Recall, Similarity Judgment, and Identification of Trees:

A Comparison of Experts and Novices

Asha C. Srinivasan Shipman*

asha.shipman@uconn.edu

Department of Anthropology

University of Connecticut

Storrs, CT 06069-2176

James Shilts Boster

james.boster@uconn.edu

Department of Anthropology

University of Connecticut

Storrs, CT 06069-2176

*Corresponding author

Abstract

We compare experts and novices in their recall, similarity judgment, and identification of trees. These tasks assess different aspects of informants' knowledge of trees. Experts agree substantially on all three tasks. However, they judge similarity of pictures of trees relying more on morphological and taxonomic features while judging similarity of names of trees relying more on utilitarian features. Novices recall few trees and accurately identify almost none of them, but substantially agree with each other in their similarity judgments of trees. Thus, we find that the natural world can be understood by experts and novices alike, but that training sharpens observations and informs about properties of species not obvious by direct observation. Moreover, the kinds of features used by experts in similarity judgments are influenced by how the species are represented. These results confirm and extend the findings of Boster and Johnson (1989) and Medin, Lynch, Coley, and Atran (1997).). They also show how it is that the domain knowledge of experts arises out of the many different ways in which they engage the trees and out of the diversity of activities and purposes in their community of practice.

Key words: Categorization, Ethnobiology, Expert-novice differences, Functional versus morphological features, Stimulus effects

Recall, Similarity Judgment, Identification of Trees

Ethnobiology is the study of how humans understand and make use of their natural environment and its resources. A central question is whether the folk understanding of nature is driven by intellectual curiosity or by practical need. On one side, authors such as Levi-Strauss and Berlin see folk classifiers as innate natural historians who seek to understand nature for its own sake, discriminating classes of organisms based on readily observable morphological features (Levi-Strauss 1966; Berlin 1973; 1992; cf. Boster and Johnson 1989). In contrast, authors such as Hunn (1982) argue that the folk impulse to classify biological kinds is driven by practical rather than intellectual considerations, and that the utilitarian features[1] of organisms (those most directly affecting the use of the plant or animal) are of primary interest to folk classifiers.

Boster and Johnson (1989) tested these alternative explanations in a study of American novice and expert judgments of the similarity of fish. They asked novices and experts to free pile sort 43 line-drawings of commonly caught marine species of fish. In a free pile sort an informant is given a set of stimuli and asked to subdivide the stimuli based on similarity. Informants can make as many groups as they wish with as many members in each group as they would like. The informant then explains the reasoning behind the groupings. At the beginning of the research project, the authors bet with each other on the results – Johnson expected that the expert similarity judgments would be based on primarily on utilitarian features of the fish, Boster expected that, since the uses of the fish would be determined by their morphology, there would be no difference in the expert similarity judgments from one purely based on morphology. Both Boster and Johnson were wrong in their expectations: They found that the features used in the judgments of similarity among organisms depended on the judges' familiarity with the organisms. Experts were more likely to judge similarity on the basis of a combination of utilitarian, ecological, and morphological features of the fish while novices based their judgments entirely on the morphological features of the fish represented in the line drawings that they sorted.

Medin, Lynch, Coley, and Atran (1997) investigated how different kinds of American tree experts (e.g., tree and park maintenance workers, taxonomists, and landscapers) judged the similarity of and inductively reasoned about trees. They asked the various tree experts to free pile sort 48 index cards labeled with the common names of trees. Informants sorted only those trees they were familiar with. They also examined whether experts' similarity judgments influenced their inductive reasoning about trees using triads tests, again, with common names. They found high within-group agreement among the three expert groups and significant between-group differences.[2] Taxonomists' similarity judgments were the most similar to the scientific taxonomy, landscapers' judgments were the most dissimilar, and maintenance workers fell in between. Taxonomists' groupings were based on morphology and landscapers' groupings were based on utilitarian criteria (e.g., trees for city streets, flowering trees, ornamental trees, and weed trees). They concluded that similarities in similarity judgment stem from the correlational structure of the organisms while differences arise when experts weigh features differently depending on the nature of their expertise.

In sum, both Boster and Johnson (1989) and Medin et al. (1997) find that informant groups vary in their attention to morphological or utilitarian features of organisms; Medin et al. (1997) show that this variation is not limited to the contrast between novices and experts but can exist among expert groups as well. They also show that this variation does not depend on the organisms being hidden from view: Trees, unlike marine fish, are in plain sight and yet are categorized in very different ways by the different expert groups. Like fish, some information regarding trees (e.g., strength of the wood, disease susceptibility, and reproductive mechanism) is not obvious on casual observation and is likely to be learned primarily through the pursuit of specific goals. The goals related to experts' interactions with the trees (e.g., selecting trees for landscape use, treating diseased trees, or assessing trees for marketability) probably determine what features they pay attention to and the categories, natural or artificial, that they form.

One question that emerges is: How fixed is an expert's attention to a particular kind of feature? Is reliance on morphological versus utilitarian features determined by one's occupation or will reliance shift depending on the nature of the task? Medin et al.'s (1997) study would seem to imply a high degree of stability of feature preferences; their groups of experts (with the exception of the landscapers) tended to rely on the same features on both a similarity judgment and an inference task. We interpret this stability as a consequence of the fact that the tasks employed by Medin et al. (1997) can be completed using the same mental strategy. However, if tasks are chosen to recruit very different kinds of knowledge, it is possible that all informants will respond using similar strategies responding to the nature of the task, over-riding the differences among expert groups. This would be consistent with the sort of mental flexibility found by Boster and D'Andrade (1989) who showed that different groups of informants all relied on one suite of features in judging songbirds and another suite in judging non-songbirds.

Research Design

An appropriate test of this question is to administer tasks with very different cognitive demands to a diverse group of domain experts and to novices. Like Medin et al. (1997), we chose the domain of locally common trees. The tasks we chose are a free-recall task, a similarity judgment task, and an identification task. Our free-recall task asked informants to recall as many trees as she or he could remember within a 15 minute period. The informants were given sheets of lined paper to write on but no other kinds of instruction other than to use only common names so that their lists could be compared with novices' lists. Because the order is unconstrained, we can test the degree to which recall is clustered on various sorts of criteria (e.g., taxonomy, morphology, use, ecology).

Our similarity judgment task, the successive pile sort (Boster 1994), asked informants to discriminate among a set of 26 trees common to Connecticut, presented as cards with either pictures or names of trees. The picture set presented large, glossy, color photographs of the leaves, branches, bark, and tree silhouette taken from a popular field guide to trees edited by Whitman (1986). The other set of cards presented the common name of each tree species used in the picture set. There is a large overlap between the trees used in both studies; 19 of the 26 trees used in the successive pile sort and identification task in this study are identical with or are members of the same genus as trees used by Medin et al. (1997). Also, although Medin et al. (1997) had a complete list of 48 trees in their study, because informants only judged trees they were familiar with, many informants would have judged fewer trees. The seven genera used in this study that were not included in Medin et al.'s (1997) list were chosen to represent species common in Connecticut forests or suburban landscapes. As a potential measure of expertise and knowledge of tree history we included the American chestnut in our study. The chestnut was an important hardwood in the Connecticut woodland until the chestnut blight in the early 1900s made mature trees quite rare.

The stimulus cards used in this task were arrayed alphabetically on a table and each informant was asked to: "Group these trees the way you think they naturally go together." Informants then moved the cards into groups and provided a rationale for their grouping. After informants created their initial piles, they were then asked to group the piles according to their similarity. For example, if an informant creates five piles, they are asked which two piles they would group together first because they go together best, and so on. Once the piles were completely grouped, the informant was asked to return to the initial piles and begin sub-dividing them. They were asked to choose the pile in which the trees go together least and to divide that pile into two piles so that the trees go together better. Informants continued in this fashion until all the trees were separated. The successive pile sort, by forcing informants to divide, group, and then subdivide all of the trees, fairly rapidly elicits far more information than is possible with a free pile sort or triads test.

Unlike the identification task, the successive pile sort task does not require any prior knowledge of morphological features of trees and unlike the identification and free-recall tasks, the successive pile sort does not require that the informant know any names of trees. A similarity judgment task using names of trees makes a very different set of demands on the informant than the same task using pictures; if names are used, the informant must mentally summon a particular suite of features of the trees in order to judge the resemblances among the trees, thus requiring knowledge of both the names of the trees and their features. However, it does not favor attention to morphological features of trees in the same way that the similarity judgment task with pictures does; the informant may use any feature she or he likes to sort the names. Half of the experts judged the similarity of trees based on pictures while the other half judged on common names (i.e., words). Novices, because of their unfamiliarity with trees, only judged the similarity of trees based on pictures. While performing the successive pile sort informants explained what features they were relying on.

Our identification task used the set of picture cards from the similarity task and asked the informant to provide a common name to identify each tree. This task requires that the informant both be able to recognize the tree on the basis of its morphological features and to remember the name that labels that particular set of features.

Insert about here: Figures 1-4

These tasks are chosen to represent some of the different kinds of thinking about trees we expect tree experts to engage in. The identification task taps an essential piece of knowledge that allows experts to associate a physical specimen with its unique name and hence with its properties. The ability to identify trees thus allows arborists to properly protect trees against disease, landscapers to anticipate how trees will mature into a landscape, and horticulturists to predict how well a graft will take. The processes of successively distinguishing groups of trees performed in the successive pile sort parallels the logic of a taxonomic key commonly used by experts to classify an unknown specimen. More generally, many expert judgments are made on the basis of similarity and the sharing of ecological, morphological, and utilitarian characteristics. For example, landscapers who are unable to procure a particular species of tree may be asked to provide a close substitute. Arborists may be asked to provide a list of trees suitable for creating a windbreak. Horticulturists categorize trees in their greenhouses. Finally, experts are often called upon to list inventories of trees suitable for a particular use or ecological setting, an ability tapped by the free-recall task.

Of course, novices are also required to identify, judge similarity, and recall things; these are not mental processes unique to the experts and, indeed, are necessary for everyday activity. In this study we expected that the differences in how expert and novice participants perform the three tasks would be due to the amount and kinds of cultural knowledge having to do with trees that they possess. We expected our expert informants to recall more trees and in a more clearly taxonomically clustered order than novices in the free-recall task. We expected that experts would rely on both utilitarian and morphological criteria and novices to rely only on morphological criteria in similarity judgments, just as Boster and Johnson (1989) found. We also expected that the type of stimulus set would affect the type of features relied upon in similarity judgments; that experts presented with pictures would rely more heavily on morphological features while experts presented with the names would rely more heavily on utilitarian features. Finally, we expected that novices would accurately identify far fewer trees than experts, but that expert groups would not vary among themselves in accuracy of identification.

Informant Recruitment and Demographics

We chose expert informants from three common professions that seemed the most specifically engaged with trees: Landscaping, horticulture, and arboroculture. The label of expert was applied to people who worked in these professions. Experts whose work included more than one of the professional categories were assigned to the category they thought best fit them. Within each occupational category experts varied according to their age (23 to 72; mean age 47), experience (ranging from 4.5 to 45 years in the field), education (including apprenticeships, licenses, certificates, and college degrees), and professional orientation (e.g., landscape design, landscape installation, pesticide application, tree removal, and stump grinding). Experts attributed their professional interest in trees to exposure working for family businesses (e.g., farms, fruit stands, landscaping companies, greenhouses), to focal individuals who took experts on as apprentices and shared their love of trees, and to a variety of quirky circumstances which led them to jobs working with trees and plants.

Recall, Similarity Judgment, Identification of Trees

The first informants interviewed were Ms. Shipman's in-laws who have owned a landscaping and nursery business for over 40 years. We recruited more participants using family members' professional contacts, and continued recruiting participants via this form of "snowball sampling." All of the participants lived in Connecticut and all were of European descent. It appears that, at least in Connecticut, minority involvement in these professions is low. Though we sought equal numbers of male and female participants we were only able to recruit four female arborists, five female horticulturists, and nine women working in landscaping, thus resulting in an unbalanced expert sex ratio of 36 males:18 females. We did recruit equal numbers of arborists, landscapers, and horticulturists (N=18).

The majority of expert participants belonged to state, regional, and national professional organizations and they drew on their fellow members in these organizations to assist us in this study. Among the organizations whose members participated in this study include the Connecticut Horticulture Society, the Connecticut Nursery and Landscape Association, the American Nursery and Landscape Association, the American Rhododendron Society, the International Society of Arboroculture, and the Connecticut Tree Protection Association.

The novice informant pool was also sex-biased, but in favor of females (10 male: 23 female). Participants were all University of Connecticut students (mean age 23); most of whom were enrolled in an upper-level anthropology class and received class credit for participating in the project. Novices had no experience working with trees, and few of them had any educational background with respect to plants or trees beyond high school biology. It is also unlikely, given the young age of this group, that any of them were home owners with primary responsibility for the landscaping around their homes. A few of these novice informants did report having family members who had guided them through the forest identifying trees. Particularly relevant was their lack of training in plant taxonomy which more than half of the experts had studied formally. All but a few novices were of European descent, proportional to their representation in the student body. All participants were native English speakers.

Data Analysis

Free-recall Task

The lists of trees that informants recalled were compared in length and in the amount of taxonomic clustering in recall. Each species of tree recalled by any informant was assigned a unique 7 digit code that reflected its taxonomic classification at the division, class, subclass, order, family, genus and species rank, based on Dirr (1998) and on the United States Department of Agriculture Plants Database maintained at http://plants.usda.gov. Responses were coded as specifically as possible (e.g., “oak” was coded down to the genus Quercus while “Christmas tree” was coded down to the order Pinales). Responses that could not be identified at all (such as “Easter tree”) were coded as not matching any other plant at any rank. We then calculated the taxonomic distances among the entirety of trees recalled by informants to create a species by species taxonomic distance matrix. Taxonomic distance is defined as the number of nodes one must ascend in the taxonomy so as to include both members of a pair of trees. Two species in the same genus have a taxonomic distance of one; two genera in the same family have a taxonomic distance of two; and so on. The path-length of an informant's recalled list of trees was calculated by summing the taxonomic distances between successive trees recalled. Unidentifiable trees were assigned the maximum taxonomic distance of seven to the adjacent trees in the list. Clustering in recall (Puff 1979; Roenker, Thompson, and Brown 1971; Romney, Brewer, and Batchelder 1993) was measured as a Z-score by randomly permuting an informant's list a million times, subtracting the average path-length of the randomly permuted lists from the path-length of the actual list, and dividing by the standard deviation of the randomly permuted lists. As the amount of clustering in recall increases, the path-lengths of the randomly permuted lists become much larger than the path-length of the actual list relative to the standard deviation of the random lists, and the calculated Z-score reaches a much larger negative value.

Similarity Judgment Task

The informant responses to the successive pile sort were aggregated to form three overall similarity matrices, one for each combination of informant group and stimulus type (i.e., novice pictures, expert pictures, expert words).

These three matrices were used to create the average-link hierarchical cluster diagrams shown in Figure 6. As a quantitative assessment of between-group differences in the similarity task, we categorized the criteria that experts provided to explain their groupings of trees as either utilitarian (e.g., used in furniture, landscaping, fruiting trees), ecological (e.g., soil preferences, typical habitat), morphological (e.g., leaf shape, tree size, tree shape), or taxonomic (e.g., native, non-native, deciduous, evergreen).[3]

Identification Task

We tabulated the number of correctly identified trees for each informant. Informants' responses to this task were checked for accuracy to the genus level so that expert and novice responses could be compared.

Results and Discussion

Free-recall Task

Experts listed significantly more trees than novices did (expert[4] mean = 68.8; novice[5] mean = 18.0; $t = 17.2$ $p < .01$). Many experts reported composing their lists by beginning with a common tree genus such as maple or oak and then listing as many species and varietals as they could recall in that genus before moving to another genus and repeating the process. Several experts reported mentally opening up their favorite tree guide and starting at the beginning with *Abies* (fir) or *Acer* (maple). Others recalled lists from tree manuals or the arrangement of their greenhouse and nursery stock. This technique of recalling items that share strong associations very efficiently unpacks information typically stored in long term memory. In contrast with the expert-novice difference and counter to Medin et al.'s (1997) findings, there were no differences among expert occupational groups in the amount of structure in their lists. Interestingly, variance in the length of lists is small but significant. On average arborists produced longer lists than horticulturists and landscapers (arborist mean = 78, landscaper mean = 63; horticulturist mean = 65; $F = 4.8$, $p < .05$; Tukey post-hoc test showed A > H = L; $p < .05$). Novices recalled few trees and based their recall on randomly organized cultural elements such as street signs, apartment complex names, song lyrics, and furniture wood.

The differences between experts and novices in the heuristics used to recall trees left an imprint on their relative amounts of clustering in recall, as shown in Figure 5. The horizontal axis indicates the number of trees recalled by each informant and the vertical axis is the Z-score measure of the structure in each list, as described above, with the most highly structured lists having the greatest negative scores. Novices form a tight cluster at the top left corner of the figure indicating that their lists were both short and not highly structured while experts produced longer, well organized lists of trees. The longest list of 106 trees had the most structure $Z = -19.3$, with an associated probability of approximately 10^{-79}. Figure 5 also illustrates a strong trend in that those who cluster more in recall and have a more highly structured mental organization of trees can remember greater numbers of trees ($r = -.94$, $p < .0005$). Experts' recalled lists of trees were significantly more highly structured than those of novices ($t = -15.6$ $p < .0005$).

Insert about here: Figure 5

Insert about here: Table 1.

Similarity Judgment Task

A comparison of the cluster diagrams (see Figure 6) shows striking overall concordance and some systematic differences[6]. All three cluster diagrams split evergreens from deciduous trees as the first distinction among the trees, but other patterns also emerge. Novices and experts strongly agree in how they successively sort evergreens; the groupings of the three pines, the three spruces, and the pairing of the eastern redcedar and eastern arborvitae are found in all three cluster diagrams. There is less agreement on deciduous trees although several groupings recur in the three cluster diagrams (e.g., black walnut, chestnut, beech and apple, pear, flowering-dogwood). The ginkgo appears to be an outlier in all three cluster diagrams.

Insert about here: Figure 6.

Qualitative analysis of informant group differences in similarity judgment

The systematic differences among the three cluster diagrams appear to be the result of differences in the kinds of features used by the informants when judging the similarity of trees, varying according to informant expertise (novice-expert) and the type of stimulus (picture-word). Lacking experience with trees, novice informants reported basing their judgments of similarity on features of the leaf and branch. For example, they grouped evergreens according to their needle shape, resulting in three main evergreen clusters: The "long-needle" evergreens (pines), the "short-needle" evergreens (spruces, hemlock, and balsam fir), and the scale evergreens (arborvitae and redcedar) which novices often referred to as "feathery." Attention to these features resulted in an aggregate similarity matrix that is closely correlated with the taxonomic relations among the evergreens, despite novices' lack of experience with them (cf. Boster and Johnson 1989). When discriminating among deciduous trees and the ginkgo, novice informants again relied on morphological features. There were three major groupings in the novice cluster diagram: Trees with "oval-shaped leaves with jagged edges" (walnut, chestnut, beech and hickory), trees with "heart-shaped leaves" (apple, dogwood, catalpa, pear, aspen, and basswood), and trees with "divided" or "lobed" leaves (pin oak, sugar maple, sycamore, tulip, and sassafrass). Novices considered the ginkgo as similar to the sassafrass and tulip tree because they all have unusually shaped leaves. However, relying on morphological features to discriminate the deciduous trees resulted in a weaker correspondence to the scientific taxonomy than was found for evergreens.

Experts, on the other hand, like expert recreational fishermen judging the similarity of fish reported by Boster and Johnson (1989), used a range of criteria to discriminate among trees, regardless of whether they judged pictures or common names of trees. When judging evergreens, although all experts created groupings that closely mirrored the taxonomy, their verbal explanations for the groupings varied. Some experts mentioned taxonomic criteria (e.g., "these are all spruces"), other times they mentioned morphological criteria (e.g., "these have short stiff needles that completely surround the branch"), ecological criteria (e.g., "these trees are tolerant to salt water and blue stain disease"), or utilitarian criteria based on uses for trees and tree products (e.g., "these make good wind-breaks").

Experts often broadened their criteria when sorting deciduous trees. Walnut, hickory, beech, and oak trees were often grouped together because experts referred to them collectively with one or more of the following criteria: "large," "native," "woodland," "nut-bearing trees with similarly shaped leaves." In the same vein, maples and oaks were often paired on the basis of one or more of the following criteria: "strong," "sturdy," "native trees that are common to both the woodland and the suburban landscapes that have lobed leaves." Quaking aspen, basswood, catalpa, sassafrass, and tulip trees all fell into a folk category that was variously described using utilitarian adjectives such as "weedy," "dirty," and "junky." Finally, the grouping of dogwood with pear and apple tree was variously described as "popular" and "small, flowering ornamental."

The relative isolation of the ginkgo makes good taxonomic sense. Belonging to an ancient lineage of trees, it is in its own division (ginkgophyta) separate from the division of evergreens (coniferophyta) and deciduous (magnoliophyta) trees. Many informants noted the ginkgo's distinct heritage, unusual bi-lobed leaf shape, and parallel venation but grouped it with deciduous trees.

Quantitative analyses of informant group differences in similarity judgment

In analyzing the criteria used by experts to discriminate trees we found that, overall, arborists mentioned far more ecological criteria and far fewer taxonomic criteria than expected, horticulturists mentioned far fewer ecological criteria and far more taxonomic criteria[7], and landscapers fell in between (χ^2 = 60.52; d.f. = 6; $p < .01$). Table 2 compares the six combinations of stimulus type (picture – word) and expert type (arborist, horticulturist, landscaper) in their frequency of use of four kinds of criteria. Experts sorting pictures reported mostly taxonomic criteria while experts sorting words reported mainly utilitarian criteria. Arborists judging pictures are exceptional; they mentioned almost equal numbers of taxonomic and utilitarian criteria. Experts sorting words mentioned slightly more than one and a half times the number of criteria mentioned by experts sorting pictures, suggesting that presentation of a visual stimulus diminishes the felt need to articulate the explanation of the groupings (i.e., the informant feels the explanation is self-evident in the pictures themselves). There was also a significant difference in the judged similarity of the trees, depending on whether experts were sorting words or pictures (QAP[8] Z = 13.2; Pearson's r = .2, $p < 10^{-6}$). However, there were no significant differences among the three types of experts doing the same task (experts sorting words QAP Z = -.1, Pearson's r = .0, $p > .05$; experts sorting pictures QAP Z = 1.0; Pearson's r = .02, $p > .05$).

Insert about here: Table 2.

To explore further differences between informant groups in their similarity judgments of trees, we created a utilitarian matrix based on the criteria used by experts to explain their groupings to capture the similarities in the uses of the trees. We then compared the taxonomic and the utilitarian matrices with the aggregate similarity matrices of novices judging pictures (np), experts judging pictures (xp), and experts judging words (xw) (See Table 3 below.) The highest correlation was between novices and experts sorting pictures, followed closely by the correlation of experts judging pictures and experts judging words. The aggregate similarity judgments of all three groups were highly correlated with the taxonomy, but the correlation of the judgments to the utilitarian matrix steadily increases from novices (the lowest), to experts judging pictures, up to experts judging words (the highest correlation).

Insert about here: Table 3.

We then repeated these analyses, to determine whether differences among aggregated responses were matched by systematic differences in individual responses. They were: Novices corresponded more closely with the taxonomy than experts (overall $F = 9.6$, $p < .0005$; Tukey posthoc np:xp, $p < .02$; np:xw $p < .0005$) while experts judging words and those judging pictures did not differ significantly in their proximity to the taxonomy (Tukey posthoc xp:xw $p > .05$). In contrast, experts judging words were closer to the utilitarian model than both novices and experts judging pictures (overall $F = 19.9$, $p < .0005$; Tukey posthoc xw:np, $p < .0005$; xw:xp $p < .0005$) while the latter two groups did not differ significantly from each other (Tukey posthoc np:xp $p > .05$). Thus we see a clear expert – novice contrast in degree of correspondence to taxonomy and a clear picture – word contrast in degree of correspondence to utility. As shown in Figure 3, novices correspond best with the taxonomic model, experts judging words correspond best with the utilitarian model, and experts judging pictures fall in the middle on both scales.

Insert about here: Figure 7.

When we compared the amount of agreement within the three informant groups, we found that novice agree the most with each other followed by experts judging pictures and experts judging words (Overall $F = 280.1$, Tukey posthoc p for all comparisons < .001)[9]. Boster and Johnson (1989) also reported greater within-group agreement among novices versus expert fishermen judging fish. Like the novice fishermen, the tree novices' collective ignorance about trees focused their attention to the morphological features presented in the pictures of the trees. Most likely it is this forced dependence on one set of features that produced the observed high level of agreement among novices. In comparison, experts judged trees using a number of different features, resulting in lower levels of within-group agreement.

Expert Occupation and Training Effects

In their study of tree experts Medin et al. (1997) concluded that occupation had a significant effect on how their expert informants categorize trees. To test the effects of occupation on similarity judgments in our study, we created a model matrix, the cells of which were "1" if the pair of informants corresponding to the row and column was in the same occupation, and "0" if they were not and compared this model matrix with the inter-informant correlation matrix. We did this separately for experts judging pictures and experts judging words. We found that there was no association between the pattern of agreement among informants and their occupation regardless of which stimulus set was used (expert picture QAP Z = 1.0, Pearson's r = .02, p > .05; expert word QAP Z = -0.1; Pearson's r = .0, p > .05)[10]. In contrast, the choice of pictures or names as stimuli did have a significant effect using the same test (QAP Z = 13.8, Pearson's r = .2, $p < .0005$).

How can we explain the lack of strongly significant occupational and training differences among experts in the similarity judgment task, particularly in light of the results found by Medin et al. (1997)? We do not believe that the difference was due to the choice of species used as stimuli because of the high overlap between our species list and Medin et al.'s (1997). The difference between our results and Medin et al.'s (1997) are much more likely due to the large difference between our experts and their experts rather than from the modest difference between our species list and their species list. All but five of the expert informants who participated in our study belonged to at least one professional organization. A large number of arborist informants belong to the Connecticut Tree Protective Association (CTPA), an association founded in 1922 just after the state began mandating licenses for "tree experts." The goals of CTPA include licensing and continuing education for arborists as well as public outreach programs. Even if arborists were not members, they all passed through the same rigorous training to become licensed in Connecticut. Similarly, a large proportion of the horticulturist and landscaper informants belong to the Connecticut Horticultural Society and the Connecticut Nursery and Landscape Association (CNLA). Both of these professional organizations provide classes and lectures on issues in the field. The CNLA is renowned for its licensing program and having the CNLA license is prestigious. While Medin et al. (1997) describe their experts as representing “a continuum of types of tree expertise [p. 55],” the informants in our study probably represent the high-end of the continuum. Forty of the experts had four-year college degrees and eight had associate’s degrees; most of these degrees were directly job-related. Perhaps even more important, these experts on average had 23.5 years of experience working with trees and plants. It may be that the extensive training and long experience of our sample of experts wiped out any occupational differences that would have been detectable with a less expert sample of experts.

Identification Task

Most novices were unfamiliar with the trees they were asked to identify (mean correct = 3.7; sd = 2.2). Many novices named each of the evergreen trees with the label *pine*, so that, just as a broken watch is correct twice a day, 15 novices correctly identified at least one of the 3 pines. Novices were also unfamiliar with deciduous trees; 15 novices correctly identified the oak tree and 28 correctly identified the maple, but several novices confused the two. By contrast, most experts were able to correctly identify nearly all of the trees (mean correct = 20.4; sd = 3.9). Many experts expressed their preference for identifying evergreen trees on the basis of live specimens because features such as needles per bundle (for the pines) and softness (to differentiate between spruces and firs) are not available in pictures. One arborist informant who runs a stump-grinding operation spoke about using scent to identify evergreen species. Almost half of the experts (21) correctly identified the American chestnut, suggesting that their knowledge of trees extended beyond trees they commonly saw or worked with.

If experts judging the similarity among trees using the common names as stimuli had realized that the pictures in the identification task represented the same trees, they might have been expected to correctly identify more of the trees than experts judging pictures. However, there were no significant differences in identification between these two groups; the type of stimuli used in the similarity judgment task did not affect performance in the identification task ($t = 1.26$; $p > .05$).

The high degree of accuracy in tree identification among experts may partly be explained by the extensive tree identification training that arborists, horticulturists, and landscapers receive in Connecticut and to the long years of experience of our sample. The Connecticut state arborist licensing exam includes a tree identification test in which examinees must identify trees on the grounds of the Tree Extension Service Experiment Station in New Haven. Several arborists reported spending a great deal of time studying for this test and pride themselves in their ability to identify trees quickly and accurately. One arborist proudly reported his habit of identifying the trees his crews needed to cut while driving by job sites in his truck at 40 miles per hour. Horticulturists are familiar with trees that they cultivate and nursery catalogs, used by horticulturists and landscapers, are filled with glossy photographs and lists of trees labeled with both common and Latin names. Landscapers, depending on their education, may not have spent much time memorizing local trees; their attention is very much drawn to the showy cultivars their customers desire. However, some colleges and universities require landscape and horticulture students to take courses that incorporate tree identification. One of the landscaping companies whose employees participated in this study has an annual tree identification competition on their web site. Again, the fact that we sought out truly expert experts may have both masked potential differences among the occupational groups and also heightened their difference from novices.

If we found expert experts, we also found exceptionally arboreally ignorant novices. Any self-respecting five year old Mayan or Aguaruna Jivaro is able to identify far more species of trees and plants than the much older novices in our sample, thus confirming an observation made by Dougherty (1978:77-78) some years ago that "expertise and indifference are not atypical states for human beings with respect to particular domains" and that the depth of recognition and classification of domain members is relative to the "interest and attention of human groups and individuals" to the domain. Novices' difficulty with identifying common state trees likely reflects the minimal attention they give to plants. Unless they are avid gardeners, trees fade into a sort of green background fog, ignored because discriminating amongst the trees does not aid them in any task they typically confront. Owning a home may increase attention to trees, whether to "problem trees" which drop a lot of leaves and branches or to comforting sources of shade or blossoms. Most of the novices we interviewed were students living in apartments, dormitories, or with their parents, few were home owners. Some novices regarded their inability to identify more than a handful of trees with genuine regret, and stated that they "ought to know" the trees in this study. However, even if they did not know the names of many of the trees, they were still able to recognize and agree on the pattern of resemblances among the trees, particularly evergreens. This parallels the findings of Boster (1987) and Boster and D'Andrade (1989) for birds and Boster and Johnson (1989) for fish similarity judgments.

Conclusions

Overall, we found that there are significant differences in what novices and experts know about trees and how they think about them, as well as some commonalities. Experts, regardless of their occupation, share a great deal of knowledge about trees. They are able to recall a great number of trees, listing them roughly in taxonomic order and are able to identify most common Connecticut trees. Novices, in contrast, are unable to recall more that a few trees and correctly identify only the most common ones. Given that the part of Connecticut where the research was carried out is thickly forested and the trees are easy to observe, these large discrepancies between experts and novices do not stem from a lack of opportunity on the part of the novices to observe the trees, and suggest that the ability to recall and identify trees hinges on their personal interest in the domain. However, the fact that a lack of knowledge about the trees did not hamper their ability to strongly agree in their similarity judgments of the trees suggests that they certainly share with the experts the natural ability to discern the features that distinguish the trees. Again, this parallels the findings of Boster (1987) who showed that diverse groups of informants (two groups of South American Amerindians, U.S. college students, and professional ornithologists) agree in their recognition of patterns of resemblance among a collection of South American bird specimens. Boster and D'Andrade (1989) and Boster (1996) interpret this result as evidence that cross-cultural agreement in biological classification is the outcome of a pan-human perceptual strategy that selects those characters of a collection of organisms that yield the most informative classification.

The variety of kinds of knowledge that experts can draw upon to judge the similarity of trees (e.g., taxonomic, morphological, ecological, utilitarian) paralleled Boster and Johnson's (1989) findings on recreational fisherman judging the similarity of line drawings of fish and Nolan's (2002) findings on expert herbalists judging the similarity of both pictures and common names of plants.

Arborists, for example, were more likely to think about trees based on where they are "used" in a landscape (e.g., street trees) than on morphological features. However, unlike Boster and Johnson (1989), the tree experts' groupings of trees were very similar regardless of the criteria they mentioned as important. They also relied on taxonomy and morphology to a far greater degree than Boster and Johnson's (1989) recreational fishermen. The fact that experts attend more to utility than novices (largely because they know about it and novices don't) but that the expert fishermen relied on utilitarian features more than the tree experts is probably due to differences both between the tree specialists and the fishermen and between trees and fish.

The expert informants in our study are highly educated, skilled, and experienced and are probably at the high end of the expertise continuum Medin et al. (1997) described. To be successful, arborists, horticulturists, and landscapers need to be, and are, aware of both biological and utilitarian aspects of trees because biology often determines the use of trees. Because utility and biology are intertwined, tree experts' knowledge of the taxonomy of the trees helps them predict the form, growth features, wood strength, and preferred habitat of a particular species. That is, the tree experts participating in this study had much greater training in taxonomy than the expert fishermen interviewed by Boster and Johnson (1989), and the utilitarian characteristics of fish are less predicted by their placement in the taxonomy (e.g., experts regard species from a number of families of fish as members of the category labeled “trash fish”).

Our finding that the way that the trees were presented in the similarity judgment task (either as picture or as common name) had a much greater influence on expert responses than their particular occupational specialty also has important implications. We dissent from Medin et al’s (1997) certainty that one form of stimulus item is intrinsically superior to other forms (i.e., it is better to present words than pictures), and argue that a comparison of the ways that informants respond to the same items presented in different forms can offer important insights into their categorization and reasoning. Our findings suggest that visual stimuli tend to increase informant agreement by providing the morphological features for similarity judgment and thereby somewhat reduce the relative importance of utilitarian features: Experts who judged pictures of trees emphasized taxonomic criteria while experts who judged common names of trees put a greater emphasis on utilitarian criteria. Further, experts who judged pictures provided far fewer justifications during this task compared with experts who judged common names. These findings suggest that future ethnobiological research addressing folk classifications of the natural world ought to incorporate different types of stimuli in order to capture the diverse ways in which people perceive and mentally organize elements of their environment.

We believe that an important contribution of this research is to reinforce the lesson taught by recent work in cognitive anthropology (e.g., Lave, 1988;

Lave and Wegner, 1991; Chaiklin and Lave, 1993; Hutchins, 1995; Keller and Keller, 1996) that knowledge is acquired in action, learned as it is needed to solve different problems. A Brazilian child street-vendor, who shows no difficulty providing correct change, may not be able to solve a simple arithmetic problem when it is presented as a manipulation of numerals as it would be in school (Guberman, 1999). Here, performance on a school exam would not be a valid measure of the child's mathematical ability because it would fail to capture the context in which the knowledge had been acquired – it fails to capture cognition on the move. Each different problem space involving a domain has the power to motivate the acquisition of wildly different bodies of knowledge about the domain. This helps explain the great richness of the knowledge of the tree experts consulted in this study. They must deal with the trees as problems or pests in landscaping, as possible ornaments, as sources of useful fruits or wood, as suited to particular soils, drainage types, and light conditions, as shapes to be identified, as individuals whose taxonomic relationships should be discerned, etcetera. And they also must interact with each other as they collectively form a community of practice (*sensu* Lave and Wegner, 1991). Each of these different ways of engaging the trees produces a different body of knowledge and each body of knowledge may influence the judgments of similarity and difference in a different way. If anything, their different modes of interaction with the trees is far richer than the ways in which the recreational fishermen studied by Boster and Johnson (1989) interacted with the fish they sought to catch and eat. The tree experts have acquired these diverse bodies of knowledge because knowledge of all of them is necessary in their practice. They must offer solutions that are ‘spherically sensible” – sensible when seen from any perspective – and each different body of knowledge is a source of a different set of constraints on the solution to a problem.

References Cited

Berlin, Brent

1973 Folk Systematics in Relation to Biological Classifications and Nomenclature.Annual Review of Ecology and Systematics 4:259-271.

Berlin, Brent

1992 Ethnobiological Classification: Principles of Categorization of Plants and Animals in Traditional Societies. Princeton: Princeton University Press.

Boster, James S.

1987 Agreement Between Biological Classification Systems is Not Dependent on Cultural Transmission. American Anthropologist 89(4):914-920.

Boster, James S.

1994 The Successive Pile Sort. Cultural Anthropology Methods 6(2):7-8.

Boster, James S.

1996 Human Cognition as a Product and Agent of Evolution. *In* Redefining Nature: Ecology, Culture, and Domestication. Roy Ellen and Katsuyoshi Fukui, eds. Oxford: Berg.

Boster, James and Roy D'Andrade

1989 Natural and Human Sources of Cross-Cultural Agreement in Ornithological Classification. American Anthropologist 91(1):132-142.

Boster, James S. and Jeffrey C. Johnson

1989 Form or Function: A Comparison of Expert and Novice Judgments of Similarity Among Fish. American Anthropologist 91(4):866-889.

Chaiklin, Seth and Jean Lave

1993. Understanding Practice: Perspectives on Activity and Context. Cambridge: Cambridge University Press.

Dirr, Michael A.

1998 Manual of Woody Landscape Plants. Their Identification, Ornamental Characteristics, Culture, Propagation, and Uses. Champaign, IL: Stipes Publishing, Inc.

Dougherty, Janet W.

1978 Salience and relativity in classification. American Ethnologist 5(1):66-80.

Guberman, Steven R.

1999 Cultural aspects of young children's mathematics knowledge. In J. V. Copley (Ed.), *Mathematics in the early years* (pp. 30-36). Reston, VA: National

Council of Teachers of Mathematics.

Hubert, Lawrence J. and Schultz, James

1976 Quadratic assignment as a general data analysis strategy. British Journal of Mathematical and Statistical Psychology 29:190-241.

Hunn, Eugene

1982 The Utilitarian Factor in Folk Biological Classification. American Anthropologist 84(4):830-847.

Hutchins, Edwin

1995 Cognition in the Wild. Cambridge, MA: MIT Press

Keller, Charles and Janet Keller

1996 Cognition and Tool Use: The Blacksmith at Work. Cambridge University Press.

Jean Lave

1988 Cognition in Practice. Cambridge: Cambridge University Press.

Lave, Jean, Etienne Wenger

1991 Situated Learning. Cambridge: Cambridge University Press.

Lévi-Strauss, Claude

1966 The Savage Mind. Chicago, IL: The University of Chicago Press.

Medin, Douglas L., Elizabeth B. Lynch, John D. Coley, and Scott Atran

1997 Categorization and Reasoning among Tree Experts: Do All Roads Lead to Rome? Cognitive Psychology 32(1):49-96.

Nolan, Justin M.

2002 Wild Plant Classification in Little Dixie: Variation in a Regional Culture. Journal of Ecological Anthropology 6(1):69-81

Puff, C. R. (ed.)

1979 Memory Organization and Structure. New York: Academic Press.

Roenker, D. L., Thompson, C. P., and Brown, S. C.

1971 Comparisons of measures for the estimation of clustering in free-recall. Psychological Bulletin 76:45-48.

Romney, A. Kimball, Brewer, Devon D., and Batchelder, William H.

1993 Predicting semantic clustering from semantic structure. Psychological Science 4:28-34.

Recall, Similarity Judgment, Identification of Trees

Whitman, Ann H. (ed.)

1986 National Audubon Society Pocket Guide to Familiar Trees of North America: Eastern Region. New York: Alfred A. Knopf, Inc.

Notes

[1] We use the term "feature" in this paper to refer to a characteristic of the plant used in identification or similarity judgment tasks, in contrast to the term "criterion" which we use to refer to a verbal explanation offered by an informant of a feature.

[2] According to the authors, these differences were both statistically different and meaningfully different.

[3] While the terms *deciduous* and *evergreen* describe whether or not trees retain their leaves during a dormant period, informants seemed to be using these terms in a taxonomic sense rather than ecologically.

[4] An expert free list closest to the average length of 69 was as follows: pignut hickory, shellbark hickory, red oak, white oak, black oak, northern red oak, sweet cherry, sour cherry, black cherry, choke cherry, American beech, European beech, sugar maple, silver maple, Japanese maple, red maple, Eastern white pine, red pine, black pine, mago pine, apple, crabapple, pear, peach, butternut, black walnut, chinese chestnut, American chestnut, hawthorn, cedar - Juniper, cedar, English maple, crimson bark maple, horsechestnut, buckeye, white birch, gray birch, river birch, yellow birch, poplar, red spruce, white spruce, Norway spruce, balsam fir, white fir, douglas fir, locust, honey locust, Canadian hemlock, Carolina hemlock, zelkova, star magnolia, saucer magnolia, Virginia magnolia, black spruce, laurel oak, swamp oak, weeping willow, catalpa, dawn redwood,

---

flowering dogwood, kousa dogwood, black gum, sycamore, American elm, Chinese elm, American arborvitae, and Japanese arborvitae. The "cedar - Juniper" tree was coded as *Juniperus virginiana* because these trees in the Northeast sometimes go by the common name cedar juniper and belong to the Juniper family. The "cedar" was coded as *Cedrus*, belonging to the cedar family. This is another example of how common names may lead to groupings of species which are not very closely related.

[5] A novice free list of the average length of 18 was as follows: willow, oak, maple, elder, pine, evergreen, walnut, apple, pear, cherry, peach, poplar, palm, orange, olive, magnolia, mangrove, and cypress. We coded the "elder" tree as an alder and "evergreen" down to order *Pinales*.

[6] The cluster diagrams very accurately represent the similarity data. The correlations for the novices, experts judging pictures, and experts judging words are .99, .98, and .97 respectively.

[7] The coding of criteria as ecological, morphological, taxonomic, and utilitarian was done separately by Shipman and Boster and their codings combined. Although we do not have a quantitative assessment of the concordance of our codings, our agreement was high.

[8] QAP refers to the Quadratic Assignment Program (Hubert and Schultz 1976). It allows one to assess the similarity of a pair or sets of matrices, where ordinary statistical tests would be rendered invalid by virtue of the non-independence of

---

cases. Here the computation of QAP Z has been supplemented in each case by Monte Carlo permutation tests, in which pairs of matrices are randomly permuted, a similarity measure computed for each trial, and the proportion of times randomly permuted matrices are as or more similar than the actual matrices is calculated. The estimation of probability by Monte Carlo simulation accords reasonably well with that given by a Z distribution.

[9] Insert here: Table 4.

Table 4 shows the means and standard deviations for agreement within and between informant groups (np = novices judging pictures, xp = experts judging pictures, and xw = experts judging words) for the similarity judgment task. Overall, the agreement is higher within groups than between groups.

[10] We also found no difference between experts with advanced degrees and other experts (QAP Z = -1.0, Pearson's r = -0.5, p > .05).