

# Categorization and Reasoning among Tree Experts: Do All Roads Lead to Rome?

DOUGLAS L. MEDIN, ELIZABETH B. LYNCH, AND JOHN D. COLEY

*Northwestern University*

AND

SCOTT ATRAN

*Centre National de la Recherche, France, and University of Michigan*

To what degree do conceptual systems reflect universal patterns of featural covariation in the world (similarity) or universal organizing principles of mind, and to what degree do they reflect specific goals, theories, and beliefs of the categorizer? This question was addressed in experiments concerned with categorization and reasoning among different types of tree experts (e.g., taxonomists, landscape workers, parks maintenance personnel). The results show an intriguing pattern of similarities and differences. Differences in sorting between taxonomists and maintenance workers reflect differences in weighting of morphological features. Landscape workers, in contrast, sort trees into goal-derived categories based on utilitarian concerns. These sorting patterns carry over into category-based reasoning for the taxonomists and maintenance personnel but not the landscape workers. These generalizations interact with taxonomic rank and suggest that the genus (or folk generic) level is relatively and in some cases absolutely privileged. Implications of these findings for theories of categorization are discussed. © 1997 Academic Press

## DO ALL ROADS LEAD TO ROME?

Human categorization processes necessarily reflect an interplay between the *world*—information available in the environment (Berlin, 1992; Rosch, Mervis, Gray, Johnson, and Boyes-Braem, 1976)—and the *mind*—needs, goals, and theories of the categorizer (Atran, 1990; Gelman & Coley, 1991;

This work was partially supported by NSF Grant 9110145 and NIH Grant MH55079 to Douglas Medin and NSF Grant 931798 to Scott Atran. Special thanks to Florence Sales, Dennis Cepelcha, and Steve Ames for help in participant recruitment, and to all our expert informants from Morton Arboretum, Chicago Botanic Gardens, Evanston Ecology Center, Bartlett Tree Service, Glenwood Tree Experts, Little Acres Tree Service, Broadleaf Tree Care, Evanston Parks and Forestry, Skokie Park District. We also thank Judy Han for help with data entry, and Larry Barsalou, Sandra Waxman, Dan Sperber, and two anonymous reviewers for helpful comments on the paper. Correspondence should be sent to Douglas Medin, Department of Psychology, Northwestern University, 2029 Sheridan Rd., Evanston, IL 60208. E-mail: medin@nwu.edu.

Murphy & Medin, 1985). Perhaps nowhere is this interplay more evident than in the case of folkbiological categories—naive groupings of plant and animal species. The pervasiveness and complexity of plant and animal species guarantee that the world is a rich source of folkbiological information. Likewise, the variety of human uses for plants and animals guarantees that human interests and goals are also rich and varied with respect to living things. We begin this article with some observations involving cross-cultural comparisons of folkbiological categories and we use observations to motivate our within-culture comparison of different goals and interests associated with distinct types of expertise.

Although “carving nature at its joints” is an unpalatable metaphor on several grounds, it does convey the notion that nature comes in chunks and that some ways of organizing it are more likely to take hold than others. This view represents a prominent position in current research on folkbiology. One instantiation of this view is the idea that folk taxonomies are direct reflections of the structure of the natural world (Berlin, 1992). Another is that human theories and interests have universal components that make their influence on categorization indistinguishable from that of world structure (Atran, 1990). For example, people may have a universal tendency to treat biological kinds as having an underlying essence (e.g., Atran, 1987; Keil, 1989; Medin and Ortony, 1989; Gelman, Coley, & Gottfried, 1994; Gelman and Wellman, 1991).

There are well-documented correspondences between folk taxonomies from different cultures and between those folk taxonomies and scientific classification (Atran, 1990; Berlin, 1992; Boster, 1987; Hunn, 1975; see Malt, 1995, for a general review) that could be taken to reinforce the view that biological classification is universally constrained. Of course, the correspondences between folk and scientific taxonomies are hardly perfect, and there is no compelling reason to expect them to be. One problem in comparing two folkbiological categorization schemes is that the distribution of plants and animals may differ across cultures. Likewise, comparing scientific and folk taxonomies is limited by distributional differences between worldwide and local populations. Still, a correlation of even  $+0.70$  between two categorization schemes is accounting for only half of the variance. It is an act of courage to assume that the remaining variance is simply noise induced by measurement problems and distributional differences. In the present studies we investigate the correspondence of scientific categories with the folk categories of three different types of experts when the domain is held constant, thus avoiding distributional problems.

We also investigate the influence of folk and scientific categories on inductive reasoning. One explanation for the close convergence that has been observed between scientific and folk taxonomy is that both aim toward capturing the correlational structure of the environment. Gilmour (1940) argued that scientific categories were “general purpose” and based on overall similarities of many features rather than just a few. Capturing the correlational

structure of the environment in one's categories is important because categories are often used as the basis of inductive inference: information discovered about a particular individual can often be generalized to the entire category to which that individual belongs. Scientific and folk categories may converge because they are both constrained by the goal of maximizing inductive potential (Anderson, 1990). Neither the cross-cultural nor the psychological work on categorization has addressed the question of how goals and interests may affect reasoning. Thus, as well as studying whether and how type of expertise might affect the groupings of trees, we look at the effects of type of expertise on reasoning about trees.

In addition to the question of whether type of expertise affects categorization and reasoning, there is the question of where. According to scientific taxonomy, biological kinds form a deep hierarchy (species, genus, family, order, and so on), and some of these ranks or levels may be much more salient than others. Goals may have an influence only at certain levels of a taxonomy. That is, there may be a "privileged" level at which the categories and reasoning patterns of the different expert types converge. Indeed, many have suggested that in both scientific and folkbiological classification systems there is one level that is especially salient. For scientists, it is the *genus* level (e.g., oak or *Quercus*). Genera have been the fundamental building blocks of all scientific classification systems since Theophrastus (Atran, 1990; Greene, 1909/1983) and form "the most natural and useful level of [biological] classification" (Bartlett, 1940, p. 351). For ethnobiologists, the privileged level in categorization is the "folk-generic" level, which corresponds to the scientific genera or species (Berlin, 1992; Hunn, 1977; Malt, 1995). Folk-generic categories also appear psychologically privileged with respect to inductive inference (Coley, Medin, & Atran, submitted). If the folk-generic/scientific genus rank is truly privileged, we might expect the highest correspondence to science at that level. Differing goals, knowledge, and activities might in turn exert their influence at other levels in the taxonomic hierarchy.

The notion of a privileged level can be thought of in absolute or in relative terms. If a level is *absolutely* privileged, then categories at that level should be extremely salient, virtually "crying out to be named" (Berlin, 1992, p. 53). Such categories might well figure into other, special-purpose taxonomies as primitives and would seldom if ever be broken up. Alternatively, a level might be *relatively* privileged, in that categories at a given level are more likely to appear across subjects, are more inductively powerful, and are more coherent than categories at other levels, but are not reified or universal to the extent that absolute privilege would imply.

All of the experts that participated in the current studies have substantial knowledge and experience concerning trees. However, they differ in their characteristic activities and type of expertise. Specifically, we focus on landscapers, parks maintenance personnel, and taxonomists. Landscapers tend to be concerned with utilitarian aspects of trees and with placing suitable (i.e.,

low-maintenance, disease-resistant) trees in appropriate (e.g., aesthetically pleasing) settings. Parks maintenance workers are primarily involved with removing damaged and diseased trees, planting new trees, and pruning and treating trees in public parks and along streets. Finally, taxonomists tend to be involved in research, consulting, and a variety of educational activities.

There is a great deal of natural diversity among tree species; thus, there are compelling natural similarities and discontinuities among kinds of trees. By using the same set of objects across all the experts we avoid the problem of separating true differences in reasoning and categorization from the effects of stimulus distribution that are inherent in cross-cultural comparisons. By looking at the categorization and reasoning of different types of experts we hope to address the issue of whether there are universal constraints upon biological cognition. We do not expect to find a straightforward yes or no answer to this question. Rather, we hope that our studies will specify particular aspects of categorization and reasoning that are universally constrained, or alternatively, influenced by goals and interests.

### EXPERIMENT 1: CATEGORIZATION OF TREES

In Experiment 1, we examine the groupings of tree species, and explanations for those groupings, provided by three types of tree experts. Are there differences in spontaneously generated taxonomies? If systematic differences exist, the precise nature of these differences might specify the ways in which human interests and goals impact folkbiological taxonomic systems. There are several distinct theoretical positions with respect to differences in categorization as a function of type of expertise.

One possibility is that all roads do lead to Rome—differing goals and concerns may not penetrate the compelling spontaneous categorization that biological kinds afford. But one does not have to be a structure-in-the-world theorist to anticipate the absence of group differences. Goals and associated differences in patterns of interaction may not affect biological classification, because experts interact with biological kinds in so many different ways that a general purpose organizational scheme has the greatest utility. In either case we would expect a high correspondence between folk and scientific taxonomy, justifications that are uniform and predominantly taxonomic, and similarly structured taxonomic hierarchies across groups.

Alternatively, it may be that goals and activities do affect biological classification, but that their influence on categorization is difficult to detect because biological objects have highly correlated features. For example, if something has feathers it is highly likely to have a beak and wings as well (Rosch, 1975, 1978; Rosch and Mervis, 1975). To the extent that features are correlated within categories, it will increase the likelihood that classifiers with different orientations will create the same categories for different reasons. For example, perhaps feathers are important to Bob because he uses feathers to make pillows. Bob has a category of all the objects that he can use to procure

feathers. Alternatively, perhaps wings are important to Jane because she studies flight of vertebrates. Thus, Jane has the category "winged vertebrates." Bob and Jane have almost identical categories. However, their rationale for placing the objects into these categories differs dramatically. In short, similar categories in terms of content may mask differences in the features relevant to those categories in the minds of the categorizers (see also Boster & D'Andrade, 1989). If this were the case, we would expect differences in justifications across groups in the face of high correspondences between folk and scientific taxonomies.

Yet another possibility is that different experts form categories using the same principles, but the taxonomies in which those categories occur are organized differently. For example, perhaps goals and theories do not influence the *kinds* of distinctions that are made by the categorizer but rather the *number* of distinctions (Berlin, 1972). For example, it could be that the taxonomists will have a deeper, more highly differentiated taxonomy than the other types of experts, because they are accustomed to working with trees at a multiple levels of specificity. On this view, experts might agree substantially on the content of categories, but differ on structural aspects of the taxonomy such as the number of levels in the taxonomy or the number of categories at any given level.

Of course, there is also psychological research that would, by extension, support more substantial differences as a function of type of expertise. Different goals, activities, and knowledge may well lead to distinct hierarchical structures—fundamentally different ways of organizing nature. This possibility represents the strongest and most pervasive way in which human goals might shape folkbiological taxonomy.

Models of categorization allow for selecting weighting of features or dimensions (e.g., Medin and Schaffer, 1978; Nosofsky, 1984; see Wisniewski and Medin, 1994, for the more radical possibility that theories may influence the construction and construal of features), and even children show dynamic changes in attention or feature weighting as a function of context (e.g., Landau, Smith, and Jones, 1988). Distinct types of expertise may lead to different patterns of attention to and weighting of morphological properties of trees. One could think of these differences as yielding a variable stretching and shrinking of dimensions of some multidimensional similarity space (e.g., Boster and D'Andrade, 1989; Nosofsky, 1992). On this account the same general similarity space would describe all experts, and any differences might be attributed to variable weighting of a set of common dimensions. For example, taxonomists may differ from landscape and parks maintenance groups in giving extra weight to features associated with reproduction. If differences in the weighting of dimensions were strong enough to overcome the influence of correlated dimensions, one might observe group differences in sorting as well as differences in the justifications.

The final position we shall examine is motivated by Barsalou's analyses

of ad hoc or goal-derived categories (Barsalou, 1982, 1983, 1985, 1991). He argues that categories constructed in the service of goals (e.g., things to take on a camping trip) violate the correlational structure of the environment. In addition, he contrasts context-independent properties of concepts (e.g., basketballs are round) and context-dependent properties (basketballs can float), and he notes that goals often serve to activate context-dependent properties. Especially important for our purposes, Barsalou suggests that frequently used goal-derived categories may develop well-established category representations much like those of common taxonomic categories (see Barsalou and Ross, 1986).

Barsalou's analysis suggests that extensive use of goal-derived categories may lead to the following patterns of spontaneous sorting: (1) the correspondence between sorting and scientific taxonomic distance should be low, (2) the justifications for sortings should be utilitarian (linked to goals), and (3) the folk-taxonomic structure may look more like a series of goal-relevant cross-cuts than a true hierarchical taxonomy. Finally, goals do not necessarily partition the full set of entities in a domain. This raises the possibility that sorting may reflect a mixture of goal-derived and taxonomic categories.

The job descriptions of our three sets of experts suggest that landscape personnel are most likely to frequently use goal-derived categories. Because much of their work involves placing the right tree in the appropriate context, utilitarian functions (providing shade, aesthetic qualities, etc.) are salient to landscapers. On the other hand, the day-to-day activities of parks maintenance personnel involve caring for an existing population of trees, rather than using trees to achieve particular goals. Finally, we assume that the taxonomists will use scientific, rather than goal-derived categories, because most of them frequently use the scientific taxonomy in their work.

In summary, there are theoretical motivations to expect anything from no effect of type of expertise on sorting to a radical restructuring in the service of goals. The justifications for sorts provide data bearing on the idea that the same categories may be created for different reasons. The degree to which expert groups differ in their clustering of tree species should indicate the degree to which human goals and interests influence folkbiological classification. Moreover, the exact nature of differences and similarities among groups will help us to specify the precise aspects of folkbiological classification affected by such goals.

## METHOD

### *Participants*

The participants were 24 men and women (mean age 42.5, range 27 to 72 years old) having occupations related to trees. The average number of years with experience dealing with trees was 16.1, and the range was between 5 and 53 years. Eight participants had at most completed high school, five had

some college work, eight had completed college, and three had advanced degrees (including two with Ph.D.s). Eleven had received at least some formal education regarding trees.

Participants were drawn from a variety of sources and occupations. Sources included two Chicago-area parks districts (Evanston and Skokie), the Morton Arboretum, the Chicago Botanical Gardens, and several private tree-maintenance companies. Participants were paid for their participation. Participants fall into three broad groups: taxonomists, landscapers, and maintenance workers. Taxonomists are principally engaged in research, teaching, and other educational activities. Landscape workers focus on design, aesthetic, and utilitarian aspects of trees. Maintenance workers focus on planting, pruning, and generally maintaining city trees. We will use these groupings as a form of shorthand, as this partitioning represents a considerable oversimplification. In practice, these categories are not necessarily mutually exclusive. For example, one of our experts who ultimately fell into our landscape group has been active in research and teaching for the last 40 years. In addition, workers for private tree-maintenance companies often engage in a mixture of maintenance and landscape activities. In short, we are employing categories when it would be more accurate to say that we have a continuum of types of tree expertise.

### *Design and Materials*

The stimulus materials consisted of names of 48 tree species typed on index cards. Name cards were used instead of pictures to avoid biasing participants toward visual features for their taxonomic judgments and to instead mobilize their overall knowledge of trees. The tree species used are listed in Fig. 1. We used a classical scientific taxonomy (i.e., one based on a combination of morphological and evolutionary considerations), based on information culled from several field guides (Little, 1980; Mohlenbrock, 1992; Petrides, 1988).

There are approximately 80 species of trees native to the Chicago area, and another 30 to 40 species have been introduced. After consulting with numerous reference sources and a few experts on local trees, we selected 48 tree species as broadly representative. Our set consisted of 29 native and 19 introduced species. The sample of 48 trees was biased toward the most common trees as indexed by a combination of city surveys of street trees and in the case of native species, by Swink and Wilhelm (1994). The set of 48 also included trees that often are used in landscaping (e.g., grey dogwood, sweet crabapple, and star magnolia) which would rarely if ever appear on city streets.

Figure 1 also shows that the 48 trees were chosen to represent a broad spectrum of scientific taxa. At the highest taxonomic level, there are 42 angiosperms and 6 gymnosperms. In folk terminology gymnosperms are commonly referred to as "conifers" or "evergreens," and angiosperms as "broad-leafed" trees. In this regard, ginkgo is of particular interest because it is a broad-leafed gymnosperm. These two divisions break down into 3

<i>Species*</i>	<i>Genus</i>	<i>Family</i>	<i>Order</i>	<i>Subclass</i>	<i>Class</i>	<i>Division</i>
Weeping Willow	<i>Salix</i>					
Eastern Cottonwood						
White Poplar	<i>Populus</i>	<i>Salicaceae</i>	<i>Salicales</i>			
Little-leaf Linden	<i>Tilia</i>	<i>Tiliaceae</i>	<i>Malvales</i>	<i>Dillenidae</i>		
Black Walnut	<i>Juglans</i>					
Shagbark Hickory	<i>Carya</i>	<i>Juglandaceae</i>	<i>Juglandales</i>			
European Black Alder	<i>Alnus</i>					
Paper Birch						
River Birch	<i>Betula</i>					
American Hornbeam	<i>Carpinus</i>	<i>Betulaceae</i>				
American Beech	<i>Fagus</i>					
Bur Oak						
Northern Red Oak						
Pin Oak						
White Oak	<i>Quercus</i>	<i>Fagaceae</i>	<i>Fagales</i>			
American Elm						
Siberian Elm	<i>Ulmus</i>					
Hackberry	<i>Celtis</i>	<i>Ulmaceae</i>				
White Mulberry	<i>Morus</i>	<i>Moraceae</i>	<i>Urticales</i>			
American Sycamore						
London Planetree	<i>Platanus</i>	<i>Platanaceae</i>	<i>Hamamelidales</i>	<i>Hamamelidae</i>		
Star Magnolia	<i>Magnolia</i>					
Tuliptree	<i>Liriodendron</i>	<i>Magnoliaceae</i>	<i>Magnoliales</i>	<i>Magnoliidae</i>		
Catalpa	<i>Catalpa</i>	<i>Bignoniaceae</i>				
Green Ash						
White Ash	<i>Fraxinus</i>	<i>Oleaceae</i>	<i>Scrophulariales</i>			
American Elder	<i>Sambucus</i>	<i>Caprifoliaceae</i>	<i>Dipsicales</i>	<i>Asteridae</i>		
Sweet Crab Apple	<i>Malus</i>					
American Mountain-Ash	<i>Sorbus</i>					
Black Cherry	<i>Prunus</i>					
Washington Hawthorn	<i>Crataegus</i>	<i>Rosaceae</i>	<i>Rosales</i>			
Kentucky Coffeetree	<i>Gymnocladus</i>					
Honeylocust	<i>Gleditsia</i>	<i>Caesalpinaceae</i>	<i>Fabales</i>			
Amur Maple						
Boxelder						
Norway Maple						
Silver Maple						
Sugar Maple	<i>Acer</i>	<i>Aceraceae</i>				
Horsechestnut						
Ohio Buckeye	<i>Aesculus</i>	<i>Hippocastanaceae</i>				
Tree of Heaven	<i>Ailanthus</i>	<i>Simaroubaceae</i>	<i>Sapindales</i>			
Grey Dogwood	<i>Cornus</i>	<i>Cornaceae</i>	<i>Cornales</i>	<i>Rosidae</i>	<i>Magnoliopsida</i>	<i>Angiosperms</i>
Austrian Pine						
Eastern White Pine						
Scotch Pine	<i>Pinus</i>					
Colorado Spruce	<i>Picea</i>	<i>Pinaceae</i>	<i>Pinales</i>			
Eastern Redcedar	<i>Juniperus</i>	<i>Cupressaceae</i>	<i>Taxales</i>	<i>Pinidae</i>	<i>Pinatae</i>	
Ginkgo	<i>Ginkgo</i>	<i>Ginkgoaceae</i>	<i>Ginkgoales</i>			<i>Gymnosperms</i>

Note: Common folk names, rather than Latin names are given for Species.

FIG. 1. Scientific taxonomy of tree species used in Experiments 1 and 2.



classes, 7 subclasses, 16 orders, 21 families, and 33 genera. There are 9 families represented by more than 1 genus, and 9 genera represented by more than 1 species.

Finally, our tree set includes variations in patterns of folk nomenclature. Berlin (1992) and others argue that common monolexemic names typically (but not always) mark folk-generic categories, which in turn often correspond to scientific genera. For example, the four representatives of the genus *Quercus* are all commonly called "oaks." Further examination of Fig. 1 reveals that in our set, five of the genera with multiple members have the same folk-generic names (birch, oak, elm, ash, pine) and four have different folk-generic names (cottonwood/poplar, sycamore/planetree, maple/boxelder, and horsechestnut/buckeye). In addition, the American mountain-ash is marked as if it were a member of the ash genus (*Fraxinus*), but it is not. These variations provide some opportunities to uncouple the role of scientific and folkbiological names in expert categorization.

### *Procedure*

Participants were tested individually. They were told that we were interested in how experts organize their knowledge about trees. First, participants were shown the 48 tree cards one at a time and asked to tell the experimenter a bit about each kind of tree. Responses were tape recorded and later transcribed. If a participant indicated that he or she was unfamiliar with a particular tree, that card was set aside and not used for the sorting task.

After this initial phase, the remaining cards were spread out in front of the participant, who was asked to "put together the trees that go together by nature into as many different groups as you'd like." When the initial sorting was complete the participant was asked to describe or justify the basis for each grouping. Participants were then invited to combine sorts by "putting together those groups of trees that go together by nature into as many larger groups as you'd like." Again, justifications for each category were solicited. Participants were free to combine as many (or as few) groups as they liked on any round of sorting. The successive pile sorting was repeated until the participant indicated that no further grouping seemed natural. At this point the experimenter restored the categories created during the initial free sort and invited the participants to "split as many of the groups as you'd like into smaller groups of trees that go together by nature." The subpile sorting was repeated until participants indicated that no further subdivisions seemed sensible.

*Measurement.* A mathematic tool well suited to comparing individual taxonomies is the cultural consensus model (CCM) of Romney, Batchelder, and Weller (1986). The model is implemented by means of a factor analysis of a matrix of intersubject agreement. If the factor analysis results in a single factor solution, then according to Romney et al. a single competence or consensus underlies the responses. The dimension can be thought of as re-

flecting the degree to which each participant shares in the consensual knowledge. Competencies are provided by the first factor score of each participant. Failure of consensus takes the form of either negative factor loadings on the first dimension or a multiple factor solution. The CCM has been widely applied to the categorization of folkbiological kinds to reveal both consensus and residual agreement (e.g., Atran, 1994; Boster and Johnson, 1989; Coley, 1995; Johnson, Mervis, & Boster, 1992; López, Atran, Coley, Medin, & Smith, in press). Significant residual agreement indicates the existence of subgroups who agree among themselves beyond levels predicted by the general consensus. Examining both consensus and residual agreement is one way to determine whether there are systematic differences among the experts.

## RESULTS

Because of the richness of information provided by the sortings and justifications, we present a variety of measures of category organization. We first describe how performance on the sorting task was quantified. Then we ask whether there exists a single consensual folk taxonomy for tree experts, and we present evidence that supports partitioning experts into groups. Next we present analyses focusing on two main questions: (1) How closely does expert folk taxonomy correspond with scientific taxonomy? (2) How well do the folk taxonomies of groups of experts correspond to each other? Analyses addressing these questions include qualitative characterization of the consensual taxonomies of each group along with justifications, structural comparisons of the breadth and depth of taxonomies for different kinds of experts, and an analysis of categories common to all three groups. To anticipate, our analyses reveal commonalities but also systematic differences.

### *Scoring*

Each participant's tree taxonomy was obtained by translating the groupings made during the initial free sorting, successive superordinate sorts, and successive subordinate sortings into a taxonomic hierarchy. An example of such a taxonomy is presented in Fig. 2. On the first round of sorting (level 3 in this case), this participant sorted the trees into eight "common trees" (e.g., No. 14 Colorado spruce, No. 37 silver maple, and No. 31 paper birch), "seven trees used in the same landscape setting" (e.g., No. 8 Austrian pine, No. 20 green ash, and No. 6 american sycamore), "four trees that grow independently" (e.g., No. 11 boxelder and No. 44 weeping willow) and "elms" (No. 3 and No. 36, American and Siberian elm), among others. On the first round of combining these initial groups (level 4), the participant combined the Colorado spruce–silver maple–birch group and the Austrian pine–ash–sycamore group into one group of "everyday" or "typical" trees. At the same time, the elms were combined with the "independent" group. On the second round of combining (level 5), the participant combined these two groups into a single group of "popular" trees. At that point, with a total of three groups

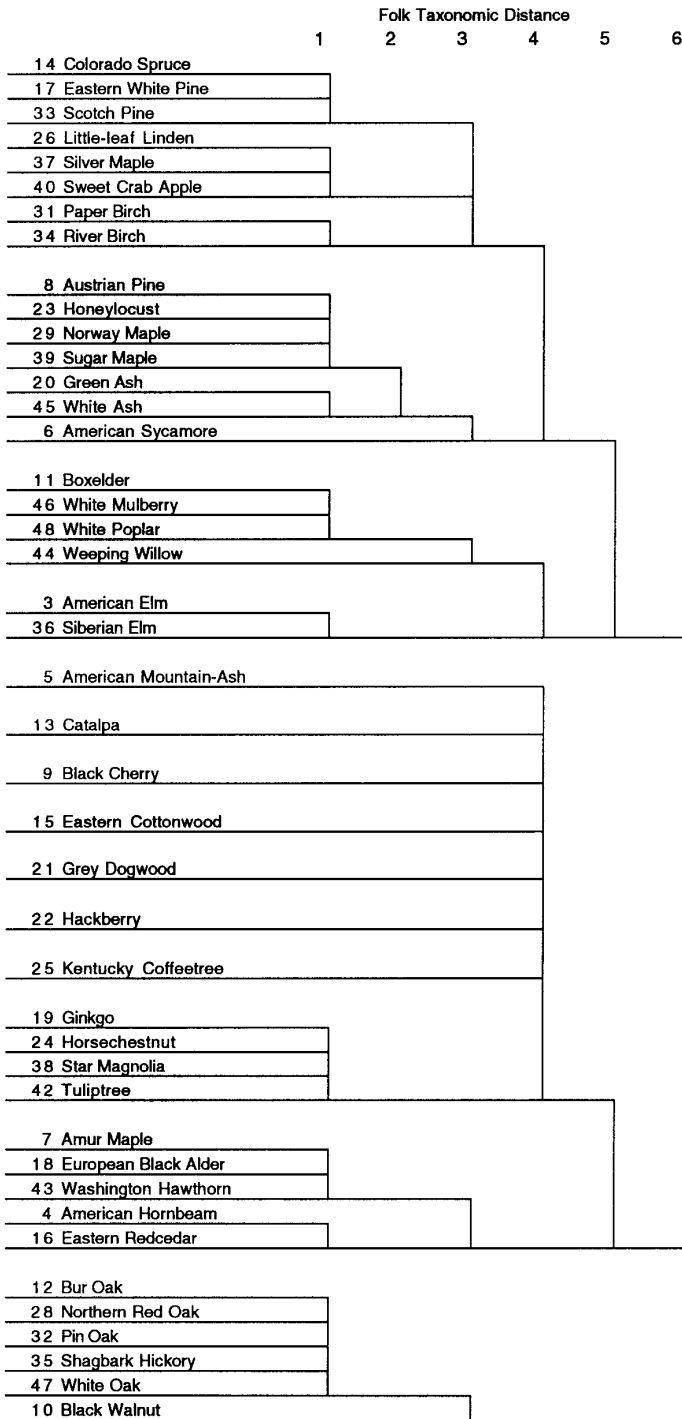


FIG. 2. Example of a tree-sorting taxonomy.

remaining, the participant declined to combine groups further, so the remaining groups were combined at level 6. The participant's original groups were then restored; on the first round of splitting (level 3), the participant divided the "common" trees into three groups: "conifers" (spruce and pines), "birches," and "the rest" (linden, silver maple, sweet crab apple). The participant also divided the "trees used in the same landscape setting" into two groups: American sycamore ("more eye-catching") versus the rest. On the next round of splitting (level 2), the participant declined to further divide the "common" subgroups, but did split the Austrian pine-green ash cluster into two groups: "ash trees" and "the rest." At this point the participant declined to split any more groups, so they were separated into individual species at level 1.

For each taxonomy we derived a pairwise tree-by-tree folk-taxonomic distance matrix by calculating the distance between all possible pairs of trees in the taxonomy. The lowest level at which two given trees go together represents the distance between them. Low folk-taxonomic distance corresponds to high folk-biological relatedness. The distance between any species and itself is 0. If two tree species were placed together at the lowest level (most specific) sort their taxonomic distance was 1. (In the rare event that an expert suggested that two folk taxa were coextensive, we set their distance at 0.) Two species combined at the second most specific level would have a distance of 2, and so on. For example, in Fig. 2, No. 39 sugar maple is distance 1 from No. 29 Norway maple, 2 from No. 20 green ash, 3 from No. 6 American sycamore, 4 from No. 34 river birch, 5 from No. 11 boxelder, and 6 from No. 5 American mountain-ash. The result was a  $48 \times 48$  distance matrix with the upper diagonal being equivalent to the lower diagonal (because distance is symmetrical). Unfamiliar trees were scored as missing data for all pairs involving the unfamiliar tree.

Overall, the number of trees known ranged from 33 to 48 ( $M = 43.9$ ,  $SD = 4.8$ ). Nineteen trees (40% of the set) were familiar to all participants, and 34 trees (71%) were familiar to over 90% of participants. All taxonomists were familiar with all 48 trees; landscapers ( $M = 44.8$ ,  $SD = 5.2$ ) and maintenance workers ( $M = 41.4$ ,  $SD = 4.0$ ) did not differ in terms of mean number of familiar trees,  $t(18) = 1.64$ , n.s. These numbers are no doubt underestimates. We used a single common name for each species, and our choice may not have matched the common name used by an expert. For example, seven experts (five in the maintenance group and two in the landscape group) indicated that they were unfamiliar with both the American elder and the American hornbeam. But the American elder is also known as elderberry, and American hornbeam is commonly referred to as blue beech or as ironwood. In other cases our common names may have been too specific. Five experts (three in the maintenance subgroup and two in the landscape group) indicated unfamiliarity with European black alder. This is surprising given that European black alder is not unusual and appears as a street tree.

The problem may be that this tree is referred to as black alder or even just alder. However, there may be differences between landscapers and maintenance workers in terms of which trees were deemed unfamiliar. Specifically, Fisher's exact tests revealed that maintenance workers were more likely than landscapers to be unfamiliar with grey dogwood, Amur maple, and eastern redcedar ( $ps < .05$ ). This is understandable, given that these trees are commonly used as ornamental trees but rarely planted in parks and along city streets.

### *Cultural Consensus Model and Identification of Expert Subgroups*

*Overall consensus on tree categorization.* Each expert's tree-distance matrix was correlated with that of every other expert, yielding a  $24 \times 24$  matrix in which entries correspond to agreement among experts on pairwise tree distances derived from their individual taxonomies. A principal component factor analysis was then performed on the intersubject correlation matrix to determine how well it fits the Romney et al. (1986) cultural consensus model. A fit to the CCM is supported by a single factor solution in which (1) the first latent root or eigenvalue is relatively large compared to the rest, (2) all scores on the first factor are positive, and (3) the first factor accounts for most of the variance. A strong fit would indicate a single consensus among experts in their sorting of tree names.

The overall results provide moderate support for the CCM. The first three eigenvalues were 8.04, 2.45 and 1.42, accounting for 33.5, 10.2, and 5.9% of the variance, respectively. The first latent root is 3.3 times the size of the second root; in addition, all expert scores on the first factor were positive ( $M = .554$ , range: .212–.804). This suggests a moderate fit of the model. However, the second factor accounted for over 10% of the variance, and more significantly, the loadings on the second factor appeared to correspond to type of experience, roughly based on self-described occupational activities. Specifically, maintenance workers and taxonomists tended to have positive loadings on the second factor, whereas landscapers tended to have negative values. Thus, the cultural consensus model indicates moderate consensus among experts but also hints at systematic variation by subgroup.

*Analysis of residual agreement.* In order to test the hypothesis of systematic group differences in patterns of categorization, a subject-by-subject residual agreement matrix was prepared, as described in Nakao and Romney (1984). First, the products of first-factor consensus scores were obtained for each pair of subjects. This represents agreement predicted by each participant's knowledge of the consensus. Next, this predicted agreement matrix was subtracted from the observed agreement matrix, yielding a residual agreement matrix, which was standardized. This standardized residual agreement matrix was compared to a model matrix corresponding to subgroup membership (i.e., a  $24 \times 24$  matrix with entries of "1" if the corresponding participants belong to the same group, otherwise "0"). These two matrices were then compared

using Monte Carlo simulations, as described in Hubert and Shultz (1976), to assess whether residual agreement is higher among subgroups than among randomly chosen pairs of participants. If patterns of agreement are completely described by the consensus component, there should be no appreciable residual agreement among experts belonging to the same subgroup (see Boster, 1986; Coley, 1995; Johnson, Mervis, & Boster, 1992). In contrast, systematic agreement within subgroups would lead to significant association between the model and residual matrices. Indeed, this turned out to be the case. The degree of association between the model matrix and the residual agreement matrix was significant, as assessed by the quadratic assignment program (Hubert & Schultz, 1976),  $z = 10.69$ ,  $p < .001$ , indicating significant residual within-group agreement not explained by the overall consensus factor. Thus, the residual within-group agreement among taxonomists, landscapers, and maintenance workers, respectively, is higher than expected by chance.

*Summary.* Results of the cultural consensus model suggest a shared component to the taxonomies of the participants, accounting for roughly one-third of the variance. But residual analyses also revealed systematic differences within our expert sample, differences that fall roughly along occupational lines. Given this significant residual agreement among subgroups, we will report subsequent analyses by group, although it is important to bear in mind that this represents an approximation that amounts to a continuous dimension rather than discrete clusters.

### *Relations to Scientific Taxonomy*

The relation between folk and scientific taxonomy was analyzed in several ways. First, in order to derive an overall index of the relation of scientific and folk taxonomy, we directly correlated folk and scientific taxonomic distance for each group of experts. Second, we addressed the question of whether folk taxonomy corresponds especially well to scientific categories at a particular rank by analyzing the mean folk taxonomic distance between species at each scientific rank. We also examined scientific groupings hypothesized to be especially salient—genus and family—by how often they were preserved or broken up in the sortings of individual participants. Finally, we examined the role of folk nomenclature in determining whether genera were preserved or not.

*Correlations between folk and scientific taxonomic distance.* In order to correlate folk distance with scientific distance, we averaged the pairwise folk taxonomic distance for each expert subgroup (taxonomists, landscapers, maintenance workers) and correlated the resulting matrices with an analogous matrix representing the “scientific” taxonomic distance between each species. Correlations between folk and scientific taxonomy were all significant, but differed greatly between groups. Not surprisingly, taxonomists’ sortings correlated highly with science ( $r = 0.85$ ). The correlation for maintenance workers was lower ( $r = 0.48$ ), and that for landscapers was even lower ( $r = 0.37$ ),

reflecting only moderate correspondence with scientific taxonomy. The low correlations for landscapers and maintenance workers appear to leave room for a generous contribution of goals and interests.

*Folk distance as a function of scientific rank.* Another question was whether there were any scientific ranks (e.g., the genus level) that were especially salient with respect to folk taxonomy. To address this, for each participant we averaged the folk-taxonomic distances between trees related at each scientific rank. For example, an expert's folk distances between all the pairs of trees related at the genus level were averaged, and likewise for all pairs related at the family, order, subclass, class, and division level. These scores were analyzed using a 3 (Expert Group: Maintenance Workers, Landscapers, Taxonomists)  $\times$  6 (Level of Scientific Relation: Genus, Family, Order, Subclass, Class, Division) mixed ANOVA. To the degree that folk distance mirrors scientific distance, the folk distance between more closely related trees (e.g., those related at the genus or family level) should be lower than the folk distance between more distantly related trees (e.g., those related at the class or division level). As predicted, folk distances increased systematically with scientific distance, ( $F(5,105) = 108.64$ ,  $MSe = 0.196$ ,  $p < .001$ ). Tukey HSD analyses showed that folk distances were lowest for trees related at the genus level, lower at the family level than at all other higher-order levels, and lower at order, subclass, and class than at division ( $p < .01$ ).<sup>1</sup> However, this pattern interacted with expertise type, ( $F(10,105) = 7.79$ ,  $MSe = 0.196$ ,  $p < .001$ ). This interaction is shown in Fig. 3.

To further explore this interaction, separate one-way ANOVAs were performed for each expert group; folk distances at each scientific rank were compared using Fisher's least-significant difference tests ( $p < .05$ ). For the taxonomist group both genus and family were reliably lower than every higher-order rank. Beyond that, order was reliably lower than class and division; subclass and class were also reliably lower than division. Given the relative lack of power associated with the smaller sample size ( $n = 4$ ) the taxonomist group showed considerable differentiation. For the maintenance workers, folk distances between genus members were lower than all other ranks, and distances between family members were lower than all other ranks, save genus. No other differences were reliable. For landscapers, members of the same genus were not rated any closer than members of the same family; however, folk distances for genus and family were lower than for any other higher-order ranks. Finally, landscapers rated subclass and class as reliably

<sup>1</sup> The rank at which ginkgo is related to other gymnosperms (pine, spruce, redcedar) was excluded from this analysis because there was only a single species represented at this distance. Moreover, an artifact of our procedure may have biased responses at this rank, because to reproduce the normative scientific distance the expert would have had to decline to do any lumping for two levels and then agree to lump at the third. However, our procedure stopped at the first point where an expert declined to do further lumping.

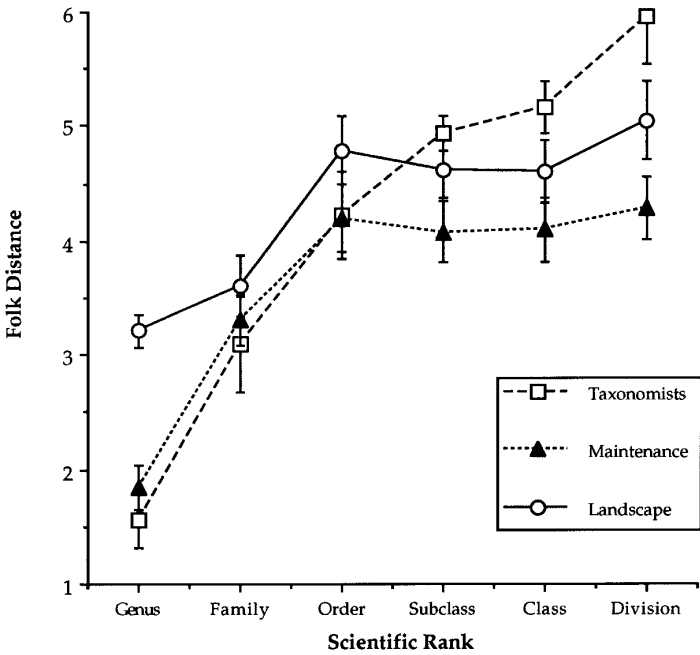


FIG. 3. Folk distance as a function of scientific rank.

closer than division. Overall, the pattern of results is clear. For taxonomists and maintenance workers (but not landscapers), the genus-level groups were closer than family-level groups; for all three groups of experts, trees related at the genus and family levels were also closer in terms of folk-taxonomic distance than higher-order groups. Maintenance workers made no further higher-order differentiations, taxonomists made many, and landscapers differentiated between the rank of division (corresponding to the distinction between gymnosperms and angiosperms) and lower ranks. Overall, genus and family—relatively low-level scientific relations—were salient for all groups; higher-order scientific relations were much less so.

*Individual preservation of genus and family groups.* Given that trees related at the genus and family level were closer in folk taxonomy than more distantly related trees, a related question is to what degree participants' groupings actually correspond to these scientific groupings. Note that the finding that members of the same genus are deemed taxonomically closer on average than members of higher-order scientific groups does not require that any participant produce a grouping that directly corresponds to any genus. An individual taxonomy may show one of three patterns with respect to a category such as the genus *Acer* (which in our sample of trees is represented by Norway maple, Amur maple, sugar maple, silver maple, and boxelder; see Fig. 1): (1) it may



be preserved precisely, (2) it may appear as a subset of a more inclusive category, or (3) it may be broken up. These possibilities are not mutually exclusive if one looks across levels, but we scored each category for each expert at the level that maximized its correspondence to scientific taxonomy. Thus, if a category was preserved at one level and was a subset at a higher level or rank, we scored it as preserved. If it was preserved at one level and was broken up at a more specific level, it was scored as preserved. Finally, if the category was a subset of one level and broken up at a more specific level, it was scored as a subset.

Overall, genus-level categories were preserved much more often than family-level categories. Specifically, across all subjects, genus-level categories were preserved on 42.1% of all possible occasions; family-level categories were preserved on only 19.3% of all possible occasions. Family-level categories were also much more likely than genus-level categories to be broken up, 62.3% versus 30.2%. Genus-level categories appeared as a subset 27.7% of the time compared with 18.4% for family-level categories.

Not surprisingly, expert subgroups preserved genus and family groups to different degrees. Taxonomists never broke up a genus-level category; indeed, genus-level categories were preserved 88.9% of the time. In contrast, landscapers broke up genus-level categories 50.6% of the time, reliably more than maintenance workers ( $M = 22.6\%$ ,  $t(18) = 2.70$ ,  $p < .02$ ). Likewise, landscapers preserved genus-level groups 26.7% of the time; maintenance workers preserved them 38.9% of the time. Taxonomists broke up family-level categories only 16.7% of the time and preserved them 66.7% of the time. In contrast, landscapers broke up family-level categories 65.9% of the time and preserved them 5.6% of the time; the corresponding figures for maintenance workers were breaking 78.3% and preserving 13.3% of the time. Neither of these latter subgroup differences were statistically reliable.

Overall, taxonomists were much more likely to preserve scientific taxa at the genus and family level; neither maintenance workers nor landscapers tended to respect scientific families, but maintenance workers were more likely than landscapers to respect genera. Genus-level scientific groupings are much more likely to show up intact in folk taxonomies than in family-level scientific groupings. However, even genus-level scientific groupings are preserved less than half the time. Finally, landscapers were twice as likely as maintenance workers to break up genus-level categories.

*Role of language.* Because genus-level groupings were preserved only about half the time, it is of interest to examine correspondence of folk and scientific classification at the genus level as a function of folk nomenclature. Our stimuli include a subset of genus categories where folk-generic marking and scientific taxonomic status coincide perfectly. For example, both representatives of the birch genus *Betula* are also marked as different folk specifics of the same folk-generic birch (paper birch and river birch). This is also true for pine, oak, and elm. There are also four situations where folk nomenclature

fails to support scientific genera. American sycamore and London planetree are both members of the genus *Platanus*, Ohio buckeye and horsechestnut are both *Aesculus*, eastern cottonwood and white poplar are both *Populus*, and four of the maples plus boxelder are all *Acer*.

In order to examine the effects of patterns of folk nomenclature on preservation of these generic taxa, we compared how often supported genera versus nonsupported genera were preserved and broken up. Results show a consistent effect of folk naming patterns. Maintenance workers preserved 56.7% of genera supported by folk names, but only 35.0% of genera not supported by folk names ( $t(9) = 3.71, p < .005$ ). Likewise, maintenance workers broke up 37.5% of genera not supported by folk names, but only 13.3% of genera supported by folk names ( $t(9) = 2.57, p < .05$ ). Landscapers preserved 40.0% of genera supported by folk names, but only 13.3% of genera not supported by folk names ( $t(9) = 2.74, p < .05$ ). In short, both groups preserved the genus level more often when it was supported by folk terminology. Interestingly, five experts preserved the subset of the genus *Acer* marked by folk nomenclature (that is, they put Amur maple, silver maple, sugar maple, and norway maple together but excluded boxelder).

Finally, there is one situation among our tree species where folk nomenclature is misleading with respect to scientific classification. As mentioned earlier, white ash and green ash are members of the genus *Fraxinus*; American mountain-ash (genus *Sorbus*) is only distantly related. Nevertheless, 5 of the 10 maintenance workers and 2 of the 10 landscapers preserved this category (white ash, green ash, American mountain-ash). None of the taxonomists did so. This lends further support to the idea that folk terminology influences folk-taxonomic sorting, at least among maintenance workers and landscapers.

*Summary.* Overall, comparisons of folk and scientific taxonomy suggest modest correspondence accompanied by differences based on type of expertise. Correspondence is higher at the rank of genus; genus groups are split up less often and preserved more often than higher-order groups. However, correspondence to science differed appreciably by expertise type. Although taxonomists did not break up genus-level categories (nor for the most part, family-level categories), the maintenance workers and especially landscapers did so frequently. In general, landscapers' sortings conformed least well with scientific taxonomy. Folk patterns of naming also influenced responses; apparently sharing a common name reinforced genus-level commonalities, although as we shall point out in the discussion, it is difficult to pinpoint cause and effect.

### *Relations among Folk Taxonomies of Expert Groups*

Results so far clearly suggest systematic differences between classification systems for different kinds of experts. Therefore, it becomes important to specify the nature of these differences. In order to do so, we directly compare taxonomies for the three groups of experts. First, we present comparisons of

structural aspects of the taxonomies. Then we present qualitative descriptions of consensual clusters for each group, analyses of justifications, and an analysis of categories common to all experts. By pinpointing how these taxonomies diverge and converge, we gain insights into how human goals and interest impact folkbiological classification.

*Structural aspects of group taxonomies.* Systematic differences emerged between experts in terms of the number of groups they formed initially and the number of levels their taxonomies contained. For their initial sorts the experts averaged 13.5 categories (range = 5 to 31). The subgroups once again showed distinct differences. Taxonomists averaged 19.0 (range = 18 to 20), maintenance workers 16.5 (range = 8 to 31), and landscapers 8.2 (range = 5 to 12) groupings on their initial sorting. Landscapers had reliably fewer initial groups than either maintenance workers ( $t(18) = 3.59, p < .01$ ) or taxonomists ( $t(12) = 9.37, p < .01$ ).

The average number of final groupings (i.e., the point at which participants declined to combine piles any further) was 6.1. Again, experts showed distinct differences. Taxonomists averaged 3.5 (range = 2 to 8), maintenance workers averaged 9.4 (range = 2 to 24), and landscapers averaged 3.9 (range = 2 to 11) final categories. Landscapers had reliably fewer final categories than maintenance workers, ( $t(12) = 2.21, p < .05$ ), but differences between maintenance workers and taxonomists fell short of reliability.

Differences were also evident in the number of levels or depth associated with individual taxonomic structures. Overall, taxonomies had an average of 4.0 levels (range = 2 to 6) but again there were reliable subgroup differences. The average for maintenance workers ( $M = 3.3$  levels) was reliably smaller than that for the landscapers ( $M = 4.5, t(18) = 3.50, p < .01$ ) or taxonomists ( $M = 4.8, t(12) = 3.24, p < .01$ ). Strikingly, although maintenance workers created more initial groups than landscapers, they produced *fewer* superordinate levels of sorting ( $t(18) = 2.61, p < .02$ ).

In sum, the three groups produced structurally distinct taxonomies. Maintenance workers' taxonomies were broad and shallow, with many initial groupings and modest subsequent collapsing of those groups into more general superordinate classes. Landscapers and taxonomists both produced deeper taxonomies with a small number of highly inclusive classes, but landscapers started with relatively few initial groups, whereas taxonomists (and maintenance personnel) started with many. These differences coupled with differences in the content and justifications reinforce the conclusion that the category organization of the landscapers diverges from the other two groups.

*Cluster analysis.* Cluster analyses were run separately for the taxonomists ( $n = 4$ ), landscapers ( $n = 10$ ), and maintenance workers ( $n = 10$ ). A matrix representing mean pairwise distance between all trees was computed for each group of experts. This was then subject to cluster analysis, using the average link method (Sneath & Sokal, 1963), yielding the tree diagrams shown in Figs. 4–6. Not surprisingly, the taxonomists' consensual clustering (Fig. 4)

corresponds quite closely with scientific taxonomy. There is a cluster corresponding to each of the nine genera and to six of the nine families represented in our set of tree species. Moreover, the ginkgo (a broad-leaf deciduous gymnosperm) is classified with the other gymnosperms, rather than with other broad-leafed deciduous angiosperms.

Cluster analysis for landscapers (Fig. 5) reveals a clear contrast. Six of nine genera and only three of nine families are preserved. The departures from scientific taxonomy are striking. The Siberian elm and the American elm are in different clusters. Members of the genus *Acer* (maples) appear in three different clusters. The ginkgo is grouped with angiosperms. Furthermore, several of the clusters reveal utilitarian concerns. For example, one cluster consists of Kentucky coffeetree, hackberry, ginkgo, little-leaf linden, honeylocust, white ash, and green ash; these are all desirable trees to plant along city streets. Similarly, Washington hawthorn, sweet crab apple, star magnolia, and grey dogwood form a cluster of flowering ornamental trees. European black alder, Amur maple, American hornbeam, and American mountain-ash are also small trees useful for ornamental roles. American elm, American beech, black walnut, and shagbark hickory are considered to be "specimen" trees—useful in stand-alone contexts or as the centerpiece of a landscape layout. Finally, the cluster consisting of white mulberry, boxelder, tree of heaven, Siberian elm, eastern cottonwood, white poplar, and weeping willow correspond to undesirable "garbage" or "weed" trees, according to landscapers. In short, the aggregate landscaper taxonomy reflects utilitarian as well as taxonomic groupings.

The aggregate taxonomy for maintenance workers (Fig. 6) is distinct from that of both the taxonomists and the landscapers. Like landscapers, six of the nine genera and three of the nine families were preserved. Also like landscapers, the ginkgo was placed with other broadleaf trees rather than with the conifers. Amur maple is isolated, perhaps because it is a relatively unfamiliar tree for the maintenance group or perhaps because it is significantly smaller than other maples. Grey dogwood is also isolated, in part because it is considered by many maintenance workers to be a shrub rather than a tree. Notably absent are the utilitarian groupings seen for landscapers; many of the maintenance workers' nontaxonomic clusters appear to be based on morphological rather than utilitarian properties. For example, shagbark hickory, black walnut, Ohio buckeye, and horsechestnut form a group of trees with nuts; hackberry, black cherry, sweet crab apple, and white mulberry are fruit-bearing trees. Six of the 10 maintenance workers did label one of their initial categories as "weed" trees, but the consensual cluster analysis did not preserve this category. For maintenance workers, cluster analysis reveals primarily taxonomic or alternative morphologically based groupings. Finally, there is some evidence that folk terminology influenced maintenance workers' clustering; American mountain-ash forms a group with the white ash and green ash even though in scientific terms the American mountain-ash is only distantly related.

**TAXONOMISTS**

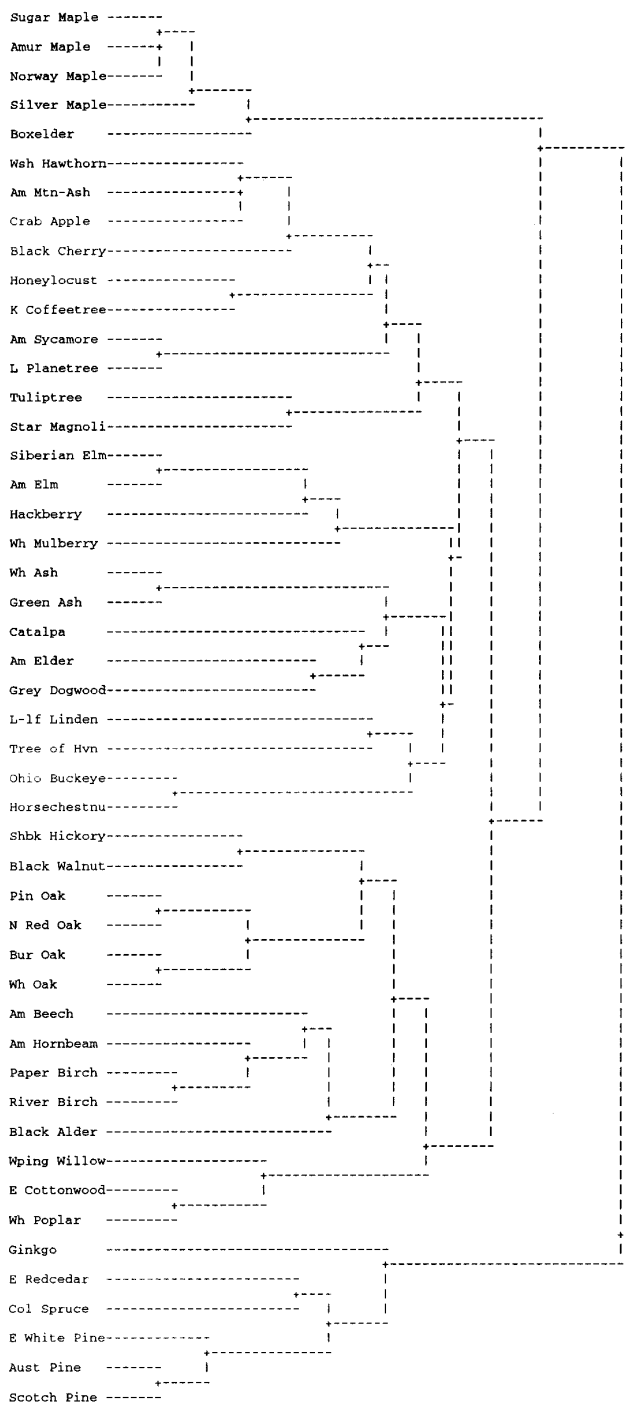


FIG. 4. Consensual cluster: Taxonomists.

## LANDSCAPERS

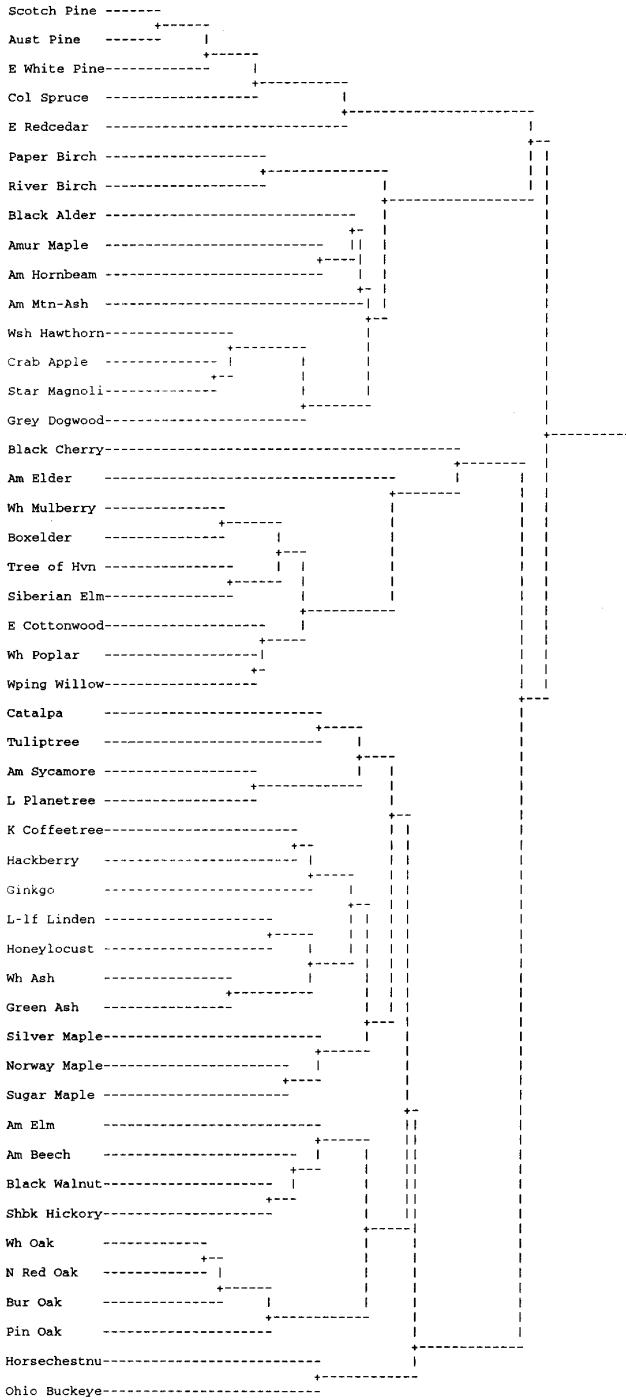


FIG. 5. Consensual cluster: Landscapers.

**MAINTENANCE WORKERS**

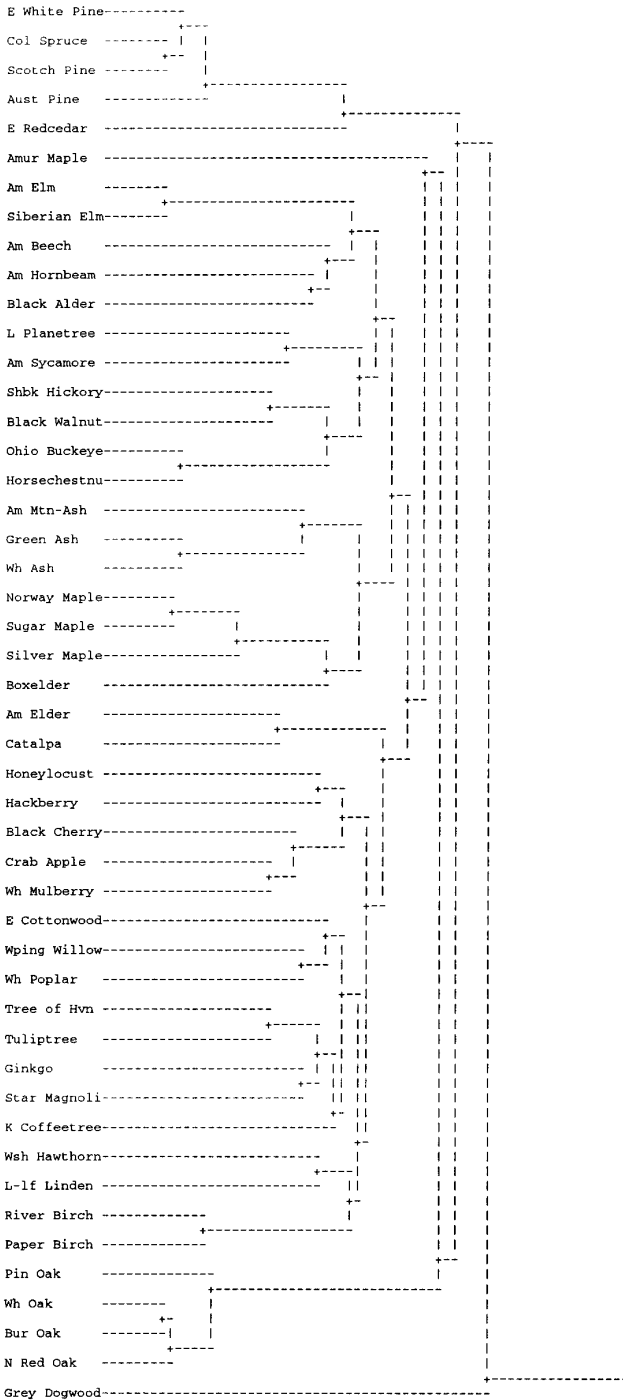


FIG. 6. Consensual cluster: Maintenance workers.

TABLE 1  
Proportion of Experts in Each Subgroup Using the Various Types of Justifications  
for Any Category in Their Initial Sort

Type of justification	Taxonomist	Expert group maintenance	Landscaper
Taxonomic	1.00	0.90	0.90
Morphological	0.25	0.80	0.40
Weed	0.00	0.60	1.00
Landscape utility	0.00	0.10	0.80
Aesthetic	0.00	0.40	0.60
Size	0.00	0.20	0.60
Distribution	0.00	0.20	0.40
Native/nonnative	0.00	0.10	0.50

Together, the cluster analyses reveal interesting qualitative differences and similarities among the expert groups (see analysis of “common clusters” below). Not surprisingly, the taxonomists’ aggregate taxonomy corresponded closely with science. In contrast, the landscapers’ aggregate taxonomy showed many clearly utilitarian groupings. Finally, the maintenance workers’ aggregate taxonomy showed the influence of morphological similarity and folk nomenclature.

*Justifications.* A close examination of how experts explained their groupings helps to clarify and highlight the patterns seen in the cluster analyses. We developed eight categories to capture the most common types of sorting justifications: taxonomic (in which a group of trees were said to belong to “the same family” or were names, e.g., “oaks”), morphological, “weed,” landscape utility, aesthetic, size, distribution, and native/nonnative. Most of these are self-explanatory and none were treated as mutually exclusive. For example, the category description “large, native, landscape trees” would be scored as size, native/nonnative, and landscape utility. The only difficulty in scoring concerned whether “ornamental” is a utilitarian or an aesthetic description. We resolved this by assuming that both categories were implicated. Distribution refers to either relative abundance or characteristic location of trees (e.g., forest).

Table 1 summarizes the justifications given on initial sortings broken down by subgroup (justifications on other sorts show essentially the same pattern). Again there are distinct subgroup differences. Unsurprisingly, taxonomists used taxonomic justifications; all of them also mention morphological properties at some point in the sorting task (though only one did so on the initial sort). Maintenance personnel also employed taxonomic and morphological justifications, but they depart from taxonomists in mentioning weed status and (sometimes) aesthetic aspects of trees. Distributional information was also mentioned by six maintenance workers but generally for justifications after first sort. All



landscapers used weed status as a justification on their initial sort and, in most cases, on other sorts. For the initial sort, landscape utility, size, aesthetic quality, and native/nonnative status appear more frequently as justifications than morphological properties (morphological justifications are given at least once in the task by 8 of the 10 landscape members). In short, the landscape group used a broader variety of justifications than the other two groups. Eighty percent of the landscapers used taxonomy, morphology, weed status, aesthetic qualities, landscape utility, and size at some point in their sorting. Overall, justifications reveal different concerns driving the classifications. They also reveal that for landscapers and maintenance groups, classifications were not carried out by exhaustively sorting trees on a single basis. Rather, within a single round of sorting, different dimensions, as indexed by different types of justifications, were used side by side.

*Common clusters across expert groups.* So far, results indicate distinct taxonomies for the three groups of experts. To look for potential commonalities, we directly compared the aggregate taxonomies to determine the extent to which they overlapped (see Hunn, 1975). If all and only the members of a category in one taxonomy are the members of a category in the other taxonomy, those categories are in correspondence. For example, in Fig. 6, the maintenance workers have a cluster containing green ash and white ash. In Fig. 5, the landscapers also have a cluster containing all and only green ash and white ash. Thus, the two clusters correspond. Note that the correspondence need not be at the same level in the two taxonomies, nor does it require identical internal relations. For instance, another cluster corresponding between maintenance workers and landscapers is the pine–spruce–redcedar cluster (again see Figs. 5 and 6). However, the internal organization of that cluster differs between the two groups.

An exhaustive analysis of the three aggregate taxonomies revealed eight clusters, encompassing 22 trees (46% of the sample of trees) that corresponded perfectly across all groups. These clusters are listed in Table 2. Because these “common clusters” were derived from aggregate taxonomies, it is important to show that they reflect the categorizations of individual experts and they are not just an artifact of averaging. To address this question we counted the number of times each expert broke up a common cluster (using the criteria we described before). Results clearly are consistent with the aggregate analysis. Taxonomists literally never broke up a common cluster. Maintenance workers broke up these categories 17.6% of the time, and the landscapers broke up these categories 20.8% of the time. The latter two did not differ in the number of times they broke up common clusters,  $t(18) = 0.52$ , n.s. Apparently, the common clusters have some psychological reality on the level of individual experts.

A closer look at these common clusters reveals that they are not all cut out of the same folkbiological cloth. Three of the clusters appear to be “folk-generic” taxa in Berlin’s (1976; 1992) sense: the birch, ash, and oak clusters have the nomenclatural and taxonomic properties that Berlin outlines for folk-

TABLE 2  
Clusters Common to All Three Expert Groups

"Folk-generic" clusters	Other clusters
Green ash White ash	White poplar Weeping willow Eastern cottonwood
Paper birch River birch	Black walnut Shagbark hickory
Northern red oak White oak Bur oak Pin oak	Ohio buckeye Horsechestnut  American sycamore London planetree  Scotch pine Austrian pine Eastern white pine Colorado spruce Eastern redcedar

generic categories. Thus, these three might represent compelling discontinuities in nature: categories that are truly "beacons on the landscape of biological reality" (Berlin, 1992, p. 53). Patterns of folk nomenclature seem to disqualify the other clusters as folk-generics; all contain members named with multiple monolexemic labels. Thus, in Berlin's scheme they might qualify as "intermediate" taxa. If so, what is their basis?

Two of them, Ohio buckeye-horsechestnut and American sycamore-London planetree, correspond to the scientific genera and therefore may represent salient natural discontinuities which for some reason are unsupported by folk nomenclature. The pine-spruce-redcedar cluster is the most scientifically diverse group, but it also may represent a folkbiologically salient cluster by virtue of its taxonomic isolation. As is clear from Fig. 1, although the pines and Colorado spruce are not very closely related to eastern redcedar, all three are far from the rest of the trees in our sample. Indeed, needle-bearing coniferous trees are morphologically quite distinct from broadleaf deciduous trees and in general are much less common in the greater Chicago area. Thus, the pine-spruce-redcedar cluster might represent a salient group because despite having only moderate within-cluster similarity, it has great between-category distinctiveness (Hunn, 1975).

Finally, the two remaining clusters are composed of black walnut-shagbark

hickory and white poplar—weeping willow—eastern cottonwood. Both of these clusters span two genera within a single scientific family and thus might represent salient morphological clusters. However, members of both clusters also share salient functionally relevant properties that—albeit perhaps correlated or even causally tied to morphology—might well be the actual basis of classification for these clusters. Specifically, black walnut and shagbark hickory are both striking, stand-alone specimen trees, both have nuts, and both have valuable wood. In contrast, white poplar, weeping willow, and eastern cottonwood are frequently nominated as “weed trees.” They tend to have weak wood, to drop a great deal of twigs and other plant matter, and to be generally difficult to maintain. In short, it is not easy to pinpoint the basis of these clusters without looking at individuals’ justifications.

To tally these justifications, each individual’s taxonomy was examined. For each expert, the justification for the smallest grouping containing every tree in the common cluster was coded. If an expert was unfamiliar with some trees in a cluster but grouped all remaining trees from that cluster together, then a justification was counted (if the expert was familiar with at least half of the trees in the cluster, and the cluster contained more than two trees). If an expert cross-classified trees in a common cluster, no justification was recorded. Results are presented in Fig. 7.

Results for the oak, ash, and birch cluster support the notion that they are folk-generic categories. Overwhelmingly, these groupings were justified on purely taxonomic grounds; they were put together because they were “birches” or “the oak family.” The same was true for the conifer cluster, often named as “evergreens,” “conifers,” or “pines.” Note that this is not a necessary consequence of patterns of nomenclature: neither maples (Amur maple, Norway maple, sugar maple, silver maple) nor elms (American elm, Siberian elm) were classified with comparable unanimity.

Patterns were quite different for the other four clusters. For these clusters, morphological and landscaping-related justifications competed with taxonomic justifications. For the white poplar—eastern cottonwood—weeping willow cluster, “weediness” was the modal justification, whereas for the Ohio buckeye—horsechestnut cluster, morphological similarities won out over landscape use. For the other two clusters, there was no clear consensus. Modal justifications vary by expert group. For example, maintenance workers justified the American sycamore—London planetree cluster on morphological grounds (e.g., “they’re very similar”), landscapers explained the cluster in terms of landscaping usefulness (e.g., “compatible for use as city trees or specimen trees”), and taxonomists justified the same cluster on taxonomic grounds (e.g., “members of the sycamore family”). Thus, for these four common clusters, different groups of experts seem to pick up on different features of information available in the environment, but the correlation of those features leads them to form the same clusters.

In summary, despite a number of differences in the aggregate taxonomies

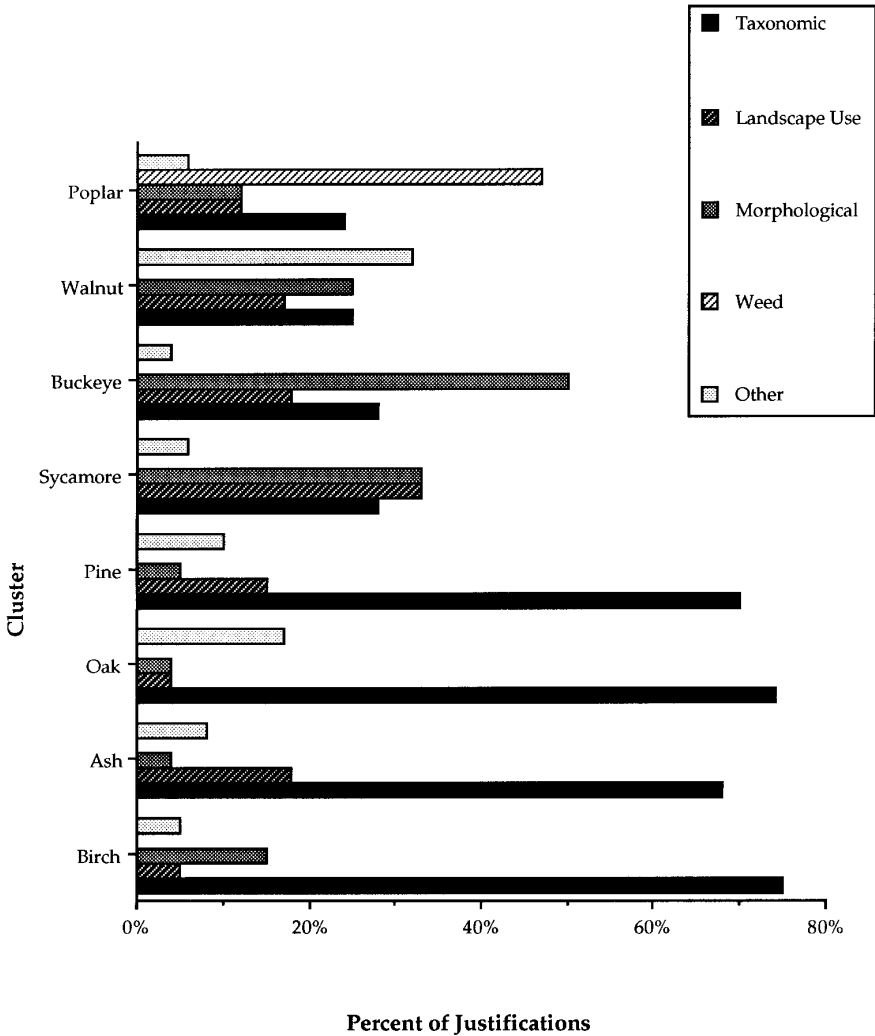


FIG. 7. Justifications for clusters common to all three expert groups.

of taxonomists, landscapers, and maintenance workers, we did find a subset of clusters of tree species for which there was a consensus in classification. For some of these clusters, consensus also exists with respect to justifications provided for the groupings. These may well represent salient discontinuities given by the world, and in some cases, perhaps classic folk generics. For other clusters, justifications differed as a function of type of expertise. This demonstrates that agreement in classification does not necessarily deny the impact of human goals and interests.

## DISCUSSION

We have presented so many results and analyses that perhaps it is best to begin by summarizing them. One factor aiding understanding is that the pattern of results across measures is remarkably convergent. We begin by looking at the differences that emerged across the three groups of experts.

The taxonomist group provides a useful comparative standard. Their sorting justifications are very highly correlated with scientific taxonomy. For ranks ranging from genus to family up to division, as taxonomic distance increased, mean distance in their consensual clustering increased. Furthermore, the taxonomist's sorting never broke up the genus level. Finally, justifications were almost exclusively in terms of taxonomy and morphology.

The maintenance group is similar in certain key respects to the taxonomists. Sorting and justifications were primarily in terms of taxonomic and morphological properties. The correlation between scientific taxonomic distance and folk distance was lower, in part because in several cases the morphological properties used in sorting are not necessarily properties given high weight in scientific taxonomy. For example, the maintenance consensual categories based on type of fruit or nut partially crosscut scientific taxonomy. The most salient noncorrespondence with science was that the ginkgo was placed with other broad-leafed, deciduous trees rather than with the other gymnosperms. Still, the overall correlation of folk taxonomy and scientific taxonomy was of a magnitude comparable to that observed by Lopez et al. (in press) in their studies of folkbiology of mammals. The maintenance group also differs from the taxonomist group in its inclusion of a clear utilitarian category, weed trees (fast-growing, weak-wooded trees that create maintenance problems). Although there was not a clear consensus among maintenance workers as to the members of this category, and thus it does not appear in the consensual cluster, 60% of maintenance workers had some grouping referred to as a "weed" or "garbage" category.

Where does the maintenance group fall with respect to the question of whether the genus level is psychologically privileged? The picture is mixed, but perhaps the best summary statement is that genus is relatively but not absolutely privileged. It is relatively privileged in that the genus level was more likely to be preserved and less likely to be broken up than the family level. In addition, there was a much larger jump in folk distance between the genus level and the family level than between the family level and higher taxonomic ranks. Maintenance group members appear to have broad shallow taxonomic structures where genus and family are the key ranks. Genus is not absolutely privileged. Even in the most favorable condition where folk terms and science converge on the genus level, the sorting revealed a corresponding category only about half the time. Furthermore, maintenance personnel commonly employed a weed-tree category on their initial sorts, right along with taxonomic and morphological categories.

Genus-level categories were more likely to be preserved when marked in folk terminology than when not marked. In addition, pseudomarkings appeared to exert an influence. Unfortunately, it is difficult to tease apart the causal force of this correlation. One is tempted to conclude that folk terminology affects categorization, but an equally plausible argument is that folk terminology *reflects* morphological similarity. That is, it may be that (for example) American mountain-ash is so-called because of its similarity to ashes (e.g., they both have compound leaves). One possible way to unconfound morphological similarity from folk terminology would be to look for regional differences in naming to see if they are associated with corresponding differences in sorting.

The differences between taxonomists and maintenance personnel find ready interpretation in terms of relative weighting of a shared set of dimensions. Both groups primarily focus on morphological properties. These are, however, weighted somewhat differently, one salient example being the maintenance group members classifying the ginkgo with other broad-leaf trees rather than with the other gymnosperms. The justifications for sortings reveal some group differences that can be tied to interest and goals. Membership in the “weed” category is directly linked to maintenance problems, but there are also morphological properties correlated with weediness (e.g., weak wood, brittle branches, etc.). Finally, the taxonomists had deeper hierarchies on average than the maintenance workers. We suspect that this reflects the greater concern on the part of science with worldwide distributions of plants and animals and associated evolutionary relationships. In general, scientific taxonomies have more levels than folk taxonomies (Berlin, 1992).

The pattern of sorting and justification for the landscape group showed the largest influence of utilitarian factors and the largest departure from scientific taxonomy. The correlation of landscapers’ folk distance with scientific distance was low, accounting for only about 14% of the variance. Genus-level categories were broken up over half the time, and there was a greater change in folk distance between the family level and higher level taxa than between the genus and family level. The justifications of landscape group members also reflect utilitarian concerns. Weed status, landscape utility, size, and aesthetic value were mentioned more frequently as justifications than morphological properties. The landscape group created considerably fewer categories on their initial sorting than did the other two groups, and they were also much more likely to provide a combination of descriptors as a justification (e.g., large, native, shade trees). Despite using fewer categories in their initial sort, landscape group members nonetheless created more superordinate ranks than did the maintenance personnel. In many cases superordinate categories were created by collapsing over one of a set of multiple descriptors. In short, relative to the maintenance group, the landscape group created a deeper but more narrow hierarchical structure.

It is hard to escape the impression that the landscape sorting is a utilitarian

structure imposed on nature rather than a reflection on categories salient in the environment. If membership in a utilitarian category depends on meeting multiple criteria, then one might expect the landscape group to show more cases where sorting justifications include more than one of type. Eight members of the landscape group, two of the maintenance group, and none of the taxonomists gave multiple justifications (e.g., “exotic, meritorious, landscape trees”) for at least one category on their initial sort. This pattern also suggests one explanation for why landscapers created more superordinate levels; they could simply collapse over a multidimensional distinction. For example, the categories “native street trees” and “nonnative street trees” suggests “street tree” as a superordinate category.

Overall, the sorting and justification profile for the landscape group corresponds remarkably well with Barsalou’s (1983) analysis of goal-derived categories. The sortings crosscut the correlational structure of the environment (as embodied in scientific taxonomy), were justified in terms of multiple utilitarian factors, and created a hierarchy based on combinations of goal-relevant factors. The modest agreement with science and with the other two groups of experts appears to derive from a combination of goals incompletely partitioning the set of trees and a convergence of morphological and utilitarian criteria on the same categories. Apparently years of experience lead the goal-derived categories of landscape workers to become well-established in memory. The next experiment addresses the question of whether these categories are also used in reasoning.

To avoid any ambiguity, we are not arguing that our sorting task reveals *the* way that the experts categorize trees. Doubtless, any of our experts could come up with alternative bases for categorization. For example, some members of our landscape group have had considerable formal training and they very likely could produce a taxonomic sorting corresponding closely to science. Instead, our claim is that the sorting data indicate one salient organizational scheme. We would have thought that our instructions (“put the trees together that go together by nature”) would have prejudiced the results in favor of taxonomic sorting. Therefore, the fact that sorting departed from science as far as it did (except for the taxonomists) is more than a little surprising.

Although we found little evidence that scientific genera are absolutely privileged categories for our tree experts, the aggregate taxonomies did reveal eight clusters common to all three expert groups. Five of these common clusters correspond to genera and two correspond to families. Although the justifications for some of these sortings differed across groups, the overall pattern is consistent with the idea that the genus level (and secondarily the family level) may be relatively but not absolutely privileged.

There are two reservations about our procedures and results that we wish to address. One is that the relative privilege of the genus level may simply

reflect the fact that, by chance, smaller categories are more likely to be preserved than larger categories. The most obvious counterargument is that category size and rank are only partially correlated; in our sample some of the families are smaller than other genera and yet these genera are more likely to be preserved. The differences are also qualitatively larger than one would expect by chance. The other main reservation is that our procedure forced participants to generate both mutually exclusive categories and a taxonomic hierarchy. This issue is best addressed by future research but we offer a few relevant comments here. First, none of the experts in this sample raised any concern about forming mutually exclusive groups. (This is in sharp contrast to tree experts who are ecologists, a group we are currently studying.) Second, very few of our results hinge on taxonomies per se; almost all are evident on the initial sorting (the depth of hierarchies being an obvious exception). We reserve additional comments for the general discussion.

*Summary.* Taken together, these comparisons reveal specific group commonalities and differences. There seem to be some clusters of trees that either are similar enough or convergent across goals to be put together regardless of type of expertise. Against this background, differences stand out as salient. Most strikingly, what were organized as taxonomic categories by maintenance and taxonomist groups appeared as goal-derived categories for the landscape group. This does not show that this is the only organization available for the landscape group (it surely is not), but it underlines the accessibility of a utilitarian organization. It remains to be seen whether the nontaxonomic bases for categorization carry over into other conceptual tasks. The second experiment addresses this issue.

## EXPERIMENT 2: REASONING ABOUT TREES

The second experiment provides a potentially converging measure of category organization in the form of a category-based reasoning task (Rips, 1975). In this experiment, we examine competing bases for extending novel properties from one tree species to another. Experts are told that a tree has some novel property and are then asked which of two alternative trees is most likely to have that property. Alternatives vary in their scientific (and folk) distance from the target. The question of interest is the degree to which different kinds of categories promote inferences about physiological properties. Will categories formed in the service of particular goals and interests reflect a deeper aspect of natural structure?

Several justifications for categories that crosscut science appeared frequently enough across experts (one might mention “shade trees” and another “indigenous desirable shade trees”) to allow us to consider them relatively stable categories. These categories were *ornamentals*, *street trees*, and *weed trees*. Ornamentals are generally smaller trees that have special decorative qualities rendering them useful in a landscape setting.<sup>2</sup> Street trees are gener-

<sup>2</sup> The most frequent ornamentals mentioned were Washington hawthorn, sweet crab apple, star magnolia, grey dogwood, Amur maple and American hornbeam.



ally strong, disease-resistant, low-maintenance trees that do not grow high enough to interfere with overhead wires.<sup>3</sup> The weed-tree category has been discussed above; this category was mentioned by 16 of the 24 experts. Members of this category based on six or more nominations (from the sorting justifications of the first experiment) include: white mulberry, tree of heaven, boxelder, eastern cottonwood, white poplar, Siberian elm, weeping willow, and catalpa. In general, the landscape group appears to consider more trees as weed trees than do the maintenance personnel.<sup>4</sup> Finally, we looked at categories that included the native versus nonnative distinction as part of their justification. Given that different experts modified this category in various ways (local versus nationwide versus foreign), the consensus was weak. The nominations were, however, almost without exception, accurate.

Using these salient folk categories, we attempt to pit scientific distance against membership in a folk category as a converging measure for the sorting results presented in Experiment 1. For example, one item might have boxelder for a target and the choices might be Norway maple (same scientific genus) and tree of heaven (fellow weed tree). If experts believe that categories such as *weeds* and *ornamentals* form coherent clusters, then they should use them to infer new properties, i.e., they might infer that boxelder would share a novel property with tree of heaven. This would suggest that the sorting data reflect deeper conceptual organization. Alternatively, if such categories only represent a few relevant shared features but fail to capture important underlying similarities, they might not promote inferences as well as scientific categories do; i.e., the expert might still prefer to draw an inference from boxelder to Norway maple.

In about half the cases, the scientific match to the target is in the same scientific genus as the target. In turn for half of these cases the scientific match and the target are marked by folk nomenclature as members of the same category (e.g., Norway maple and sugar maple), while in the other half of the trials folk nomenclature fails to convey this relation (e.g., Ohio buckeye and horsechestnut). By comparing trials in which the scientific match is or is not supported by folk nomenclature, we can assess the role of folk nomenclature on induction patterns.

<sup>3</sup> Street trees nominated by more than one person were green ash, white ash, honeylocust, Norway maple, and little-leaf linden.

<sup>4</sup> Justifications associated with Table 1 reveals both general agreement and subgroup disagreement. White mulberry, tree of heaven, boxelder, Eastern cottonwood, and (probably) white poplar are salient weed trees for both subgroups. Siberian elm and weeping willow are highly nominated by the landscape group but not the maintenance group. Weeping willow, paper birch, and black cherry are rarely used as street trees and maintenance personnel may have little interaction with them. A look at the initial descriptions of trees given by the experts suggests that the maintenance workers are well aware of problems associated with the Siberian elm, but apparently the taxonomic grouping (placing it with the American elm) was a more salient basis for categorization.

We say that the reasoning task provides *potential* converging measure because there is no reason to think that experts are limited to a single form of category organization (see also Heit and Rubinstein, 1994). Indeed, we used three distinct types of so-called “blank” properties to see if different kinds of properties evoked different forms of category organization. Also of interest is the idea that the genus level may be psychologically privileged (e.g., Coley et al., submitted). If so, then when the target and one alternative are of the same genus, that alternative should be selected to the exclusion of other forms of category organization. In addition, taxonomic choices should favor the genus level over higher levels more than a family or order match is chosen over correspondingly higher taxonomic levels.

If results converge with those of Experiment 1, taxonomists' responses should closely follow scientific relations, whereas those of landscape and maintenance personnel should not necessarily do so. Furthermore, the maintenance experts' inferences should be more likely to follow scientific taxonomy at the genus level than at higher levels. Finally, landscapers who deviated most from scientific taxonomy in sorting should be most likely to use folk categories to promote inferences.

## METHODS

### *Participants*

The participants were a subset of the tree experts used in the first experiment. Specifically, the reasoning task was given to 3 of the 4 taxonomists, 8 of the 10 landscapers, and all 10 maintenance workers. Selection was based on availability, and there is no reason to think that the sample for the second study is in any way unrepresentative of the population used in the first experiment.

### *Design and Materials*

The triads used in the reasoning task (available upon request) were constructed midway through the first experiment and some of the anticipated contrasts were not realized when the full set of data became available. Nonetheless, the 43 tests provide a range of pairings pitting taxonomic distance (at the genus level and at higher levels) against native status, weediness, and landscape utility. The remaining tests sampled a range of taxonomic distance but did not pit taxonomic distance against any consensual salient nontaxonomic category. The overall aim was to use a broad enough set of tests to set up at least some contrasts between folk taxonomic and scientific distance. For landscapers this proved to be easy; for maintenance personnel we were only partially successful because their sortings mirrored scientific taxonomy more closely than those of landscapers.

### *Procedure*

Experts were run individually. The general procedure was to present triads of trees on cards where one tree (at the top of the card) was said to have

some property, and then participants were asked to judge which of the other two trees would be most likely to share that property. Participants were told that there were no right or wrong answers, but rather that we were interested in how they thought about the trees.

The 43 triads were presented one at a time in a random order for each of the three types of properties to yield a total of 129 trials. Properties were chosen to be relatively unfamiliar but nevertheless to plausibly promote inferences from one tree species to another. The triads were blocked by property but property order was randomized for each participant. The instructions were as follows for the reproduction, disease, and physiological property reasoning tasks:

(1) Reproduction: “Say we discovered a new kind of tree that could be hybridized with [target tree]. If this new tree could also be hybridized with other trees, is it more likely that the new tree would be hybridized with [tree A] or with [tree B]?”

(2) Disease: “Say we discovered a new disease that infected [target tree]. If this disease could also infect other trees, is it more likely that the new disease would infect [tree A] or [tree B]?”

(3) Physiology: “Say we discovered a new enzyme inside the cells of [target tree]. If this new enzyme were found inside the cells of other trees, would it be more likely to be found in the cells of [tree A] or in the cells of [tree B].”

Participants were invited (but not required) to explain their answers. The entire task took about 30 minutes to run.

## RESULTS

The reasoning data may be analyzed at a number of levels of generality. First, choices did not vary with type of property; generally the option selected for one property (e.g., disease) was also selected for the other two properties. Next, consider agreement with scientific versus folk taxonomic distance. Breaking responses down by subgroups, the mean proportion of scientific matches were 94.7, 84.0, and 79.3% for taxonomists, landscapers, and maintenance workers, respectively (the overall mean was 82.8%). Agreement with folk-taxonomic distance was 65.5% for the landscapers compared with 82.9% for the maintenance workers. Overall then, inductions of the landscapers agreed more with science than with their own consensual distance, but maintenance personnel showed the opposite pattern. As we shall see, these effects and interactions are reliable.

### *Scientific versus Folk-Consensual Distance*

For the maintenance and landscape subgroups one can directly contrast scientific taxonomic distance with distance in the consensual folk taxonomy (see Figs. 5 and 6). Specifically, for landscapers there were 54 items (18 triads  $\times$  3 properties) that directly contrasted scientific versus folk-consensual

distance; for the maintenance workers, there were 12 such items (4 triads  $\times$  3 properties). For each group, we assessed whether that group showed a reliable preference for science versus folk consensus as a basis for induction (i.e., whether the proportion of responses favoring folk consensus differed reliably from .50). On average, landscapers based only 28.5% of inductions on their consensual sorting when it conflicted with science; in other words, their inductions were better predicted by scientific taxonomy than by their own consensual sorting ( $t(7) = 6.248, p < .001$ ). Interestingly, the landscapers inductions were slightly better (but not reliably) predicted by the *maintenance* consensual sorting than by scientific taxonomy. In contrast, maintenance workers reliably gave responses on conflicting trials consistent with folk consensus ( $M = 75.8\%$ ,  $t(9) = 6.146, p < .001$ ). These data support the generalization that maintenance workers use a single (folk) taxonomy for sorting and reasoning whereas landscapers use one conceptual organization for sorting but shift to another for reasoning.

### *Taxonomic Level*

The above generalization concerning landscapers must be qualified by an effect of taxonomic level. For slightly more than half of the test items the scientific match was at the genus level. On these items landscapers overwhelmingly preferred the scientific match (96.7%). For levels above genus, overall agreement with science was significantly lower ( $M = 67.1\%$ ,  $t(7) = 7.983, p < .001$ ), albeit still above chance ( $t(7) = 3.651, p < .01$ ). Furthermore, landscapers were much less likely to choose scientific matches above the genus level if they conflicted with their folk taxonomy ( $M = 47.2\%$ ) than if science agreed with their folk taxonomy ( $M = 90.6\%$ ,  $t(7) = 7.811, p < .001$ ). In fact, when in conflict with folk taxonomy, choice levels for above-genus scientific matches did not differ from chance ( $t(7) = 0.493, n.s.$ ). Overall, landscapers followed scientific distance over their own consensual distance at the genus level; at higher levels, the conflict between science and consensual taxonomy led to chance performance. Thus, there is a suggestion that folk taxonomy had some influence on inferences when scientific matches were above the generic level. The greater agreement with scientific than with folk distance for the landscape group was confined to triads where the science option matched the target in genus.

On three of the eight trials in which the scientific match was above the genus level and conflicted with the folk match, landscapers reliably differed from chance. On one of these items the folk match seems to be quite salient. That is, when the target was Norway maple, landscapers chose little-leaf linden (83% of the time) rather than the scientifically closer horsechestnut, likely because Norway maple and little-leaf linden are frequently used street trees. On another item they chose little-leaf linden to go with ginkgo (again, 83% of the time) rather than the scientifically closer Scotch pine. Presumably this was because both ginkgo and little-leaf linden share the salient morpho-

logical property of being broad-leaved rather than needle-leaved. Finally on the third item that differed from chance, landscapers thought a property true of the black cherry would also be true of scientifically related sweet crab apple, rather than Ohio buckeye. Though this choice corresponds to science, it could also be due to utilitarian concerns: both sweet crab apple and black cherry are small flowering trees that can be used as ornamentals. Though the remaining five items on which scientific and folk matches conflicted were at chance, the fact that there were so many nonscience choices indicates that the folk taxonomy had some influence on inferences.

It is not possible to compare the influence of genus-level scientific matches with folk matches for the maintenance subgroup, because for all triads with a genus match scientific and folk distance were in agreement. (Note that this also means that the landscape versus science contrasts could as well be thought of as landscape-consensual sorting versus maintenance-consensual sorting.) For items above the genus level, overall agreement with scientific distance was still above chance ( $M = 59.7\%$ ,  $t(9) = 2.952$ ,  $p < .05$ ). Nevertheless, maintenance workers were much less likely to choose the scientific match when it conflicted with their folk taxonomy ( $M = 24.2\%$ ) than when it agreed with folk taxonomy ( $M = 71.5\%$ ,  $t(9) = 8.243$ ,  $p < .001$ ). When science and folk agreed, scientific responses were above chance levels, ( $t(9) = 4.349$ ,  $p < .005$ ); as mentioned before, however, when science and folk taxonomy disagreed, *folk-taxonomy based* responses were above chance levels. Finally, for five items where folk distance was equal for both choices, responses were above chance favoring the scientific alternative ( $M = 63.3\%$ ,  $t(9) = 2.481$ ,  $p < .05$ ). Overall, maintenance workers' reasoning was better predicted by their folk taxonomy than by scientific distance.

As a further analysis of taxonomic level, we looked at agreement with science when the closer alternative matched the target at the level of genus, family, and order, and the other option matched at two levels higher (order, subclass, and class, respectively). These results are summarized in Table 3, broken down by subgroups. Level differences were striking for the maintenance and landscape subgroups. An analysis of variance for the maintenance and landscape subgroups yielded a highly reliable effect of rank ( $F(2,32) = 39.038$ ,  $MSe = 0.122$ ,  $p < .001$ ), and Tukey HSD analyses reveal that genus matches yielded greater agreement with science than family matches, which in turn yielded greater agreement with scientific taxonomy than order matches ( $p < .01$ ). Moreover, planned comparisons reveal that landscapers were more likely than maintenance workers to agree with science when the scientific match was at the family level,  $F(1,16) = 5.622$ ,  $MSe = 0.185$ ,  $p < .05$ . There were no group differences at the genus or order level. In brief, the genus level appears to be relatively (and absolutely) privileged on the reasoning tests.

### *Genus Marking*

We noted in the first study that genera marked by folk terminology were more likely to appear intact in sorting than genera not marked in folk terminol-

TABLE 3  
Proportion of Scientific Choices at Different Scientific Ranks, by Group

Expert group	Genus	Family	Order
Maintenance workers	89.2%	64.4%	58.9%
Landscapers	96.4%	80.6%	57.4%

ogy. When one looks at the subset of triads where the distance of the alternative (nonscience) is held constant, a similar pattern appears in reasoning. For the maintenance subgroup the genus match is selected 97.9% of the time when marked in language versus 83.3% of the time when it is not marked ( $t(9) = 2.912, p < .05$ ). The corresponding figures for the landscape subgroup are 98.4% and 93.2%, respectively ( $t(7) = 1.616, n.s.$ ). By far the largest contributor to this marking effect comes from two triads involving the maple not marked in common name, the boxelder, where agreement with science was 55% for maintenance workers and 81.2% for the landscapers. In brief there is a clear association between marking and reasoning for the maintenance group, and a similar but weaker trend for the landscapers. Note that the data do not permit inferences concerning whether marking directly affects reasoning; alternatively, morphological similarity may affect both folk nomenclature and patterns of reasoning.

### *Special-Purpose Folk Categories*

*Native status.* Although our intention was to assess effects of weed status, native status, and landscape function on reasoning, in practice the task is complex because it is difficult to rule out alternative covarying factors. Consider the tests where the alternative to the scientific match shared native or nonnative status with the target. There were strong departures from science on three of these tests (for each triad, the first tree is the target, the second is more closely related to the target scientifically, and the third matches the target with respect to native status): (1) northern red oak, European black alder, tuliptree; (2) tree of heaven, silver maple, white mulberry; (3) ginkgo, eastern redcedar, European black alder. Agreement with science was only 66.3, 20.0, and 13.3%, respectively, for maintenance workers, and 41.7, 33.3, and 58.3%, respectively, for landscapers. Maintenance workers as a group did not differ from chance for item (1), but chose the nonscience alternative at above-chance levels for items (2) and (3),  $t_s > 2.946, p_s < .02$ . Landscapers as a group did not differ from chance for any item ( $t_s < 1.528, n.s.$ ), but nonetheless show a relatively large percentage of nonscientific choices. These patterns seem to suggest that native status exerted a considerable influence on reasoning, but we very much doubt that that is the case. One expert justified his inference from northern red oak to tuliptree over European black alder

by noting that both trees are tall (unlike the alder). Similarly, although tree of heaven, silver maple, and white mulberry are all considered weed trees, the first and third are more frequently described as weeds than silver maple; thus, weediness rather than native status could be driving responses on that item. Finally, for ginkgo, European black alder is likely preferred over eastern redcedar because the former two are both broadleaf deciduous trees (the same preference is also observed when ginkgo is the target and Scotch pine and little-leaf linden are alternatives—only 7.4% of parks group choices and 16.7% of landscape group choices favored the taxonomically nearer Scotch pine). Thus, covarying factors that are seemingly independent of native status may be driving response patterns.

*Weed trees.* We have already mentioned some of the results concerning the weed category. Overall, maintenance workers and landscapers did not differ in mean proportion of nonscientific (weed-based) responses ( $M = 14.7$  and 10.4%, respectively). To further explore the use of this category in induction, we compared how often participants drew inferences on the basis of weed status as a function of the nature of the scientific match (marked genus match, unmarked genus match, higher match). For maintenance workers, there was a clear effect of the nature of the scientific match,  $F(2,18) = 7.221$ ,  $p < .005$ . Based on Scheffe's post hoc test, maintenance workers were much more likely to make weed-category-based inferences for unmarked genus match items (i.e., those involving boxelder,  $M = 45.0\%$ ) than for marked genus-match items ( $M = 1.7\%$ ,  $p < .05$ ). Response levels were intermediate for higher matches ( $M = 23.3\%$ ), differing from neither of the first two. Thus, when genus is marked choices overwhelmingly favor the scientific (and folk) match; when it is not marked the choices are more equivocal, especially for nonmarked genus items (i.e., those involving boxelder, a classic weed tree).

A parallel analysis for landscapers also revealed differences as a function of item type ( $F(2,14) = 4.809$ ,  $p < .05$ ), but the pattern did not match that of the maintenance workers. For landscapers, choice of weed matches was reliably higher for matches above the genus ( $M = 29.2\%$ ) than for marked genus matches ( $M = 1.4\%$ ,  $p < .05$ ). Weed matches for nonmarked genus items fell in-between ( $M = 18.8\%$ ). For the landscape group the departure from science was particularly evident for the triad involving two prototypical garbage trees, white mulberry and eastern cottonwood. Here 54.2% of choices favored the nonscientific match. In short there is some evidence that weed status influenced reasoning as long as a marked genus match is not available. Moreover, this influence is evident on different items for landscapers versus maintenance workers.

*Landscape utility.* Finally, groupings of trees based on functional landscape utility may have influenced reasoning. Landscapers and maintenance workers did not differ in overall proportion of nonscientific choices based on landscaping usefulness ( $M = 46.7$  and 43.7%, respectively). These numbers derive primarily from trials where the scientifically closer

match was above the genus level. (Note also that they are not reliably above chance.) Only one item pitted scientific distance at the genus level: mean nonscientific choices for this item were 4.2% for landscapers and 6.7% for maintenance workers.

As in the case of native status, the actual bases of inductions on landscape utility trials are somewhat ambiguous. Specifically, morphological dimensions that may ultimately determine landscape utility—rather than landscape utility per se—may drive inferences. For example, on one item, the landscape choice, sweet crab apple, and the target, American hornbeam, are both small ornamentals, unlike the scientifically closer-related white oak. However, reasoning on the basis of height alone would also lead to an inference from American hornbeam to sweet crab apple. Height is certainly one factor in whether a tree is considered a “small ornamental,” but it is not the only consideration. In short, because landscape function is based on morphological characteristics it is difficult to tease apart the contribution to reasoning of these characteristics from landscape utility per se.

Overall, clear influences of folk taxonomy, linguistic markings, taxonomic level, and at least some special-purpose folk categories emerged. Furthermore, landscapers and maintenance workers differed in their reasoning patterns, and in the relations between reasoning patterns and sorting. We take up these findings below.

## DISCUSSION

Results of the reasoning experiment support two main conclusions. First, genus-level categories are privileged for induction. Second, patterns of reasoning, and the relations between reasoning and sorting, differed across subgroups. Participants' inferences overwhelmingly agreed with science when the scientific match was at the genus level. The sole exception to this trend was that levels of scientific responses dropped to 55% for maintenance workers for the unmarked-genus trials on which boxelder was the target. Almost without exception, the influence of folk-consensual taxonomy was apparent when contrasted with scientific matches at ranks superordinate to genus. Clearly, on this task, the genus level was privileged for induction. This fits with recent findings by Coley et al. (submitted) that folk-generic categories (roughly equivalent to marked genus categories in the present study) are privileged for induction.

Despite the unanimity with which inferences agreed with science at the genus level, clear differences emerged between subgroups. For both the maintenance and taxonomist subgroups one could say that the same category organization revealed by sorting carries over into reasoning. For the taxonomists this is essentially equivalent to scientific taxonomy. Some of their justifications were in terms of known patterns of disease but these are highly correlated with taxonomic distance.



For maintenance workers, the taxonomy guiding reasoning appears to be their consensual folk taxonomy. When folk and scientific taxonomy made different predictions, maintenance workers were more likely than landscapers to reason in accord with their own consensual taxonomy. When folk and scientific taxonomy conflicted and the scientific relation was above the genus level, maintenance worker responses were consistent with consensual folk taxonomy as a basis for inferences. The correlation of morphological characters with utilitarian function makes it difficult to pinpoint the exact basis for reasoning, but the data strongly suggest that at least weed status, and perhaps landscape utility, influenced choices. Furthermore, folk nomenclature also appeared to have an effect on reasoning patterns among maintenance workers. Overall, maintenance workers' reasoning patterns were more or less in accord with their consensual sorting.

In sharp contrast to maintenance workers, landscapers used a reasoning strategy that departed substantially from their consensual sorting. Overall, landscapers were more likely than maintenance workers to make inferences that agreed with scientific taxonomic distance. In sorting, landscapers were most likely to cross-classify genus-level groups of trees; in reasoning, landscapers almost never drew inferences on the basis of folk categories when the scientific choice was at the genus level. As mentioned above, when folk and scientific distance conflicted, landscapers were more likely than maintenance workers to base inferences on science. We hasten to point out that the landscape reasoning data agreed more with maintenance consensual sorting than with either their own consensual sorting or with science. The fact that maintenance sorting and science predicted landscaper reasoning about equally well makes the basis for landscape reasoning ambiguous. Folk nomenclature had no reliable effect on patterns of reasoning for landscapers. The landscape folk categories *weed* and *ornamental* had some effect on landscapers' reasoning, but only at levels above the genus. In short, the folk categories of the landscapers as derived from the sorting task were less inductively powerful than the folk categories of the maintenance personnel.

Overall, results support the claim that genus-level categories are inductively privileged. Furthermore, distinct types of expertise again led to distinct performances on the task. Specifically, taxonomists and maintenance workers' inferences were predicted by their consensual sortings, whereas landscapers' inferences were not. These patterns of match and mismatch between sorting and reasoning suggest unanticipated complexities, which we explore next.

## GENERAL DISCUSSION

To what degree do folkbiological conceptual systems reflect universal patterns of feature covariation in the world, or universal habits of mind, and to what degree do they reflect specific goals and needs of the categorizer? There is no simple answer to that question but the complexities are themselves intriguing. We began by considering the role of goals, theories, and belief

systems in categorization as manifested by distinct occupational activities with respect to the folkbiological kind, *tree*. With the domain held constant, any differences in categorization and reasoning related to type of expertise reflect nonuniversal contributions of the mind to the understanding of the biological world. Similarities, in contrast, suggest universal tendencies in the structure of mind and/or world. Our results reveal systematic similarities *and* differences, which we will briefly review before discussing their implications.

### *Profiles of Expert Groups*

*Taxonomists.* The taxonomist subgroup essentially reproduced classical scientific taxonomy. Taxonomists tended to produce taxonomies that were both broad and deep. Their consensual clustering was highly correlated with science (+.85), reflecting scientific grouping at higher ranks (order, class) as well as lower ranks (genus, family). Likewise, their inferences were almost always based on science, showing a great deal of consistency with their sorting. This is perhaps not surprising, given their formal training in scientific taxonomy.

*Maintenance personnel.* In general, maintenance workers' taxonomies were broad but shallow, with few higher-order groupings. They showed a moderate but lower correlation with science deriving from both the use of the utilitarian category, *weed trees*, and from the use of distinctive morphological features that are given less weight by science. Departures from science reflect, in part, different patterns of weighting of morphological features. A striking example of the latter is that of the ginkgo, which is a broad-leafed, deciduous gymnosperm that was not grouped with other (needled, evergreen) gymnosperms. For maintenance workers, sorting and reasoning data are in strong concordance with each other. In both tasks agreement with scientific taxonomy was maximized at the genus level. Maintenance workers also appeared to be influenced by patterns of folk nomenclature.

*Landscape personnel.* Finally, the landscape subgroup's sorting displayed only a weak correlation with science, showing much less respect for genus-level groupings than other experts, and included a number of utilitarian categories such as *weed trees*, *ornamentals*, *specimens*, and *street trees*. Their sorting was less a reflection of nature than an imposition of utilitarian categories onto it. Taxonomies tended to be narrow but deep, with groupings at a given level based on a variety of different considerations. Landscaper's reasoning, in contrast, agreed much more with scientific taxonomy than their own consensual folk taxonomy, and this was especially true when there was a genus match available. Landscapers appear to use one category organization (goal-derived and utilitarian) for sorting and a different one for reasoning. One candidate for the alternative categorization scheme is the maintenance group consensual taxonomy—the reasoning data of the landscape group were actually slightly more in agreement with maintenance folk distance than with scientific distance.

*Explanations for Similarities and Differences among Experts*

*Common groupings across experts.* Two pieces of evidence support the observed partial consensus across groups: the finding of absolute inductive privilege of genus-level categories, and the finding that certain clusters of trees were preserved across all groups of experts. This latter result is somewhat qualified by the fact that for some of the common clusters, experts gave justifications that differed systematically by group. One possible explanation for similarities among experts is that nature occurs in strikingly discontinuous bundles that impose themselves on human cognition. Alternatively, it might be that similarities are due to universal cognitive tendencies rather than structure in the world per se. A third idea is that language in the form of genus marking may play an important role in consensual sorting and reasoning.

Experiment 2 demonstrates that both scientific and folk taxonomies can be used to predict category-based reasoning. When triads contained a scientific match at the genus level, inferences almost always agreed with scientific taxonomy. At higher levels, folk taxonomy had a moderate effect for landscapers, and a strong influence on maintenance workers. This finding of absolute privilege of the genus level in induction is consistent with the findings of Coley et al. (submitted) and also suggests some universal structure in folkbiological thought. Genus categories may capture salient “general-purpose” discontinuities in nature that support inductive inferences (Berlin, 1992). Alternatively, genera may cohere due to the presumption of a shared essence (Atran, 1990), which itself may depend on linguistic marking of kinds and genus-level categories. In any case, folk generics seem to form coherent and inductively useful categories regardless of one’s goals or interests.

Experiment 1 revealed that approximately half of the trees were classified into clusters common to all three groups. For the most part, these were genus- or family-level groupings, some of which were supported by folk nomenclature (ash, oak) and some of which were not (Ohio buckeye–horsechestnut, weeping willow–white poplar–eastern cottonwood). It is tempting to conclude that these common clusters represent salient discontinuities in the natural world; however, implications of this finding are not so straightforward. In the case of the three “classic” folk-generic groupings (oak, ash, birch), morphological similarities, scientific relations, and folk naming patterns all seem to conspire to highlight these classes. Agreement on these classes could be attributed to structure in the world, or to a common tendency for the human mind to form classes based on such information, or even to common goals with respect to these subsets. Another possibility is that goals influence categorization decisions, but the correlated structure of the world often results in agreement about the content of categories despite the use of different criteria to arrive at those categories. Indeed, for several common clusters, experts differed systematically in the justifications they gave for those clusters. For these clusters, it appears that agreement reflects different

experts relying on different criteria—based on different goals, and priorities—but agreeing because of the nonrandom, correlated nature of the relevant features.

Of course these common patterns of sorting must be considered against the background of differences. For the other half of the trees consensual groups did not emerge, even when the genus level was marked by language (as in the case of elms). Goals and characteristic activities do have a role to play and theories of categorization must address these expertise-related differences.

*Structural differences in taxonomies.* One difference between the groups was in the shape of their taxonomies, a difference that may reflect the different purposes for which each group classifies trees. Maintenance workers' taxonomies tended to be shallow, with a large number of initial classes justified on the basis of morphological and sometimes utilitarian concerns. This is what one would expect from a classic folk taxonomy within a single life form, given Berlin's (1976, 1992) characterization of folkbiological classification. Taxonomists, like maintenance workers, had a relatively large number of initial classes, but their taxonomies had many more levels, especially superordinate levels, which were justified predominantly in taxonomic terms. This reflects the taxonomist's quest for evolutionarily significant similarities across broad classes of plants, a concern that is presumably irrelevant to maintenance workers. Finally, the landscapers' taxonomies present yet a third profile. They tended to be narrow, with fewer initial classes, but deep, with relatively many higher- and lower-order groups. Landscapers also used the widest variety of justifications, including many conjunctive justifications ("large native-specimen trees"). These findings are consistent with a "taxonomy" based on a number of semi-independent utilitarian dimensions being imposed over the set of trees. We return to this observation below in a comparison of the landscapers' profile to Barsalou's account of goal-derived categories. Overall, differences in taxonomies may well represent one way in which the goals of each expert group influence categorization. How might one address these differences?

*Changes in feature weighting.* One potential explanation for group differences proposed in the introduction was that classification is based on the same set of features, but with differential weighting given those features by different groups. Differences between taxonomists and maintenance workers may be captured within the framework of such similarity-based models of categorization, provided that one allows goals and functions to infuse, stretch, and shrink that similarity space (see Goldstone, 1994). In particular, taxonomists closely mirrored science; maintenance workers' departures from science, in part, reflect different patterns of weighting morphological features. But they also reflect some contribution of utilitarian concerns, such as weed status. These concerns partition the set of trees only incompletely, and moreover, they are correlated with morphological properties such that sorting

consists of an integration of morphological and functional considerations. In short, maintenance experts can be characterized in terms of heavy emphasis on morphological properties seamlessly integrated with moderate attention to goal-related utilitarian factors. The same features, weighted differently (e.g., maintenance workers tend to weight features like wood strength and leaf shape more heavily than do taxonomists), appear to drive the classification systems of both taxonomists and maintenance workers. Likewise, both groups “stuck to their taxonomic guns” in reasoning, drawing inferences which, for the most part, agreed with their sortings.

The landscapers, in contrast, violated the correlational structure of the stimulus set to such a degree that structure-based accounts would be severely strained. No stretching or shrinking of a common space would transform their consensual clustering to conform to that of either of the other two groups (save for the degenerate case of a set of dimensions with zero weighting for various subgroups).

*Goal-derived categories.* A final possibility considered in the introduction was that essentially goal-derived categories might lead to radical disagreement among experts. For landscapers this account fits the classification data quite well; function was a much more salient consideration for them than for the other two groups. Barsalou’s analysis of goal-derived categories and his conjecture that frequently used goal-derived categories may become well established in long-term memory (Barsalou, 1982, 1983, 1991) predict the pattern of landscape sorting in striking detail. The various landscape desiderata comprise a set of criteria that various species do or do not fall into and these criteria crosscut the correlational structure that scientific taxonomy presumably captures. Categories correspond to different goal structures where the goals dictate the categories. Consequently, sorts have only weak correlation with scientific taxonomy, weak enough that it may derive from goals failing to completely partition the set of trees or from correlations of morphology with utility. In short, landscape expertise leads to a highly accessible category organization that concedes little to nature.

Indeed, the mismatch of landscapers’ sorting and reasoning data provides further support for the above analysis. Assume that the landscapers’ consensual taxonomy reflects goal-derived, special-purpose categories rather than general-purpose categories expected to maximally capture similarities in the world. If so, there is no reason to expect these goal-derived categories to support inferences about any properties except those related to the goals. The reproductive, disease, and physiological properties used in Experiment 2 have little relation to the functional, utilitarian goals that seemed to drive the landscapers’ sortings from Experiment 1. Therefore, landscapers sensibly abandoned their salient but inductively less useful (in this case, at least) consensual taxonomy in favor of a general-purpose taxonomy (either science or an organization approximating that of the maintenance workers).

Follow-up research is needed to assess the idea that landscapers (and per-

haps the other groups as well) have multiple forms of category organization available to them. So far we have not established reliability of sorting across occasions. One clear prediction from Barsalou's framework is that landscapers should show less test-retest reliability than the other subgroups because multiple utilitarian criteria do not constitute a true hierarchy and can be applied in any order. Our preliminary observations from tests underway are consistent with this prediction.

### Conclusions

It is unrealistic to expect that questions at the broad level of "Do all roads lead to Rome" will have a simple yes or no answer. Rather than taking a middle-ground interactionist position, we have presented a detailed study revealing some aspects of folkbiological reasoning in which all roads do appear to lead to Rome, and other aspects for which some roads might lead to Athens or Alexandria instead. With respect to the former, groups of tree experts with different backgrounds, training, and goals show agreement on classification of at least a subset of trees, and on patterns of inductive inferences to closely related trees. On the other hand, we also found systematic, expert-group-based differences in the structure of taxonomies, changes in feature weighting, and convincing evidence of specific goal-derived categories.

Advocates of the view that variable goals, theories, and belief systems lead to variable patterns of conceptual organization can take encouragement from our results. Each group of experts showed a profile distinct from the other two groups. At the same time, to identify a variable role for goals and theories in biological categorization is not to completely undermine the universalist hypothesis. The universalist position is supported by the reasoning data that reinforce the idea that the genus level is psychologically privileged. A priori domain-specific forms of mental organization may not determine a single folkbiological structure but may tend to privilege certain types of categories and category-based reasoning over others. The present studies constitute only a small piece of the puzzle. However, they do license the conclusion that though not all roads lead to Rome, in many cases they may lean toward Rome.

### REFERENCES

- Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.
- Atran, S. (1987). Ordinary constraints on the semantics of living kinds: A commonsense alternative to recent treatments of natural-object terms. *Mind and Language*, **2**, 27–63.
- Atran, S. (1990). *Cognitive foundations of natural history*. New York: Cambridge University Press.
- Atran, S. (1994). Core domains versus scientific theories: Evidence from systematics and Itza-Maya folkbiology. In L. Hirschfeld & S. Gelman (Eds.), *Domain specificity in cognition and culture* (pp. 316–340). New York: Cambridge University Press.
- Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory and Cognition*, **10**, 82–93.
- Barsalou, L. W. (1983). Ad-hoc categories. *Memory and Cognition*, **11**, 211–227.
- Barsalou, L. W. (1985). Ideas, central tendency, and frequency of instantiation as determinants

- of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **11**, 629–649.
- Barsalou, L. W. (1991). Deriving categories to achieve goals. In G. H. Bower (Ed.), *The psychology of learning and motivation: Vol. 27. Advances in research and theory*. New York: Academic Press.
- Barsalou, L. W. & Ross, B. H. (1986). The role of automatic and strategic processing in sensitivity to superordinate and property frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **12**, 116–134.
- Bartlett, H. (1940). History of the generic concept in botany. *Bulletin of the Torrey Botanical Club*, **67**, 349–362.
- Berlin, B. (1972). Speculations on the growth of ethnobotanical nomenclature. *Language in Society*, **1**, 51–86.
- Berlin, B. (1976). The concept of rank in ethnobiological classification: Some evidence from Aguaruna Folk Botany. *American Ethnologist*, **3**, 381–399.
- Berlin, B. (1992). *Ethnobiological classification: Principles of categorization of plants and animals in traditional societies*. Princeton, NJ: Princeton University Press.
- Boster, J. S. (1986). Exchange of varieties and information between aguaruna manioc cultivators. *American Anthropologist*, **88**, 429–436.
- Boster, J. S. (1987). Agreement between biological classification systems is not dependent on cultural transmission. *American Anthropologist*, **89**, 914–919.
- Boster, J. S., & D'Andrade, R. (1989). Natural and human sources of cross-cultural agreement in ornithological classification. *American Anthropologist*, **88**, 569–583.
- Boster, J. S., & Johnson, J. C. (1989). Form or function: A comparison of expert and novice judgments of similarity among fish. *American Anthropologist*, **91**, 866–889.
- Coley, J. D. (1995). Emerging differentiation of folkbiology and folkpsychology: Attributions of biological and psychological properties to living things. *Child Development*, **66**, 1856–1874.
- Coley, J. D., Medin, D. L., & Atran, S. *Does privilege have its rank? Folkbiological taxonomy and induction in two cultures*. Manuscript submitted for publication.
- Gelman, S. A., & Coley, J. D. (1991). Language and categorization: The acquisition of natural kind terms. In J. P. Byrnes & S. A. Gelman (Eds.), *Perspectives on language and thought: Interrelations in development* (pp. 146–196). Cambridge: Cambridge University Press.
- Gelman, S. A., Coley, J. D., & Gottfried, G. M. (1994). Essentialist beliefs in children: The acquisition of concepts and theories. In L. A. Hirschfeld and S. A. Gelman (Eds.), *Mapping the Mind: Domain Specificity in Cognition and Culture*. New York: Cambridge Univ. Press.
- Gelman, S. A., & Wellman, H. M. (1991). Insides and essences: Early understandings of the non-obvious. *Cognition*, **38**, 213–244.
- Gilmour, J. (1940). In J. Huxley (Ed.), *The new systematics*. Oxford: Clarendon Press.
- Goldstone, R. L. (1994). The role of similarity in categorization: Providing a groundwork. *Cognition*, **52**, 83–162.
- Greene, E. L. (1983 [1909]). In F. N. Egerton (Ed.), *Landmarks in botanical history*. Stanford, CA: Stanford University Press.
- Heit, E., & Rubinstein, J. (1994). Similarity and property effects in inductive reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **20**, 411–422.
- Hubert, L., & Schultz, J. (1976). Quadratic assignment as a general data analysis strategy. *British Journal of Mathematical and Statistical Psychology*, **29**, 190–241.
- Hunn, E. (1975). A measure of the degree of correspondence of folk to scientific biological classification. *American Ethnologist*, **3**, 309–327.
- Hunn, E. (1977). *Tzeltal folk zoology: The classification of discontinuities in nature*. New York: Academic Press.
- Johnson, K. E., Mervis, C. B., & Boster, J. S. (1992). Developmental changes within the structure of the mammal domain. *Developmental Psychology*, **28**, 74–83.
- Keil, F. C. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.

- Landau, G., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, **3**, 229–321.
- Little, E. L. (1980). *The Audubon Society field guide to North American trees: Eastern region*. New York: Alfred A. Knopf.
- López, A., Atran, S., Coley, J. D., Medin, D. L., & Smith, E. E. (in press). The tree of life: Universals of folkbiological taxonomies and inductions. *Cognitive Psychology*.
- Malt, B. C. (1995). Category coherence in cross-cultural perspective. *Cognitive Psychology*, **29**, 85–148.
- Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical processing* (pp. 179–195). New York: Cambridge University Press.
- Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, **85**, 207–238.
- Mohlenbrock, R. H. (1992). *Forest trees of Illinois*. Springfield: Department of Conservation.
- Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, **92**, 289–316.
- Nakao, K., & Romney, A. K. (1984). A method for testing alternative theories: An example from English kinship. *American Anthropologist*, **88**, 313–338.
- Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **10**, 104–114.
- Nosofsky, R. M. (1992). Exemplars, prototypes, and similarity rules. In A. F. Healy, S. M. Kosslyn, & R. M. Shiffrin (Eds.), *From learning theory to connectionist theory: Essays in honor of William K. Estes* (Vol. 1) Hillsdale, NJ: Lawrence Erlbaum Associates.
- Petrides, G. A. (1988). *Peterson field guides: Eastern trees*. Boston: Houghton Mifflin.
- Rips, L. J. (1975). Inductive judgments about natural categories. *Journal of Verbal Learning and Verbal Behavior*, **14**, 665–681.
- Romney, A. K., Weller, S. C., & Batchelder, W. H. (1986). Culture as consensus: A theory of culture and informant accuracy. *American Anthropologist*, **88**, 318–338.
- Rosch, E. (1975). Universals and cultural specifics in human categorization. In R. Breslin, N. Lonner, & S. Bochner (Eds.), *Cross Cultural Perspectives on Learning*. New York: Sage.
- Rosch, E. (1978). Principles of categorization. In E. Rosch and B. B. Lloyd (Eds.), *Cognition and Categorization*. Hillsdale, NJ: Erlbaum.
- Rosch, E. & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, **7**, 573–605.
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, **8**, 382–439.
- Sokal, R. R., & Sneath, P. H. A. (1963). *Principles of numerical taxonomy*. San Francisco: Freeman.
- Swink, F., & Wilhelm, G. (1994). *Plants of the Chicago region*. Indianapolis: Indiana Academy of Science.
- Wisniewski, E. J., & Medin, D. L. (1994). On the interaction of theory and data in concept learning. *Cognitive Science*, **18**, 221–281.

(Accepted March 20, 1996)