

Ontology and Geographic Objects: An Empirical Study of Cognitive Categorization

David M. Mark¹, Barry Smith², and Barbara Tversky³

Abstract: Cognitive categories in the geographic realm appear to manifest certain special features as contrasted with categories for objects at surveyable scales. We have argued that these features reflect specific ontological characteristics of geographic objects. This paper presents hypotheses as to the nature of the features mentioned, reviews previous empirical work on geographic categories, and presents the results of pilot experiments that used English-speaking subjects to test our hypotheses. Our experiments show geographic categories to be similar to their non-geographic counterparts in the ways in which they generate instances of different relative frequencies at different levels. Other tests, however, provide preliminary evidence for the existence of important differences in subjects' categorizations of geographic and non-geographic objects, and suggest further experimental work especially with regard to the role in cognitive categorization of different types of object-boundaries at different scales.

KEYWORDS: Geographic ontologies, geographic categories, prototypes, spatial cognition, mereotopology, human-subjects testing, geographic information systems, GIS.

1. Introduction

1.1 Research on Cognitive Categorization

The world is filled with an uncountable number of different things. Yet even infants identify recurrences, treating not quite identical things as if they were the same. What kinds of things are regarded as recurrences? How are they organized? Research on cognitive categorization and research in ontology have addressed these questions in complementary ways.

One inspiration for the modern work in cognitive categorization came from field work on folk taxonomies of indigenous plants, animals, diseases, and the like, categories central to daily existence (Berlin 1978; Lopez, et al. 1997). For many of these categories, one level of abstraction, termed the "basic" level, seemed more

¹ Department of Geography, NCGIA and Center for Cognitive Science, University at Buffalo. Email: dmark@geog.buffalo.edu

² Department of Philosophy, NCGIA and Center for Cognitive Science, University at Buffalo. Email: phismith@acsu.buffalo.edu

³ Department of Psychology, Stanford University. Email: bt@psych.stanford.edu

useful than others: categories on this level are more frequent in use, morphologically simpler, and earlier to enter a language. For taxonomies of people living close to the earth, this is the level of *oak* and *trout* rather than the level of *bird* or *tree* or of types of oaks or trout. Rosch and her collaborators (see Rosch 1978 for a review) undertook to study the structures of taxonomies of common categories in North-American culture—taxonomies of fruits, tools, plants, clothing, furniture, and the like—concentrating especially on categories of objects (with closed object-boundaries) at surveyable scales. Both vertical and horizontal structures were investigated, the former relating to the level of inclusiveness in a taxonomy, especially the question of a privileged, basic level, the latter to the relations among category members at a particular level of inclusiveness.

To study the structure of natural taxonomies, the Rosch group used a number of tasks: generating instances of categories at several levels, generating attributes of categories at several levels, evaluating the “goodness” or “typicality” of exemplars of a category, using reaction times to identify instances as category members, and more. Two general findings emerged. First, as in folk taxonomies, there is a privileged level of inclusiveness: the basic level. In North America, this is the level of *fish*, *tree*, *table*, *shirt*, and *screwdriver* rather than the more general level of *animal*, *plant*, *furniture*, *clothing*, or *tool* and rather than the more specific level of *trout*, *oak*, *coffee table*, *dress shirt*, or *Phillips head screwdriver*. The defining criterion for the basic level was the informativeness of the category, indexed by attribute listings, relative to the number of category distinctions required by the level of specificity. People list few features in common to all category members at the superordinate level of *tool* or *clothing*; in contrast, they list many features at the level of *hammer* or *sweater*, but few additional ones at the level of *ballpeen hammer* or *v-neck sweater*. In addition, the basic level is favored in a number of cognitive tasks reflecting appearance, function, and naming. The feature that best characterizes the basic level is parts, such as *handles*, *legs*, *tops*, and *arms*. These are simultaneously features of appearance and of function, and serve as a bridge for making inferences from one to the other (Tversky and Hemenway 1984).

1.2 Prototypicality and Category Unification

The second major finding of the Rosch enterprise was that the internal structure of categories seems to be better characterized as a typicality structure rather than in terms of necessary and sufficient conditions. That is, category members seem not to share a set of criterial properties that non-members do not share; rather, there seems to be a group of properties characteristic of a category that typical or good members of the category are more likely to have more of. Thus, chairs, typical of the *furniture* category, have more of the properties listed for most furniture than carpets. Chairs have arms, legs, and a seat, in common with sofas, and partly in common with beds. Carpets have none of these features. The same holds for robins as opposed to penguins (among birds) and for shirts as opposed to ties (among clothing).

The nature of categorical glue remains controversial. In spite of the evidence that people are unable to produce necessary and sufficient features for inclusion in a category, there are those who maintain that, even though categories may not have easily definable necessary and sufficient features, users seem nonetheless to have implicitly acknowledged certain underlying conceptual cores or pertinent theories. For natural kinds, the conceptual core would be based on molecular or genetic structure

(Bloom 1996; Gelman and Wellman 1991; Keil and Batterman 1984); for artifacts, the conceptual core would be based on the intentions of the designer or user (Bloom 1996). Others take issue with the notion of conceptual core, citing evidence against it for such common categories as *water* (Malt 1994, 1995; Malt and Johnson 1992, 1998).

1.3 Beyond Objects

Can the principles established by Rosch be extended beyond the realms of living things and artifacts used to establish them? For the most part, the examples of living thing and artifact categories studied are small enough to be manipulable by humans, and for most of them shape seems to be a highly salient factor for purposes of categorization. This is not the case for some other kinds of categories. In spite of differences, some evidence for a basic level and for typicality has been found for categories of scenes, such as *store* or *beach* (Tversky and Hemenway 1983); for events such as *going to the movies* (e.g., Abbott, et al. 1985; Morris and Murphy 1990; Rifkin 1985), for colors (e.g., Rosch 1975), for emotions (e.g., Ekman 1992; Izard 1992; Johnson-Laird and Oatley 1992; Plutchik 1993), and for social stereotypes (Cantor and Mischel 1979). The extension of Rosch's principles to geographic categories is the focus of the research reported in the present paper.

1.4 Taxonomy and Partonomy

Taxonomic or kind-of hierarchies familiar to us from the Linnean classification of plants and animals is only one way to organize knowledge. Another way to organize knowledge is in terms of a cognitive partonomy (e.g., Miller and Johnson-Laird 1976; Tversky 1990; Tversky and Hemenway 1984). Like a taxonomy, a partonomy is a hierarchy, but based on a part-of relation rather than on a kind-of relation. A foot is part of a leg which is part of a body; a floor is part of a room which is a part of a building. Some of the utility of taxonomies is that they validate inferences of properties and class inclusion. If a robin is a bird, it is also an animal. If an animal breathes and reproduces, then so does a bird; if birds fly and lay eggs, then so do robins. Partonomies do not permit property inferences; there is even some controversy over whether they permit part-of inferences (i.e., whether part-of as organizing principle of cognitive partonomies is a transitive relation). However, as noted earlier, they do seem to promote (though not promise) inferences from appearance to function.

Especially relevant to our present concerns, is that both temporal and spatial concepts associate naturally with partonomies as well as with taxonomies. A year consists of months which consist of days which in turn can be divided into minutes, then seconds. Even more relevant, the earth can be partitioned into land and water; land into continents; continents into countries, and so on. It will be important to bear in mind in what follows the distinction between geographic taxonomies and geographic partonomies.

2. An Ontological Framework

2.1 Elicited Ontologies

The focus of our ontological work has been on what Smith has called *elicited ontologies* (Brogaard et al. 1999), which is to say: theories about given object domains designed to model how these domains are conceptualized by given individuals or cultures (Egenhofer and Mark 1995a, Smith and Mark 1999). This

study of elicited ontologies is distinct from epistemology: it focuses not on knowledge and belief in general but rather on the ontological content of certain domain-specific representations. It draws to some degree on the work of Quine (1953), who showed how we can use logical methods to study the ontological commitments embodied in scientific theories, and extends this idea to the study the folk-ontological theories embodied in beliefs of human subjects (and to the study of the taxonomies embodied in information systems and in associated bodies of data).

2.2 Geographic Objects, Their Parts and Boundaries

The domain of ontology comprehends objects, relations, boundaries, events, processes, qualities, and quantities of all sorts. Our focus here is on geographic objects in the narrow sense, which comprehends regions, parcels of land and water-bodies, topographic features such as bays, promontories, mountains and canyons, hills and valleys, roads, buildings, bridges, as well as the parts and aggregates of all of these. Geographic objects are thus in every case spatial objects on or near the surface of the earth. They are objects of a certain minimal scale, they are typically complex, and they have parts. An adequate ontology of geographic objects must therefore contain a theory of part/whole, or *mereology*. The latter must comprehend within its orbit the parts represented in cognitive paronomies and coded for in natural language; but it must include also other sorts of parts, for example the products of arbitrary delineation within an extended spatial whole, perhaps also negative parts (holes), such as the interior of a canyon.

Geographic objects also have boundaries. They are prototypically connected or contiguous, but they are sometimes scattered or separated. They are sometimes closed (e.g., lakes), and sometimes open (e.g., bays). The concepts of boundary, contiguity and closure are topological notions. Thus an adequate ontology of geographic objects must contain a *qualitative topology*, a theory of boundaries and interiors, of connectedness and separation, that is integrated with a mereological theory of parts and wholes to form a *mereotopology*. (Smith 1996)

An object is 'closed' in the mereotopological sense, if it includes its outer boundary as part; it is 'open' if this outer boundary is included rather in its complement. Ordinary material objects (pets, tools) are closed in this sense. They have bona fide boundaries, which is to say, boundaries that correspond to genuine discontinuities in the world. Regarding geographic objects, however, matters are not so simple. Consider the mouth of a river, or the boundary of hill where it meets a valley. Where we place the boundary is here a matter of fiat, rather than a choice dictated by physical discontinuities. A variety of different types of fiat boundaries are projected into geographic space at locations wholly or partly independent of such discontinuities, some of them crisp, some graded (Burrough and Frank 1996). It follows that geographic objects may often overlap (consider the overlap between hill and valley, or between river and mouth), in a way that is normally excluded for objects at sub-geographic scales such as dogs and apples (though not for *events* such as walking or shopping, and not for *states* such as anger or envy).

As Smith and Varzi have argued (1997), a topological theory with the resources adequate to deal with geographic objects must accordingly be two-sorted, embracing both a theory of bona fide or physical boundaries, on the one hand, and a theory of fiat boundaries based on non-classical topological principles, on the other.

We hypothesize, now, that the opposition between classical and fiat boundaries is cognitively salient in the geographic realm, even for non-experts. A complete ontology adequate to the purposes of modeling non-expert geographic cognition must thus be constructed on the basis of a two-sorted topology of the sort described. In contrast to the ontologies underlying most geographic information systems, which rest on discretized metric world models, such an ontology must have the resources to represent the *qualitative* conceptual categories conveyed by natural language, along the lines set out in Smith 1995, Smith and Varzi (in press), Casati et al. 1998, Casati and Varzi 1994, 1996. Here we concentrate on ontological issues pertaining to geographic *objects*; our work thus parallels the studies of spatial *relations* set out in Mark and Egenhofer 1994, 1994a, 1995; Egenhofer and Mark 1995.

2.3 Categorical vs. Accidental Predications

Because geographic objects are (paradigmatically) immovable, it follows that they are not merely *located* in space, they are tied intrinsically to space in such a way that they inherit from space many of its structural (mereological, topological, geometrical) properties.

To see more precisely what is involved here, let us distinguish *categorical predications*: is a man, is a fish, is a lake, etc., from *accidental predications*: is suntanned, is swimming, is hungry, etc. The former tell us under what category an object falls (what an object is, enduringly). The latter tell us how an object is *per accidens* at a given moment, what state the object is in, what process it is undergoing; thus they pertain to ways in which objects change from occasion to occasion. We hypothesize that it is a distinguishing mark of geographical objects that—because they change and move at best very slowly—size, location, and even position may for them be matters for categorical prediction. Hence good candidate basic-level geographic categories will often form pairs or series, as illustrated by cases such as: *bay – cove*, *mountain – hill*, *pond – lake – sea – ocean*. For manipulable objects of surveyable scale, in contrast, predications of location, size and position are almost always accidental. (Objects at surveyable scales do not change category, for example, when they are upside down). We hypothesize therefore that terms for few basic-level categories of objects at surveyable scales will code for size, location, and position.

We hypothesize further that where, in the world of natural objects at surveyable scales, it is the interiors of a thing which are salient for purposes of categorization, for geographic objects both sides of the boundary are significant. A chipmunk is a chipmunk whether it is in water or on land or abutting another chipmunk. An identical piece of sandy ground is a riverbank or a bluff depending on what it is next to.

2.4 Water in Geographic Space

As an illustration of the types of more specific hypotheses which must be tested via ontology elicitation experiments, consider the category *lake*. Is a lake a three-dimensional body of water in geographic space, or a two-dimensional sheet of water, or is it a depression in the Earth's surface (possibly) filled with water? Dry lakes exist, but are they lakes when they are dry, or merely places where lakes were, and might be again? The definitions of 'lake' and similar terms contained in geographic or cartographic data standards and in dictionaries represent the consensus among experts as to the meanings of such terms. The U.S. Spatial Data Transfer Standard (SDTS) defines lake as "any stand[ing] body of inland water" (see Mark 1993, 1993a, Smith

and Mark 1998, 1999). A lake, then, on this ontology is a body of water of a certain sort.

Consider, however, the arguments of Hayes, in his general treatise on the ontology of liquids (1985a), to the effect that a lake is a contained space defined by geographical constraints (surrounding mountains, the lake bed) that is full of water. The contained space ends at the surface of the water and a lake is full by definition. If it contains only half of its usual volume of water, then its level is low. A reservoir behind a dam, in contrast, can be half full, or empty, just as a cup on the table can be half full. On Hayes' view, a lake is a fixed object in geographical space. On the bodies-of-water ontology, in contrast, it would be constantly changing in virtue of water flow: it would be a phenomenon, not an object. This is a matter of the ontology of lakes.

3. Previous Research on Geographic Categories

Although much scientific effort has been invested in the classification of geographic objects and phenomena, including work on dictionaries of geographic terms (cf. Moore 1978; Mayhew 1997), on cartographic data standards such as the U.S. standard, SDTS (see Fegeas et al. 1992), there have been few empirical studies of geographic categories that have involved testing with human subjects. Four exceptions, reviewed below, are: Battig and Montague's (1968) study of category norms, Tversky and Hemenway's (1983) research on cognition of indoor and outdoor scenes, testing of cartographic feature codes by the National Committee for Digital Cartographic Data Standards (Rugg and Schmidt, 1986), and Lloyd et al.'s (1996) study of basic-level geographic categories.

3.1 Battig and Montague's Research on Category Norms

Battig and Montague (1968) elicited examples of 56 categories from more than 400 students in Maryland and Illinois. Students were given 30 seconds for each category. Of interest is the frequency and order of listing of exemplars. Cross-site correlations were generally high. Of the categories tested by Battig and Montague, one was "a natural earth formation." A total of 34 different "earth formations" were listed by at least 10 of the subjects. Here, the ten most frequently-listed terms, with their frequencies among 442 subjects, are listed in Table 1 (where N is the number of subjects who listed the given feature):

a natural earth formation	<u>N</u>
mountain	401
hill	227
valley	227
river	147

rock	105
lake	98
canyon	81
cliff	77
ocean	77
cave	69

Only one individual was listed (the Grand Canyon), 14 times. The only movable items listed were iceberg and glacier.

3.2 Tversky and Hemenway's Research on Indoor and Outdoor Scenes

Tversky and Hemenway (1983) applied Rosch's research methods to objects of geographic scale, which in their paper are called '(outdoor) environmental scenes'. Their goal was to provide a taxonomy of kinds of environmental scenes and to identify a basic level of scene categorization, the level not only most commonly used, but also 'apparently most useful in other domains of knowledge concerned with environments, for example, architecture and geography'. The first set of experiments established a basic level for scenes. One set of participants generated categories and subcategories for indoor and outdoor scenes. A total of 210 other participants generated attribute, activity, and part norms for categories at different levels of abstraction. For all the norms, the level of beach, mountains, city, and park for outdoor scenes and store, school, restaurant, and home for indoor scenes was especially informative. In a second set of studies, measures of language use in identifying photographs of scenes and in completing sentences such as, "The Kingstons furnished their _____ with furniture they built themselves" converged on the basic level. The environmental categories thus obtained may contain geographical objects such as lakes and rivers, but they do not in and of themselves constitute a taxonomy of geographic objects. Rather, they serve as the *settings* for objects and activities of various kinds.

3.3 National Committee for Digital Cartographic Data Standards

Empirical work on cartographic feature definitions was conducted during development of the standard that eventually became SDTS, and is reported by Rugg and Schmidt (1986). They performed a "consistency test," a "completeness test," and an "ease of use" evaluation. The consistency test circled the names of 51 features on a standard US Geological Survey topographic map. Subjects were given the map, a list of interim feature definitions from the draft standard, and other materials, and were asked to assign a feature code to each feature. A total of 21 subjects from 8 different mapping organizations completed the test, and the mean consistency score across all features and subjects was 85.4 %; that is, the percentage of subjects putting each feature in its respective most frequent category was a little over 85 %. This demonstrates that the subjects, all expert or professional cartographers, could effectively and consistently

use the feature definitions to code the features on the map. However, it did not test whether the subjects agreed with the definitions, or thought they were correct. The completeness test asked a different group of 17 expert subjects to assign every symbol in the legends of a USGS topographic map and a nautical chart to some feature code in the draft standard. About 98 % of the symbols on the topographic map, and 85 % for the nautical chart, were successfully assigned to feature types.

3.4 Lloyd et al.'s Research on Basic-Level Geographic Categories

Lloyd et al. (1996) proposed that the common categories of administrative units in the United States (*country, region, state, city, neighborhood*) are all basic level categories under the superordinate category *place*. Their model however populates the subordinate category layer not with subclasses but with instances that are *particular cases*, such as *the South* or *Georgia* or *Charleston*. Hence their work (in spite of its title) refers primarily not to categories at all, but rather to the cognition of objects.

4. Report of Experimental Results

In what follows we report some of the results of pilot research on the basis of seven experimental protocols which were administered to freshman-level philosophy students in Buffalo in December 1999. Later phases of empirical work will include application of refined experimental question-sets to English-speaking subjects in three other regions and to subjects from countries speaking German, French, Spanish, Hungarian and Finnish.

4.1 Experiment 1: Elicitation of Category Norms for Geographic Kinds

This is a partial replication of Battig and Montague (1968) on 33 students at Buffalo, using 11 of their non-geographic categories and 6 new geographic ones: *a kind of geographical feature*; *a kind of water feature that would be shown on a map*; *a kind of geographic feature made by humans (not 'natural')*; *a kind of human settlement (populated place)*; *a political entity*; and *a kind of geographic object that typically has an indeterminate (fuzzy, graded, or uncertain) boundary*. Non-geographic categories were included to provide a basis for comparison with earlier results and with future cross-linguistic results to allow evaluation of the hypothesis that geographic categories are more culturally variable than living thing and artifact categories.

The current results were similar to previous ones and are displayed in Tables 2-5. As is evident from Table 2, "geographical feature" elicited solely natural and not artificial geographic features. Indeed, when "a kind of geographic feature made by humans" was asked for explicitly (Table 4), there were few exemplars and low consensus, suggesting that this category lacks a clear core or essence. The category "political entity" did not produce geographic objects at all (subjects listed rather: *president, democrat, republican, senator*, etc.). The category "geographic object that typically has an indeterminate boundary" also yielded little consensus. Both of these categories will be re-tested with alternative wordings.

"Mean", in the middle column of the following tables, is the average rank of the feature in the listings by those subjects who listed that feature.

Table 2
<i>a kind of geographical feature</i>

Table 3
<i>a kind of water feature</i>

<u>N</u>	<u>Mean</u>	<u>Feature</u>
31	1.65	Mountain
19	4.79	Lake
17	4.00	Ocean
17	4.82	Plain
16	4.44	River
15	4.00	Hill
12	6.83	Desert
8	3.50	Plateau
7	4.43	Valley
6	5.50	Canyon
6	6.83	Forest
6	4.17	Volcano

<u>N</u>	<u>Mean</u>	<u>Feature</u>
30	2.27	River
30	2.53	Ocean
29	2.31	Lake
11	3.91	Sea
10	3.70	Stream
10	4.70	Bay
9	4.22	Pond
7	6.00	Gulf
6	5.00	Waterfall
3	5.33	Sound
3	6.67	Swamp

Table 4
geographic feature made by humans

<u>N</u>	<u>Mean</u>	<u>Feature</u>
10	1.60	Man-made lake
10	1.40	Dam
7	2.00	Canal
6	2.50	Road
5	2.40	Some ponds
4	1.75	Buildings

Table 5
a kind of human settlement (populated place)

<u>N</u>	<u>Mean</u>	<u>Feature</u>
30	1.83	City
22	2.23	Town
18	3.00	Village
12	4.50	Country
12	3.75	State
6	4.83	county
6	4.00	Suburb

4.2 Experiment 2: Good Examples of Categories

Rosch (1973) found that goodness-of-example ratings correlated with Battig and Montague frequencies. Here, we replicate her procedure, selecting 6 instances varying in frequency for elicitation of goodness-of-example ratings from 46 participants, displayed in Table 6.

Table 6
Goodness-of-example results for “a natural earth formation”

<u>Feature</u>	<u>F, B&M</u>	<u>F, this study</u>	<u>Goodness-of-example</u>
----------------	-------------------	----------------------	----------------------------

Mountain	401	28	1.43
River	147	5	2.11
Lake	98	6	2.39
Gully	45	1	3.07
Crater	16	1	3.19
Iceberg	3	2	2.05

A total of 46 subjects rated the terms, and the results are very similar to those obtained by Rosch for non-geographic categories: goodness-of-example ratings produced the same order as the Battig and Montague frequencies, with the exception of *iceberg*, which was given as an example by only 3 of Battig and Montague's 442 subjects and two of ours, yet was rated the second best example of a *natural earth formation* in the goodness-of-examples ratings. This protocol will be applied to each geographic category included in our Experiment 1, and in each language to be tested. Results will tell us what subjects count as best examples of geographic objects. Common features of these best examples will then be used as baseline data for purposes of ontology elicitation.

4.3 Experiment 3: Numbers of Characteristics, Activities, and Parts

Following Lloyd et al.'s (1996) use of another of Rosch's experimental protocols, we will ask subjects to list as many *characteristics* they associate with given geographic categories as they can in 90 seconds, and ask other subjects to list *parts*, or *activities*, that they associate with a given geographic categories, also in 90 seconds. For comparison with Lloyd et al., we will include *city* in the set of terms tested. The other categories tested will include *lake*, *pond*, *reservoir*, *bay*, *river*, *hill*, and *mountain*. Like Experiment 2, this task is designed primarily for general purposes of ontology elicitation. We hypothesize that natural geographic categories will be thought to have many associated characteristics and activities and few lexically foregrounded parts.

4.4 Experiment 4: Tests of the Nature of Boundaries

We hypothesized that geographic objects have a distinct ontology in part due to characteristic ontological features of their boundaries. We will test whether fiat and bona fide boundaries are commonly considered to be different, and in what ways. We will test whether the distinction between crisp and graded boundaries is cognitively salient. And we will also test the ways in which fiat boundaries at geographic scales may or may not differ from counterpart phenomena found e.g. in our categorization of body parts such as *knee* or *belly* or *chin*.

Sample pre-test questions that address how boundaries are conceptualized include the following: "In what ways are the boundary of a country and the boundary of an apple similar, and in what ways are they different?" and "List some ways in which the edge of a wetland differs from the edge of a park." Respondents focussed on the objects and their attributes rather than the boundaries or edges, suggesting that the questions need to be altered to elicit the desired information. A third question, "Who do you think owns the boundary between two adjacent land parcels in the area of your home town?" yielded more illuminating answers. Almost all thought that the two

parcel owners had an equal kind of ownership of the boundary, a result that is consistent with principles governing fiat boundaries in our ontological theory.

Frequency	Question: Who do you think owns the boundary between two adjacent land parcels in the area of your home town?
18.50	d. the boundary is jointly owned by both parcel owners
14.00	e. the boundary has no owner
10.50	c. the two parcel owners each own their half of the boundary
2.00	a. the owner of the older parcel
0.00	b. the person who has owned one of the parcels the longest

4.5 Experiment 5: Tests Related to Definitions

In this experiment, subjects are asked to rate definitions of geographic feature terms. In a pre-test, we asked subjects to “indicate which of the following definitions of “a lake” best corresponds to what you think a lake is.” Again, several subjects marked more than one response, and in those cases, if N answers were marked, each received 1/N credit in the tally.

16.97	A large inland body of water
7.19	Water contained within a predominantly natural shoreline that exhibits no appreciable current
6.47	An extent of water larger and deeper than a pond
5.02	A natural depression in the Earth’s surface that can be more or less full of water
3.44	A part of the earth’s surface, other than the ocean, that is covered by still water
3.19	A closed loop formed by a shoreline, with a water surface inside it
1.95	A large inland natural sheet of water
1.70	Any standing body of inland water

4.6 Experiment 6: Attributes, Differences, and Use

Several questions explore further ways in which geographic categories are defined and distinguished by non-expert subjects. We presented the instruction “List 5 things that are generally true of Xs,” where “X” is a place-holder for geographic terms such as *lake, pond, river, mountain, hill, etc.*, to 46 subjects, each of whom was asked to list truths about one of {lake, pond, river, creek, lagoon}. 13 subjects were asked about lakes, and 7-9 about each of the other categories. 41 of 46 subjects responded to the question with reasonable answers, listing an average of 3.86 (for pond) to 4.57 (for river) truths per subject. However, there was little consensus.

Questions of the type: “What are some of the main differences between an X and a Y?” were given to 46 subjects, who were asked to give differences between randomly selected pairs from the group {lake, pond, river, creek, lagoon}. Subjects listed an average of only 2.3 differences, despite the fact that there were places to list five differences on the questionnaire. Twelve subjects did not respond at all to this question. Although the sample size was small, it is interesting that water flow was the most frequently cited difference for 3 of the 4 pairs that included a flowing (river, creek) and non-flowing (lake, pond, lagoon) feature, and size was the most frequent river-creek and lake-pond difference.

5. Conclusions.

These results are preliminary, but they suggest a number of ways in which ontological differences between geographic objects and objects at surveyable scales may induce corresponding differences in the cognitive categorizations of non-expert subjects. For geographic objects, boundaries give rise to important and complex issues—for example pertaining to the oppositions: fiat vs. bona fide, crisp vs. graded—which do not arise, or have not been studied, in relation to the artifacts and living things on which most work on object-categorization has been focused hitherto. Functions and conceptual cores, on the other hand, are important in conceptions of artifacts and living things, but they seem less important in relation to objects at geographic scales. In artifacts, function is related to designed use; in living things, sometimes to intended use (e.g., food, work-animals) and sometimes to the behavior of organisms. Our future research will be designed to provide more precise formulations of such differences, to subject the resultant hypotheses to empirical testing, and to use the results of such testing to generate a stable framework for purposes of ontology elicitation in the geographic realm.

6. Acknowledgments

This paper began as a project of Research Initiative 21, “Formal Models of Common-Sense Geographic Worlds,” of the U.S. National Center for Geographic Information and Analysis (NCGIA), supported by the National Science Foundation (SBE-8810917). It is also an activity of the “Cognitive Models of Geographic Space” committee of NCGIA’s Project Varenus (SBE-9600465). Support by NSF is gratefully acknowledged.

7. References

- Abbott, V. Black, J. H. and Smith, E. E. 1985 “The Representation of Scripts in Memory”, *Journal of Memory and Language*, 24, 179-199.
- Battig, W. F., and Montague, W. E., 1968 “Category Norms for Verbal Items in 56 Categories: A Replication and Extension of the Connecticut Norms”, *Journal of Experimental Psychology Monograph*, 80, No. 3, Part 2, pp. 1-46.
- Berlin, B. 1978 “Ethnobiological Classification”, in: E. Rosch and B. B. Lloyd (eds.), *Cognition and Categorization*, Hillsdale, NJ: Erlbaum, 9-26.
- Bloom, P. 1998 “Theories of artifact categories”, *Cognition*, 66, 87-93.

- Brogaard, B., Peuquet, D. and Smith, B. 1999 *Objects and Fields. Report on the Specialist Meeting of Varenus Research Initiative*, Buffalo/Santa Barbara/Maine: National Center for Geographic Information and Analysis.
- Burrough, Peter A. and Frank, Andrew U. (eds.) 1996 *Geographic Objects with Indeterminate Boundaries*, London and Bristol, PA: Taylor and Francis.
- Cantor, N. and Mischel, W. 1979 "Prototypes in Person Perception", in: L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, 12, New York: Academic Press.
- Casati, R., and Varzi, A. C. 1994 *Holes and Other Superficialities*, Cambridge, MA, and London: MIT Press (Bradford Books).
- Casati, R., and Varzi, A. C. 1996 "The Structure of Spatial Location", *Philosophical Studies* 82, 205-239.
- Casati, R., Smith, B., and Varzi, A. 1998 "Ontological Tools for Geographic Representation", in: Guarino (ed.), 77-85.
- Egenhofer, M. J., and Mark, D. M. 1995 "Modeling Conceptual Neighborhoods of Topological Relations", *International Journal of GIS*, 9, No. 5, pp. 555-565.
- Egenhofer, M. J., and Mark, D. M., 1995a "Naive Geography," in Frank and Kuhn (eds.), 1-15.
- Ekman, P. 1992 "Facial Expression and Emotion", *American Psychologist*, 48.
- Fegeas, R. G., Cascio, J. L., and Lazar, R. A. 1992 "An Overview of FIPS 173, the Spatial Data Transfer Standard", *Cartography and Geographic Information Systems* 19 (5), 278-293.
- Frank, A. U. and Kuhn, W. (eds.) 1995 *Spatial Information Theory: A Theoretical Basis for GIS*, Berlin: Springer-Verlag, Lecture Notes in Computer Sciences No. 988.
- Gelman, S. A. and Wellman, H. M 1991 "Insides and Essences: Early Understandings of the Non-Obvious", *Cognition*, 38, 213-244.
- Gunalik, I. B. and Friend, J. H. (eds.) 1966 *Webster's New World Dictionary of the American Language*, College Edition, Cleveland and New York: The World Publishing Company.
- Hayes, P. 1985 "The Second Naive Physics Manifest," in: Hobbs and Moore (eds.), 1-36.
- Hayes, P., 1985 Naive Physics I: Ontology of Liquids. in: Hobbs and Moore (eds.), 71-108.
- Hobbs, J. and Moore, R. (eds.), *Formal Theories of the Commonsense World*, Norwood, NJ: Ablex.
- Izard, C. E. 1992 "Basic Emotions, Relations among Emotions, and Emotion-Cognition Relations", *Psychological Review*, 99, 561-565.

- Johnson-Laird, P. N. and Oatley, K. 1992 "Basic Emotions, Rationality, and Folk Theory", *Cognition and Emotion*, 6, 201-223.
- Keil, F. C. and Batterman, N. 1984 "A Characteristic-to-Defining Shift in the Development of Word Meaning", *Journal of Verbal Learning and Verbal Behavior*, 23, 221-236.
- Lloyd, R., Patton, D., and Cammack, R. 1996. Basic-level geographic categories. *Professional Geographer*. 48: 181- 194.
- Lopez, A., Atran, S., Coley, J. D., Medin, D. L, and Smith, E. E. 1997. "The tree of life: Universal and cultural features of folkbiological taxonomies and inductions," *Cognitive Psychology* 32, 251-295.
- Malt, B. C. 1994 "Water is not H₂O." *Cognitive Psychology*, 27, 41-70.
- Malt, B. C. 1995 "Category Coherence in Cross-Cultural Perspective." *Cognitive Psychology*, 29, 85-148.
- Malt, B. C. and Johnson, E. C. 1992 "Do Artifact Concepts have Cores?" *Journal of Memory and Language*, 31, 195-217.
- Malt, B. C. and Johnson, E. C. 1998 "Artifact Category Membership and the Intentional-Historical Theory." *Cognition*, 66, 79-85.
- Mark, D. M. 1993 "A Theoretical Framework for Extending the Set of Geographic Entity Types in the U.S. Spatial Data Transfer Standard (SDTS)", *Proceedings, GIS/LIS'93*, Minneapolis, November 1993, 2, pp. 475-483.
- Mark, D. M. 1993a "Toward a Theoretical Framework for Geographic Entity Types", in: Frank, A. U., and Campari, I (eds.), *Spatial Information Theory: A Theoretical Basis for GIS*, Berlin: Springer-Verlag, Lecture Notes in Computer Sciences No. 716, pp. 270-283.
- Mark, D. M., and Egenhofer, M. J. 1994 "Calibrating the Meanings of Spatial Predicates From Natural Language: Line-region Relations", *Proceedings, Spatial Data Handling 1994*, 1, 538-553.
- Mark, D. M., and Egenhofer, M. J. 1994a "Modeling Spatial Relations Between Lines and Regions: Combining Formal Mathematical Models and Human Subjects Testing", *Cartography and Geographic Information Systems*, 21 (4), 195-212.
- Mark, D. M., and Egenhofer, M. J. 1995 "Topology of Prototypical Spatial Relations Between Lines and Regions in English and Spanish", *Proceedings, Auto Carto 12*, Charlotte, North Carolina, March 1995, 245-254.
- Mayhew, S., 1997. *A Dictionary of Geography*. Oxford, New York : Oxford University Press.
- Miller, G. A. and Johnson-Laird, P. N. 1976 *Language and perception*, Cambridge, MA: Harvard University Press.
- Moore, W. G., 1978. *A Dictionary of Geography: Definitions and Explanations of Terms Used in Physical Geography*. New York: Barnes and Noble Books (2nd Edition).

- Morris, M. W. and Murphy, G. L. 1990 "Converging Operations on a Basic Level in Event Taxonomies", *Memory and Cognition*, 18, 407-418.
- Plutchik, R. 1993 "Emotions and Their Vicissitudes: Emotions and Psychopathology", in: M. Lewis and J. M. Haviland (eds.), *Handbook of Emotions*, New York: Guilford, 53-65.
- Quine, W. V. O. 1953 "On What There Is," *From a Logical Point of View*, Cambridge, MA: Harvard University Press.
- Rifkin, A. 1985 "Evidence for a Basic Level in Event Taxonomies", *Memory and Cognition*, 13, 538-556.
- Rosch, E., 1973 "On the internal structure of perceptual and semantic categories", in: T. E. Moore (ed.), *Cognitive Development and the Acquisition of Language*, New York, Academic Press.
- Rosch, E. 1975 "The Nature of Mental Codes for Color Categories", *Journal of Experimental Psychology: Human Perception and Performance*, 1, 303-322.
- Rosch, E., 1978 "Principles of categorization", in: E. Rosch and B. B. Lloyd (eds.) *Cognition and Categorization*. Hillsdale, NJ: Erlbaum.
- Rugg, R., and Schmidt, W. 1986 "Testing the interim proposed standard for cartographic features", in: H. Moellering, (ed). *Issues in Digital Cartographic Data Standards: Report #7, Digital Cartographic Data Standards: A Report on Evaluation and Empirical Testing*. Columbus, Ohio: National Committee for Digital Cartographic Data Standards.
- Smith, B., 1995. "On Drawing Lines on a Map." In Andrew U. Frank and Werner Kuhn (eds.), *Spatial Information Theory. A Theoretical basis for GIS*. Berlin/Heidelberg/New York, etc.: Springer (1995): 475-484.
- Smith, B., 1996. "Mereotopology: A Theory of Parts and Boundaries". *Data and Knowledge Engineering*, 20, 287-303.
- Smith, B., and Mark, D. M., 1998 "Ontology and Geographic Kinds", *Proceedings, Eighth International Symposium on Spatial Data Handling*, 308-320: <http://www.geog.buffalo.edu/ncgia/i21/SDH98.html>.
- Smith, B. and Mark, D. M. 1999 "Ontology with Human Subjects Testing: An Empirical Investigation of Geographic Categories", *American Journal of Economics and Sociology* (in press).
- Smith, B., and Varzi, A., 1997 "Fiat and Bona Fide Boundaries: An Essay on the Foundations of Geography", in S. C. Hirtle and A. U. Frank (eds.), *Spatial Information Theory. International Conference COSIT '97. Laurel Highlands, Pennsylvania, October 1997* (Lecture Notes in Computer Science 1329), Berlin/New York: Springer Verlag, 103-119. Revised version forthcoming as "Fiat and Bona Fide Boundaries", *Philosophy and Phenomenological Research*. <http://wings.buffalo.edu/philosophy/faculty/smith/articles/fiat.pdf>.
- Smith, B., and Varzi, A., (in press) "The Niche," *Noûs*, <http://wings.buffalo.edu/philosophy/faculty/smith/articles/niches.html>.

- Tversky, B. 1990 "Where partonomies and taxonomies meet", in: S. L. Tsohatzidis (ed.), *Meanings and prototypes: Studies on Linguistic Categorization*, London: Routledge, 334-344.
- Tversky, B. and Hemenway, K. 1983. "Categories of Environmental Scenes." *Cognitive Psychology*, 15, 121-149.
- Tversky, B. and Hemenway, K. 1984 "Objects, Parts and Categories" *Journal of Experimental Psychology: General*, 113:169-193.