Tall is typical: Central tendency, ideal dimensions, and graded category structure among tree experts and novices

ELIZABETH B. LYNCH, JOHN D. COLEY, and DOUGLAS L. MEDIN
Northwestern University, Evanston, Illinois

Many accounts of categorization equate goodness-of-example with central tendency for common taxonomic categories; the best examples of a category are average members—those that are most similar to most other category members. In the present study, we asked 24 tree experts and 20 novices to rate goodness-of-example for a sample of 48 trees and found (1) that the internal structure of the category *tree* differed between novices and experts and (2) that central tendency did not determine goodness-of-example ratings for either group. For novices, familiarity determined goodness-of-example ratings. For experts, the "ideal" dimensions of height and weediness, rather than average similarity to other trees, were the primary predictors of goodness-of-example ratings for experts. The best examples of *tree* were not species of average height, but of extreme height. The worst examples were the weediest trees. We also found systematic differences in predictors of goodness-of-example as a function of type of expertise. We argue that the internal structure of taxonomic categories can be shaped by goal-related experience and is not necessarily a reflection of the attributional structure of the environment. Implications for models of category structure and category learning are discussed.

All members of a category are not equal. Rather, categories exhibit *graded structure*, in which some members represent the category better than others. For example, for American undergraduates, robins are considered better members of the category *bird* than penguins; and Labrador retrievers, better members of the category *dog* than poodles. Graded structure seems to be a universal property of categories. It has been shown to exist in common taxonomic categories (see, e.g., Rosch & Mervis, 1975; Smith, Shoben, & Rips, 1974), abstract categories (Hampton, 1981), goal-derived categories like *things to eat on a diet* (Barsalou, 1985), ad hoc categories like *things that could fall on your head* (Barsalou, 1983), and even formal categories like *odd number* and *square* (Armstrong, Gleitman, & Gleitman, 1983).

How representative an exemplar is of its category (the typicality of the exemplar) turns out to have an important influence on many cognitive tasks. For example, experimental participants are faster at verifying the category membership of typical exemplars as opposed to atypical exemplars (McCloskey & Glucksberg, 1979; Smith, 1978): people are faster at saying that a robin,

rather than a penguin, is a bird. Typical exemplars are more likely to be generated in category-listing tasks (Barsalou, 1983, 1985; Mervis, Catlin, & Rosch, 1976). Typical category members are learned faster than atypical category members (Mervis & Pani, 1980; Rosch, Simpson, & Miller, 1976). The typicality of a category member also affects its inductive potential (Osherson, Smith, Wilkie, López, & Shafir, 1990; Rips, 1975); people are more likely to think that all members of a category share a novel property attributed to a typical member of the category rather than to an atypical member of the category.

Thus, graded structure is an aspect of concept representation that any model of categorization must address. Not surprisingly, a fair amount of research has been devoted to understanding the determinants of the graded structure of categories. By far the most popular view is that typicality is driven by *central tendency*—that an example is typical to the extent that it is similar to fellow category members and not similar to members of alternative categories. Similarity-based models of categorization provide a natural account of typicality based on central tendency. Rosch and Mervis (1975) showed that an exemplar was rated a better example of a category to the extent that it was similar to other category members and dissimilar to noncategory members. Using feature listings for natural superordinate and basic categories, they showed that category exemplars that had more listed features in common with other category members were also rated as the most typical. In two experiments with artificial categories, Rosch and Mervis again showed that exemplars that shared more features with other category members and fewer with noncategory members were

rated as more typical of the category. These typical examples were also easier to learn and quicker to be identified after learning.

Organizing categories around central tendencies facilitates the general function of categorization: to reflect and emphasize the uneven distribution of attributes in the environment. As Rosch and Mervis (1975) put it,

> if natural categories of concrete objects tend to become organized so as to render the categories maximally discriminable from each other, it follows that the maximum possible cue validity of items within each category will be attained ... The principle of family resemblance relationships [central tendency] can be restated in terms of cue validity since the attributes most distributed among members of a category and least distributed among members of contrasting categories are, by definition, the most valid cues to membership in the category in question. (p. 575)

Thus, the notion of central tendency as the determinant of graded structure is attractive because it is based on the same principles that can be used to explain category learning. That is, both the internal and the external structures of categories serve to maximize within-category similarity and between-category distance.

But central tendency does not determine the graded structure of all categories. Barsalou (1985) found that although central tendency was the major determinant of the graded structure of taxonomic categories, it played no role in the graded structure of goal-derived categories, such as *clothes to wear in the snow*. Rather, the graded structure of goal-derived categories was organized around ideals. For instance, exemplars are better members of the category *clothes to wear in the snow* to the extent that they keep people warmer. The best example of this category (a down jacket) is not the exemplar that is most like other category members; rather, it is the exemplar with the most extreme value on the goal-related dimension. The better the exemplar serves the goal of the category, the more typical the exemplar of the category.

Trait categories also have an internal structure organized around goals or ideals rather than central tendency. Borkenau (1990) showed that the typicality of acts for traits and the similarity between acts and traits, as well as among acts and among traits, were higher to the extent to which the items shared extreme values on dimensions of the similarity space. Similarly, Read, Jones, and Miller (1990) showed that the similarity and typicality ratings of behaviors can be predicted by the extent to which the behaviors achieve the same goal. Chaplin, John, and Goldberg (1988) also found that the typicality of traits and states was organized around ideals, or extreme values on attribute dimensions.

In sum, research on typicality indicates that the determinants of typicality structure vary according to whether the category is goal-based or taxonomic. For goal-based categories, category members are more typical to the extent that they share extreme values on a goal-related dimension. For taxonomic categories, objects are more typical to the extent that they are similar to an average member

of the category. This is not always the case, however. Atran (1999) has observed that among the Itzaj Maya of Peten, Guatemala, the graded structure of taxonomic natural kind categories is determined on the basis of ideals. Using a set of 104 pictures of Central American birds, Atran compared direct ratings of which birds were "true" birds with central tendency ratings defined as average similarity derived from a sorting task. For the Itzaj Maya and for University of Michigan undergraduates, passeriformes (small perching birds, like sparrows) were high in central tendency. The American students also saw passerines as ideal ("true") birds. The Itzaj, however, nominated galliformes (turkeys) as ideal, despite their low central tendency ratings. The Itzaj said that the wild turkey was a "true" bird because it had tasty meat and was extremely beautiful. Thus, for the American participants, central tendency and ideals converged on the same birds; for the Itzaj, they did not.

Atran's (1999) study raises the possibility that the amount or type of experience with category members can affect the internal structure of a category. That is, contrary to the suggestion of Rosch and Mervis (1975), the internal structure of taxonomic categories is not solely a reflection of the correlational structure of the environment. In the present study, we tested the possibility, raised by Atran, that different kinds of experience with a taxonomic category can lead to differences in the determinants of typicality. We compared the graded structure of the category *tree* for three kinds of Chicago area tree professionals and a group of American college students. The three types of tree professionals were taxonomists, landscapers, and parks maintenance personnel. We were interested in whether the determinants of the graded structure of the taxonomic category *tree* would differ for novices who had seen millions of trees but had little goal-related experience with them and for tree experts, a population that has salient goals in relation to the category as well as extensive experience with its attributional structure. Further, we were interested in whether or not the different types of goals that characterized our three groups of tree experts would differentially affect graded structure.

We knew from our previous research on experts' categorization and reasoning about trees (Medin, Lynch, Coley, & Atran, 1997) that experts organize their knowledge about trees along both goal-related and morphological dimensions. In that study, the distribution of goal-related and taxonomic groupings varied systematically across our expert groups. The predominant type of grouping across experts was not goal-related, but taxonomic (or morphological). The landscapers displayed the greatest number of goal-related groupings on a free sorting task, whereas the taxonomists displayed no goal-related groupings and the maintenance workers' primary goal-related grouping was *weed trees*. *Weed trees* was the most common nontaxonomic grouping among the experts in our free categorization (sorting) task.[1] The weediness of a tree is an impediment to a maintenance

worker or landscaper who is trying to keep a neat and organized landscape. Weed trees tend to be fast growing, with weak wood, and they require high maintenance. In addition, the landscapers produced other goal-related groupings in the sorting task, such as *street trees*, *shade trees*, and *ornamental trees*.

In the present study, the attributional structure of trees was available to our novice sample, but they probably had virtually no goal-relevant experience with trees (except, perhaps, for having climbed them as children). We assumed that any knowledge that the novice group possessed about trees had been acquired passively.

In the present study, we investigated whether goal-related experience would affect the graded structure of the category *tree* by comparing novice and expert typicality ratings of trees. Further, including different types of experts enabled us to compare the ways in which different types of experience affect graded structure. To answer these questions, we asked the novices and the three groups of tree experts to rate the extent to which they thought 48 tree species found in the Chicago area were good examples of the category *tree*. We then investigated the degree to which several possible predictors of tree typicality, including *central tendency*, *familiarity*, *height* (a potential ideal dimension), *weediness* (a goal-related dimension, and a potential "negative ideal"), and *frequency* actually predicted these ratings.

Given the definition above, in the present study central tendency was a measure of the average similarity of each tree to each other tree. We tested central tendency measures derived from both the Medin et al. (1997) sorting task (which incorporates goal-relevant dimensions) and scientific central tendency, which measures the similarity of trees according to scientific classification. Trees prototypical by virtue of being similar to many other trees tend to have high central tendency scores. We also tested the goal-related dimension of weediness. As discussed above, trees that exhibit any of a cluster of undesirable properties can be considered weed trees. Our previous work had suggested that this was the most salient of a set of functional, goal-related dimensions that might be present in experts' tree classifications (Medin et al., 1997). Thus, to the degree that such considerations might influence graded category structure, any tree species that exhibited these undesirable properties would be considered a poorer example of *tree*. In addition, the frequency with which a category instance occurs has been shown to be related to perceived typicality (e.g., Barsalou, 1985; see also Boster, 1988). Another variable that we thought might influence goodness of example ratings, especially for the novices, was familiarity (Malt & Smith, 1982; Worthen & Nakamura, 1995). Finally, post hoc observations suggested that the characteristic height of a tree species should have a strong influence on its representativeness.

If, as previous research has shown, overall attributional structure is the major determinant of the graded structure of taxonomic categories, with goals only peripherally affecting category structure, we could expect central tendency to be the major predictor of rated typicality of trees for both novices and experts. On the other hand, experience might render goal-related attributes like weediness salient enough to override central tendency as the structural basis even of taxonomic categories. Because the tree experts in this study were highly knowledgeable about both the attributional structure of the category *tree* and its many functional roles, they were an excellent population with which to determine both the role of perceptual attributes on the graded structure of a taxonomic category and the extent to which goals could affect graded structure.

METHOD

Participants

Our participants were 24 tree experts from the Chicago area and 20 Northwestern University undergraduates participating for class credit. The tree experts ranged in age from 27 to 69 and worked mainly with trees in their primary occupations. Our sample consisted of three types of tree experts: taxonomists working mainly in research and education ($n = 5$), landscapers who focus on placing trees in appropriate settings ($n = 9$), and parks maintenance personnel who plant new trees, prune, treat disease, and remove dead or dying trees ($n = 10$). All expert participants were involved in a larger, long-term study of categorization and reasoning among tree experts.

Materials

Two familiarity scales and a typicality scale were created, using the tree species listed in Table 1. Approximately 80 species of trees are native to the Chicago area, and another 30–40 species have been introduced. After consulting with numerous reference sources and a few experts on local trees, we selected 48 tree species, including 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 the Swink and Wilhelm (1994) description of plants of the greater Chicago area. The set of 48 also included a few landscaping trees that would rarely if ever appear on city streets. The 48 trees were chosen to represent a broad spectrum of scientific taxa and to include variations in patterns of folk nomenclature. Consensual patterns of sorting were used to establish separate measures of central tendency that were specific to each expert group (see Medin et al., 1997).

Procedure

Goodness-of-example (GOE) ratings with respect to the category *tree* were collected from each participant for all 48 tree species. The instructions given to the participants in this study were identical to those used by Rosch and Mervis (1975). These instructions told the participants to rate how good an example each tree species was of the category *tree* on a seven-point scale, with 1 representing a *very good example* of the category and 7 representing a *very poor example* of the category or *not a member* of the category.

We also gathered familiarity ratings for the 48 tree species in two ways. Both familiarity scales asked each participant to rate the familiarity of each tree on a seven point scale, with 1 indicating *very familiar* and 7 indicating *unfamiliar*. The scales differed in their axis labels. One scale (ID familiarity), designed primarily for undergraduates, defined high familiarity as the ability to identify the tree (a 1 on the scale), while lack of familiarity meant they had never heard the name of the tree before (a 7 on the scale). This scale was completed by the full sample of undergraduates and a subset of the tree expert sample ($n = 19$). The second scale, which was completed only by the subset of tree experts, measured familiarity with

Table 1
Tree Species Used, With Goodness-of-Example Rankings and Predictor Values

Tree Species	Expert GOE*	Novice GOE	Novice ID Fam	Expert ID Fam	Expert Work Fam	Expert Folk CT†	Height	Weed	Freq‡
American beech	13	25	0.26	−0.32	−0.96	−0.07	60	0	1.68
American elder	47	33	−0.47	−1.62	−1.41	0.1	9	3	0
American elm	5	3	0.82	0.26	0.90	−0.02	70	1	3.77
American hornbeam	40	39	−0.97	−0.74	−0.52	−0.05	25	0	1.72
American mountain ash	38	31	−0.46	−0.21	−0.65	0.01	20	1	0.7
American sycamore	12	10	0.66	0.28	0.29	−0.1	63	2	2.38
Amur maple	46	1	0.41	0.08	0.05	−0.08	17	0	1.46
Austrian pine	26	8	0.45	0.04	0.23	0.24	55	0	1.6
Black cherry	35	26	0.47	−0.57	−0.45	0.06	55	3	0.78
Black walnut	7	17	0.55	−0.06	−0.49	−0.05	63	0	1.08
Box elder	45	35	−0.13	0.67	0.75	−0.02	40	13	2.07
Bur oak	3	6	0.65	0.11	0.28	−0.17	75	0	2.04
Catalpa	31	41	−0.92	0.53	−0.34	−0.07	50	6	1.91
Colorado spruce	19	16	0.32	0.19	0.33	0.3	45	0	0
E. cottonwood	29	32	−0.20	0.53	0.83	−0.02	88	11	1.82
E. redcedar	42	18	0.06	−0.45	−0.24	0.31	45	1	1.57
E. white pine	17	12	0.53	−0.12	−0.02	0.29	65	0	1.93
European black alder	44	42	−0.84	−0.39	−0.57	0.02	30	0	1.72
Ginkgo	21	45	−0.52	0.75	0.20	0.15	65	0	2.33
Green ash	10	36	−0.64	−0.17	0.57	−0.09	55	0	3.28
Grey dogwood	48	34	−0.41	−0.97	−0.92	0.11	13	0	0
Hackberry	23	40	−0.95	0.41	0.42	−0.05	50	0	3.03
Honey locust	14	37	−0.82	0.67	1.13	−0.03	50	0	3.34
Horsechestnut	11	28	−0.58	0.37	−0.14	−0.04	63	0	1.96
Kentucky coffee	28	47	−0.73	0.54	−0.08	−0.03	68	0	2.59
Little-leaf linden	22	46	−0.79	0.50	0.59	0.04	85	0	2.05
London planetree	20	44	−0.97	−0.49	−0.15	−0.07	65	0	1.2
N. red oak	4	2	0.72	−0.46	0.23	−0.18	45	0	3.4
Norway maple	6	11	0.35	0.36	0.97	−0.14	68	0	2.62
Ohio buckeye	8	27	0.45	0.27	−0.29	−0.05	30	0	1.99
Paper birch	32	24	0.36	0.09	−0.51	0.06	60	2	0.48
Pin oak	15	4	0.83	−0.05	0.12	−0.17	65	2	2.2
Scotch pine	30	9	0.93	−0.48	−0.08	0.26	45	0	1.69
River birch	25	22	0.09	0.47	0.05	0.03	55	0	0.48
Shagbark hickory	16	20	−0.16	0.34	−0.53	−0.05	70	0	1.08
Siberian elm	33	15	−0.08	0.35	0.57	0.01	60	11	2.34
Silver maple	27	14	0.63	0.63	0.95	−0.14	60	3	2.89
Star magnolia	39	30	−0.51	−0.41	−0.71	0.08	18	0	1
Sugar maple	2	7	0.78	0.18	0.54	−0.16	68	0	3.07
Sweet crabapple	34	19	0.71	−1.34	−0.46	0.02	20	0	2.53
Tree of heaven	41	43	−0.79	0.54	−0.17	0.06	50	14	1.49
Tuliptree	18	48	−0.68	0.35	−0.53	0.04	43	1	2.07
Washington hawthorne	36	38	−0.74	−0.59	−0.16	0.06	28	0	2.61
Weeping willow	24	13	1.68	0.51	0.14	0.06	35	10	2.49
White ash	9	23	−0.05	−0.19	0.43	−0.09	65	0	0.7
White mulberry	43	29	−0.24	−0.42	−0.05	0.09	40	12	0
White oak	1	5	0.93	0.04	0.35	−0.16	65	0	0
White poplar	37	21	0.25	0.21	−0.48	0.02	55	11	2.27

Note—GOE, goodness-of-example; ID Fam, familiarity based on ability to identify tree; Work Fam, familiarity based on how frequently participant worked with tree; CT, central tendency. *GOE ratings were standardized for each participant and averaged across all participants for each species, and the mean rank GOE scores are presented here. †Central tendency scores were standardized and multiplied by −1 (high scores equal low average sorting distance to other tree species). ‡Raw frequencies (number of individual trees) are reported. In analyses, log frequency scores were used.

the trees in terms of how often they worked with each tree: a rating of 1 indicated that the participant worked with the tree *very frequently* and a 7 indicated that the participant worked with the tree *rarely*. The second scale was designed to better approximate familiarity among the experts, who were expected to be able to identify most of the trees.

Predictors of Goodness-of-Example

We used five predictors of typicality: central tendency, familiarity, weediness, height, and frequency. The values of these predictors for each tree species are given in Table 1.

Central tendency.[2] We measured central tendency in two ways. Our first measure comprised the expert central tendency and expert group specific central tendency measures from the hierarchical sorting task of Medin et al. (1997). Twenty-four tree experts, including the entire sample used in the present study, performed a successive pile sorting task using name cards representing the same 48 trees used in the present experiment. The participants were asked to "put together the trees that go together by nature into as many groups as you like." After the initial sort, the participants were asked to successively group their initial piles of trees into higher level groups. Following the higher order grouping phase, the par-

ticipants were invited to divide the original groups into subgroups. The participants were asked to explain their groups at each level.

The sorting task yielded a taxonomy of trees for each participant. We assumed that the proximity of the trees in an individual's taxonomy represented the similarity of those trees for that individual. From these taxonomies we derived pairwise distances—as measured by the number of links that must be traversed in the hierarchy in order to reach a common link—between the 48 trees. The measure of central tendency used in this study is the standardized mean distance of each tree from every other tree averaged across all participants (overall expert central tendency) and averaged across expert subgroups (expert group specific central tendency). The Cultural Consensus Model (Romney, Weller, & Batchelder, 1986) verified that there was consensus among the experts as to the similarities among the trees. (See Medin et al., 1997, for further details.)

Our second measure of central tendency was scientific central tendency—the pairwise distance of each tree from every other tree in the scientific taxonomy. The central tendency measures above incorporated experience-derived goal-relevant information to some degree (except in the case of the taxonomists, whose sorts correlated .97 with scientific taxonomy) and so might not be appropriate predictors of graded structure for novices. The scientific central tendency measure focused on overlap of attributes. Research has shown that scientific central tendency is a fairly good measure of similarity relations among biological organisms (Berlin, 1992; Boster, 1987, 1988; Malt, 1995). Scientific central tendency served as a measure of tree central tendency for the novices, for whom we did not have a direct measure of tree similarity.

Weediness. Our measure of weediness, like our measure of central tendency, was derived from the sorting task in Medin et al. (1997). For each tree we derived a weediness score by summing the number of times that it was placed in an explicitly identified *weed tree* category by an expert.[3] Although other goal-related groupings emerged for some of our experts (landscapers especially), their frequencies were too low to serve as reliable predictors of typicality.

Height. Our measure of *height* was derived from a number of different field guides to Chicago-area trees (Dirr, 1978; Little, 1980; Petrides, 1988). Where a range was given for characteristic adult height, we used the midpoint and then averaged estimates across sources.

Frequency. To index frequency, we summed recent tree inventories from two Chicago-area municipalities (Evanston and Wilmette). These inventories were limited because they included trees only in the parkway (the area between the street and the sidewalk) and thus did not include trees in parks or on private property, and because not all of our experts lived or worked in these areas. However, this measure provided at least a rough index of frequency-of-occurrence for each tree species in our sample. Because the populations for various species often differed by several orders of magnitude, we subjected these data to log transformations before entering them into the correlations reported below.

RESULTS

GOE ratings were standardized for each participant to account for differences in their use of the 7-point scale. Table 1 presents rank GOE scores for experts and novices.

First, we will discuss raw and partial correlations of each of our variables with novice and expert average GOE ratings. Then we will compare predictors of GOE across our three expert subgroups.

Expert and Novice Correlations

Raw correlations are presented in the upper right-hand section of Table 2, which shows large differences in the predictors of novice and expert GOE ratings. Expert GOE ratings are highly correlated with all predictors except scientific central tendency, whereas novice GOE ratings are highly correlated with familiarity and mildly correlated with expert work familiarity. Height accounts for the most variance in expert GOE ratings, with the other factors accounting for slightly less.

Because of intercorrelations among predictors, second-order partial correlations were performed to assess the relation between GOE and each of the predictors with the other predictors held constant (following Barsalou, 1985). The second-order partial correlations for each group are shown in the first two columns of Table 2.

As can be seen in Table 2, when other factors are held constant, weediness and height explain much more variance in expert GOE ratings (59% and 54%, respectively) than does central tendency (35%). The other factors do not predict a reliable amount of variance. Thus, for the expert group as a whole, though central tendency informs tree typicality to a significant extent, the graded structure of trees is primarily organized around the ideals of height and weediness. A multiple regression analysis confirmed the results of the partial correlations. When expert central tendency, weed, height, frequency, and the two familiarity measures were entered into the analysis,

Table 2
Partial and Raw Correlations Among Indices for All Participants

| | Partial | | Raw | | | | | | | | |
	Expert GOE	Novice GOE	Novice GOE	Expert CT	Science CT	Novice ID Fam	Expert ID Fam	Expert Work Fam	Height	Weed	Freq
Expert GOE			.38†	−.44†	−.06	.38†	.39†	.47†	.65‡	−.40†	.37†
Novice GOE				−.11	.04	.88‡	.00	.29*	.17	−.10	.08
CT folk (expert)	−.35*				.82‡	−.06	−.20	−.27	−.26	.05	−.34
CT science	−.24	−.02				.04	−.02	−.01	.03	−.13	−.12
Nov ID Fam		.88‡					.00	.19	.15	.00	.07
Exp ID Fam	.11							.62‡	.58‡	.26	.36*
Exp Work Fam	.14								.55‡	.12	.54‡
Height	.54‡	.10								.02	.26
Weediness	−.59‡	−.21									−.03
Frequency	.09	.02									

Note—GOE, goodness-of-example; ID Fam, familiarity based on ability to identify tree; Work Fam, familiarity based on how frequently participant worked with tree; CT, central tendency. *p < .05. †p < .01. ‡p < .001.

the regression was reliable ($R = .83, R^2 = .69, F = 15.4$, $p < .001$) with expert central tendency, height, and weed as predictors. When expert central tendency was replaced by scientific central tendency, the regression accounted for virtually the same amount of variance ($R = .82, R^2 = .66$), but scientific central tendency did not emerge as a predictor. In fact, when central tendency was removed from the analysis, the amount of variance accounted for remained about the same ($R = .81, R^2 = .65$).

For novices, GOE ratings are accounted for almost entirely by familiarity. None of the other factors predict any reliable GOE rating variance. Again a multiple regression analysis was performed including scientific central tendency, weediness, frequency, height, and the novice familiarity measure. The regression was reliable ($R = .89$, $R^2 = .79, F = 31.6, p < .0001$) with only the familiarity measure as a reliable predictor.

Differences in Perceived Typicality by Type of Expertise

The previous analyses were based on GOE ratings averaged across all experts. We were interested in whether different types of professional experience (i.e., landscaping, tree maintenance, and taxonomy) would have differential effects on the internal structure of *tree*. Thus, we analyzed determinants of typicality separately for each of our three tree professional subgroups. Table 3 shows the partial correlations for each expert subgroup, as defined by Medin et al. (1997).

For the maintenance personnel weediness ($r = -.55$), and to a lesser degree, folk central tendency ($r = -.31$) predicted GOE ratings. All factors were entered into a multiple regression ($R = .72, R^2 = .52, F = 7.33, p < .001$) which confirmed that only weed and to a lesser extent folk central tendency were reliable predictors. The correlation between maintenance worker and landscaper GOE ratings was .75, and between maintenance worker and taxonomist, .76.

Landscapers and taxonomists exhibited a different pattern. For both groups, height ($r_l = .57, r_t = .60$) and weediness ($r_l = -.45, r_t = -.50$) predicted GOE ratings (see Table 3). Again, multiple regressions confirmed these partial correlations (landscapers, $R = .81, R^2 = .66, F =$

Table 3
Partial Correlations With GOE for Subgroups, for All Trees and for Nonweed Trees

	Maintenance GOE	Taxonomist GOE	Landscaper GOE
Subgroup CT[a]	−.31*	−.09	−.22
CT science[b]	−.12	−.09	−.16
Expert ID Fam	.12	.19	.09
Expert Work Fam	.06	.08	.14
Height	.26	.60‡	.57‡
Weediness	−.55‡	−.50‡	−.45†
Frequency	.04	.18	.14

Note—GOE, goodness-of-example; ID Fam, familiarity based on ability to identify tree; Work Fam, familiarity based on how frequently participant worked with tree; CT, central tendency.   *$p < .05$.   †$p < .01$.   ‡$p < .001$.   [a]This correlation does not control for subgroup CT.   [b]This correlation does not control for scientific CT.

13.02, $p < .0001$; taxonomists, $R = .79, R^2 = .61, F = 10.55, p < .0001$).[4] The correlation between landscaper and taxonomist GOE rankings was .81.

DISCUSSION

In this study we examined the extent to which familiarity, central tendency (both folk and scientific), frequency of occurrence, height, and perception of trees as "weeds" influenced goodness-of-example ratings collected from novices and three kinds of tree experts. We observed three major results. First, we found that different factors determined the internal structure of the category tree for novices and experts. Thus, we found an "expertise effect" on typicality structure. These findings are consistent with Atran's (1999) cross-cultural findings. Second, we found that, although *tree* is a taxonomic category with a coherent and visible attributional structure, central tendency was not the primary determinant of graded structure for either novices or experts. Ideal dimensions—rather than central tendency—were primary determinants of graded category structure for experts. Third, we found that the determinants of graded structure varied among our expert subgroups. Below we discuss these findings in turn and then consider their implications for models of category structure and acquisition.

Expertise Effects on Graded Category Membership

The determinants of graded category membership differed for novices and experts. For novices, the extent to which they were familiar with a tree emerged as the sole determinant of its prototypicality. For experts, the story was more complicated. Overall, tree typicality was determined primarily by ideals on the dimensions of height and weediness and secondarily by folk central tendency. Thus, knowledge of trees appears to bring about a qualitative shift in the internal organization of the category.

Malt and Smith (1982) found that subjects tended to produce fewer properties for atypical words, indicating that atypicality may be determined in part by familiarity. However, they found that the relationship between familiarity and typicality might be due in part to a response bias in which subjects give words low typicality ratings because they are unfamiliar, rather than because they perceive the referent itself to be atypical of the category. Consistent with this, Worthen and Nakamura (1995) found that familiarity predicted typicality ratings better than central tendency did for a low knowledge group, whereas the two factors predicted typicality equally well for a higher knowledge group. These studies suggest that level of familiarity in itself does not determine novice tree GOE ratings. Rather, familiarity may be a marker of something else—say, cultural significance—that more directly affects the graded structure of tree categories.

As a group the maples and oaks are the most typical trees for the novices. This is not surprising, and it may be due to the prevalence in American culture of products deriving from these trees (e.g., maple syrup; oak floors

and furniture). Revealing cases indicate that item familiarity is not the sole determinant of tree typicality. An interesting case is the Amur maple, of which few novices had ever heard (ranked 17 in familiarity) but which they ranked first in typicality, along with the other highly ranked maples. This may be related to Coley, Medin, and Atran's (1997) finding that novices assume that properties of trees cohere at the genus level (e.g., maple), so that an unknown maple is bound to be very similar to better known maples.

Other interesting cases are the weeping willow, which was ranked 1 in familiarity but only 13th in typicality, and the sweet crabapple, which was quite familiar to the novices (ranked 8) but not very typical. These two cases suggest some influence of perceptual characteristics on novice tree graded structure. Both trees are somewhat unique perceptually, the weeping willow with its hanging wispy branches, and the sweet crabapple with its flowers and small stature. In short, although these trees are relatively familiar, their distinctiveness makes them atypical examples of *tree*.

Given the examples above, along with research suggesting that scientific central tendency is a good measure of naive similarity judgments among biological organisms (Berlin, 1992; Boster, 1987, 1988; Malt, 1995), it is interesting that scientific central tendency did not have any influence on the graded structure of the novice tree category. Perhaps if we had measured novice similarity directly (by showing pictures to the participants), we would have found an influence of central tendency on typicality ratings. However, it is difficult to measure novice similarity of unfamiliar trees because trees are not easily represented visually. It is not clear what part of the tree should be in the photograph, and perhaps more importantly, tree form can vary dramatically as a function of environmental conditions.

In sum, the novice results are difficult to interpret because of the novices' lack of familiarity with the domain. Interestingly, although trees are salient and ubiquitous objects, most people in our culture show a surprising lack of knowledge about them. Despite this lack of knowledge and the consequent difficulty of testing novices, trees are an interesting domain in which to study the effects of goals on taxonomic structure. Trees are natural, rather than constructed, objects. They are not designed to achieve one particular goal, but are complex and linked to a variety of goals and knowledge structures in our expert population. The performance of our expert sample suggests some important ways in which knowledge and goals might affect the internal structure of this complex category.

Ideals as Determinants of Graded Category Membership

This study is the first to demonstrate that ideals can serve as the *primary* determinants of the graded structure of a natural kind taxonomic category. Although Barsalou (1985) found that ideals influenced the graded

structure of both goal-derived and taxonomic categories, central tendency was always a stronger predictor for taxonomic categories. Our results, however, indicate that for experts, ideals are consistently the strongest predictor of GOE ratings for trees. That is, the most representative trees—those which are the best examples of the category *tree*—are not trees of average height, but rather trees of extraordinary height. Likewise, weediness served as a negative ideal for the experts. Thus, for the experts in our sample, *tree* appears to be structured around positive and negative ideal dimensions rather than the average of several relevant dimensions.

Weediness as negative ideal. Barsalou (1985) demonstrated that ideals for goal-derived categories are defined in terms of goals. Our finding raises the possibility that goals influence the graded structure of taxonomic categories as well. The negative ideal of weediness certainly fits this description. A tree is weedy to the extent that it causes problems for humans.[5] That is, from a maintenance standpoint, weed trees may have weak wood (and therefore drop branches), or may have large leaves, flowers, or fruit that drop and clog gutters, stain sidewalks, and generally create a mess. Often weed trees grow easily from roots and therefore pop up where they are unwanted. For example, the tree of heaven, one of the most pernicious weed trees, grows everywhere, including cracks in sidewalks (as in *A Tree Grows in Brooklyn*). Rarely planted due to its weak wood which can be cut with a pocket knife, it is never considered a desirable tree (except, perhaps, in Brooklyn).

Barsalou (1985) claims that "ideals may often be determined independently of exemplars, being acquired through the process of planning how to achieve goals before exemplars are ever encountered" (p. 631). This explanation of ideals as being goal-related seems appropriate in the case of tree weediness. Some of our experts think of trees in terms of which kinds of trees can best fulfill some particular goal. For example, landscapers must satisfy a number of constraints when deciding what kind of tree to plant in a particular part of a landscape. Similarly, when deciding which trees to plant along the street of a city, it is important to plant trees that do not interfere with electrical wires, are salt-tolerant (in Chicago at least), and are not too messy. Therefore, in the case of weediness, as Barsalou states above, the ideal may come before the exemplars.

Height as positive ideal. In contrast to weediness, it is not clear what expert goal is facilitated by tree height. One possibility is that increased height allows a tree to provide better shade. Although this is plausible, shade does not seem to be a major concern of our tree experts. It has rarely been mentioned by our tree experts as a dimension of concern, so it would be odd for it to play such a major role in organizing the tree category.

If height does not help tree experts achieve any goal, why does height in particular organize the category *tree*? Organizing the category *tree* around the ideal of height may increase the distinctiveness of *tree* relative to other

related categories (e.g, *shrub, plant*). Hunn (1999) suggests that *tree* differs from other high-level natural kind categories such as *bird* and *fish* in that it is not defined in nature by a clear perceptual gap. That is, large trees fade imperceptibly into small trees, which then fade into shrubs. We found, however, that the correlation between height and typicality was evident throughout the full range of heights and not simply near some presumed transition point between shrubs and trees. Still, the fact that height distinguishes trees from other plants may render it an essential, or defining dimension of "treeness."

Thus, there are reasons why height in particular might be an ideal dimension and help to organize the concept *tree*. In general, however, we do not believe that *tree* is idiosyncratic in being a taxonomic category organized around ideals. Furthermore, the degree to which ideals organize taxonomic categories may increase with expertise. Atran (1999) argues that the Itzaj, who are familiar with Central American birds, organize bird categories around ideals, whereas American undergraduates unfamiliar with these birds do not. Moreover, K. E. Johnson (personal communication, June 26, 1997) has also found evidence that American bird experts are influenced by ideal dimensions. Taken together, these findings hint that experts' categories may be organized around ideals more than novices', and that, more generally, taxonomic categories may be organized around ideals to a much larger degree than previously believed.

Expert Group Differences in Determinants of Typicality

The third interesting finding is that the experts were not uniform in the manner in which they organized their *tree* category. We found that the graded structure of "tree" for the maintenance workers was determined mainly by weediness and to a lesser extent folk central tendency. The taxonomists and the landscapers both organized their category *tree* around the ideal dimensions of height and weediness. These results are consistent with Medin et al. (1997), who have shown that the maintenance workers use the same organizational structure in two different tasks, whereas the landscapers use two different structures. That is, one could interpret the influence of folk central tendency on the graded structure of the maintenance worker tree category as a reflection of the same organizational structure that underlies the maintenance worker sorting and reasoning. The landscapers, on the other hand, used goal-derived dimensions on the sorting task (including weediness and size) but scientifically relevant dimensions to reason about novel tree properties. The present study shows that the internal structure of the landscapers' tree category is organized around goal relevant dimensions. The taxonomists, like the landscapers, organize the structure of the tree category around the dimensions of weediness and height. Attentiveness to these dimensions by the taxonomists was not demonstrated on either of the Medin et al. tasks. Thus, the present study provides further evidence for the characterization of our

expert groups whereby the maintenance workers rely on a single organizational structure across multiple tasks, while the landscapers, and this time, the taxonomists, employ different types of tree organization, depending on the task.

Implications

We have shown that ideal dimensions may outweigh central tendency as determinants of graded category membership in common taxonomic categories. This finding has important implications for models of categorization and for conceptual acquisition and development.

Similarity-based models of categorization frequently equate graded category membership, typicality, and central tendency. For example, in their investigations of the effect of typicality on category-based induction, Osherson et al. (1990) inferred typicality from participants' similarity judgments. They assumed that the categories with the highest central tendency were the most typical examples of the category. We have shown that for a diverse group of tree experts central tendency is not the primary determinant of graded category membership, and that for a subgroup of those experts, it plays almost no role at all. Rather, both height and weediness are dimensions that function like ideals in organizing the graded structure of trees. This suggests that experts are not merely computing similarity over members of the category *tree*, but instead considering specific dimensions that are important for that particular category—some of which may be goal-derived, some not—and choosing instances that maximize these ideals as good instances of *tree*. It is important not to overstate our case; central tendency was not irrelevant to graded category membership. Still, models that equate typicality with central tendency must take into account the fact that extreme, rather than average, values on salient dimensions may better predict goodness of example.

The importance of ideals in determining graded category structure also has implications for category acquisition. One question raised is at what point in the learning process ideal dimensions play a role. The answer may well vary with the nature of the ideal dimension. A goal-derived (negative) ideal dimension such as weediness may depend on prior knowledge of category instances in order to be activated as an organizing principle. One must learn a certain amount about what kinds of dimensions tree species vary on, and to what ends people plant trees, to appreciate what makes a tree a weed. Indeed, coming to consider a dimension such as weediness as relevant to a category might involve a certain amount of reorganization of existing knowledge about the category. Solomon (1997) presents similar evidence for reorganization of knowledge among wine experts. In contrast, as we argue above, height is a perceptually available ideal that may help distinguish trees from other organisms, and so may be useful as a organizing principle early in the learning process. Height was not correlated with undergraduate tree GOE ratings in the present study. How-

ever, this may be because the undergraduates were not familiar enough with the trees to know their height. One clue that height may play a role for undergraduates is the fact that the sweet crabapple was highly familiar, yet received a low GOE ranking. We suspect that this might have been due to its height. In sum, ideal dimensions may play a role throughout the learning process, both by guiding initial category learning and by reorganizing expert knowledge.

Finally, we have been referring to the category tree as taxonomic because it is a natural category with a coherent attributional structure. Average similarity among members of the tree category *could have* served to distinguish trees from related categories, which is not the case for goal-derived categories. In what have commonly been called goal-derived categories (e.g. traits, *things to remove from your house in case of a fire*), the attributional structure of the members of the category could not serve to distinguish that category from others. Goal-related dimensions provide the necessary basis of similarity for those highly attributionally dissimilar category members. All of our expert groups have demonstrated that goal-related dimensions structure the *tree* category at least on some tasks. Importantly, this study provides important evidence that goals do not serve as the basis for category organization only in the absence of attributional similarity.

Conclusions

We have presented evidence that the ideal dimensions of height and weediness play a stronger role in predicting graded category structure among tree experts than does central tendency. We have also presented evidence for systematic differences regarding which dimensions predict goodness-of-example for *tree*. These results suggest that ideal dimensions are important in determining category structure not only for goal-derived categories lacking correlated attributes, but also for taxonomic categories with strong attribute structure. Given these results, current models of categorization and conceptual acquisition that equate goodness-of-example with central tendency may well misrepresent a basic determinant of human category structure.

REFERENCES

ARMSTRONG, S. L., GLEITMAN, L. R., & GLEITMAN, H. (1983). What some concepts might not be. *Cognition*, 13, 263-308.

ATRAN, S. (1999). Itzaj Maya folkbiological taxonomy. In D. L. Medin & S. Atran (Eds.), *Folkbiology* (pp. 119-204). Cambridge, MA: MIT Press.

BARSALOU, L. W. (1983). Ad hoc categories. *Memory & Cognition*, 11, 211-227.

BARSALOU, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 11, 629-649.

BERLIN, B. (1992). *Ethnobiological classification: Principles of categorization of plants and animals in traditional societies*. Princeton, NJ: Princeton University Press.

BORKENAU, P. (1990). Traits as ideal-based and goal-derived social categories. *Journal of Personality & Social Psychology*, 58, 381-396.

BOSTER, J. S. (1987). Agreement between biological classification systems is not dependent on cultural transmission. *American Anthropologist*, 89, 914-919.

BOSTER, J. S. (1988). Natural sources of internal category structure: Typicality, familiarity, and similarity of birds. *Memory & Cognition*, 16, 258-270.

CHAPLIN, W. G., JOHN, O. P., & GOLDBERG, L. R. (1988). Conceptions of states and traits: Dimensional attributes with ideals as prototypes. *Journal of Personality & Social Psychology*, 54, 541-557.

COLEY, J. D., MEDIN, D. M., & ATRAN, S. (1997). Does privilege have its rank? Inductive inferences within folkbiological taxonomies. *Cognition*, 63, 73-112.

DIRR, M. A. (1978). *Photographic manual of woody landscape plants*. Champaign, IL: Stipes.

HAMPTON, J. A. (1981). An investigation of the nature of abstract concepts. *Memory & Cognition*, 9, 149-156.

HUNN, E. (1999). Size as limiting the recognition of biodiversity in folkbiological classifications: One of four factors governing the cultural recognition of biological taxa. In D. L. Medin & S. Atran (Eds.), (pp. 47-69). *Folkbiology*. Cambridge, MA: MIT Press.

LITTLE, E. L. (1980). *The Audubon Society field guide to North American trees: Eastern region*. New York: Knopf.

MALT, B. C. (1995). Category coherence in cross-cultural perspective. *Cognitive Psychology*, 29, 84-148.

MALT, B. C., & SMITH, E. (1982). The role of familiarity in determining typicality. *Memory & Cognition*, 10, 69-75.

McCLOSKEY, M., & GLUCKSBERG, S. (1979). Decision processes in verifying category membership statements: Implications for models of semantic memory. *Cognitive Psychology*, 11, 1-37.

MEDIN, D. L., LYNCH, E. B., COLEY, J. D, & ATRAN, S. (1997). Categorization and reasoning among tree experts: Do all roads lead to Rome? *Cognitive Psychology*, 32, 49-96.

MERVIS, C. B., CATLIN, J., & ROSCH, E. (1976). Relationships among goodness-of-example, category norms, and word frequency. *Bulletin of the Psychonomic Society*, 7, 283-284.

MERVIS, C. B., & PANI, J. R. (1980). Acquisition of basic object categories. *Cognitive Psychology*, 12, 496-522.

OSHERSON, D. N., SMITH, E. E., WILKIE, O., LÓPEZ, A., & SHAFIR, E. (1990). Category-based induction. *Psychological Review*, 97, 185-200.

PETRIDES, G. A. (1988.) *Peterson field guides: Eastern trees*. Boston: Houghton Mifflin.

READ, S. J., JONES, D. K., & MILLER, L. C. (1990). Traits as goal-based categories: The importance of goals in the coherence of dispositional categories. *Journal of Personality & Social Psychology*, 58, 1048-1061.

RIPS, L. J. (1975). Inductive judgements about natural categories. *Journal of Verbal Learning & 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., & MERVIS, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7, 573-605.

ROSCH, E., SIMPSON, C., & MILLER, R. S. (1976). Structural bases of typicality effects. *Journal of Experimental Psychology: Human Perception & Performance*, 2, 491-502.

SMITH, E. E. (1978). Theories of semantic memory. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 6, pp. 1-56). Potomac, MD: Erlbaum.

SMITH, E. E., SHOBEN, E. J., & RIPS, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, 81, 214-241.

SOLOMON, G. E. A. (1997). Conceptual change and wine expertise. *Journal of the Learning Sciences*, 6, 41-60.

SWINK, F., & WILHELM, G. (1994). *Plants of the Chicago region*. Indianapolis: Indiana Academy of Science.

WORTHEN, J. B., & NAKAMURA, G. V. (1995). Roles of familiarity and family resemblance in determining representativeness in social categories. *American Journal of Psychology*, 108, 195-211.

NOTES

1. In general, weed trees are weak-wooded trees whose limbs tend to break in the wind, often causing them to have an asymmetric and un-

aesthetic overall shape. In addition they may be "dirty," often dropping fruits (e.g., mulberry) and large leaves (e.g., catalpa) that clutter the sidewalks and streets and clog gutters. Weed trees may also reproduce easily and grow where they are not wanted (e.g., box elder, tree of heaven).

2. Boster (1988) found that similarity based on scientific relationships predicted the graded structure of *bird* for a taxonomist, a South American indigenous population, and a group of North American undergraduates. Thus, there is some reason to suspect that ratings of scientific central tendency (i.e., mean similarity in a scientific taxonomy) might predict goodness-of-example ratings. We initially included such a measure, but ultimately eliminated it because (1) it was significantly correlated with "folk" central tendency ($r = .54, p < .001$), and (2) it did not predict goodness-of-example on any analysis.

3. Not every expert produced a *weed tree* category.

4. All expert subgroup regressions were repeated replacing the subgroup central tendency measure from Medin et al. (1997) with the scientific central tendency measure. This made little difference in the amount of variance explained for any group. Although the folk central tendency measure was a reliable predictor of maintenance personnel GOE ratings, scientific central tendency was not.

5. Incidentally, ecologists, a group of tree experts not included in this study, do not agree with the experts in the present study about which trees are weeds. Rather than focus on the relation of a tree to human goals, they focus on the role that each tree plays in the ecosystem for which it evolved. For the ecologists, the weed trees are the imported trees that invade and often eventually destroy natural ecosystems. For example, the *sugar maple*, a highly typical and widely used and respected tree among the experts in the present study, is considered a destructive weed by local ecologists because it can ruin native oak woodlands by blocking oak seedlings from the sun (it can grow in the shade and oaks cannot). Thus, weediness is in the eye of the beholder.