognitive effects of noncognitive regions, especially administrative regions, on spatial judgments such as location, distance, and direction (Stevens and Coupe 1978, Maki 1981, Carbon and Leder 2005). Other psychological work has looked at the implications of cognitive regions as such on the mental organization of spatial and thematic knowledge. Hirtle and Jonides (1985) demonstrated that semantic associations between landmark buildings on a college campus

created spatial 'clusters' (i.e., regions) that had implications for the organization and recall of locational knowledge on the campus (see also Hund and Plumert 2003, Wiener *et al.* 2004). McNamara and his colleagues (McNamara 1986, McNamara *et al.* 1989) further provided evidence that people store locational knowledge in a partially hierarchical organization, whether in an artificial layout in a large laboratory room or a spatial map-like graphic. Their work also emphasized that cognitive regional organization occurs in the absence of physical or perceptual boundaries (but see Simmering and Spencer 2007).

Friedman and her colleagues began an extensive program of research over a decade ago that has examined the effect of cognitively regionalizing spatial information at scales ranging from portions of countries to continents (Friedman and Brown 2000, Friedman *et al.* 2005, Friedman and Montello 2006, Friedman 2009). This research has demonstrated several important aspects of regionally organizing spatial knowledge at these relatively large scales (i.e., small cartographic scales). Regional organization influences systematic errors people make in judging the latitudes of cities as well as distances between them (as was reported at smaller scales in some of the work reported above). These errors are not just a perceptual effect of looking at shapes on world maps but are conceptual, arising from both spatial and thematic beliefs people have about areas around the world, garnered from a variety of direct and indirect sources. People understand a large portion of the earth's surface almost entirely in terms of coarse regions, although when people do have more knowledge about an area (perhaps the region in which they live), they can often demonstrate spatial knowledge more precisely and metrically, albeit still approximately. Spatial cues, such as the latitudes of anchor cities, can improve the accuracy of people's spatial judgments. And while the general influence of regional organization can be found across cultural groups (e.g., Xiao and Liu 2007), specific patterns of regional beliefs do vary somewhat across cultural groups (Friedman *et al.* 2006). Friedman and her colleagues have shown that most of these conclusions are relatively robust, revealing themselves in the face of different methodologies for collecting estimates of spatial properties.

Geographers and geographic information scientists have also investigated the nature of cognitive regions. Like their work on regions in general (e.g., Jones 1959), much geographic work on cognitive regions has focused more on boundaries than has the work of psychologists. A good example comes from Aitken and Prosser (1990), who were interested in residents' ideas about their neighborhoods. Although neighborhoods are often officially 'defined' by city planning offices and the like, these are often not residents' conceptions of the neighborhoods. That is, neighborhoods are typically good examples of vernacular cognitive regions, even if administrative versions of these regions also exist, created by fiat. Aitken and Prosser had participants draw boundary lines on aerial photographs to indicate areas they conceived of as neighborhoods in San Diego. As the authors noted, such a direct boundary-drawing task has frequently been used by geographers to study people's informal regions. To collect data on familiarity, they superimposed a regular grid of square cells over the photos, at a resolution of about one cell per two city blocks. Participants rated their level of familiarity with the city area in each cell, on an 11-point scale, and recorded how many times they had passed through that area during the previous week. These cell data were used to create visual representations of familiarity and interaction surfaces to compare to neighborhood boundaries.

Montello *et al.* (2003) examined the assessment of the cognitive neighborhood known as 'downtown.' They stopped city residents on the streets and asked them to draw boundaries around the area they considered to be downtown. These were cartographically averaged to create probabilistic surfaces of the nonuniform downtown region with vague

boundaries. The researchers then had residents indicate the boundary again, but this time with varying degrees of confidence, a technique intended to assess the vagueness of cognitive regions directly at the level of an individual person's conceptions.

2. Research goals

In this article, we compare a traditional means of obtaining regional information (drawing boundaries) that only allows vagueness to be measured as averages across research participants with a novel task in which boundary vagueness can be observed directly within participants. We report two studies of pre-existing cognitive regions and their boundaries at the scale of US states/Canadian provinces (i.e., the largest administrative units below the country level). We examine the degree of homogeneity or heterogeneity of these regions across individual participants. Our interest is in the nature of the boundaries participants use to mentally organize cognitive regions, including their location, shape, and degree of vagueness. We contrast the role of spatial factors in determining cognitive regions, such as evenly dividing administrative units into northern and southern regions, with the role of relevant thematic factors, such as unevenly dividing administrative units according to culturally shared ideas about 'northern' and 'southern.'

To explore vagueness in cognitive boundaries, we assess their nature empirically using a conventional boundary-drawing task in Study 1 and a novel high-resolution rating task in Study 2. For the latter, participants directly indicate a variable degree of region membership independently for individual cells in a raster we overlaid on the state/province. We replicate all data collection and analysis for each study both in California, United States, and Alberta, Canada, with samples of student participants from the universities in Santa Barbara, California, and Edmonton, Alberta. We chose these two administrative regions for convenient access to participants, but they do increase the generalizability of our results and potentially allow us to contrast the nature of cognitive regions in places with and without widely held vernacular regionalizations of the administrative units.

3. Study 1: measuring cognitive regions with an unstructured boundary-drawing task

In our first study, we explored participants' cognitive regions within California and Alberta using a direct and relatively unstructured drawing task. Like Aitken and Prosser (1990), we simply asked participants to draw boundary lines around cognitive regions on outline maps. This straightforward task was easy to understand and placed few a priori constraints on participants' expressions of their regional conceptions. It provided evidence of how readily people perform informal regionalization not obviously based on either administratively or physically defined existing boundaries, and it provided evidence of the degree of agreement about these boundaries among individual participants.

Because we wanted to explore the widely used vernacular regionalization of Northern and Southern California sometimes colloquially referred to as 'NorCal' and 'SoCal,' we asked our California participants to draw a boundary separating the state into exactly two regions, Northern and Southern California. For comparison, we similarly asked our Alberta participants to draw a boundary separating the province into two regions, Northern and Southern Alberta. This comparison provides a valuable contrast to California, insofar as there is little or no vernacular distinction made between 'NorAlb' and 'SoAlb.'

3.1. Methods

3.1.1. Participants

Two samples of participants, one at the University of California, Santa Barbara (UCSB) and one at the University of Alberta (UA), took part in Study 1. A total of 93 students in a geography class participated from UCSB, as part of a classroom exercise on regions in the United States (but not specifically about California regions). The class was made up of undergraduate students from all levels and most majors on campus, with only about 11% majoring in geography. We examined responses only from participants who had spent most of their lives growing up in California (assessed on the response sheet). Our final sample of 80 UCSB participants consisted of 35 females and 45 males. The UA sample consisted of 24 students (13 females and 11 males) from the introductory psychology research pool, who participated in partial fulfillment of a class requirement. Like the UCSB sample, this class was made up of undergraduate students from all levels and most majors on campus, with very few majoring in earth sciences, which includes geography at that school. All members of this sample had spent most of their lives growing up in Alberta; we collected responses only from such participants. In fact, almost all UA participants grew up in or near the city of Edmonton, where the research took place.

3.1.2. Materials

Each sample of participants drew region boundaries on an outline 'map' (an outline polygon) of either California or Alberta printed on a sheet of paper (the collected sketches in Figures 1 and 3 include the outline polygons). In addition, the top of the response sheet asked participants to report their gender, birth date, and where they spent the most time growing up.

3.1.3. Design and procedure

Responses from both samples were collected in single groups, one at UCSB and one at UA. After filling in information about themselves on the response sheet, participants were instructed to draw boundaries between regions in California (or Alberta). All participants read that 'Regions are pieces of the earth's surface that enclose fairly similar or homogeneous areas. The similarity may be based on natural or cultural variables, or any combination.' They then read instructions asking them to draw a single boundary line separating the polygon into exactly two regions, Southern California (Alberta) and Northern California (Alberta). 'On the outline map of California (Alberta) below, draw a line separating "Southern California (Alberta)" from "Northern California (Alberta)." This line does not need to be flat or even straight, if you don't think it should be.'

3.2. Results and discussion

We consider the California regions first, the complete set of sketches of which are shown in Figure 1. We coded the sketched regions with respect to the number and nature of their boundary lines. Of the 80 UCSB participants, six actually drew two California region boundaries, thus three regions, in all cases indicating a band or transition zone between Southern and Northern California instead of a line (and in four cases, labeled as 'Central California'). Of the 74 participants who drew a single boundary line, 54% drew a straight line. Boundary lines classified as straight were coded for general slant; they were

Figure 1. All California sketch maps from Study 1.

classified as flat if they followed a latitude parallel[3], orthogonal if they were at a right angle to the state's east–west midline, or neither. In California, orthogonal boundary lines south of the state's 'elbow' at Lake Tahoe on the eastern border do not follow latitude lines (i.e., are not flat). In fact, all straight boundary lines in this condition were drawn at or south of the state's 'elbow' at Lake Tahoe on the eastern border, and thus no orthogonal lines could be flat. Revealingly, three participants drew boundaries that did not transect the entire state east–west – they were drawn so as to depict Southern California as just the southwestern portion of the state (around the Los Angeles/San Diego area).

We graphically averaged the Northern and Southern California regions for the 80 maps. To average, we digitized each boundary line at several points along its extent and averaged the latitudes of these points according to their location with respect to the east–west midline running from the northern to the southern edge of the state. For the six maps depicting a third region, we averaged the line drawn through the middle of the central region. The resulting *empirical mean boundary line* is shown in Figure 2, along with 95% confidence interval lines that run roughly parallel on either side. For comparison, we include two dashed lines showing versions of the *median latitude boundary*, the straight line that divides the state's latitude extent exactly in half. The orthogonal version transects the state at a right angle to the east–west midline; the flat version transects the state parallel to a latitude line. Figure 2 shows that the mean line drawn by participants is neither straight nor flat and is considerably south of both versions of the median latitude

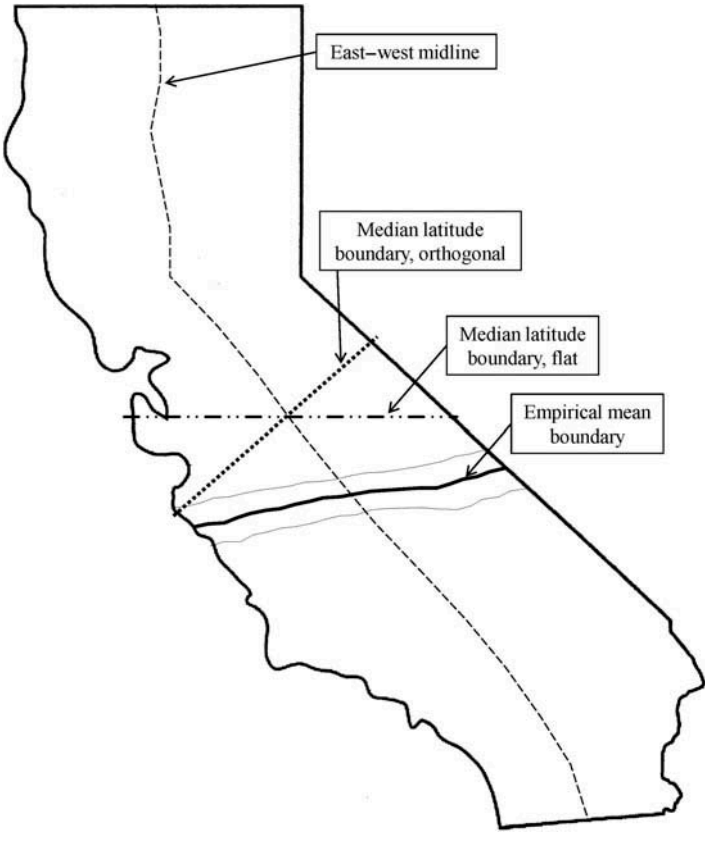

Figure 2. Northern–Southern California boundary lines from Study 1. The median latitude boundary is the straight line that divides the state's latitude extent exactly in half; the orthogonal version transects the state at a right angle to the east–west midline, and the flat version transects the state parallel to a latitude line. The empirical mean boundary is the arithmetic mean boundary line averaged over the latitudes of all boundary lines drawn by participants (thin lines on either side indicate 95% confidence interval).

boundary. In fact, one-sample *t*-tests of each digitized point along the mean line indicate that all points significantly differ from the corresponding flat median latitudes (all $t[79]$'s > 5.00, p's < 0.0001), as suggested by the fact that the mean line is entirely outside the 95% confidence interval. Similarly, all points significantly differ from the corresponding orthogonal median latitudes, except for the westernmost point, where the orthogonal line falls furthest south ($t[79] > 1.56$, $p < 0.15$; all other $t[79]$'s > 2.00, p's < 0.05).

We next consider the Alberta regions, the complete set of sketches of which is shown in Figure 3. All 24 participants actually drew one Alberta region boundary, thus two regions. Unlike California, no one drew three regions. Of these 24 participants, 67% drew a straight line. Again, straight boundary lines were coded for general slant, classified as flat if they followed a latitude parallel and as diagonal otherwise. Unlike California, Alberta has no general bend in its shape; all flat boundaries are thus equivalent to orthogonal boundaries, at least north of the irregular southwestern provincial border along the Rockies. In fact, only six participants drew boundaries at the intersection of

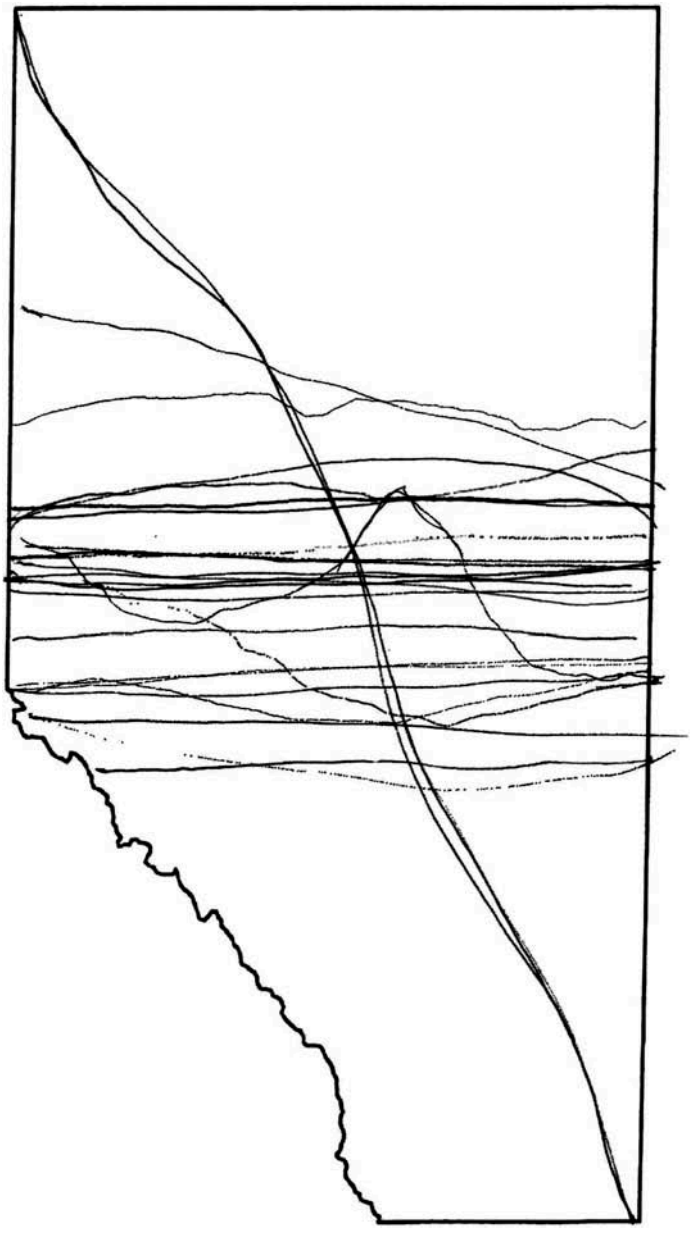

Figure 3. All Alberta sketch maps from Study 1.

this southwestern border with the straight northwestern border; all but two of the rest were clearly north of this point. Of the straight boundaries, all but one were coded as flat. Also unlike California, all boundaries drawn transected the entire province east–west. (This may relate to the fact that the eastern and western borders of the province, in the northern portion, follow the 110° and 120° meridian lines.)

As with the California data, we graphically averaged the regions for the 24 maps. The empirical mean boundary line is shown in Figure 4, again along with 95% confidence

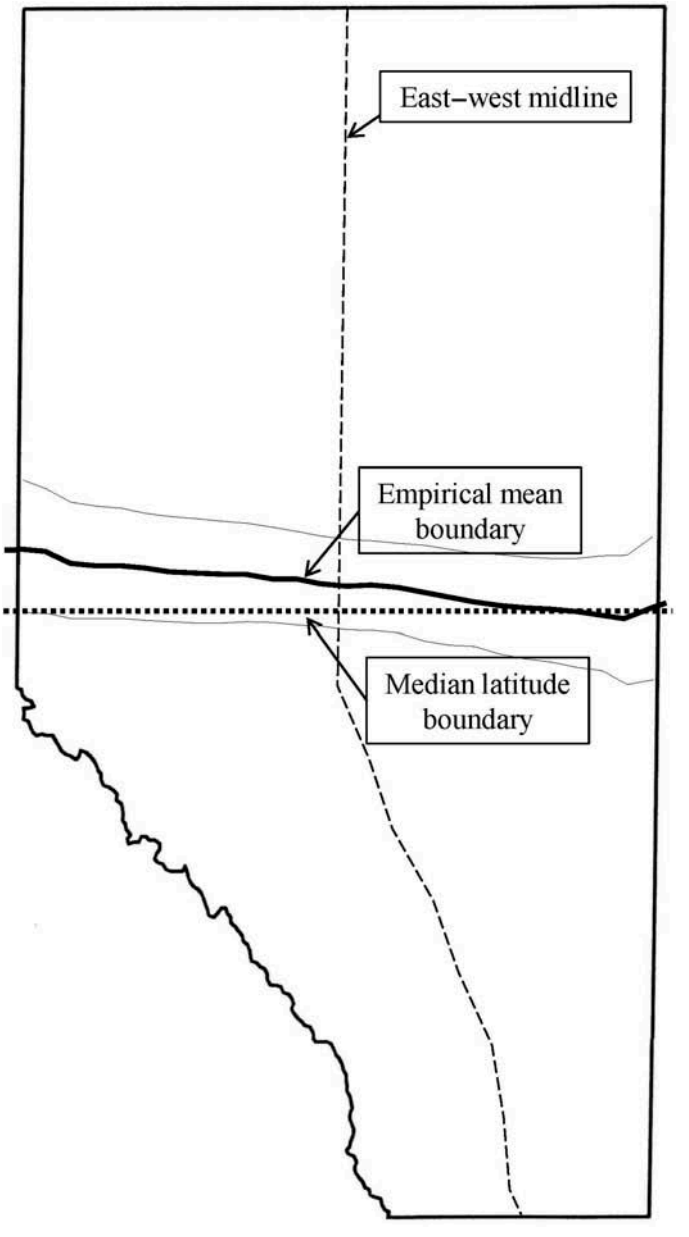

Figure 4. Northern–Southern Alberta boundary lines from Study 1. The median latitude boundary is the straight line that divides the state's latitude extent exactly in half and transects the state at a right angle to the east–west midline. The empirical mean boundary is the arithmetic mean boundary line, averaged over all boundary lines drawn by participants (thin lines on either side indicate 95% confidence interval).

interval lines. For comparison, we include a dashed median latitude boundary. For Alberta, the mean line drawn by participants is fairly straight and flat and is close to the latitude of the median latitude boundary at all digitized points. None of the points along the mean line significantly differed from the corresponding median latitudes by one-

sample *t*-tests, even the westernmost point that deviates furthest from the median line (all t[22]'s <2.04, p's > 0.05), as suggested by the fact that the mean line is entirely inside the 95% confidence interval.

Taking the California and Alberta maps together, we see that both groups of participants found the task of drawing Northern and Southern regions to be straightforward, consistent with the idea that lay people commonly do regionalize the earth's surface. We believe these patterns of conceptual regionalization in the two administrative units are consistent with an important difference between these administrative units and samples of participants. There is a common vernacular regionalization of California into Northern and Southern California. Sometimes this classification includes a third 'Central' region, which may in fact be the vernacular reference for the extended vague boundary between Northern and Southern California. There is no such common vernacular regionalization of Alberta, although there is a common distinction between the mountains of southwest Alberta and the non-mountainous remainder of the province; also, 'Northern Alberta' is sometimes identified in association with the vast oil sands there.

Further support for this conclusion is provided by the nature of the regions participants provided. Albertans followed this instruction almost 'geometrically,' all drawing exactly one boundary (two regions), almost 70% of these boundaries being straight and over 90% of these straight boundaries being flat. In contrast, Californians showed much more variety in their approach to drawing a boundary line. Not only did several draw two boundary lines (three regions) when asked to draw one, but only 54% of the single lines were straight and only 55% of the straight lines were flat, with an additional 22% of the straight boundary lines following an orthogonal pattern. Over 20% of the straight lines in California were neither flat nor orthogonal. And a few Californians did not even transect the entire state with their single boundary line; no Albertans did that. When we averaged the boundary lines, we found further support for differences in the region concepts of Albertans and Californians. The empirical mean boundary line in Alberta is nearly straight and flat and does not significantly differ in location from the median latitude boundary, consistent with the notion that Albertans simply bisected the outline of the province and cut it into two equal pieces, north and south. (The distinction between mountainous and non-mountainous Alberta could explain the small northern tilt of the empirical mean boundary line to the west, assuming additional data would show it to be a statistically significant tilt). Quite differently, the empirical mean boundary line in California is clearly not straight or flat and is located significantly south of the median latitude boundary, whether that is defined as flat or orthogonal to the state's midline. Perhaps the fact that Edmonton is close to the median latitude of Alberta, while Santa Barbara is fairly far south of the median latitude in California, could produce a 'southern anchor' for the UCSB participants. However, while UA participants are all from Edmonton and the surrounding area, the UCSB students we included in our data set are from around the state of California, including many from the San Francisco Bay area. Similarly, the difference between the California and Alberta mean lines might reflect, in part, the greater complexity of California's shape. This might play a role, but we see it as more relevant to the idea that the mean line would not be flat than that it would not be monotonic. The latter is more consistent with the idea that the cognitive regions of Northern and Southern California are not mere subdivisions of the state according to latitude, but reflect informal regional organization based on believed thematic properties, whether cultural, environmental, or both. Our next study addresses this more directly.

4. Study 2: measuring cognitive regions with a structured rating task

In Study 1, we solicited participants' ideas about cognitive regions in California and Alberta with a direct and relatively unstructured drawing task. This was simple and straightforward and placed few a priori constraints on participants' expressions of the locations of regions. But it methodologically forced participants to imply that they think of the cognitive regions as being internally uniform with boundaries made up of sharp lines that are one-dimensional. Nonuniform region memberships and vague boundaries can only be inferred with the direct drawing task by aggregating over participants and taking discordance among them as reflecting vagueness (such as the confidence interval around the mean line). Although disagreement among people is one legitimate expression of region vagueness, the concept of cognitive regions suggests they are probabilistically graded within individual people as well as within cultural and subcultural groups. That is, vagueness should also be reflected in uncertainty or imprecision within an individual person's conceptualization of a region.

In Study 2, therefore, we directly assessed potential nonuniformity and boundary vagueness within individual participants. To do this, we overlaid a relatively high-resolution regular grid of cells on top of the state/province outline polygons used in Study 1, reminiscent of Aitken and Prosser's (1990) use of regular grid cells to collect measures of familiarity and interaction with areas of urban neighborhoods. We asked participants to rate the landscape within each cell (on a scale from 1 to 7) in terms of how much that cell is 'Northern California' ('Northern Alberta') or 'Southern California' ('Southern Alberta'). In this way, we allowed participants to directly express any beliefs they might have about nonuniform region membership and vague boundaries between neighboring cognitive regions, presumably at a resolution sufficiently fine enough to capture nearly all genuine probabilistic variation across space.

4.1. Methods

4.1.1. Participants

We again tested two samples of participants, one at UCSB and one at UA. A total of 47 students participated from UCSB, earning a small amount of extra credit in their geography class for participating. The classes were made up of undergraduate students from all levels and most majors on campus, with only about 4% majoring in geography. We again examined responses only from participants who had spent most of their lives growing up in California. Our final sample consisted of 44 participants (30 females and 14 males). The UA sample consisted of 48 students (33 females and 15 males) from the introductory psychology research pool, who participated in partial fulfillment of a class requirement. Like the UCSB sample, this class was made up of undergraduate students from all levels and most majors on campus, again with very few majoring in earth sciences. All members of this sample had spent most of their lives growing up in Alberta (almost all near the city of Edmonton); again, we collected responses at UA only from such participants.

4.1.2. Materials

Data were collected on the same outline 'maps' of either California or Alberta used in Study 1 but overlaid with a grid of regular hexagonal cells (Figure 5). Participants at UCSB provided ratings on sheets of paper; those at UA responded on a computer screen.

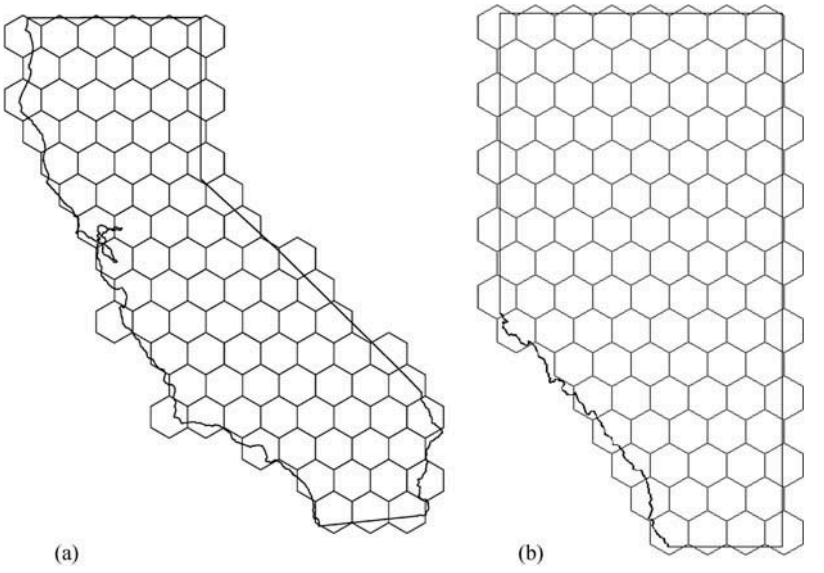

Figure 5. Outline 'maps' (polygons) with hexagonal grids on which participants rated each cell in Study 2: (a) California, (b) Alberta (California and Alberta not to mutual scale).

Unlike Aitken and Prosser (1990), who used square cells, we used hexagonal cells because they are regular shapes that completely tessellate the land surface but with less prominent directional biases or tight corners than squares or equilateral triangles, the other two regular shapes that completely tessellate two-dimensional surfaces. The grid was of fairly high resolution – the California map was overlaid with 90 cells and the Alberta map with 110. Grid cells covered virtually all of the land area of each map. To achieve this coverage, some of the cells went outside the polygon boundary. In terms of actual earth surface, each cell in California covered a bit less land area (1900 mi^2 [4920 km^2]) than did each cell in Alberta (2500 mi^2 [6475 km^2]). As in Study 1, the response sheet also asked participants to report their gender, birth date, and where they grew up.

4.1.3. Design and procedure

Responses from both samples (California and Alberta) were collected individually for Study 2. After reading and signing a consent form, participants were given the response sheet (at UCSB) or sat in front of the computer screen (at UA) to start data collection. After filling in the information about themselves, they read instructions to 'tell us how you think of different parts of California (Alberta) as being part of Northern or Southern California (Alberta).' We explained that Northern and Southern did not just mean 'areas corresponding to the compass directions, but what people informally mean when they think of and talk about Northern California (Northern Alberta) and Southern California (Southern Alberta). That is, not just spatial location, but attitudes, feelings, lifestyles, and so on that people often associate with the places and people of the two regions. Thus, just because an area is in the southern part of the state (province), it may not necessarily express Southern California (Southern Alberta) very much, if at all, and just because an area is located in the northern part of the state, it may not necessarily express Northern California (Northern Alberta) very much, if at all. In the same way, an area to the north

could very much express Southern California (Southern Alberta) and an area to the south could very much express Northern California (Northern Alberta).'

Participants then read they would be shown an outline map of the state of California (the province of Alberta) with a grid of hexagonal cells laid over it. They were instructed to write (California) or type (Alberta) a number from 1 to 7 in each cell to indicate whether the area in that cell was more Northern or Southern. This numerical rating scale was printed to the right of the outline map, with 1 labeled 'very northern California (Alberta),' 2 'moderately northern California (Alberta),' 3 'slightly northern California (Alberta),' 4 'equally northern and southern California (Alberta),' 5 'slightly southern California (Alberta),' 6 'moderately southern California (Alberta),' and 7 'very southern California (Alberta).' Participants were asked to make sure to put a rating in each and every cell. Where a cell at the edge of the map polygon was partially outside the state boundary, participants were told to base their cell rating only on the portion that was inside the state/province. They were asked to take their best guess if they were not sure. After completing their ratings, participants were thanked and marked down for credit.

4.2. *Results and discussion*

Taking California first, we map the mean cell ratings in Figure 6. This map also shows standard deviations across participants by printing the numerals of the mean ratings as

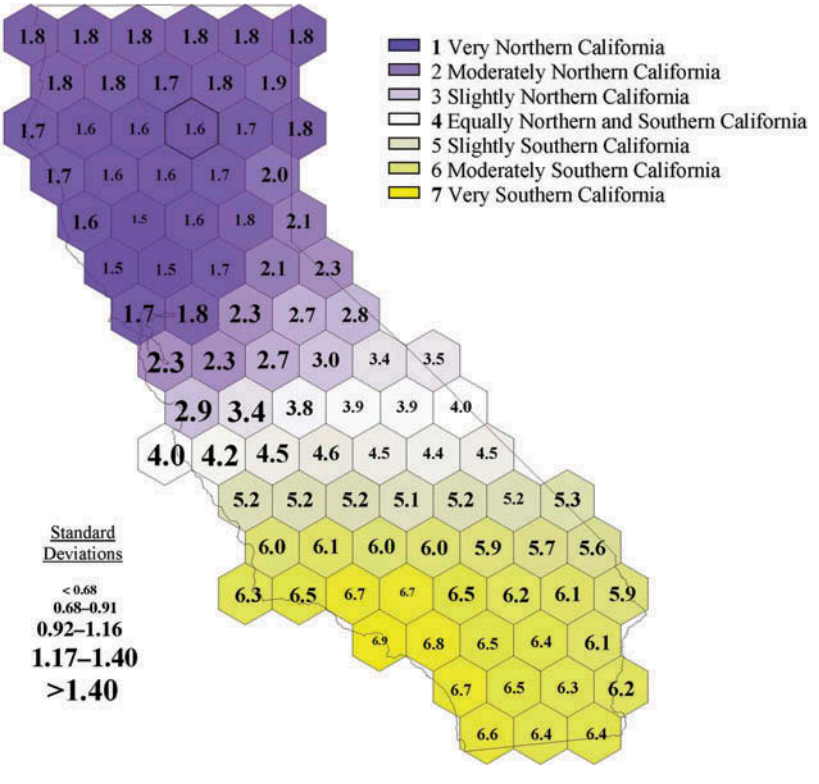

Figure 6. Means and standard deviations of ratings of 'Northern' and 'Southern' for hexagonal cells in California, Study 2. The grouped classes of standard deviations indicated in the legend are based on a single equal-interval division of the range of all values from each cell in both administrative regions, California and Alberta.

smaller (less variability) or larger (more variability). Boundary vagueness and nonuniform membership functions are evident; the boundary can accurately be described as a two-dimensional transition zone rather than a one-dimensional line. These properties are revealed both within participants, as reflected by individual ratings varying across the range from 1 to 7, and across participants, as reflected by rating standard deviations that are nonzero and vary across locations. Moreover, this boundary is not uniformly vague across its extent but is more precise near the western border (the Pacific coast) of California than the eastern border. Likewise, the regions display nonuniform membership functions and not just near their boundaries. The cells rated as most 'Northern' are in cells near the Pacific coast that are five or six rows south of the northern border. That is, the area of California rated as most Northern is more than 200 miles south of the northern edge of the state. Areas due east (i.e., at the same latitude) of this most Northern-rated area are rated nearly a full scale point less Northern. The core of Northern California, according to these ratings, is within Lake County, a little southeast of the city of Ukiah. Similarly, the cells rated as most 'Southern' are three rows north of the southern border, also toward the Pacific coast. The area of California rated as most Southern is at least 100 miles north of the southern edge of the state. Areas due east of this most Southern-rated area are rated nearly a full scale point less Southern. Thus, the core of Southern California is more or less around the city of Los Angeles and Orange County.

The pattern of cell means provides evidence of boundary vagueness and nonuniform membership functions across the set of participants. This pattern could derive from averaging over patterns of cell ratings that vary across participants but do not show much evidence of gradation in region membership, taken individually. This is not the case, however. Out of 44 participants, 38 used all seven points of the rating scale and no one used fewer than five. What's more, over 45% of participants rated some cells in the northernmost row with values greater than '1' (all of them used values of '1' further south), and over 40% rated some cells in the southernmost row with values less than '7' (all of them used values of '7' further north). This provides strong evidence for vague cognitive regionalization within individuals, not just across them.

Considering Alberta, we show the mean cell ratings in Figure 7. As with California, boundary vagueness is evident, but the vagueness appears to be more or less uniformly wide across its extent. Likewise, the regions display nonuniform membership functions, but unlike California, there is a clear linear transition from the most northern row of cells being rated as most Northern to the southernmost being rated as most Southern. The area of Alberta rated as most Northern is along the northern edge of the province. Similarly, the cells rated as most 'Southern' are two rows of cells along the southern edge of the province, with a slight tendency for cells east of center to be rated as most Southern. Like California participants, a clear majority (39/48) of Albertans used all seven points of the rating scale; only one used as few as three. Again, several participants rated some cells in the northernmost row with values greater than '1' (just under 40%), but a third of these did not use 1's in any other rows. Similarly, just under 40% of participants rated some cells in the southernmost row with values less than '7,' but again, almost a third of did not use 7's in any other rows. So while this again provides strong evidence for cognitive vagueness within individuals, not just across them, it differs from the California data in adhering more closely to a pattern of rating the northernmost cells as most northerly and rating the southernmost cells as most southerly. Furthermore, as compared to California, the pattern of mean cell ratings suggests cognitive regionalization based more on a simple spatial interpretation in terms of latitude and less on thematic properties of the natural or cultural environment. Perhaps

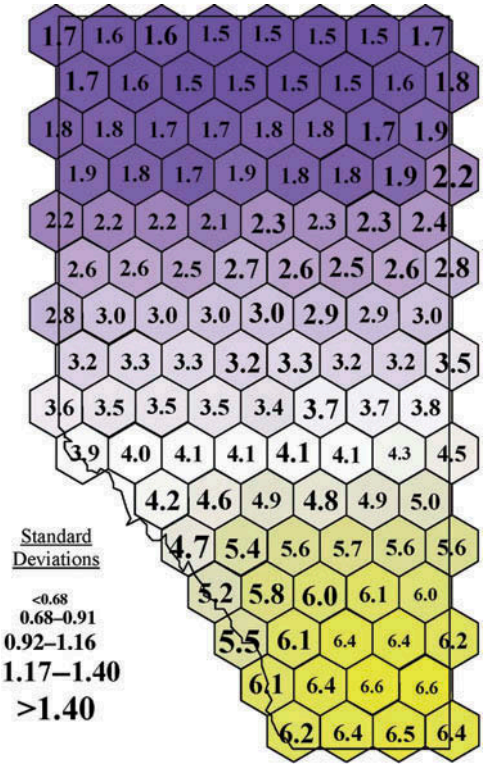

Figure 7. Means and standard deviations of ratings of 'Northern' and 'Southern' for hexagonal cells in Alberta, Study 2. The grouped classes of standard deviations indicated in the legend are based on a single equal-interval division of the range of all values from each cell in both administrative regions, California and Alberta.

the 'bent' shape of California, with a large portion not aligned north and south, might also contribute to this difference with the Alberta ratings.

5. General discussion

The two studies we report in this article empirically explore cognitive regionalization at the scale of a large US state and a large Canadian province. We collected data on cognitive regions with both a traditional boundary-drawing task (Study 1) and a novel task involving interval-level ratings on a 7-point scale within hexagonal cells of a high-resolution grid (Study 2). Participants from both samples found it quite straightforward to draw boundary lines, as has been demonstrated several times in published literature and probably countless times in informal classroom exercises and demonstrations over the years. However, although the rating task used in Study 2 was undoubtedly novel to participants, they apparently did not find it confusing or difficult to understand. The concept of spatial vagueness is something that lay people appear to grasp readily (when phrased in nonspecialist terms) and are prepared to express when given a task that allows it to be expressed. Montello *et al.* (2003) reported that people who were stopped along the streets of a city readily comprehended the idea of a downtown region varying in size (and somewhat in location) about which they were either 100% or only 50% confident was

actually downtown, and they could readily depict these different regions with a boundary-drawing task. Our cell rating task allowed participants to express variable degrees of membership throughout the Northern or Southern regions of a larger administrative region (California or Alberta). Given the opportunity, individual respondents readily expressed boundary vagueness and nonuniform region membership functions via their cell ratings.

Of course, there is no limit to the number, location, and shape of potential regionalizations of the earth surface and that certainly includes cognitive (informal) regionalizations. Many participants drew boundaries in California that were not straight, or if they were, they were not 'flat', that is, did not follow a latitude line. In fact, the empirical mean boundary line in California deviated significantly to the south of a potential straight boundary line based on median latitude. The empirical mean boundary line in Alberta, in contrast, was rather flat and did not significantly fall north or south of the median latitude boundary.

In Study 2, both boundary vagueness and nonuniform membership functions were evident. Interestingly, the vague boundary in California was not uniformly thick across its extent, being noticeably vaguer to the east than to the west. The vague boundary in Alberta, in contrast, was uniformly thick across its east–west extent. Even more dramatically, the membership function in California transitioned non-monotonically across latitudes for regions considered as Northern or Southern California. Our UCSB participants rated cells as most Northern that were quite far from the northern border of California, and while cells rated as most Southern were not far from the southern border, they were clearly not at that border. In both cases, cells rated as most extremely Northern or Southern were in the western part of the state, not in the center. Our UA participants, in contrast, showed a distinctly different pattern of nonuniformity. The cells they rated as most Northern were at the northern border of Alberta, those rated as most Southern were at the southern border, and those rated as most extremely northern or southern were mostly equally distributed east–west. That is, the mean membership function in Alberta transitioned continuously and monotonically across latitudes for regions considered as Northern or Southern.

We think the various differences in the Northern and Southern cognitive regions of California and Alberta are interesting and important. Although California may objectively have more natural and cultural geographical diversity than Alberta (although Alberta has its share), we find that people who grew up in California are much more likely to organize the state into two regions that have thematic meaning beyond simply latitude. This undoubtedly reflects the common vernacular regionalization of California into 'NorCal' and 'SoCal,' with a Central region occasionally included. These are vernacular cognitive regions, in the standard sense, that exist as a common way many people in California organize their understanding of both spatial and thematic properties of the state and its inhabitants. As such, when identifying Northern and Southern California, respondents feel no compunction about dividing the state unequally with respect to latitudinal extent or identifying the boundary zone between the two in a way that does not follow a line of latitude across the entire state. Because such a vernacular regionalization of Alberta is much less common, we speculate that UA participants interpreted our request to indicate Northern and Southern Alberta as concerning vague *spatial relations* rather than vague cognitive regions. As such, UA participants tended to divide Alberta evenly into northern and southern regions, defined latitudinally, more or less equivalently across the east–west extent of the province. This corresponds to the use of 'Northern' and 'Southern' as relational terms. Researchers from GIScience and other disciplines are interested in vague spatial relations, such as near, around, above, and so on, and there is a great deal

of literature on the subject (Fisher and Orf 1991, Mark *et al*. 1995, Bloom *et al*. 1996, Robinson 2000).

Bucholtz *et al.* (2007) report interesting research on cognitive regionalization based explicitly on cultural factors rather than spatial factors. They asked several hundred respondents (mostly California residents) to draw boundaries on an outline map of California, as we did in Study 1, to identify cognitive regions based specifically on linguistic dialects ('Please draw a boundary around each part of California where you believe people speak differently'). They referred to their approach as *perceptual dialectology*. In the many cases where their respondents labeled dialect regions according to location terms, they most often referred to Northern and Southern California (or variants thereof). This demonstrates how commonly people use spatial location ('geographic area' in Bucholtz *et al*.'s terminology) to organize their thematic beliefs about people and their culture. Our own findings echo theirs in that many respondents expressed the belief that inland areas were not really part of the 'Northern/Southern California' distinction.

6. Conclusions

The cell rating task we used in Study 2 measures cognitive regionalization at an unusually high resolution for spatial units as large as California and Alberta. The task allowed participants to directly express spatial beliefs they might have about nonuniform region membership and vague boundaries between neighboring cognitive regions at a resolution sufficiently fine enough to capture nearly all of the probabilistic variation across space. Our method allowed us to demonstrate not only boundary vagueness but vagueness of varying extent along the boundary within individuals, and nonuniform region membership based not only on spatial relations such as north and south but thematically based regionalization (such as those implied by the terms 'NorCal' and 'SoCal'). It is clear from comparing Figures 2 and 4 to Figures 6 and 7 that the cell rating method allows participants to express much more about their individual regional (and sub-regional) beliefs than simply drawing boundaries. And it appears to be a method that is easy to understand and execute. Although direct boundary drawing would still be a preferred method in some problem contexts, such as when the problem scale calls only for low resolution, we believe the cell rating task is useful in a variety of problem contexts beyond the specific issue we explore here of cognitive regionalization of portions of states or provinces.

Besides the basic-science interest we have in characterizing cognitive regions as geographic and psychological phenomena, there are applied reasons to improve our understanding of such informal conceptualizations of space and place. GIScientists have an ongoing interest in understanding lay conceptions in order to design and implement better information systems – to make geospatial technologies more efficient, effective, and equitable (Altman 1994, Wang 1994, Schlieder and Henrich 2011, Schockaert 2011). Our work in this article contributes to the study of geospatial ontologies and attempts to formalize concepts such as place (Jones *et al.* 2001, Agarwal 2005). Indeed, we believe that the concept of place is a subset of the concept of region. Future work should further explore the geometry of cognitive regions in various contexts, allowing not only for vague instead of precise boundaries, but vague boundaries that are not uniformly vague (i.e., that vary in width) along their extent. And although people probably tend to assume planar enforcement in their regionalization, avoiding overlapping regions and non-covered areas, not everyone does in all contexts. Research and system design can be made flexible with respect to enforcing planar enforcement, contiguity, and related properties on the spatial and thematic conceptions of people thinking and communicating about geographic

phenomena. Such conceptions are an inherent and universal part not only of geographic beliefs but of geography itself, and they should be reflected in geographic information.

Acknowledgments

We thank our research participants at the University of California–Santa Barbara, the University of Alberta, Bernd Kohler for programming the experiment, and Tyler Austring for helping collect data. We also acknowledge feedback from the reviewers, which improved the manuscript. This research was partially supported by a Natural Sciences and Engineering Research Council of Canada award to Alinda Friedman.

Notes

1. We henceforth drop the qualifier 'geographic,' although the generic concept of regions includes spatial categories that are not conspicuously geographic, for example, the region between your eyes.
2. Any of these terms may also be used in other ways that give them status as administrative, thematic, or functional regions – a property known as *region polysemy*.
3. We ignore here deviations of latitude lines from spherical straightness and other discordances between a spherical earth and a flat map projection, which in any case are small at this scale.

References

Agarwal, P., 2005. Operationalising 'sense of place' as a cognitive operator for semantics in place-based ontologies. *In*: A.G. Cohn and D.M. Mark, eds. *Spatial information theory*. Berlin: Springer, 96–114.

Aitken, S.C. and Prosser, R., 1990. Residents' spatial knowledge of neighborhood continuity and form. *Geographical Analysis*, 22 (4), 301–325. doi:10.1111/j.1538-4632.1990.tb00213.x

Altman, D., 1994. Fuzzy set theoretic approaches for handling imprecision in spatial analysis. *International Journal of Geographical Information Systems*, 8 (3), 271–289. doi:10.1080/02693799408902000

Bloom, P., *et al.*, eds., 1996. *Language and space*. Cambridge, MA: The MIT Press.

Bucholtz, M., *et al.*, 2007. Hella nor cal or totally so cal? The perceptual dialectology of California. *Journal of English Linguistics*, 35 (4), 325–352. doi:10.1177/0075424207307780

Burrough, P.A. and Frank, A.U., eds., 1996. *Geographic objects with indeterminate boundaries*. London: Taylor & Francis.

Carbon, C.-C. and Leder, H., 2005. The wall inside the brain: overestimation of distances crossing the former iron curtain. *Psychonomic Bulletin & Review*, 12 (4), 746–750. doi:10.3758/BF03196767

Didelon, C., *et al.*, 2011. A world of interstices: a fuzzy logic approach to the analysis of interpretative maps. *The Cartographic Journal*, 48 (2), 100–107. doi:10.1179/1743277411Y.0000000009

Fisher, P.F. and Orf, T.M., 1991. An investigation of the meaning of near and close on a university campus. *Computers, Environment, and Urban Systems*, 15 (1–2), 23–35.

Friedman, A., 2009. The role of categories and spatial cuing in global-scale location estimates. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35 (1), 94–112.

Friedman, A. and Brown, N.R., 2000. Reasoning about geography. *Journal of Experimental Psychology: General*, 129 (2), 193–219. doi:10.1037/0096-3445.129.2.193

Friedman, A. and Montello, D.R., 2006. Global-scale location and distance estimates: common representations and strategies in absolute and relative judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32 (2), 333–346.

Friedman, A., *et al.*, 2005. Cross-cultural similarities and differences in North Americans' geographic location judgments. *Psychonomic Bulletin & Review*, 12 (6), 1054–1060. doi:10.3758/BF03206443

Hirtle, S.C. and Jonides, J., 1985. Evidence of hierarchies in cognitive maps. *Memory & Cognition*, 13 (3), 208–217. doi:10.3758/BF03197683

Hund, A.M. and Plumert, J.M., 2003. Does information about what things are influence children's memory for where things are? *Developmental Psychology*, 39 (6), 939–948. doi:10.1037/0012-1649.39.6.939

Jones, C.B., Alani, H., and Tudhope, D., 2001. Geographical information retrieval with ontologies of place. *In*: D.R. Montello, ed. *Spatial information theory: foundations of geographic information science*. Berlin: Springer, 322–335.

Jones, S.B., 1959. Boundary concepts in the setting of place and time. *Annals of the Association of American Geographers*, 49 (3), 241–255. doi:10.1111/j.1467-8306.1959.tb01611.x

Kronenfeld, B., 2005. Gradation as a communication device in area-class maps. *Cartography and Geographic Information Science*, 32 (4), 231–241. doi:10.1559/152304005775194719

Leung, Y., 1987. On the imprecision of boundaries. *Geographical Analysis*, 19 (2), 125–151. doi:10.1111/j.1538-4632.1987.tb00120.x

Maki, R.H., 1981. Categorization and distance effects with spatial linear orders. *Journal of Experimental Psychology: Human Learning and Memory*, 7 (1), 15–32.

Mark, D.M. and Csillag, F., 1989. The nature of boundaries on 'area-class' maps. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 26 (1), 65–78. doi:10.3138/D235-3262-062X-4472

Mark, D.M., *et al.*, 1995. Evaluating and refining computational models of spatial relations through cross-linguistic human-subjects testing. *In*: A.U. Frank and W. Kuhn, eds. *Spatial information theory: a theoretical basis for GIS*. Berlin: Springer, 553–568.

Martin, G.J., 2005. *All possible worlds: a history of geographical ideas*. 4th ed. New York: Oxford University Press.

McNamara, T.P., 1986. Mental representations of spatial relations. *Cognitive Psychology*, 18 (1), 87–121. doi:10.1016/0010-0285(86)90016-2

McNamara, T.P., Hardy, J.K., and Hirtle, S.C., 1989. Subjective hierarchies in spatial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15 (2), 211–227.

Minshull, R., 1967. *Regional geography: theory and practice*. Chicago, IL: Aldine.

Montello, D.R., 2003. Regions in geography: Process and content. *In*: M. Duckham, M.F. Goodchild, and M.F. Worboys, eds. *Foundations of geographic information science*. London: Taylor & Francis, 173–189.

Montello, D.R., *et al.*, 2003. Where's downtown? Behavioral methods for determining referents of vague spatial queries. *Spatial Cognition and Computation*, 3 (2–3), 185–204.

Prescott, J.R.V., 1965. *The geography of frontiers and boundaries*. Chicago, IL: Aldine.

Robinson, V., 2000. Individual and multipersonal fuzzy spatial relations acquired using human-machine interaction. *Fuzzy Sets and Systems*, 113 (1), 133–145. doi:10.1016/S0165-0114(99)00017-2

Schlieder, C. and Henrich, A., 2011. Spatial grounding with vague place models. *SIGSPATIAL Special*, 3 (2), 20–23. doi:10.1145/2047296.2047301

Schockaert, S., 2011. Vague regions in geographic information retrieval. *SIGSPATIAL Special*, 3 (2), 24–28. doi:10.1145/2047296.2047302

Semple, E., 1907. Geographical boundaries–I. *Bulletin of the American Geographical Society*, 39 (7), 385–397. doi:10.2307/198931

Simmering, V.R. and 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 (4), 871–894.

Stevens, A. and Coupe, P., 1978. Distortions in judged spatial relations. *Cognitive Psychology*, 10 (4), 422–437. doi:10.1016/0010-0285(78)90006-3

Wang, F.J., 1994. Towards a natural language user interface: an approach of fuzzy query. *International Journal of Geographical Information Systems*, 8 (2), 143–162. doi:10.1080/02693799408901991

Wang, F.J. and Hall, G.B., 1996. Fuzzy representation of geographical boundaries in GIS. *International Journal of Geographical Information Systems*, 10 (5), 573–590. doi:10.1080/02693799608902098

Wiener, J.M., Schnee, A., and Mallot, H.A., 2004. Use and interaction of navigation strategies in regionalized environments. *Journal of Environmental Psychology*, 24 (4), 475–493. doi:10.1016/j.jenvp.2004.09.006

Xiao, D. and Liu, Y., 2007. Study of cultural impacts on location judgments in Eastern China. *In*: S. Winter, *et al.*, eds. *Spatial information theory*. Berlin: Springer, 20–31.