A bird's eye view: biological categorization and reasoning within and across cultures

Jeremy N. Bailenson[a,*], Michael S. Shum[b], Scott Atran[c,d],
Douglas L. Medin[b], John D. Coley[e]

[a]Research Center for Virtual Environments and Behavior, University of California, Santa Barbara, CA, USA
[b]Northwestern University, Evanston, IL, USA
[c]Centre National de la Recherche Scientifique, Lyon, France
[d]University of Michigan, Ann Arbor, MI, USA
[e]Northeastern University, Boston, MA, USA

Abstract

Many psychological studies of categorization and reasoning use undergraduates to make claims about human conceptualization. Generalizability of findings to other populations is often assumed but rarely tested. Even when comparative studies are conducted, it may be challenging to interpret differences. As a partial remedy, in the present studies we adopt a 'triangulation strategy' to evaluate the ways expertise and culturally different belief systems can lead to different ways of conceptualizing the biological world. We use three groups (US bird experts, US undergraduates, and ordinary Itza' Maya) and two sets of birds (North American and Central American). Categorization tasks show considerable similarity among the three groups' taxonomic sorts, but also systematic differences. Notably, US expert categorization is more similar to Itza' than to US novice categorization. The differences are magnified on inductive reasoning tasks where only undergraduates show patterns of judgment that are largely consistent with current models of category-based taxonomic inference. The Maya commonly employ causal and ecological reasoning rather than taxonomic reasoning. Experts use a mixture of strategies (including causal and ecological reasoning), only some of which current models explain. US and Itza' informants differed markedly when reasoning about passerines (songbirds), reflecting the somewhat different role that songbirds play in the two cultures. The results call into question the importance of similarity-based notions of typicality and central tendency in natural categorization and reasoning. These findings also show that relative expertise leads to a convergence of thought that transcends cultural boundaries and shared experiences. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Biological categorization; Reasoning; Expertise; Cross-cultural differences

 * Corresponding author. Department of Psychology, University of California, Santa Barbara, CA, 93101-9660, USA.
  E-mail address: bailenson@psych.ucsb.edu (J.N. Bailenson).

1. Introduction

Two bedrock cognitive processes are categorization (how do we decide what objects are the same kind of thing) and inductive reasoning (given that one object or class exhibits a property, how do we decide whether other related objects or classes do). Although researchers have increasingly examined these issues in real world vs. artificially contrived domains, research participants are usually introductory psychology undergraduates. Attention to real-world domains, however, raises important questions concerning the way knowledge and culture may affect these processes.

In the present studies we focus on the domain of folkbiology for two reasons: there is a rich literature concerning how humans categorize and reason about plants and animals; and there is significant variability in folkbiological knowledge within and between cultures. If it should turn out that different knowledge systems, goals and activities differentially affect people's ways of conceptualizing the natural world, then lopsided attention to a single participant pool risks biasing interpretation (Atran, 1995). In the worst case, undergraduate performance becomes something of a standard and when comparisons are made with different populations, any differences may be wrongly interpreted as either 'failing' a given experimental task or being under the influence of 'extraneous' factors when performing it.

Although studies in social psychology and decision making have increasingly brought issues concerning populations and context to the forefront of the research agenda (Gigerenzer, 1996; Hsee & Weber, 1999; Markus & Kitayama, 1991; Nisbett, Peng, Choi, & Norenzayan, 2001), there is no comparable cautionary perspective in cognitive categorization and reasoning studies. This is perhaps even more surprising given that Rosch's pioneering research on category structure was firmly rooted in cross-cultural comparisons (Heider, 1971, 1972; Rosch, 1975, 1977). In contrast to cognitive psychology, one of the most relevant subfields of anthropology, ethnobiology, has by its very nature focused on cultural comparisons of classification systems (e.g. Atran, 1998; Berlin, 1992; Boster, 1988).

An endorsement of cross-group comparisons is not without its problems. Comparisons are a reasonable strategy when there is a clear theoretical rationale for expecting some difference (e.g. Markus & Kitayama, 1991). But consider two other cases. If one compares two groups and finds no difference, then the generality of the results is on firmer ground. Here the limitation is that the results may not be considered particularly newsworthy. If one compares two groups and finds clear differences, then interpretative problems quickly emerge. Which of the many ways in which the two groups differ are crucial? For example, López, Atran, Coley, Medin, and Smith, (1997) found that US undergraduates and Itza' Maya of Guatemala showed a different pattern of responding on a category-based inductive reasoning task involving mammals. Although this undermines the universality of the particular reasoning phenomenon, the two groups differ in myriad ways (e.g. age, education, literacy, livelihood, language, cosmology and so on). Furthermore, it is, practically speaking, very likely impossible to disentangle these various factors because (cultural) groups cannot be found that represent orthogonal combinations of these variables. In short, without a clear theory, one may be confronted by the dilemma of finding either weakly informative similarities or uninterpretable differences.

As a partial remedy our research group has followed what we informally refer to as a *triangulation strategy*. The general idea is to use observations from a third group to get at least modest leverage for understanding initial group differences. The third group should resemble one group in some potentially important ways and the second group in other ways. If the third group performs very much like one of the groups and different from the other group, then the variables shared by the third group and the group it mimics become candidates for critical variables. For example, two of the groups in the present studies, US undergraduates and Itza' Maya adults, differ in all the ways faced by the López et al. (1997) comparisons. The third group, US bird experts, is similar to US undergraduates in many ways but shares with the Itza' extensive biological knowledge. If US bird experts and Itza' show similar performance and differ from the third group, then one can have confidence that one has identified a key variable (knowledge) and that it is robust over all the other differences between the two groups. If US bird experts perform comparably to undergraduates, one has at least ruled out domain knowledge as mediating the difference.

This so-called triangulation strategy obviously is not a cure-all. For example, it may be difficult to find third groups that share variables with the first two groups without introducing further extraneous variables. Even if used successfully, it is like playing "twenty questions" but only being able to ask two ("Are A and B different and if so, is C more like A or B?"). For the strategy to be effective one must either make good guesses about relevant variables or be able to collect data from additional groups to further clarify the pattern of similarities and differences.

At first glance, it might appear that our triangulation strategy is just a 2×2 design with one cell missing. But a 2×2 design presumes what the triangulation strategy is intended to discover, namely, which factors are crucial to group differences. In fact, the logic of triangulation implies compression of any number of possible 2×2 designs that together entail a host of possible explanations for group differences. Instead of running 2^Nth conventional controlled designs, each of which allows inference to a single factor, a carefully chosen third group, C, that resembles the first group, A, in a number of ways and the second group, B, in a number of other ways deliberately confounds a number of variables. This is to enable discovery of the relative importance of the set of culturally-confounded variables by which C differs from A vs. those by which C differs from B.

A 2×2 design also implies more precise matching and control of variables than is feasible in cross-cultural comparisons. In the present studies the US birders and the Itza' elders are both expert with respect to US undergraduates but it can hardly be said that the birders and the Itza' are "matched" in expertise. As we will shortly point out, there is literature that demonstrates differences in categorization as a function of relative amounts (Johnson & Mervis, 1997, 1998) and type of expertise. We view the triangulation strategy as having the potential to be applied iteratively at different levels of resolution. For example, suppose we were to find that US experts resembled Itza' in some ways but differed from them in other ways. Rather than attributing any differences to culture, one might well attempt to develop another triangular comparison involving Itza', US experts with goals and activities resembling those of the Itza' and US experts with goals and activities distinct from the Itza'. Again, it would be unlikely that one could obtain a precise match on goals and activities but might well be able to produce greater cross-cultural than within culture similarity in goals and activities. To do this, one potentially might need to

"reverse the triangle" and compare types of Itza' (e.g. bird hunters and non-hunters) with US experts. As we shall see, in the present studies the Itza' often cluster with the US experts, findings which suggest that similarity with respect to knowledge is a more important factor than similarity with respect to culture.

Before turning to the present experiments, we briefly review some relevant literature on folkbiological categorization and reasoning. This prior work will be used to motivate several hypotheses concerning the role of expertise and culture in conceptual behavior. Our focus is on the assumptions that underlie general models of categorization and category-based reasoning, in particular typicality and diversity.

1.1. Expertise effects in folkbiology

Boster and Johnson (1989) examined knowledge and sorting patterns among expert and novice fishermen. They note that morphological information (the stimuli were pictures of fish) is available to any observer, but cultural knowledge of functional and utilitarian properties of fish requires experience. Therefore, experts and non-experts should differ not only in the amount of information they possess, but also in the *kinds* of information. If this information is used in classification, then experts should be more likely to classify along functional and utilitarian lines. As expected, Boster and Johnson found that non-experts relied more on morphological information than did experts. Although similarity judgments of all groups correlated with scientific taxonomy, non-experts correlated more highly than experts. Experts' similarity judgments in turn were more highly correlated with functional similarity.

Medin, Lynch, Coley, and Atran (1997) addressed the question of inter-expert differences. Instead of comparing the conceptual structure of experts and non-experts, they examined similarities and differences among experts with different specialties within a single domain. They looked at how different types of tree experts (maintenance workers, landscapers, and taxonomists) categorized and reasoned about familiar tree species. Their analysis revealed some common conceptual organization between different types of experts, but also that expert groups differed with respect to the structure of their taxonomies and how they justified or explained the categories they formed. In sorting, maintenance workers relied on morphological features, while the landscape workers structured their sorts around goal-derived categories based on practicable interests (taxonomists sorted largely in accordance with scientific taxonomy). The reasoning of taxonomists and maintenance workers accorded well with the similarity relationships revealed on the sorting task. Landscapers' reasoning could not be predicted from their (goal-derived) sorting; instead, like the maintenance workers they relied on morphological similarity. Thus, the acquisition of expertise in a particular domain does not necessarily lead to a standardized conceptual organization of information in that domain, though reasoning tasks may reveal more agreement (see Proffitt, Coley, & Medin, 2000 for further similarities and differences among tree experts as a function of type of expertise).

The above studies of expertise clearly indicate an influence of experience in a domain with respect to reasoning and sorting about members of that domain. The way in which people structure their concepts about particular domains depends on both their level of knowledge and the *kind* of knowledge they possess by virtue of their characteristic goals

and activities. In the following section we discuss the ways in which an individual's cultural experience may effect the way he or she reasons.

1.2. Cross-cultural comparisons of folkbiology

In addition to expertise, we are also principally interested in how cultural experience may affect folkbiology. Roughly, any given "culture" is a causally distributed assemblage of mental representations, their public expressions and resultant behaviors that become relatively pervasive, enduring and interconnecting among individual members. Representations and behaviors become more or less "cultural" to the extent that they spread and stabilize within a population of minds over time (Atran, 2001). Because the current set of cross-cultural studies do not really explore the underlying causal processes of within-group variation and culture formation, for present purposes we restrict our use of the notion of culture to an imperfect community of evidentiary symptoms. We use the term in a commonsensical manner to refer to a host of confounded variables that include symbolic (e.g. language and legends), ideational (e.g. values and mores), social structural (e.g. kinship and community organizations), technological (e.g. arts and crafts), and environmental (e.g. landscape and local ecology) factors. "Culture" as such cannot be reduced to an independent variable. For our purposes, however, it suffices to say that populations that have radically different languages, legends, values, mores, community and kinship organizations, economic and ecological practices and so on belong to different "cultures". Examples include the populations that ethnobiologists have traditionally studied (e.g. Berlin, 1992).

In general, cross-cultural research in folkbiology has pointed to similarities between different cultural groups in their categorization and reasoning about natural kinds. Ethnobiologists studying systems of classification in small-scale societies (e.g. Atran, 1999; Berlin, 1978, 1992; Berlin, Breedlove, & Raven, 1973, 1974; Brown, 1984; Bulmer, 1974; Hays, 1983; Hunn, 1977) have argued that taxonomies of living kinds are organized into ranked systems. Not only are categories related to each other via class inclusion, but categories (taxa) at a given level in the system also share taxonomic, linguistic, biological, and psychological properties with other categories at that level. Moreover, these regularities in folkbiological classification and nomenclature can be seen in disparate cultures throughout the world. Indeed, these common principles observed in culturally diverse populations are often taken as evidence for universal cognitive constraints on folkbiological thought (Berlin, 1992).

These claims are supported by research that extends beyond the cataloging of folk taxonomies. For example, Boster, Berlin, and O'Neil (1986) examined disagreement between Aguaruna and Huambisa Jivaro natives by having the groups identify prepared bird specimens. Although the groups are both from the same region in northern Peru, their cultural milieux are distinct in the sense that members of the communities are not in direct contact and speak different but related languages. These authors demonstrate that the two groups exhibit similar disagreement patterns during identification that can be predicted by taxonomic relatedness; that is, both groups are more likely to confuse species that are more closely related scientifically.

Boster (1987) extended the results of these experiments to include a condition where the

participants sorted unfamiliar birds. He chose a subset of the birds used in the Boster et al. (1986) study and presented them in a sorting task to US undergraduates. He then compared the data to previous results and found that specimens the college students found perceptually similar corresponded with closely related birds according to scientific taxonomy, and also tended to be the ones confused by Peruvian natives. Although this finding is impressive in that it shows two completely distinct cultural groups performing quite similarly with the same biological kinds, the data would be more compelling had the same type of task been used for the subjects in both cultural settings.

Boster's findings support Berlin's claim that diverse groups of informants discern the same sorts of biological kinds in the same ways because of the inter-correlated structure of the biological world (see also Hunn, 1976). Of course, if features are highly inter-correlated, then two individuals (or groups) may attend to different features but produce more or less the same sorts. That is, different people may attend to very different features of a given organism, but because the features themselves are covariant, the same structure is discerned. Alternatively, there may be universal classificatory principles that interact with the world's correlational structure with the result that diverse groups of informants choose the same salient features of specimens to construct and distinguish biological kinds. Either alternative is also consistent with findings by Medin et al. (1997) that when different groups of tree experts produce similar sortings of species they may justify them differently (see also Boster & D'Andrade, 1989).

1.3. Similarity-based models of categorization and reasoning in folkbiology

The cross-cultural findings reported thus far are all interpretable in terms of 'similarity-based models' (Smith & Medin, 1981), which organize perceptually identifiable categories on the basis of correlation or covariation of stimulus attributes. Category-based induction models of taxonomic reasoning in folkbiology also rely on computations over similarity judgments or their presumed underlying features (Osherson, Smith, Wilkie, López, & Shafir, 1990; Sloman, 1993). More recent studies, however, call into question the generality of similarity-based models of categorization and reasoning in folkbiology.

López et al. (1997) explored how University of Michigan students and Itza' Maya sorted and reasoned about local mammal species. The Itza' were shown mammals from Central America while the Michigan students were shown mammals native to the Great Lakes region. López et al. found several consistent differences and similarities in reasoning and sorting. Both groups tended to justify their sorts by focusing on behavior and morphology, thereby creating groups consisting of predators and non-predators. Taxonomic sortings for both groups correlated to the same degree with one another and with scientific (evolutionary) taxonomies. However, Americans and Maya systematically diverged from science in relying on size and ferocity as salient sorting dimensions. The two groups also systematically diverged from one another, with Itza' relying more on ecological considerations (e.g. in grouping arboreal mammals) and less on size than the Michigan students.

López et al. also tested for three category-based induction phenomena as defined by Osherson et al. (1990): similarity, typicality and diversity. Similarity predicts that the stronger inference should be the one in which the premise category is deemed more similar

to the conclusion category. For example, according to sorting data, Itza' believe that *fox* and *cat* are more similar than *fox* and *dog*. US undergraduates believe that *fox* is more similar to *dog* than *cat*. Consistent with these beliefs, Itza' were more likely to attribute a novel property of *fox* to *cat*, whereas US students were more likely to attribute it to *dog*. Both populations used similarity to guide induction.

López et al. also derived a metric for typicality from the sorting data. Items that were more "typical" by virtue of having higher *central tendency* (i.e. high average similarity to other mammals) supported inferences better than items that were less typical (i.e. less similar on average to other mammals). Items that are more typical thus provide greater "coverage" of the category than items that are less typical. Both the Americans and Maya reliably performed inference tasks in conformity with this algebraic measure of typicality.

The most notable differences between the two groups centered on diversity. According to the diversity principle, an argument with taxonomically dissimilar premises should be evaluated as stronger than an argument with taxonomically similar premises, all other things being equal. For example, people should be more willing to generalize a fact that is true of mice and bears to all mammals than a fact that is true of wolves and coyotes. US undergraduates showed high levels of diversity-based reasoning; however, the Itza' did not. In fact, in one condition, the Itza' showed a reliable dis-preference for the more diverse arguments. The Itza' justified their responses on the basis of specific ecological knowledge, often leading them to conclude that the more diverse premise was the weaker of the two choices.

Subsequent work on direct typicality judgments among Itza' (Atran, 1999) shows that inductively useful notions of typicality may be driven more by notions of idealness than central tendency. In each case for which we have direct Itza' ratings, the 'truest' or 'most representative' living kind categories are large, perceptually striking, culturally important, and ecologically prominent. The dimensions of perceptual, ecological and cultural salience all appear necessary to a determination of typicality, but none alone appears to be sufficient. The three most representative birds are all large, morphologically striking and highly edible Galliformes (wild fowl): the ocellated turkey, the crested guan, and the great curassow.

In an earlier study, Barsalou (1985) argued that idealness rather than central tendency predicts typicality in goal-derived categories (e.g. foods not to eat on a diet, things to take from one's home during a fire, camping equipment), although central tendency does predict typicality in 'taxonomic' categories (e.g. furniture, vehicles), including folkbiological categories (e.g. birds). The Atran (1999) results indicate that the story may be still more complicated. For mammals, direct Itza' ratings of typicality and derived measures of central tendency largely coincide. Thus, what López et al. (1997) interpreted as confirmation of the role of central tendency in the Osherson et al. (1990) model may have as much or more to do with ideals than typicality in the model's sense (with the added complication that, for mammals, idealness may somehow be driving central tendency). For Itza' birds, however, preliminary studies indicated a clear distinction between idealness ratings and central tendency, which facilitates an exploration of the respective contributions to induction of these different conceptions of typicality.

Given the bulk of cross-cultural research in folkbiology that stresses universals in conceptual organization, cultural differences with respect to typicality and diversity are

striking. However, the source of the differences is not clear. The groups differ in terms of cultural membership, and also in terms of relative expertise about the natural world. So, the question becomes, do the differences in categorization and reasoning about mammals between the Itza' and US undergraduates owe to different information requirements that derive from cultural preference, relative expertise, or some combination of the two? More generally, how do cultural experience and relative familiarity or knowledge interact to drive conceptual differences?

1.4. Sorting out knowledge and cultural effects: expertise effects

Although there is not an extensive literature, the current evidence suggests that knowledge may be more important than culture in determining typicality ratings and category-based induction. Lynch, Coley, and Medin (2000) found that tree experts based their typicality judgments on ideals (e.g. height, absence of undesirable characteristics) and that central tendency was uncorrelated with judgments. They found no effects of type of expertise. The fact that US experts and Itza' (experts) both show effects of ideals undermines concerns about the wording of the typicality instructions in Itza' Maya somehow conveying a different notion of typicality. Lynch et al. used instructions that followed verbatim those by Rosch and Mervis (1975) in their original studies showing central tendency-based typicality effects.

Proffitt et al. (2000) examined typicality and diversity effects in reasoning among the three types of US tree experts studied by Medin et al. (1997). They found that typicality and diversity effects were weak or absent. Only taxonomists showed above chance diversity responding and maintenance personnel were actually below chance on some diversity probes. Justifications point to the three types of experts using strategies based on ecological and causal factors, much like the Itza' reasoning about mammals. Diversity was one justification among many and never constituted a majority of the justifications for any of the groups. Direct appeals to typicality were rare in one experiment and totally absent in another.

1.5. Integrating culture and expertise in typicality and reasoning: hypotheses

This brief review helps motivate the current triangulation strategy. The studies of typicality and induction with experts suggest that type of expertise matters little and that causal/ecological reasoning may be far more common than abstract, similarity-based reasoning strategies. Combining the studies of expertise with work in ethnobiology suggests the following hypothesis: there is a natural or default categorization scheme sensitive to the structure of nature (relative to the human perceptual system) but experts may develop special-purpose classification schemes as a function of characteristic activities and the additional goal-related knowledge they acquire. Whether experts develop special-purpose categorization schemes may depend on the variety of goals they have with respect to some domain and the degree to which their goals are compatible with the correlational structure of that domain.

Applying this hypothesis to the present studies suggests several predictions. First of all, we should note that we employed both birds found in Guatemala and those found in the US. The predictions hinge critically on whether or not US bird experts or Itza' have

specialized goals. Given that the primary goal of birders is to identify birds, we see no reason for expecting that their goals violate the correlational structure of bird taxonomy. Both bird identification and traditional taxonomic systems are based on morphological similarities and differences. If this analysis is correct then both US experts and US novices should base their categorization of both US and Mesoamerican birds on the natural or default taxonomy and show good agreement with scientific taxonomy.

It is less clear what to predict for the Itza'. On the one hand, they do have specialized goals with respect to some birds (e.g. hunting them for their meat) but, on the other, our prior work suggests that they have rich ecological knowledge concerning relationships between particular kinds of birds and both plants and other animals (Atran et al., 1999). The latter observation suggests a general-purpose (default) representation. If the Itza' have both special-purpose and general-purpose representations, then we would expect that the special-purpose representation would be much more evident in their sorting of familiar Mesoamerican birds than the unfamiliar US birds (one cannot safely predict that a special-purpose representation would not be used at all for US birds because although the species in the two settings had essentially no overlap, at the higher levels of genera and families there was modest to extensive overlap).

In sum, we offer the following hypotheses: (1) typicality judgments will be based on central tendency for US novices but not for US experts or Itza'; (2) typicality judgments will be based on ideals in the two groups of relative experts; (3) US novices will be far more likely to rely on abstract, similarity-based reasoning strategies in category-based induction than either of the two groups of relative experts and consequently will be more likely to show typicality and diversity effects in reasoning. This prediction should be qualified by the possibility that US experts may have a greater focus on taxonomy than Itza' (by virtue of explicitly learning scientific taxonomy) and may be somewhat more likely to show taxonomic diversity effects.

One might also expect that the results for experts may vary as a function of whether the probes involve the familiar vs. unfamiliar set of birds. If experts are going to use abstract reasoning strategies at all, it should be with the unfamiliar set, because experts should be less likely to employ ecological reasoning with unfamiliar birds (though, of course, they may do so by analogy from birds that they know).

Overall, then, our hypotheses based on prior literature lead us to expect few, if any, differences on the sorting task. The two qualifications to this prediction are that we should be more likely to find differences in justifications for sorts than in the sorts themselves and that Itza' may use a goal-derived organization for Mesoamerican birds. (Again, based on the Atran et al. (1999) findings we consider the latter possibility unlikely.) In contrast, we expect large differences for both typicality ratings and category-based inductions. With respect to our triangulation strategy, we expect expertise effects to dominate cultural influences.

To our knowledge, our study is the first to simultaneously examine the effects of culture and relative expertise on category-based reasoning. The organization of the remainder of the paper is as follows. In Experiment 1, we examine how the three groups sort two sets of bird species (birds of Central America and birds of Illinois), and how they explain or justify their sorts. The sorting data provide a similarity metric used to guide the interpretation of typicality judgments (Experiment 2) and inductive reasoning (Experiment 3).

2. Experiment 1: sorting

Experiment 1 involved asking participants to sort pictures of local and exotic birds into groups that "go together by nature". One purpose of this study was to use sorting distance as a similarity metric (i.e. objective measures of central tendency and pair-wise distance) for the second and third experiments. A second key goal was to compare within and across group patterns of sorting. In that regard there are two important questions to be addressed: (1) Do people within a group agree sufficiently in their sorting that it is sensible to claim that there is a consensual cultural or group model? (2) Are the patterns of sorting reliably different across populations? To address these questions we employ the cultural consensus model (CCM) of Romney, Batchelder, and Weller (1986).

The CCM is a factor-analytic method for computing levels of agreement and disagreement in the structure and distribution of information within and across populations. Of course, there are other possible ways of evaluating reliable patterns of inter-informant agreement. Nevertheless, the CCM has proven to be especially suited to comparisons with small-scale populations by providing formal criteria for assessing conformity among small numbers of informants who produce relatively large data sets of predetermined size (Atran et al., 1999; Boster & Johnson, 1989; López et al., 1997; Medin et al., 1997; Romney et al., 1986). In our methods, we elicit complex information from a number of informants. Rather than assume an underlying model for all informants in a population, we test whether consensus exists. Finding consensus justifies further study of group-wide patterns and establishes a somewhat higher threshold for reliability than mere statistical significance (e.g. *P* values), insofar as we do not consider cultural patterning to be established unless there is consensus in addition to significance.

The CCM is essentially principal components factor analysis. Using this analysis, a group consensus is indicated if (1) the first eigenvalue is substantially larger than the second and accounts for much of the variance, and (2) the first factor scores for each individual are positive. Each informant's first factor score represents the degree to which that person's responses agree with the consensus. That is, the pattern of correlations among informants should owe entirely to the extent to which each knows the common (culturally relative) 'truth'.

A fit to the CCM provides a number of important indicators of within-group reliability in bird classification. For example, it would indicate: (1) that there is sufficient justification for combining individual taxonomies into an aggregate group taxonomy; (2) how each subject's taxonomy compares to the aggregate taxonomy (based on how each subject ranks on a given factor).

The CCM also allows us to compare cultural taxonomies. To do this we look at patterns of residual agreement. If there is a single consensus across groups then the CCM should provide a good fit to the aggregate data and the agreement of any two informants should be just the product of their first factor scores. If the groups differ, however, then individuals within a group should be in agreement with each other to a greater extent than is predicted by the overall consensus analysis. In the present paper, we compare within vs. between group residual agreement and if the former is reliably greater than the latter, we conclude that there are group differences in consensual models. Note that use of the CCM allows us to treat cultural phenomena in terms of emergent patterns that are statistically derived and

aggregated from individual cognitions without a priori assumptions about degree of cultural homogeneity.

Before describing the experimental procedure in detail, we first provide a note on our general methodology. In the present study we asked informants to "sort together the things that go together by nature" rather than giving more abstract instructions to sort by similarity. One important lesson of the last decade of our cross-cultural experimental research with adults and children in widely different cultural settings is that experiments are not transparent windows onto reality. They are interactive communication settings wherein the experimenter contrives to get the participant to produce data that informs what the experimenter seeks to know. We have found that the production of data that are equivalently informative, or at least directly comparable, across cultural settings often requires incorporating differences in the experimental setup.

Of course, the attempt is made to keep the differences to a minimum. In fact, our pilot studies almost always use instruction sets that are as close as possible to fluent translations of one another. For example, in earlier sets of studies we explain why in biological sorting tasks we use the instruction to sort kinds that "go together by nature" rather than "are most similar" (López et al., 1997). Although sorting by "similarity" is by far the preferred instruction in psychology experiments in our society, there are often no ready translations in other societies. When we used the command "go together by nature" with the Itza' (as well as with the US subjects) we elicited the sort of general-purpose taxonomy described for cultures across the world (Atran, 1990; Berlin, 1992).

2.1. Method

2.1.1. Subjects

2.1.1.1. US experts The US experts were ten men and ten women (mean age = 50.8 years, range: 18–83 years) having either occupations or extensive experience related to birds. The average number of years spent watching and studying birds (termed 'birding' hereafter) was 22.4 years. Three participants had completed high school, one had at least some college work, six had completed college, and ten had advanced degrees, four of those being PhDs. Fourteen had received at least some formal education regarding birds. On a seven-point continuous scale with 1 indicating 'very little knowledge about birds' and 7 indicating 'total expertise', the mean self-reported rating for US experts was 5.1 (SD = 1.10). The US experts were drawn predominantly from the four following organizations: the Evanston Birding Club, the North Shore Birding Club, the Morton Arboretum, and the Chicago Ornithological Society. Occupations of informants were extremely varied and included college professors, salesmen, engineers, retired workers, economists, high school teachers, botanists, and library clerks. Most of them viewed birding as an extremely involving hobby, often dedicating their vacation time to traveling to places where they could find birds that they had never seen before.

2.1.1.2. US non-experts The novices were eight men and eight women (mean age = 20.8 years, range: 18–40 years) who were recruited through the university and paid for their participation. Out of the 16, only two had any special experience with birds: one of the

novices had raised pet parrots, and another had taken an informal nature course. On the seven-point rating scale mentioned above, the mean self-reported rating for our non-experts was 2.33 (SD = 0.90).

2.1.1.3. Itza' Maya Ten Itza' Maya (eight men and two women) living in the village of San José, Petén, Guatemala participated in the study. The Itza' are Maya Amerindians living in the Petén rain forest region of Guatemala. Men devote their time to shifting agriculture and hunting, and women concentrate on household maintenance. The Itza' were the last independent native polity in Mesoamerica to be conquered by the Spaniards, and have preserved virtually all ethnobiological knowledge recorded for Lowland Maya since the time of the initial Spanish conquest. Informants ranged in age from 50 to 78 years (mean age = 66.14 years). All were bilingual in Itza' and Spanish, although experimental instructions and responses were in the Itza' language. Some Itza' participated as part of a larger project on comparative folkbiology, while others were recruited especially for this study. There were no a priori grounds to distinguish experts among the Itza' as nearly all informants in this and other studies had continuous and extensive experience with forest plants and animals. All Itza' were well acquainted with the experimenters, and at relative ease in the session. The experimenters were part of an international, interdisciplinary team of researchers that has been studying the language, lore and natural history of Itza' and other contemporary and Pre-Columbian Lowland Maya groups for the last decade (Atran, 1993; Atran & Medin, 1997; Atran et al., 1999, in press; Atran & Ucan Ek', 1999; Lois, 1988).

2.1.2. Design and materials

Two separate groups of American participants (ten experts and eight non-experts in each) sorted the two sets (US and Tikal) of birds. These two groups were matched in terms of gender, and were comparable on other measures such as amount of formal education and knowledge concerning birds. Itza' were tested in different sessions (roughly one year apart) on Tikal and US birds, but because of the more limited pool of participants, there was some overlap in participants (five men, one woman) who sorted both US and Tikal birds.

2.1.2.1. Stimuli There were two sets of stimulus materials, each consisting of full-color illustrations of 104 bird species laminated onto index cards. One set (hereafter called 'US birds') featured only species which were either resident to the Chicago area or breed there. For this set, illustrations were taken from the Golden and National Geographic field guides, books designed to aid bird identification. The other set ('Tikal birds') were species from the Tikal region of Guatemala, and were taken from the book *The Birds of Tikal* (Smithe, 1966). The specific selection of birds was based on the inventory list assembled by the University of San Carlos (Guatemala) for the UN-sponsored Maya Biosphere Reserve.

The structure of the scientific taxonomy representing the US birds was designed to correspond maximally with that representing the Tikal bird set. Appendix A is a complete taxonomic listing of both stimulus sets. The scientific taxonomy was taken from a publication by the US Audubon Society and is consistent with both national and international

field guides to birds. The Tikal bird set consisted of 30 families and 17 orders, while the US bird set consisted of 33 families and 17 orders. One notable difference was in the number of passerines (songbirds) in the two sets. Although passerines are the numerically dominant group both in Chicagoland and Mayaland, they are a somewhat higher proportion of families in Chicagoland. Because passerines are so prevalent in the Chicago area, in order to structurally equate the two taxonomies, it became necessary to include more of them in the US set (32) than in the Tikal set (26). For similar reasons, there were more Falconiformes (birds of prey) in the Tikal set (17) than in the US set (10). Furthermore, there were some birds and taxonomic groups that were common to both sets. As Appendix A shows, there were 18 shared orders, 12 shared families, 12 shared genera, and five shared species in the two sets.

The scientific taxonomies chosen can be roughly described as "classic evolutionary taxonomies". Classic evolutionary taxonomies differ from so-called "numerical taxonomies" (Sokal & Sneath, 1973) by weighting perceptual attributes according to their phylogenetic importance (Mayr, 1965). Classical evolutionary taxonomies also differ from strictly phylogenetic classification schema, or "cladistic" taxonomies (Hennig, 1966). From a cladistic perspective, all taxonomic ranks are arbitrary (although some cladists give special status to the species level) and many classical taxa are wrongly related or delimited (e.g. there is no phyletic justification to placing birds on a par with mammals, or allowing reptiles to persist as a unitary category). Moreover, any scientific taxonomy will change over time, as new information is brought to bear (e.g. molecular attributes, fossil evidence, etc.). All this suggests that there is presently no "true" scientific standard.

Although cladistic analyses based on DNA-matching now appears to be steadily gaining credence over other strategies in biological systematics, one problem with comparative use of cladistics is that it does not furnish unique or clear-cut taxonomic resolutions of biodiversity and phylogenetic relationships. That is one reason why classical evolutionary taxonomy continues to provide a (readily accessible if only approximate) standard of reference commonly used by many scientists and most scientifically-minded amateurs (e.g. birders). From the standpoint of comparison, it also has the convenience of maintaining the larger categories (e.g. bird, mammal, fish) that we wish to explore across various informant populations (for further arguments in favor of comparative use of evolutionary taxonomy, see López et al., 1997).

2.1.3. Procedure

All participants were tested individually by the experimenters. They were told that we were interested in how they organized their knowledge about birds. First, we showed them all 104 bird cards one at a time and asked them to name them 'as specifically as possible'. Responses were tape recorded and transcribed. After this initial naming phase, all 104 cards were placed in front of the participant, who was asked to 'put together the birds that go together by nature into as many different groups as you'd like'. The experimenters recorded these initial categories and asked the informant to describe or justify their basis for each category. Subjects were then asked to combine the initial categories by 'putting together those groups of birds that go together by nature into as many larger groups as you'd like'. The experimenters then recorded the new categories and their justifications. Participants were free to combine groups as many times as they so desired. Successive

compiling was repeated until the informant indicated no further grouping seemed to be natural. At this point the experimenters restored the initial categories created during the first free sort and invited subjects to 'split as many of the groups as you'd like into smaller groups of birds that go together by nature'. The experimenters recorded the subgroupings and their justifications, and then the subpile sorting was repeated until participants indicated that no further subdivisions seemed sensible.

2.2. Results

2.2.1. Naming accuracy

The naming data are useful in providing an independent index of expertise and relative familiarity with the two picture sets. We scored each naming response on a three-point scale, with a 3 representing an exact species match, a 2 representing a correct genus or family match, and a 1 representing a match at order or higher (e.g. a 'songbird' response or a 'bird' response were both scored as 1). The three groups named all the birds from both stimulus sets except the Itza', who named only birds that they sorted from the Tikal set (e.g. an '*ix kusam*' [swallows and swifts] response that generalized to all members of the Hirundinidae family of the order Passeriformes as well as to all members of the Apodidae family of the order Apodiformes and a '*ix wirisu*' [flycatcher] response that extended to various other Passeriformes were both scored as 1). Consequently, we were unable to run a full factorial ANOVA. Instead, we ran a between-subjects ANOVA with stimulus set (US and Tikal) and subject group (US experts and non-experts) as factors and average naming score as the dependent variable. There were two main effects and a significant interaction. First, both US experts and non-experts were better at naming US birds ($M = 1.96$) than Tikal birds ($M = 1.43$) ($F(1, 32) = 93.26$, $P < 0.001$). Furthermore, US experts ($M = 2.10$) were more accurate overall than non-experts ($M = 1.20$) ($F(1, 32) = 304.38$, $P < 0.001$). In addition, the interaction demonstrated that US experts were more accurate at naming US birds ($M = 2.55$) than Tikal birds ($M = 1.66$), but that non-experts showed little difference between US birds ($M = 1.25$) and Tikal birds ($M = 1.14$) ($F(1, 32) = 56.55$, $P < 0.001$). This interaction is essentially a floor effect based on the non-experts' poor performance on both sets of birds. These results establish that the US experts were more familiar with the US birds than the Tikal birds and that their naming skills were superior to those of novices for both sets of birds.

We also ran a one-way, between-subjects ANOVA using subject group (US expert, US novice, or Itza') as the factor and average naming score on Tikal birds. The effect of group was significant ($F(2, 24) = 64.27$, $P < 0.001$). The Itza' were the most accurate ($M = 1.92$), followed by US experts ($M = 1.66$) and then the non-experts ($M = 1.14$). Tukey's HSD test (with alpha = 0.05) showed that all three groups were significantly different. Itza' were less accurate at naming passerines ($M = 1.39$) than other birds ($M = 2.11$) ($t(8) = 20.49$, $P < 0.01$). On the other hand, across US and Tikal birds, the American experts did not show any difference between trials on which they named passerines ($M = 2.13$) and trials on which they named non-passerines ($M = 2.10$). This may be a function of Itza' not having binoculars at their disposal, and not being able to see the small birds as clearly in the wild as US experts; however, it may also be a function of cultural

interest (see below). The non-experts did not show any difference in naming passerines and non-passerines on either bird set, but this is clearly a floor effect.

2.2.2. Sorting

First we discuss how performance on the sorting task was quantified. Then we consider, in turn, the degree to which performance within each group for each set of birds reflected a consensual taxonomy, the correspondence of these taxonomies with science, and finally, we compare the sorting justifications of the novice, expert and Itza' folk taxonomies.

2.2.2.1. Scoring We used each informant's hierarchical sorting to derive a bird-by-bird similarity (distance) matrix. Each informant's taxonomy was obtained by translating the groupings made during the free pile, successive pile and successive subpile sorts into a taxonomic tree. The bottom level nodes of this tree would correspond to individual bird species, and the top level node to all birds together. Intermediate levels correspond to the groupings the subject made during the sorts. Each level includes all groupings made during a given round of sorting, and the levels are rank-ordered from the last successive subpile sorting to the last successive pile sorting.

Analogous taxonomic trees were constructed for each individual. From each taxonomy, we derived a pair-wise bird-by-bird distance matrix by calculating the distance between all possible pairs of birds in the taxonomy. The lowest level at which two given birds go together in a folk taxonomy represents the distance between them. For each participant, we derived a bird distance matrix where rows and columns correspond to the sorted birds, and the cells to the distances among them.

2.2.2.2. Cultural consensus on bird taxonomies In each condition, the bird distance matrices produced by each informant were correlated with each other, yielding a single pair-wise subject-by-subject correlation matrix representing the degree to which each subject's taxonomy agreed with each other subject's taxonomy. Principal components analyses were then performed separately on each of the three subject groups' intersubject correlation matrices for both sets of birds to determine whether or not there was a 'cultural consensus' in taxonomies. Each subject's bird distance matrix was correlated with that of every other subject, yielding a 28×28 matrix in which entries correspond to agreement among subjects on pair-wise bird distances derived from their individual taxonomies.

Table 1 shows the factor solutions for the three groups combined on the two bird sets. For the analysis of US birds, the first three eigenvalues were 14.25, 1.71, and 1.21, accounting for 50.1, 6.1, and 4.3% of the variance, respectively. All subjects' scores on the first factor were positive ($M = 0.71$, range: 0.49–0.84). A one-way ANOVA on first factor scores revealed that US experts' first factor loadings ($M = 0.77$, SD = 0.05) were higher than those for non-experts and Itza' ($M = 0.64$, SD = 0.09, and $M = 0.69$, SD = 0.07, respectively) ($F(2, 25) = 8.30$, $P < 0.05$, Tukey's HSD). Non-experts and Itza' did not differ from each other. For the analysis of Tikal birds, the first three eigenvalues were 13.39, 1.34, and 1.04, accounting for 53.5, 5.4, and 4.2% of the variance, respectively. All subjects' scores on the first factor were positive ($M = 0.71$, range: 0.30–0.90). A one-way ANOVA on first factor scores again revealed differences between

Table 1
Mean factor loading scores by subject group by bird set in Experiment 1

Stimulus set	Subject group		
	US expert	Non-expert	Itza'
Tikal			
First factor	0.84	0.48	0.79
Second factor	−0.04	0.35	−0.16
Third factor	−0.04	0.05	0.03
US			
First factor	0.77	0.64	0.69
Second factor	−0.01	0.16	−0.13
Third factor	−0.21	0.12	0.16

groups $(F(2, 22) = 37.95, P < 0.05)$. Non-experts' consensus scores were lower $(M = 0.48, SD = 0.15)$ than for experts and Itza' $(M = 0.84, SD = 0.05,$ and $M = 0.79, SD = 0.04,$ respectively), who did not differ from each other. In sum, there was a strong consensus for both US birds and for Tikal birds, with the first factor accounting for most of the variance in the sorts.

2.2.2.3. Analysis of residual agreement Although we observed robust overall agreement, this was coupled with reliable group differences. To test for systematic group differences in categorization, we prepared a subject-by-subject residual agreement matrix (Nakao & Romney, 1984). First, the products of the first-factor consensus scores were obtained for each pair of subjects, representing agreement predicted by each subject's knowledge of the consensus. Next, this predicted agreement matrix was subtracted from the observed agreement matrix, yielding a residual matrix, which we then standardized. This standardized residual agreement matrix was compared to a model matrix corresponding to subject group membership (i.e. a 26 × 26 matrix with entries of '1' if the corresponding subjects belong to the group in question, otherwise '0'). These two matrices were then compared using Monte Carlo simulations to assess whether residual agreement is higher among subject groups than among randomly-chosen pairs of subjects. If patterns of agreement are completely described by the consensus component, there should be no appreciable residual agreement among subjects belonging to the same group (see Boster, 1986; Coley, 1995; Johnson, Mervis, & Boster, 1992). Systematic residual agreement within subject groups would lead to significant association between model and residual matrices. The degree of association between the model matrix and the residual agreement matrix was assessed using the Quadratic Assignment Program (Hubert & Schultz, 1976).

For the US birds all three groups showed significant residual agreement (non-experts: $z = 5.95, P < 0.05$; Itza': $z = 3.73, P < 0.05$; US experts: $z = 7.80, P < 0.05$), indicating that each group's sorts show internal consistency beyond that captured by the consensus across groups. For Tikal birds, for non-experts and Itza', there was significant residual agreement $(z = 5.48, P < 0.05,$ and $z = 2.18, P < 0.05,$ respectively). US experts, however, showed no significant residual agreement above and beyond the first factor

($z = 0.13$). Apparently, the first factor accounted for almost all consensus for US experts. These results point to differences in the taxonomies produced by each group, which we will take up in greater detail shortly.

Results of the CCM analysis suggest a shared component to the taxonomies of the subjects that accounts for more than half of the variance. Residual analyses and analysis of within-group consensus also revealed systematic differences between groups.

2.2.2.4. Correspondence to scientific taxonomy In this section we examine correlations with scientific distance. In order to compare performance from each group to science, we used the scientific taxonomy to derive a pair-wise bird-by-bird folk taxonomic distance matrix by calculating the distance between all possible pairs of birds in the taxonomy. We then compared the average matrix from each group to the science matrix. The by-subject mean correlations for each of the groups on the US birds were 0.38, 0.60, and 0.45 for non-experts, US experts, and Itza', respectively. The by-subject mean correlations for each of the groups on the Guatemalan birds were 0.34, 0.70, and 0.61 for non-experts, US experts, and Itza', respectively. Across bird sets, US experts' taxonomies corresponded most closely with science (0.65), followed by the Itza' (0.53), and then the non-experts (0.36).

To test the reliability of these group differences, we ran a 3×2 ANOVA with each individual informant's correlation with science as the dependent measure, and group (either novice, expert, or Itza') as one factor and bird set (Tikal or US) as the other. The effect of group was significant ($F(2, 47) = 48.52$, $P < 0.05$). Across groups, taxonomies corresponded with science to a greater degree for Tikal birds than for US birds (0.55 vs. 0.48) ($F(1, 47) = 9.08$, $P < 0.05$). The interaction between subject group and bird set was also significant ($F(2, 47) = 5.71$, $P < 0.05$). Tukey's HSD indicated that the interaction was due to the fact that US experts correspond highly to science when sorting foreign and local birds, while Itza' correlate more highly with science when sorting local birds and non-experts do not correlate highly with science on either set.

2.2.2.5. Cluster analysis In this section we present results from a cluster analysis to provide insights concerning the structure of our groups' taxonomies. A matrix representing mean pair-wise distance between all birds was subjected to cluster analysis, using the average link method (Sokal & Sneath, 1973), yielding the tree diagrams shown in Figs. 1–6. On both sets of birds, the three groups showed overall similarly coupled group differences. In all the taxonomies there were groups of predators, game birds, water birds, hummingbirds, and woodpeckers, to name a few. Some notable differences in the taxonomies are as follows. Whereas US non-experts and US experts generally kept all passerines (small songbirds) together in a large single group, the Itza' had them spread out more across the taxonomy in a few different clusters. For example, American subjects grouped the swallows near the flycatchers, while Itza' placed flycatchers in groups not adjacent to swallows.

We also found a difference in subjects' sorting of 'water birds'. On the US bird set, US experts had a large "water birds" cluster, featuring ducks, grebes, geese, "shore birds" and herons/egrets. This cluster was fairly isolated from the rest of the taxonomy. Although non-experts also had a water bird category, it was more spread out, was not as isolated from other birds, and was interrupted by non-water birds, such as game birds, nightjars
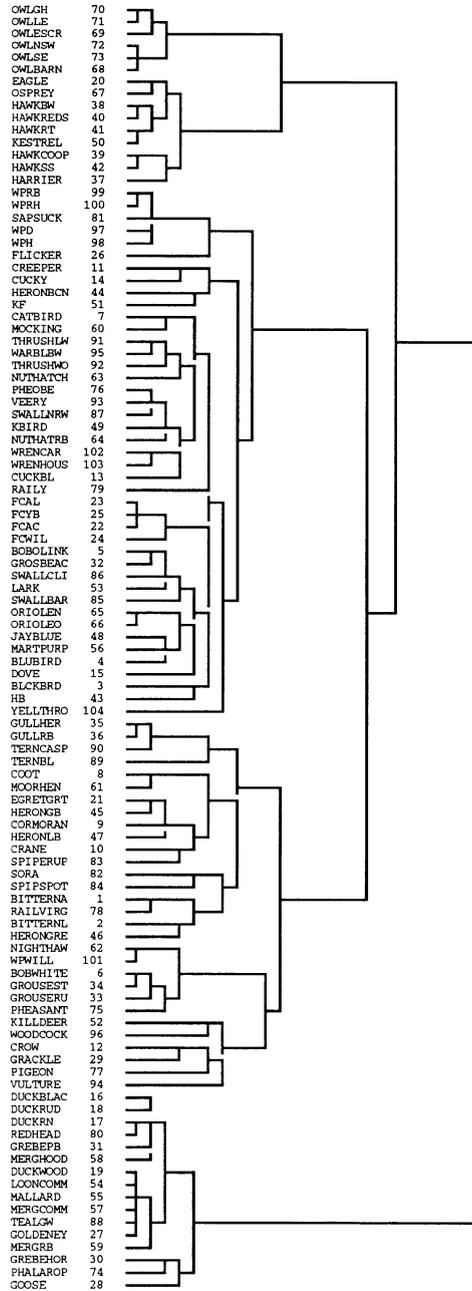
Fig. 1. Cluster analysis of non-experts sorting US birds. The numbers next to the bird names correspond to an alphabetic sort of the names.
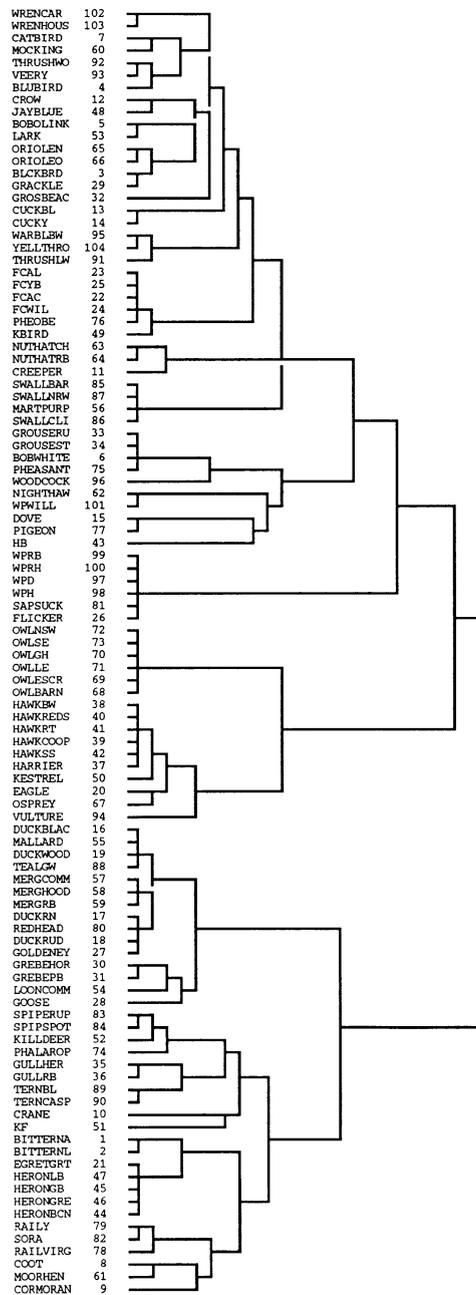
Fig. 2. Cluster analysis of experts sorting US birds. The numbers next to the bird names correspond to an alphabetic sort of the names.

Fig. 3. Cluster analysis of Itza' sorting US birds. The numbers next to the bird names correspond to an alphabetic sort of the names.

Fig. 4. Cluster analysis of non-experts sorting Tikal birds. The numbers next to the bird names correspond to an alphabetic sort of the names.

|	
VULTTURK	94
VULTBLCK	95
VULTKG	58
YELHEAD	56
OWLSC	74
OWLGH	98
OWLBW	70
OWLMOTT	71
OWLBARN	97
OWLPYG	72
FALCCOLL	12
KESTREL	96
KITESN	7
KITECY	55
KITEHB	6
HAWKEAG	11
HAWKBLAC	8
HAWKGB	9
HAWKRS	10
FALCBARR	13
HAWK2EAG	59
FALCBAT	14
FALCLAUG	57
PARROT	21
PARAKEET	63
MACAW	62
PARRBC	25
PARRRC	65
PARRWF	20
PARROC	22
PARRRL	24
PARRBH	23
WPBC	38
WPGF	39
WGGO	37
WPL	40
WPGC	41
WPCC	36
HBMANG	29
HBFAIR	30
HBSB	26
HBEMR	27
HBJACOB	28
ORPENW	91
ORPENMZ	93
ORIOYT	48
ORIOYB	49
BBIRDS	92
ROBING	64
ROBWT	90
JAYBRWN	89
TANYW	50
GROSB	52
BECARDRT	43
TANRCA	51
JAYGREEN	101
FOLGLEAN	42
TOUCNEM	75
ARACARI	78
TOUCAB	79
ANIGB	76
CUCKCS	77
TROGAN	33
QUETZAL	83
MOTMOT	32
FCBC	46
FCYC	47
FCGC	44
FCDC	45
FCGS	84
FCD	85
MARTP	87
SWALLRW	88
MARTGB	86
SWALLB	100
SWIFTG	80
SWIFTVS	99
SWIFTL	82
KFR	34
KFF	35
HERONB	5
HERONLB	102
EGRETCOM	103
HERONBB	54
QUALPART	60
QUAILST	61
TIN2	2
TING	4
TIN1	1
TINL	3
TURKEY	31
POTOOG	73
PWILLY	81
JACANAN	53
TEALBW	104
CURASG	67
GUANC	68
CHACAW	66
PIGSB	16
DOVEGRD	17
DOVEGH	19
PIGSCAL	15
DOVEQUAL	18
DOVEBG	69

Fig. 5. Cluster analysis of experts sorting Tikal birds. The numbers next to the bird names correspond to an alphabetic sort of the names.

Fig. 6. Cluster analysis of Itza' sorting Tikal birds. The numbers next to the bird names correspond to an alphabetic sort of the names.
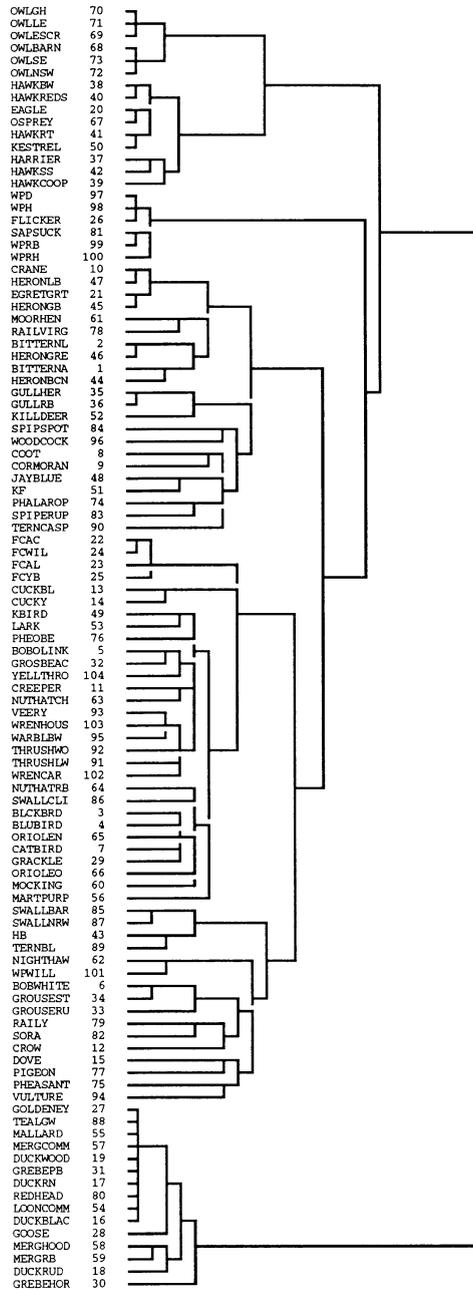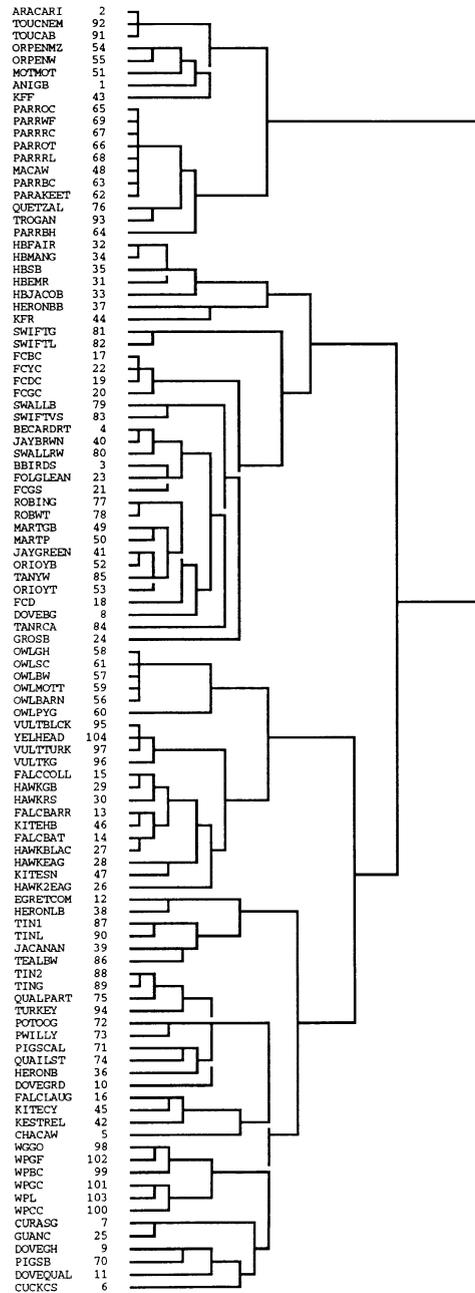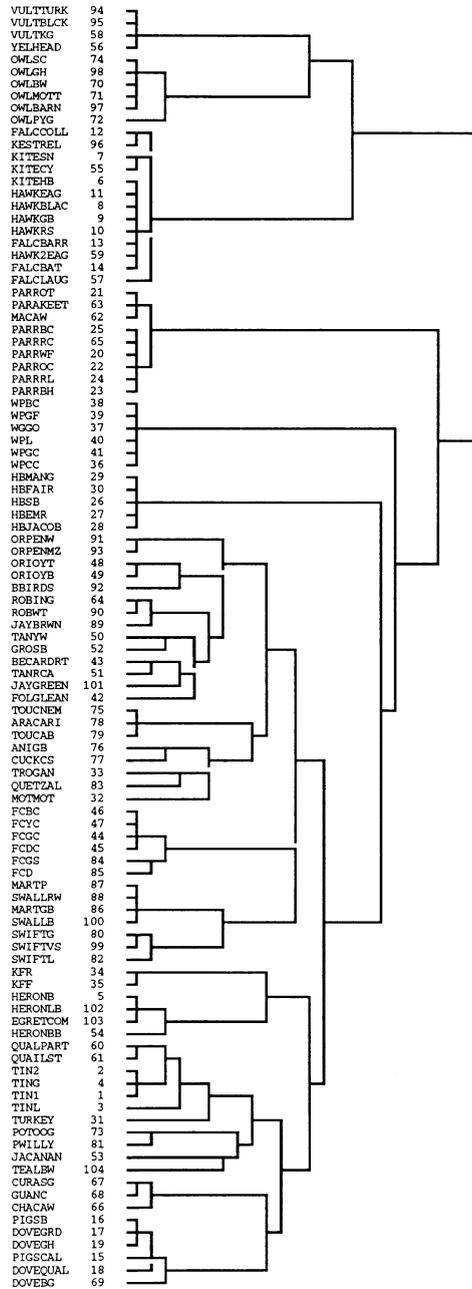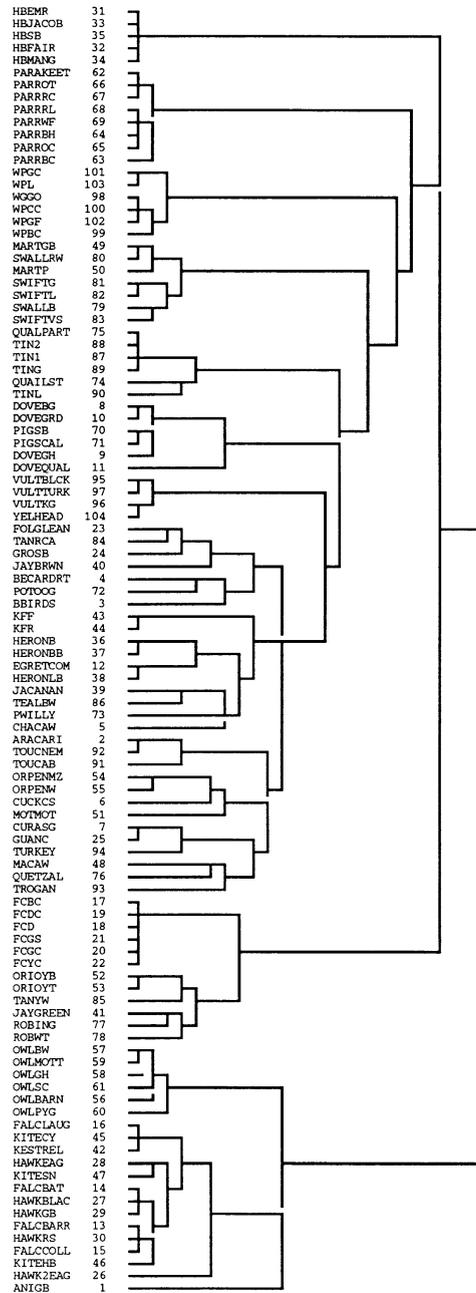
(birds that eat insects while they are flying), the pigeon, and the turkey vulture. This also reduced the correspondence of novice sorts to scientific taxonomy.

2.2.2.6. Justifications for sorts We grouped sorting justifications as taxonomic, morphological, ecological or behavioral. Because the groups did not differ substantially in their justifications we will not present them in detail. Justifications on initial sorts were mainly taxonomic, except in the cases of non-experts, who relied more on morphological justifications. The latter observation is not surprising given that the non-experts know few bird categories. Splits also tended to be justified in terms of taxonomy or morphology. Lumps were coupled with ecological and/or behavioral justifications, even for the non-experts (though the ecological justifications for the non-experts tended to be fairly abstract).

2.3. Discussion

The naming data were as expected and confirm the group differences in familiarity and expertise. Novices typically could not identify specific birds from either set. US experts were able to name US birds more accurately than Tikal birds. Itza' were better than US experts in identifying Tikal birds. Nothing about these results is surprising.

The sorting data, however, revealed some interesting and important group differences. Although the sorting results generally confirm our hypotheses as well as previous claims that different cultural groups converge in the ways they assess similarity to classify birds (and other vertebrates) this convergence still leaves over half of the variance unexplained. In almost every case, residual within group agreement was reliably higher than across group residual agreement.

Most striking are the observations concerning correlations with scientific taxonomy. We had expected that relative experts, especially Itza', might be more likely to have developed a special-purpose taxonomy and show a lower correlation with science than non-experts. To the contrary, both US expert and Itza' expert correlation with science were much higher than the non-expert with science. The US expert data are not completely surprising in that guidebooks are often organized in a way that corresponds with scientific taxonomy. The Itza' data are dramatic in that despite not being exposed to either western science in general or formal taxonomy in particular, their consensual sorting agrees more with (western) scientific taxonomy than does the consensual sort of US non-experts. This difference held for both US birds and Tikal birds. In short, not only did the Itza' fail to show a clear influence of a special-purpose taxonomy but also they join US experts in reflecting greater sensitivity to taxonomic relations than non-experts.

The results on expertise are not unprecedented. Johnson and Mervis (1998) tested bird experts, fish experts and novices on a triads task where participants were asked to pick out the two animals that were "most like the same kinds of thing". Some triads pitted overall morphological similarity against taxonomic membership. Not only were bird experts more likely to make the taxonomic choice for birds and fish experts to make the taxonomic choice for fish but also these two types of expert were substantially more likely than novices to pick the taxonomic choice for the domain where they lacked expertise. Johnson and Mervis suggested that experts had learned to weight modified parts as much as features

more related to overall similarity in contrast with novices who apparently gave the latter type of feature more weight. In short, the Johnson and Mervis findings support the idea that some combination of perceptual learning and what they referred to as "intuitive theories" (e.g. understandings of the functional significance of different features) leads experts to organize biological kinds in a manner closer to scientific taxonomy. Our results are consistent with this general interpretation in that the two groups of experts were clearly using information not reflected in the novice sorts. In short, expertise appears to involve more than a passive reception of real world structure – it includes learning to attend to the features and relationships that are most informative (see also Boster & D'Andrade, 1989) which does not necessarily correspond with overall similarity.

As noted earlier, work by Boster and Johnson (1989) indicates that expert fishermen show less consensus than non-experts. This difference in findings from our results likely owes to different interests that underlie different kinds of expertise: Boster's fishermen tended to have particular, idiosyncratic fish-catching strategies and preferences, whereas US birders may simply aspire to understand all birds better. Similarly, the Itza' appear not to employ a special-purpose taxonomy, perhaps because of their broad interest in ecological relationships and the health of the forest.

Other work by Boster (1987) suggests that passerines are more easily confused than non-passerines. Itza' were better at naming non-passerines but interestingly, experts were equally adept at naming passerines and non-passerines. US birders focus on identification as their primary goal; passerines may not be as psychologically or culturally salient for Itza'. In the next experiment we report data from typicality or goodness of example (GOE) judgments which may also bear on this notion of salience.

3. Experiment 2: GOE ratings

In Experiment 2, we sought to determine the types of birds subjects believed to be 'good examples' or 'ideal birds'. We then compared the results from Experiment 2 to Experiment 1 in order to see if central tendency was a crucial factor in assessing GOE. In addition, we examine group differences in typicality ratings and their basis. As noted previously, we hypothesized that central tendency might play a critical role in assessing GOE for non-experts only.

3.1. Method

For the US subjects, we used the instructions from Rosch and Mervis (1975), which featured a scale where 1 indicates high typicality (or GOE) and 7 indicates low typicality. Separate groups of 25 new US experts (similar characteristics as in Experiment 1, 12 on the Tikal birds, 13 on the US birds), 22 new US non-experts (similar characteristics as in Experiment 1, 12 on the Tikal birds, ten on the US birds), and nine of the same Itza' subjects from Experiment 1 rated the GOE of the birds. Each Itza' subject and US non-expert subject rated all 104 birds in the set, but each US expert rated a random sample of 26 birds from the respective set of 104 birds. Itza' subjects were asked to rank the 104 Tikal birds in terms of 'how true an example of a bird' each one was. Pilot studies indicated that, in this context, the Itza' word for "the true" (*jach*) exemplar would be

glossed with equal likelihood in Spanish as "the true/real/pure/ideal" (*el mero*) or "most typical" (*lo más tÍpico*) exemplar.

3.2. Results and discussion

Ranked typicality ratings, ranked central tendency scores, and number of birds per family for all subject groups and all birds appear in Appendix B. We correlated central tendency and rated goodness of fit for each subject group. Central tendency sorting (i.e. the average pair-wise distance between a bird and every other bird in the set) correlated with GOE ratings for the non-experts only ($Rxy = 0.52$ and 0.44 for US and Tikal birds; the number for US experts was 0.00 and 0.12; for Itza' it was -0.12). This fits with work reported by Lynch et al. (2000); typicality ratings do not seem to be based on central tendency for people relatively knowledgeable in a domain.

We next examined the different patterns in the GOE ratings for the three subject groups, keeping in mind the possibility that the groups might differ in their ideals. After exploring the data, we found the most striking difference to be the *passerine effect*. The US experts rated passerines ($M = 1.95$) as more typical than other birds ($M = 2.56$) ($F(1, 24) = 15.71$, $P < 0.01$). Furthermore, the average rank by US experts for US passerines ($M = 23.50$) was higher than Tikal passerines ($M = 44.11$) ($t(57) = 0.003$), indicating that US experts considered local passerines uniquely typical. US non-experts also rated passerines ($M = 2.56$) as more typical than other birds ($M = 3.90$) ($F(1, 21) = 21.83$, $P < 0.01$). The opposite was true for Itza'. The Itza' ranked non-passerines ($M = 29.86$) as significantly more ideal than passerines ($M = 36.60$) ($F(1, 8) = 6.45$, $P < 0.05$). While passerines account for a slightly higher proportion of the total bird population in the US than they do in Tikal, songbirds still comprise the most numerous order in Mayaland. Nonetheless, Itza' do not rate them as typical.

Previous categorization research suggests that statistical similarities or central tendencies can be determinants of prototypes or best examples (Boster, 1988; Mervis & Rosch, 1981; Rosch & Mervis, 1975; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976); however, there are many different possible measures of similarity and central tendency (Medin, Goldstone, & Gentner, 1993). We chose to use taxonomic distance as a measure of similarity in order to calculate central tendency, although we realize that alternative measures of similarity and central tendency could be used that might produce different results. Our measure has certain advantages: it makes use of natural taxonomies that are readily generated by sorting experiments in any cultural setting in which such experiments have been tried (Boster, 1991; López et al., 1997; Medin et al., 1997), and it does not depend upon explicit elicitation that may involve culturally-specific instructions (e.g. "similar to" has no ready or singular translation in Maya languages). Our results indicate that, at least in folkbiology, ideals can be made salient by factors other than similarity-based central tendency (Atran, 1999; Barsalou, 1985; Lynch et al., 2000). As we shall see, these other factors may reflect the special interests of the populations concerned.

4. Experiment 3: category-based induction

Experiment 3 builds on the results of Experiments 1 and 2 by asking how participants

use bird categories and salient examples of birds in reasoning. An important function of taxonomic classification is enabling generalizations between categories. Osherson et al. (1990) identify a set of phenomena that characterize category-based inferences in adults, and formalize a model that predicts the strength of those inferences. Consider argument (i) below:

(i)	Hyenas have an ileal vein
	Cows have an ileal vein
	Wolves have an ileal vein.

This argument is strong to the extent that belief in the premises leads to belief in the conclusion. There are two components to the Osherson et al. (1990) similarity-coverage model (SCM). The first component of the model, *similarity*, calculates the maximum similarity of the premise categories to the conclusion category; the greater this similarity, the stronger the argument. In this example, hyenas are more similar to wolves than cows are, hence similarity is calculated for hyenas. The second component – *coverage* – calculates the average maximum similarity of premise categories to members of the "inclusive category" – the lowest category that includes both premise and conclusion categories. For argument (i), the inclusive category is presumably *mammal*. In our research, the inclusive category is simply the conclusion category. The greater the coverage of the inclusive category by the premise categories, the stronger the argument. Sloman (1993) presents an alternative model; although Sloman's feature-based model differs from the SCM in important ways, for our purposes the models make the same predictions.

We focus on two phenomena: typicality and diversity. Both of these phenomena hinge on coverage. The typicality phenomenon predicts that a more typical instance promotes stronger inferences to a category than a less typical instance. Typicality in this case is computed in terms of central tendency; the typicality of an item is the average taxonomic distance of that item to all other items in the inclusive category. The higher the average similarity of that item to other members of the category, the more typical it is. Thus, more typical items provide greater coverage than less typical ones.

Like typicality, diversity is a measure of category coverage. The diversity phenomenon predicts that an argument will be inductively strong to the degree that categories mentioned in its premises are similar to different instances of the conclusion category. For example, consider arguments in (ii):

(iia)	Jaguars have protein Y
	Leopards have protein Y
	All mammals have protein Y.
(iib)	Jaguars have protein Y
	Mice have protein Y
	All mammals have protein Y.

The SCM predicts that the categories mentioned in the premise of (iib) provide greater *coverage* of the conclusion category *mammal* – i.e. are more similar to more mammals –

than the categories mentioned in the premises of (iia), thus making (iib) the stronger argument. Indeed, most American college-aged subjects agree that (iib) is stronger than (iia) (Osherson et al., 1990). In general, diversity predicts that an argument with more diverse premises will be evaluated as stronger than an argument with less diverse premises.

Before focusing on the methodological details of Experiment 3, we discuss the properties we chose for induction. Based on previous work, we decided against using identical properties for the Itza' and US induction probes. The rationale is much the same as noted in developing the sorting instructions for Experiment 1. Half of the probes involved disease and this was constant across groups. For the other half we used "enzyme" for North American subjects and "little things inside" for Mesoamerican subjects. We piloted both terms with both groups. We found that North American adult subjects are confused by "little things inside" but not "enzyme", "protein" or "disease X" (different subjects projecting different types of contents), whereas Maya subjects were confused by "enzyme" and "protein" but not by "little things inside" or "disease X". Earlier studies show that the patterns of results on different kinds of biological induction tasks for American undergraduates were statistically the same for "enzyme" and "disease", whereas the Itza' showed the same patterns of results for "little things inside" and "disease" (Atran, Estin, Coley, & Medin, 1997; Coley, Medin, Proffitt, Lynch, & Atran, 1999).

Given the results from López et al. (1997) and Proffitt et al. (2000), we hypothesize that US non-experts should exhibit more diversity responding than either of the other two groups. It would not be surprising if the US bird experts showed some modest amount of diversity responding given that they are quite familiar with the scientific taxonomy. If so, we would expect this to be more predominant among US birders who have received formal training. Overall, however, our hypothesis is that domain knowledge makes it less likely that a person will employ abstract strategies. Instead we expect to observe more concrete justifications such as the ecological/causal reasoning.

4.1. Method

4.1.1. Typicality stimuli

Eight typicality pairs (each pair consisting of two birds) were constructed for each of the stimulus sets (US and Tikal). Within each of the eight pairs of birds in the US set, the two birds were matched in terms of frequency of sighting. Birds in the Tikal set were not matched for frequency because of a lack of a source from which to rank the birds on the dimension. The stimuli were created by crossing two variables: Typicality, or central tendency (high vs. low) gauged by mean distance from other birds in the sorting task, and Family Size (high vs. low), a measure of how many members were in the same family as the bird in question. We decided to include family size as a factor because Proffitt et al. (2000) found that it was reliably used as the basis of choices on typicality probes. We had two pairs in each of these four cells for a total of eight pairs in each stimulus set. These typicality pairs appear in Appendix C.

4.1.2. Diversity stimuli

Four diversity pairs, each pair consisting of two pairs of birds (i.e. four birds in total)

were constructed for each of the stimulus sets (US and Tikal). Within each diversity pair, one pair of birds was "close" in terms of mean distance between the birds in the pair across the three subject groups; the other pair was "far" in terms of mean pair distance. To be eligible for selection for diversity stimuli, birds had to meet two requirements: (1) each bird had to be high (top fifty percentile) in mean central tendency (mean distance from all other birds in the set) across the three groups; and (2) each pair of birds had to be high (top fifty percentile) in agreement on mean pair distance for each of the groups. Of course the critical variable was coverage, and to this end we ensured that birds in the far pair had a mean distance (between each other) of at least one level higher than birds in the close pair. In addition, we computed diversity as a function of simple minimal distance based on how similar each bird in the pair was to all of the other birds in the set. Both measures will be considered in Section 4.2. Within each diversity pair, close and far pairs of birds were matched in terms of family size, central tendency, and frequency of sighting for US birds (as determined by the Audubon Society), and family size and central tendency for Tikal birds (there was no source for frequency of the Tikal birds). These diversity pairs appear in Appendix D.

4.1.3. Design

Stimulus set (US vs. Tikal) and question type (disease vs. enzyme) were crossed and blocked within subjects. Subjects saw two main blocks, one with US birds and the other with Tikal birds. Within each of these blocks, there were two question type blocks. One type of question involved induction about enzymes and the other about diseases. Each subject saw all four blocks (US-Disease, US-Enzyme, Tikal-Disease, and Tikal-Enzyme). There were four possible orders of the four blocks. Within each of the four blocks, subjects saw two diversity trials and four typicality trials. The order of the trials within each block was randomized. In addition, after the four blocks described above, each participant received one additional block consisting of four diversity trials. These trials were comprised of only Tikal birds, and were included to match some previously acquired data from the Itza'. The stimuli on these blocks were matched for Itza' central tendency but not for agreement or frequency.

4.1.4. Procedure

Informants were run individually. American participants were told that they would be shown illustrations of birds and that they would be asked to answer questions about them. For the typicality trials, we displayed both birds in each pair and asked:

> Let's assume that we discovered two new diseases (or enzymes). All we know about these diseases is that Disease A is found in these types of birds and Disease B is found in these types of birds. Which disease is more likely to be found in all types of birds?

The instructions were the same for the Itza' subjects, however instead of discussing 'enzymes', we discussed 'very small things inside of the birds'. Subjects were required to select one of the birds in the pair, and to provide a justification for their decision. Similarly, for the diversity trials, we placed one pair of birds on the left-hand side and

one pair of birds on the right-hand side, and asked the same question. The far pair appeared on the left-hand side for half the trials, and on the right-hand side for the other half.

4.1.5. Subjects

The US experts were ten men and two women (mean age = 51.8 years, range: 15–88 years) having either occupations or extensive experience related to birds. The average experience in birding was 26.1 years (range: 4–81 years). One subject had not yet completed high school, one had at least some college work, four had completed college, and five had advanced degrees, two of those being PhDs. On a seven-point continuous scale with 1 indicating 'very little knowledge about birds' and 7 indicating 'total expertise', the mean self-reported rating for US experts was 5.00 (SD = 1.04). The US experts were drawn predominantly from the same pool of subjects as the sorting task, and were paid for their participation. Occupations of subjects were extremely varied and included college professors, salesmen, insurance investigators, retired workers, lawyers, store owners, botanists, and homemakers.

The non-experts were two men and 11 women (mean age = 25.08, range: 18–41 years) who were Northwestern University students. The average number of years birding was zero. Three of the subjects had completed at least a college degree; the rest were still working towards their first degree. On the self-reported bird knowledge scale the mean rating was 1.77 (SD = 1.09). The non-experts were paid for their participation. The Itza' were 12 men and women taken from the same population as in Experiment 1, with similar ages, education experience, and knowledge concerning birds. Five of the Itza' subjects also participated in the first experiment. The time lag between the two studies was approximately 6 months.

4.2. Results

First we will describe the results from typicality probes and then turn to the diversity items. In every analysis reported below, we do not find any differences due to question type. In other words, subjects responded in a similar manner regardless of whether the question dealt with diseases or enzymes (or very small things). Consequently, we do not discuss this variable further. As we shall see the justifications for choices are as important as the choices themselves. To anticipate, in each case, we find clear group differences.

4.2.1. Typicality results

4.2.1.1. Trial selections   Table 2 shows the mean responses on the familiarity and typicality trials. On Family Size trials (those in which typicality was controlled), the dependent variable is the percentage of trials in which the high family size bird was chosen; on typicality trials (those in which family size was controlled), the dependent variable is the percentage of trials in which the high typicality bird was chosen.

As is apparent from Table 2, only the undergraduates showed any clear indication of a typicality effect (mean = 0.78 vs. 0.57 for US experts and 0.50 for Itza'). Across subject group and stimulus set, participants chose the more typical bird on 62% of the trials, reliably greater than chance ($t(36) = 20.42$, $P < 0.01$). We then ran an ANOVA with

Table 2
Mean percentages of high typicality responses on typicality trials and high family size responses on family size probes, by condition in Experiment 3

	Subject group		
	US expert	Non-expert	Itza'
Typicality trials			
Tikal birds			
Disease	0.50	0.69	0.54
Enzyme	0.71	0.85	0.67
US birds			
Disease	0.58	0.88	0.34
Enzyme	0.50	0.70	0.46
Family size trials			
Tikal birds			
Disease	0.58	0.77	0.34
Enzyme	0.67	0.77	0.45
US birds			
Disease	0.54	0.30	0.34
Enzyme	0.58	0.38	0.20

subject group and stimulus set as factors, and typicality responses as the dependent variable. None of the main effects or interactions were significant.

Using data from the unique typicality pairs within each subject group and stimulus set (20 observations), we ran a regression with bird choice as the dependent variable and GOE ratings (from Experiment 2) and central tendency as factors. For each of the two factors, we took a difference in average rank (either central tendency or GOE) of the two birds in the pair. The overall regression was not significant ($F(2, 17) = 0.46$, $R^2 = 0.05$), and none of the individual factors were significant. Next, we separated the analysis for the three subject groups. The only instance in which the model approached significance was for the US non-experts ($F(2, 5) = 2.98$, $P < 0.10$, $R^2 = 0.54$). In that analysis there was a marginal effect of GOE ratings ($t(7) = 2.09$, $P < 0.09$).

On the Family Size probes, subjects chose the bird with the larger family on 49% of the trials, essentially chance responding. Next, we ran an ANOVA on choices with subject group, stimulus set, and family size (high vs. low) as the factors. There was a main effect of stimulus set ($F(1, 34) = 10.58$, $P < 0.01$), indicating that the high family size bird was chosen more often on the Tikal bird set ($M = 0.59$, SD = 0.17) than on the US bird set ($M = 0.39$, SD = 0.14). None of the other main effects were significant. The interaction between stimulus set and subject group was significant ($F(2, 34) = 3.76$, $P < 0.05$): non-experts chose the high family size bird more on the Tikal bird set than on the US bird set, while the other two subject groups did not show as pronounced a differentiation. We looked at the items contributing to this interaction and found that it owed mainly to two trials. Interestingly, those trials that drove the interaction between stimulus set and subject group involved a choice between a passerine and a non-passerine. In both cases, non-experts were more apt to select the passerine over the non-passerine. In light of this trend,

we then looked more closely at this passerine effect to see if it extended towards the rest of the trials.

4.2.1.2. The passerine effect Of the 16 total family size and typicality trials, there were six trials where one bird in the pair was a passerine and the other was not. Using just these six trials we ran a by-item ANOVA with subject group as the independent variable and the percentage of trials that subjects chose the passerine as the dependent variable. Even with only six items, there was a significant effect of subject group ($F(2, 10) = 23.47$, $P < 0.01$). The US experts and non-experts chose the passerine over the non-passerine (66 and 86%, respectively) more than the Itza' (40%). To ensure that this difference was reliable, we compared the means of all three subject groups using Tukey's HSD (alpha = 0.01). The Itza' differed from both the US experts and non-experts, while the difference between US experts and non-experts was not reliable. In short, the Itza' avoided generalizing traits from passerines while the US subjects, especially the non-experts, preferred to generalize from them.

4.2.1.3. Typicality trial justifications We realize that explicit post hoc justifications do not always correspond to implicit and actual bases for choice and decision (Nisbett & Wilson, 1977); however, folk taxonomic justifications are often independently borne out (e.g. birds that are grouped together because they eat other birds are generally predators). We used the following justification categories for the typicality trials. (1) *Typicality* – any justification that cited a bird as being a more typical or ideal example of a bird ("These birds are more normal or more representative examples of birds"). (2) *Behavioral* – any justification that described a category's behavior, including locomotion ("hops around"). (3) *Ecological* – any justification that involved some sort of description of a category of birds and their relation to their environment, be it other birds or animals ("predator"), their diet ("eats bugs"), or their habitat ("lives in marshy swamps"). (4) *Geographical Range* – any justification that described a greater geographical range for the chosen bird ("This bird travels extremely far"). (5) *Number* – any justification that explicitly mentioned frequency of occurrence as a reason for selecting a bird ("There are a lot of this kind of bird"). (6) *Evolutionary Age* – any justification that cited an earlier emergence on the evolutionary time line as a reason for choosing a bird ("This bird was around long before the other birds").

Table 3 shows the types of justifications used on the typicality trials by group. The most striking difference is that non-experts use typicality as a reason for the choice more than half of the time, while US experts and Itza' *never* indicate typicality. Both Itza' and US experts tended to use range or ecological factors as justifications but only the US experts commonly employed evolutionary age as a justification. For each justification category, we ran a by-subject ANOVA with subject type and stimulus set as independent factors and the percentage of trials on which that particular justification was invoked as the dependent variable. Table 4 provides statistical justification for the differences apparent in Table 3.

To sum up the typicality reasoning probes, none of the groups tended to pick the bird from the larger family and only the US non-experts sowed much evidence for central tendency-based typicality effects. Instead, other factors, such as presence of passerines,

Table 3
Expert, Itza', and non-expert justifications on typicality trials by stimuli set in Experiment 3[a]

Subject group	Justification type					
	TYP	BEH	ECO	GEO	NUM	EVO
US expert						
Tikal	0	0	18	39	2	21
US	0	0	12	37	4	23
Itza'						
Tikal	0	12	60	18	1	0
US	0	7	58	26	9	0
Non-expert						
Tikal	47	2	7	4	18	2
US	56	4	3	5	18	0

[a] Categories include typicality (TYP), behavior (BEH), ecology (ECO), geographical range (GEO), number (NUM), and evolutionary age (EVO).

were more responsible for choices. Furthermore, only non-experts used typicality as a justification for their choices.

4.2.2. Diversity

4.2.2.1. Trial selection The mean percentage of diversity choices by condition appears in Table 5. Across conditions, US experts chose the far pair on 58% of the trials, the non-experts 58%, and the Itza' 45%. None of these percentages differed reliably from each other or from chance (50%). There was a trend for non-experts to show greater diversity

Table 4
Summary of reliable ($P < 0.05$) main effects found for typicality and diversity trial justifications in Experiment 3[a]

Justification category	Typicality trials		Diversity trials	
	Subject type	Stimulus set	Subject type	Stimulus set
Typicality	N > E, I	US > TIK	N > E, I	NS
Behavioral	I > N, E	NS	I > N, E	NS
Ecological	I > N, E	NS	I > N, E	NS
Geographical range	E, I > N	NS	E, I > N	NS
Number	N > E, I	NS	NS	NS
Evolutionary age	NS	NS	NS	NS
Diversity	–	–	N > I	NS

[a] Subject groups are represented by US non-experts (N), US experts (E) and Itza' (I). Subject type effects are listed in the first subcolumn. Stimulus set effects are listed in the second subcolumn, and indicate a difference between justifications based on whether the American (US) or Itza' (TIK) stimulus set was used. All of these tests are not independent.

Table 5
Mean percentages of diversity responses by stimulus set and subject group in Experiment 3

Stimulus set	Subject group		
	US expert	Non-expert	Itza'
Tikal			
Disease	65	56	43
Enzyme	54	42	46
US			
Disease	46	69	50
Enzyme	67	65	46

responding for the US birds (67%) than the Tikal birds (49%) but it was not reliable. Collapsing across question-type, we ran an ANOVA on diversity choices. With our small N, none of the main effects or interactions was significant. We then examined the data at a finer level.

4.2.2.2. Coverage models As a further test for diversity effects we ran a regression on choices using the two measures of coverage described in Section 4.1 (minimal distance from other birds in the set and the distance between the two birds in the pair) as independent variables. There were 24 diversity items in total (2 stimuli sets $\times$ 3 subject groups $\times$ 4 different pairs per stimuli set). The overall regression was not significant ($F(2, 21) = 1.16$, $R^2 = 0.20$), and neither of the individual factors was significant. Next, we ran three separate analyses for the individual subject groups. None of the individual models approached significance. We then examined our now familiar passerine effect.

4.2.2.3. The passerine effect Once again it appears that the US populations tended to choose probe pairs involving passerines while the Itza' tended to avoid them. To verify these trends we performed a by-trial analysis including all 20 trials used for the diversity task (8 diversity pairs $\times$ 2 question types = 16 conditions, plus the 4 additional pairs run at the end of the reasoning tasks = 20 total trials). We ran a regression that contained all 20 possible trials, and coded the independent variable " -1 " if there were more passerines in the far pair than the close pair (four instances), "0" if there were an even number of passerines in both pairs (seven instances), and "1" if there more passerines in the close pair than far pair (nine instances). We entered as the dependent variable the percentage of times each subject group chose the *far* pair on that particular trial.

Notice that we would expect for the US subject groups a negative relationship between our independent and dependent variables, because they should tend to prefer the far pair when it contains more passerines than the close pair (instances of " -1 "). On the other hand, for the Itza' we would expect a positive relationship between our independent and dependent variables, because they should tend to prefer the far pair when it contains fewer passerines than the close pair (instances of "1"). In all three regressions, our suspicions were confirmed. The US experts and non-experts showed a reliable ($P < 0.05$) negative relationship ($t(19) = 2.19$, $b = -13.66$, and $t(19) = 4.07$, $b = -26.45$, respectively). On

Table 6
Expert, Itza', and non-expert justifications on diversity trials by stimuli set in Experiment 3[a]

Subject group	Justification type						
	DIV	TYP	BEH	ECO	GEO	NUM	EVO
US expert							
Tikal	25	0	0	5	42	14	11
US	25	0	0	7	36	14	18
Itza'							
Tikal	1	2	6	56	28	0	0
US	0	2	10	44	37	1	0
Non-expert							
Tikal	33	37	0	12	3	13	2
US	33	23	0	19	2	23	0

[a] Categories include diversity (DIV), typicality (TYP), behavior (BEH), ecology (ECO), geographical range (GEO), number (NUM), and evolutionary age (EVO).

the other hand, the Itza' showed a marginally reliable ($P < 0.10$, two-tailed test) positive relationship ($t(19) = 1.79$, $b = 11.68$).

It seems that our groups based their choices to a considerable degree on the presence of passerines or some factor correlated with it. This passerine effect suggests that the idealness of the birds may be driving our results more than coverage. To the Itza', passerines are not considered "true birds" to the same extent as other birds in the environment.

4.2.2.4. Diversity trial justifications We coded diversity trial justifications into seven main categories: the six used for typicality trials and a seventh, *diversity*. Diversity involved any justification that invoked greater taxonomic coverage as a reason for choosing a pair of birds ("This pair is more varied and the birds come from different families").

Subjects' justifications often included more than one of these six categories (i.e. "typical birds of prey"). In these cases, justifications were broken up into their component categories and each category was assigned a share of the justification such that the total of the shares added up to 1. Table 6 displays the mean justification scores for these seven categories by group and stimulus set.

Overall, US experts gave diversity justifications on 26% of the trials. However, this proportion was driven almost entirely by two subjects who gave diversity justifications on virtually every trial. Interestingly, these two subjects had fewer years experience of birding ($M = 8.5$) than the rest of the US experts ($M = 29$). Consequently, they may have justified their choices in a manner similar to US novices. Six out of the 12 US experts did not give a single diversity justification on any of their 12 diversity trials. Non-experts gave diversity justifications on 35% of the trials. The Itza', in comparison, gave virtually no diversity justifications on any of the trials (one response total). We ran an ANOVA on diversity justifications as the dependent variable and the only significant effect was subject type ($F(2, 34) = 4.17$, $P < 0.05$). Tukey tests indicated that overall, Itza' used diversity

justifications less than US non-experts ($P < 0.05$). None of the other comparisons were significant.

There were also other notable differences in justification patterns evident in Table 6. US experts were the only group to rely on evolutionary justifications and the two expert groups were more likely to use range as a justification. The Itza' were also more likely to focus on ecology than the US groups. Table 4 provides statistical justification for these differences. This effect of ecology sheds some light on the passerine effect described above. Compared to the larger and more perceptibly distinct game birds, the smaller songbirds play relatively minor or undifferentiated roles in defining the ecological relationships that Itza' consider important to their lives and to the life of the forest that hosts them (cf. Atran & Medin, 1997). This relative lack of ecological saliency may well cause the Itza' to categorize and reason differently about them.

4.2.2.5. A "learning effect" An interesting trend in the justification data can be seen when we break down diversity and typicality justifications by block and subject group. Because the Itza' did not give either diversity or typicality justifications we did not include them in the following analysis. We partitioned the data into two blocks. The first block included the first four diversity trials (i.e. those in trials 1–12), and the second block included the four diversity trials in trials 13–24. Non-experts used a diversity justification on only 17% of trials in block 1; this proportion jumps to 43% for the trials in block 2. It appears that non-experts learn to apply the diversity justification as a strategy to respond to the diversity questions. Experts did not vary in their use of diversity justifications by block (21% in block 1, 25% in block 2). This trend also applies to non-experts' diversity choices, in that they selected the diverse pair more often in the second block ($M = 58\%$) than in the first block ($M = 54\%$). Apparently, the use of other strategies led to non-diversity so that the overall effect was one of modest diversity effects. Finally, we tested to see if diversity justifications corresponded to diversity choices. On trials with diversity justifications, the diverse pair was chosen 98.89% of the time. On trials without diversity justifications, the diverse pair was chosen 43.06% of the time, not reliably different from chance. This trend was the same for both US experts and US non-experts (the Itza' never gave diversity justifications).

To sum up the diversity results, unlike previous studies, we did not demonstrate a reliable diversity choice effect for non-experts. However, non-experts did show an increase in both diversity choices and diversity justifications as the experiment progressed. The best predictor of choices was presence of passerines – birds that tend to be 'better' examples for the US subjects and poor examples for the Itza' subjects. Finally, as in the first experiment, the justifications also illuminate differences between the two cultures: Itza', who monitor their local ecology more than the other groups, almost exclusively relied on behavioral and ecological explanations.

4.3. Discussion

Experiment 3 provides us with insight concerning the ways that cultural influence and relative expertise influence how people reason about natural kinds. Even though 'passerine' was rarely cited as a justification, American subjects tended to pick small songbirds as

generalizing to the population of all birds, whereas Itza' preferred larger, more perceptually striking birds. Given the prominent role of the larger game birds in the behavioral ecology of Mayaland, and the more interactive goals of Itza' in monitoring their ecology, then the information provided by their ideal birds would be more relevant to environmental understanding and management than information provided by songbirds. Itza' preferentially monitor those species in their ecosystem (e.g. game birds as opposed to passerines) that provide the most relevant information about the interaction of human needs with the needs of the forest.

For example, Itza' tend to have the most detailed knowledge of, and to best protect, those species that are perceived to have the most interactions both with other species and with humans (Atran et al., 1999, in press). For the Americans, whose interest in, and interaction with, the behavioral ecology is of a much reduced and altogether different order (game birds are not considered palpably crucial to survival of the human habitat), correlated perceptual information may be more relevant by default. The net effect of the passerine and the anti-passerine strategies was very modest diversity responding.

5. General discussion

The purpose of running three distinct subject groups on natural categorization and reasoning tasks in the domain of folkbiology was twofold: (1) to better determine the extent that categorization and reasoning phenomena, such as typicality and diversity, actually do generalize from their manifestation in college students to the human species at large; and (2) to disentangle effects of cultural influence from effects of relative expertise in order to better understand the scope and limits of universal cognitive principles in the specific domain of folkbiology.

With respect to the above two issues, our triangulation strategy proved to be quite useful. For a number of important phenomena the US and Itza' clustered together and contrasted with US non-experts. First of all, the expert groups sorted in closer correspondence with scientific taxonomy than did non-experts. This difference is particularly striking for the Itza' with respect to US birds because the birds, western science, and scientific taxonomy were all unfamiliar to them. US non-experts had prior exposure to the birds and to western science but their sorts corresponded less well with scientific taxonomy than did those of the Itza'.

This finding gives no comfort at all to relativists. For that matter it is also inconsistent with the opposite extreme view, namely, that everyone naturally perceives the structure of nature unless goals and activities foster a special-purpose categorization scheme. Instead our data suggest that expertise confers benefits in abstracting important relationships in nature and, as a consequence, may lead to greater correspondence with scientific taxonomy. In that regard our results are well-anticipated by the findings of Johnson and Mervis (1998) who showed that bird and fish experts were better able than novices to apprehend relational features tied to function and ecology.

Our claim is not that all types of expertise will be associated with greater correspondence with science (that flies in the face of empirical evidence). At a minimum, one would need to add the proviso that expert characteristic goals and activities not lead to

special-purpose taxonomic systems. To avoid any implied circularity, one would need to analyze the daily goals and activities of the expert group under scrutiny. In the present case, the fact that neither expert birders nor Itza' employed a special-purpose, goal-derived sorting strategy is consistent with our analysis of their characteristic activities.

Differences between the two expert groups and novices in GOE ratings and in category-based induction are equally striking. First, consider GOE or typicality ratings. Since the original Rosch and Mervis (1975) paper on the basis of family resemblance, the consensus has been that typicality ratings are driven by central tendency (similarity relations). One exception is the Barsalou (1985) evidence that ideals may play some independent role in GOE ratings for common taxonomic categories, though he also found that central tendency was a reliable predictor of GOE ratings. Our present results reinforce the speculation that central tendency may only play a significant role in the typicality judgments of non-experts. Non-experts showed a robust correlation between central tendency and GOE ratings for both the US and Tikal birds. US experts and Itza' showed no reliable correlation of ratings with central tendency for either set of birds. This lack of correlation also holds for tree experts' typicality ratings for trees (Lynch et al., 2000).

If typicality is not a proxy for central tendency, then what drives it and what functions does it serve? Berlin (1992) suggests that good examples tend to be perceptually salient and that less salient examples come to be organized around them (see also Rosch, 1975). Salient examples might be the ones first learned and presumably their distinctive characteristics are paired with important shared properties with other category members. If so, then overall similarity is not the key organizing principle but rather some function of distinctive and shared characteristics. Other work (e.g. Atran, 1999; Barsalou, 1985; Lynch et al., 2000) points to idealness as an organizing principle.

The category-based induction findings also reinforce the view that the non-experts were the 'odd group out'. Non-experts relied very heavily on familiarity or typicality as the basis of their choices on both the typicality and diversity trials. Neither the Itza' nor the US experts *ever* gave typicality as a justification for either type of probe. Instead, they used knowledge about birds that the non-experts did not possess. For example, both the Itza' and US experts frequently mentioned the geographical range of birds, an explanation that the non-experts rarely produced. This is a truly striking qualitative difference.

Both US experts and non-experts used 'diversity' as a justification for a modest minority of probes. Itza' did not. But even in this case the differences with expertise loom larger than the similarities. Two bird US experts produced virtually all the expert diversity justifications and this pattern did not change across the test session. Non-experts, in contrast, gave twice as many diversity justifications in the second half of the tests as in the first. Apparently, once they thought of it, non-experts found the diversity justification intuitively appealing.

5.1. Itza' vs. US bird experts

There are impressive similarities between the Itza' and the US birders. In Experiment 1, on both the stimulus sets, the across-group consensus scores from the Itza' were closer to

the scores from the US experts than they were from the non-experts' scores. It is quite surprising that the Itza' perform more like the US experts than the non-experts do, despite the fact that the Itza' have never seen many of the birds in the US set before. Nonetheless, we did observe some differences between the two groups of experts.

The differences we find between the groups make sense when we consider the difference between interaction and observation. In Experiment 3, the Itza' gave predominantly ecological justifications that cited a bird's relationship with other living things in their environment. The US birders tended to provide justifications about scientific taxonomy (Experiment 1) and evolutionary history (Experiment 3), two factors emphasized in the field guides on which they rely. In sum, although universal folkbiological principles are discernible across diverse human groups, different information requirements and background assumptions also influence natural categorization and reasoning. Our triangulation strategy does not allow us to determine unequivocally whether these differences are mediated by cultural factors or by differences in amount and type of expertise. To follow up the differences between the two expert groups one would need to examine other groups such as less expert Itza' or perhaps US birders for whom ecological goals are more salient. The ecological focus of the Itza' is a robust finding (Atran et al., 1999, in press), which distinguishes them from other groups living in the same area and engaged in more or less the same activities.

5.2. Cultural effects

Instances where US experts and non-experts cluster and contrast with the Itza' suggest cultural differences. The clearest cultural difference we observed is what we called the 'passerine effect'. In Experiments 2 and 3 the two cultures' perceptions regarding passerines (small songbirds) differed markedly from each other. In the GOE task, members of the two cultures clearly viewed the songbirds in a different light, with the US informants rating them high in typicality and the Itza' rating them low. US subjects were more apt to generalize from passerines than from other birds; the reverse trend is seen with the Itza', who were more likely to generalize from large, perceptually striking birds than from passerines.

The justifications for choices serve to clarify the patterns of responding to passerines. Non-experts tend to select passerines on probes and justify their choices by saying that passerines are more typical and/or familiar birds. US experts also select passerines, but they never mention typicality. Their focus tends to be on geographical range as well as evolutionary history. Itza' Maya do not so much avoid passerines as they approach non-passerines. Their justifications tend to be in terms of ecological relations and behavior. Itza' have more difficulty distinguishing passerines, perhaps because Itza' may not have as much ecological interest in passerines to bring to bear on the reasoning task. For the Itza', passerines are generally seen only as minor players in the causal processes that Itza' monitor and manipulate to sustain the human ecology (Atran, 1999).

US experts, surprisingly, do as well at naming passerines as non-passerines, perhaps because both bird groups are equally important to expert interests (that is, to identify birds for their own sake, rather than as crucial factors in human–ecosystem interaction and survival). Given that passerines may be more perceptually similar than non-passerines

(e.g. Boster, 1987), it may be that the focus of US experts on bird identification allows for compensatory perceptual learning.

5.3. Implications for other models of reasoning and categorization

In our reasoning studies, typicality strategies are reliably used only by US non-experts (undergraduates). Consequently, models invoking these principles may apply solely to situations where non-experts are reasoning about stimuli with which they have limited knowledge. Most work on the role of typicality judgments in natural categorization and reasoning stems from studies with college students. Those studies tend to support the view that similarity-based structures (e.g. central tendency, family resemblance) are the primary predictors for typicality in taxonomic categories, in general, and folkbiological categories, in particular (Barsalou, 1985; Rosch & Mervis, 1975). In this view, the mind's similarity judgments about typicality and the world's correlational structure are closely linked: typical members of categories capture the correlational structure of identifiable features in the world better than do atypical members.

This capacity to recognize correlated similarity structures in the world, such as other species types, seems to be a built in part of human as well as non-human species (Cerella, 1979; Herrnstein, 1984; Lorenz, 1966; cf. Smith & Medin, 1981). From these considerations Boster (1988:258) predicts a biological, cognitive and cultural universal:

> Passerines appear to be densely and continuously spread through the bird similarity space, whereas non-passerines are more sparsely and discontinuously distributed, leading to the choice of passerines as both more typical and more difficult to categorize than non-passerines.

But for Itza' Maya, passerines are not very typical at all. One way to follow up these findings would be to study US hunters who target game birds such as turkeys, grouse, partridges, ducks and geese. They might look more like the Itza' with respect to the passerine effects than do other US groups.

No doubt similarity structures and similarity-based typicality and diversity are important determinants in natural categorization and reasoning. Our findings suggest that, at least for American undergraduates, these may be dominant factors. But for our relative experts (US experts and Itza'), who have substantial knowledge, goals and activities about the items they classify and reason with, information other than that derived from perceptual clustering and similarity judgment is relevant to understanding natural biodiversity. Behavior and ecology, for example, appear to be crucial to the deeper and broader understanding of nature that scientists and birdwatchers seek.

Such concerns also may be critical to the way the Maya and perhaps other peoples in small-scale manage to live and survive with nature. If so, then it is practically impossible to isolate folk ecological orientation from other aspects of cultural knowledge. Thus, previous studies indicate that Itza' share with other cultural groups (e.g. Spanish-speaking Ladino immigrants, Highland Q'eqchi' Maya immigrants) an identical habitat and a similar taxonomic understanding of its flora and fauna; nevertheless, these different cultural groups cognitively model species relationships (including humans) and socially interact with the same local ecology in fundamentally different ways (Atran et al., 1999, in press). Such

findings strongly simply that culture-specific cognitions and practices – and not just biotic, demographic or other material features of the environment – reliably determine population differences in ecological orientation and folkbiological understanding.

5.4. Conclusion

Our results indicate that in the real-world cognitive domain of folkbiology, which provides our species with information for navigating its natural environment, universal, culturally specific and selective transcultural patterns of categorization and reasoning have emerged. Among three distinct subject populations – USA non-experts (under-graduates), USA experts (birdwatchers) and Itza' Maya of the Guatemalan rainforest – we found modest to strong similarities on categorization tasks despite large differences in the sorting justifications. These results provide some support for the operation of univer-sal principles of folkbiological taxonomy, as claimed by Berlin and his associates (Berlin, 1992; Berlin et al., 1973). The primary qualification on this summary is that undergraduate sorting showed fairly modest correlations with scientific taxonomy. It may be that claims about universality need to be qualified by the need for some minimal amount of exposure or experience with the domain in question. Only US non-experts show patterns of judgments that are consistent with similarity-based models of taxo-nomic inference and associated claims concerning notions of typicality and central tendency in categorization and reasoning.

We also found systematic differences associated with interests and goals. For example, Itza' Maya rely less on passerines than do USA informants to make inferences about the biological world in the face of uncertainty. The Maya, it appears, believe that passerines are causally less important than non-passerines in sustaining those human–ecosystem interactions deemed necessary to survival.

Most surprisingly, we found content-specific knowledge with respect to categorization and reasoning that both approximates scientific evolutionary taxonomy and selectively transcends cultural boundaries. The novel empirical implication in this regard is that relative expertise and interaction, rather than mere exposure and observation, with respect to natural biodiversity may be the default condition for most human groups (and for ancestral humanity). From a theoretical perspective, then, the chief interest in studying "standard groups", such as psychology undergraduates at major North American and European research institutions, may not be to establish a baseline for generalizations about folkbiological knowledge, but to explore the cognitive consequences of limited input and devolutionary cultural processes (Atran et al., in press; Wolff, Medin, & Pank-ratz, 1999).

Acknowledgements

Appendix A. Taxonomic listing of the US and Tikal bird sets

	Order	Family	Genus	Specie
US bird				
Acadian Flycatcher	Passeriformes	Tyrannidae	Empidonax	Virescens
Alder Flycatcher	Passeriformes	Tyrannidae	Empidonax	Alnorum
American Bittern	Ciconiiformes	Ardeidae	Botaurus	Lentiginosus
American Black Duck	Anseriformes	Anatidae	Anas	Rubripes
American Coot	Gruiformes	Rallidae	Fulica	Americana
American Crow	Passeriformes	Corvidae	Corvus	Brachyrhynchos
American Kestrel	Falconiformes	Falconidae	Falco	Sparverius
American Woodcock	Charadriiformes	Scolopalidae	Scolopax	Minor
Bald Eagle	Falconiformes	Accipitridae	Haliaeetus	Leucocephalus
Barn Owl	Strigiformes	Strigidae	Tyto	Alba
Barn Swallow	Passeriformes	Hirundinidae	Hirundo	Rustica
Belted Kingfisher	Trogoniformes	Trogonidae	Ceryle	Alcyon
Black Tern	Charadriiformes	Laridae	Chlidonias	Niger
Black-And-White Warbler	Passeriformes	Vireonidae	Mniotilta	Varia
Black-Billed Cuckoo	Cuculiformes	Cuclidae	Coccyzus	Americanus
Black-Crown Night-Heron	Ciconiiformes	Ardeidae	Nycticorax	Nycticorax
Blue Jay	Passeriformes	Corvidae	Cyanocitta	Cristata
Bobolink	Passeriformes	Emberizidae	Dolychonyx	Oryzivorus
Broad-Winged Hawk	Falconiformes	Accipitridae	Buteo	Platypterus
Brown Creeper	Passeriformes	Certhiidae	Certhia	Americana
Canada Goose	Anseriformes	Anatidae	Branta	Canadensis
Carolina Wren	Passeriformes	Troglodytidae	Thryothorus	Ludovicianus
Caspian Tern	Charadriiformes	Laridae	Sterna	Caspia
Cliff Swallow	Passeriformes	Hirundinidae	Hirundo	Pyrrhonota
Common Goldeneye	Anseriformes	Anatidae	Bucephala	Clangula
Common Grackle	Passeriformes	Emberizidae	Quiscalus	Quiscula
Common Loon	Gaviiformes	Gaviidae	Gavia	Immer
Common Merganser	Anseriformes	Anatidae	Mergus	Merganser
Common Moorhen	Gruiformes	Rallidae	Gallinula	Chloropus
Common Nighthawk	Caprimulgiformes	Caprimulgidae	Chordeiles	Minor
Common Yellowthroat	Passeriformes	Vireonidae	Geothlypis	Trichas
Cooper's Hawk	Falconiformes	Accipitridae	Accipiter	Cooperii
Double-Crested Cormorant	Pelecaniformes	Phalacrocoracidae	Phalacrocorax	Auritus
Downy Woodpecker	Piciformes	Picidae	Picoides	Pubescens
Eastern Bluebird	Passeriformes	Muscicapidae	Sialia	Sialis
Eastern Kingbird	Passeriformes	Tyrannidae	Tyrannus	Tyrannus
Eastern Phoebe	Passeriformes	Tyrannidae	Sayornis	Phoebe
Eastern Screech-Owl	Strigiformes	Strigidae	Otus	Asio
Evening Grosbeak	Passeriformes	Fringillidae	Coccothraustes	Vespertinus
Gray Catbird	Passeriformes	Mimidae	Dumetella	Carolinensis
Great Blue Heron	Ciconiiformes	Ardeidae	Ardea	Herodias
Great Egret	Ciconiiformes	Ardeidae	Casmerodius	Albus
Great Horned Owl	Strigiformes	Strigidae	Bubo	Virginianus
Green Heron	Ciconiiformes	Ardeidae	Butorides	Virescens
Green-Winged Teal	Anseriformes	Anatidae	Anas	Crecca
Hairy Woodpecker	Piciformes	Picidae	Picoides	Villosus
Herring Gull	Charadriiformes	Laridae	Larus	Argentatus
Hooded Merganser	Anseriformes	Anatidae	Lophodytes	Cucullatus
Horned Grebe	Podicipediformes	Podicipedidae	Podiceps	Auritus
Horned Lark	Passeriformes	Alaudidae	Eremophila	Alpestris
House Wren	Passeriformes	Troglodytidae	Troglodytes	Aedon
Killdeer	Charadriiformes	Charadriidae	Charadrius	Vociferus
Least Bittern	Ciconiiformes	Ardeidae	Ixobrychus	Exilis

(*continued*)

	Order	Family	Genus	Specie
Little Blue Heron	Ciconiiformes	Ardeidae	Egretta	Caerulea
Long-Eared Owl	Strigiformes	Strigidae	Asio	Otus
Louisiana Waterthrush	Passeriformes	Vireonidae	Seiurus	Motacilla
Mallard	Anseriformes	Anatidae	Anas	Platyrhynchos
Mourning Dove	Columbiformes	Columbidae	Zenaida	Macroura
Northern Bobwhite	Galliformes	Phasianidae	Colinus	Virgianus
Northern Flicker	Piciformes	Picidae	Colaptes	Auratus
Northern Harrier	Falconiformes	Accipitridae	Circus	Cyaneus
Northern Mockingbird	Passeriformes	Mimidae	Mimus	Polyglottus
Northern Oriole	Passeriformes	Icteridae	Icterus	Galbula
Rough-Winged Swallow	Passeriformes	Hirundinidae	Stelgidopteryx	Serripennis
Northern Saw-Whet Owl	Strigiformes	Strigidae	Aegolius	Acadicus
Orchard Oriole	Passeriformes	Icteridae	Icterus	Spurius
Osprey	Falconiformes	Accipitridae	Pandion	Halieatus
Pied-Billed Grebe	Podicipediformes	Podicipedidae	Podilymbus	Podiceps
Purple Martin	Passeriformes	Hirundinidae	Progne	Subis
Red-Bellied Woodpecker	Piciformes	Picidae	Melanerpes	Carolinus
Red-Breasted Merganser	Anseriformes	Anatidae	Mergus	Serrator
Red-Breasted Nuthatch	Passeriformes	Sittidae	Sitta	Canadensis
Red-Headed Woodpecker	Piciformes	Picidae	Melanerpes	Erythrocephalus
Red-Shouldered Hawk	Falconiformes	Accipitridae	Buteo	Lineatus
Red-Tailed Hawk	Falconiformes	Accipitridae	Buteo	Jamaicensis
Red-Winged Blackbird	Passeriformes	Emberizidae	Agelaius	Phoeniceus
Redhead	Anseriformes	Anatidae	Aythya	Americana
Ring-Billed Gull	Charadriiformes	Laridae	Larus	Delawarensis
Ring-Necked Duck	Anseriformes	Anatidae	Aythya	Collaris
Ring-Necked Pheasant	Galliformes	Phasianidae	Phasianus	Colchicus
Rock Dove	Columbiformes	Columbidae	Columbia	Livia
Ruby-Throated Hummingbird	Trochiliformes	Trochilidae	Archilochus	Colubris
Ruddy Duck	Anseriformes	Anatidae	Oxyura	Jamaicensis
Ruffed Grouse	Galliformes	Phasianidae	Bonasa	Umbellus
Sandhill Crane	Gruiformes	Rallidae	Grus	Canadensis
Sharp-Shinned Hawk	Falconiformes	Accipitridae	Accipiter	Striatus
Sharp-Tailed Grouse	Galliformes	Phasianidae	Tympanuchus	Phasianellus
Short-Eared Owl	Strigiformes	Strigidae	Asio	Flammeus
Sora	Gruiformes	Rallidae	Porzana	Carolina
Spotted Sandpiper	Charadriiformes	Scolopalidae	Actitis	Macularia
Turkey Vulture	Falconiformes	Cathartidae	Cathartes	Aura
Upland Sandpiper	Charadriiformes	Scolopalidae	Bartramia	Longicauda
Veery	Passeriformes	Muscicapidae	Catharus	Fuscescens
Virginia Rail	Gruiformes	Rallidae	Rallus	Limicola
Whip-Poor-Will	Caprimulgiformes	Caprimulgidae	Caprimulgus	Vociferus
White-Breasted Nuthatch	Passeriformes	Sittidae	Sitta	Carolinensis
Willow Flycatcher	Passeriformes	Tyrannidae	Empidonax	Traillii
Wilson's Phalarope	Charadriiformes	Scolopalidae	Phalaropus	Tricolor
Wood Duck	Anseriformes	Anatidae	Aix	Sponsa
Wood Thrush	Passeriformes	Muscicapidae	Hylocicla	Mustelina
Yellow Rail	Gruiformes	Rallidae	Coturnicops	Noveboracensis
Yellow-Bellied Flycatcher	Passeriformes	Tyrannidae	Empidonax	Flaviventris
Yellow-Bellied Sapsucker	Piciformes	Picidae	Sphyrapicus	Varius
Yellow-Billed Cuckoo Erythrophtalmus	Cuculiformes	Cuclidae	Coccyzus	

Tikal bird

(continued)

	Order	Family	Genus	Specie
Blue-Winged Teal	Anseriformes	Anatidae	Anas	discors
Amazon Kingfisher	Trogoniformes	Alcedinidae	chloroceryle	amazona
American Sparrow-Hawk	Falconiformes	Falconidae	Falco	sparverius
Banded Tiger Heron	Ardeiformes	Ardeidae	Tigrisoma	Lineatum
Barn Owl	Strigiformes	Tytonidae	Tyto	Alba
Barn Swallow	Passeriformes	Hirundinidae	Hirundo	rustica
Barred forest-Falcon	Falconiformes	Falconidae	micrastur	ruficollis
Bat Falcon	Falconiformes	Falconidae	Falco	rufiguralis
Black & White Wood-Owl	Strigiformes	Strigidae	Ciccaba	nigrolieata
Black Hawk-Eagle	Falconiformes	Accipitridae	Spizaëtus	Tyrannus
Black Vulture	Falconiformes	Cathartidae	Coragyps	atratus
Black-cheeked Woodpecker	Piciformes	Picidae	centurus	pucherani
Blue Grosbeak	Passeriformes	Fringillidae	guiraca	caerulea
Blue Ground-Dove	Columbiformes	Columbidae	Claravis	pretiosa
Blue-Crowned Motmot	Coraciformes	Momotidae	Momotus	momota
Blue-Crowned Parrot	Psittaciformes	Psittacidae	Amazona	farinsoa
Brown Jay	Passeriformes	Corvidae	Psilorhinus	Morio
Brown-Crested Flycatcher	Passeriformes	Tyrannidae	Myarchus	tyrannulus
Brown-Hooded Parrot	Psittaciformes	Psittacidae	pionopsitta	haematotis
Buff-Throated foliage-gleaner	Passeriformes	Furnariidae	automolus	ochrolaemus
Cayenne Kite	Falconiformes	Accipitridae	Leptodon	cayanensis
Cayenne squirrel Cuckoo	Culiciformes	Cuculidae	Piaya	Cayana
Chestnut-colored Woodpecker	Piciformes	Picidae	celeus	castaneus
Collared Aracari	Piciformes	Ramphastidae	Pteroglossus	Torquatus
Collared forest-Falcon	Falconiformes	Falconidae	Micrastur	semi-torquatus
Common Black Hawk	Falconiformes	Accipitridae	Buteogallus	Anthracinus
Common Egret	Ardeiformes	Ardeidae	Casmerodius	albus
Crested Guan	Galliformes	Cracidae	Penelope	purpurascens
Dusky-Capped Flycatcher	Passeriformes	Tyrannidae	Myarchus	tuberculifer
Emerald Toucanet	Piciformes	Ramphastidae	Aulacorhynchus	prasinus
Everglade Kite, Snail Kite	Falconiformes	Accipitridae	Rostrhamus	Sociabilis
Ferruginous pygmy Owl	Strigiformes	Strigidae	Glaucidium	brasilianum
Fork-Tailed Emerald	Trochiliformes	Trochilidae	chlorostilblon	canivetii
Gaumer Swift	Apodiformes	Apodidae	Chaetura	gaumeri
Giraud's Flycatcher	Passeriformes	Tyrannidae	Myiozetetes	similis
Golden-Fronted Woodpecker	Piciformes	Picidae	centurus	aurifrons
Golden-olive Woodpecker	Piciformes	Picidae	piculus	rubiginosus
Gray potoo	Caprimulgaformes	Nyctibiidae	Nyctibius	griseus
Gray's Robin	Passeriformes	Muscicapidae	Turdus	Grayi
Gray-Breasted Martin	Passeriformes	Hirundinidae	Progne	Chalybea
Gray-headed Dove	Columbiformes	Columbidae	Leptotila	plumbeiceps
Great Black Hawk	Falconiformes	Accipitridae	Buteogallus	urubitinga
Great Curassow	Galliformes	Cracidae	Cra	rubrea
Great Horned Owl	Strigiformes	Strigidae	Bubo	virginianus
Great Kiskadee	Passeriformes	Tyrannidae	Pitangus	Sulfuratus
Great Tinamou	Tinamiformes	Tinimidae	Tinamous	Major
Great-Crested Flycatcher	Passeriformes	Tyrannidae	Myarchus	crinitus
Green Jay	Passeriformes	Corvidae	Cyanocora	yncas
Green Parakeet	Psittaciformes	Psittacidae	Aratinga	nana
Green-Breasted Mango	Trochiliformes	Trochilidae	anthracolthora	prevostii
Groove-Billed ani	Culiciformes	Cuculidae	Crotophaga	sulcirostris
Hook-Billed Kite	Falconiformes	Accipitridae	Chondrohiera	uncinatus

(*continued*)

	Order	Family	Genus	Specie
Keel-Billed Toucan	Piciformes	Ramphastidae	Ramphastos	Sulfuratus
King Vulture	Falconiformes	Cathartidae	Sarcoramphus	Papa
laughing Falcon	Falconiformes	Falconidae	Herpetotheres	cachinnans
Lesser Swallow-Tailed Swift	Apodiformes	Apodidae	Panyptila	cayennensis
Lesser Yellowhead	Falconiformes	Cathartidae	Cathartes	burrovianus
Lineated Woodpecker	Piciformes	Picidae	Dryocopus	lineatus
Little Blue Heron	Ardeiformes	Ardeidae	Egretta	caerula
Little Tinamou	Tinamiformes	Tinimidae	Cryturellus	soui
montezuma Oropendola	Passeriformes	Icteridae	Gymnostinops	montezuma
Mottled Wood-Owl	Strigiformes	Strigidae	Ciccaba	virgata
Northern jacana	Charadriiformes	Jacanidae	Jacana	spinosa
Ocellated turkey	Galliformes	Phasianidae	Agriocharis	ocellata
Olive-Throated Parrot	Psittaciformes	Psittacidae	Aratinga	aztec
Ornate Hawk-Eagle	Falconiformes	Accipitridae	Spizaëtus	Ornatus
Pale-Billed Woodpecker	Piciformes	Picidae	Phloeoceastes	guatemalensis
Purple Martin	Passeriformes	Hirundinidae	Progne	Subis
Purple-Crowned fairy	Trochiliformes	Trochilidae	heliothry	barroti
Red-Crowned Ant Tanager	Passeriformes	Thraupidae	habia	rubica
Red-Crowned Parrot	Psittaciformes	Psittacidae	Amazona	viridigenalis
Red-lored Parrot	Psittaciformes	Psittacidae	Amazona	autumnalis
Resplendent Quetzal	Trochiliformes	Trogonidae	Pharomachrus	mocinno
Ringed Kingfisher	Trogoniformes	Alcedinidae	ceryle	torquata
Roadside Hawk	Falconiformes	Accipitridae	Buteo	magnirostris
Rose-Throated Becard	Passeriformes	Tyrannidae	Pachyramphus	aglaiae
Rough-Winged Swallow	Passeriformes	Hirundinidae	Stelgideptery	Ruficollis
Ruddy Ground-Dove	Columbiformes	Columbidae	Columbina	talpacoti
Ruddy Quail-Dove	Columbiformes	Columbidae	Geotrygon	montana
Rufescent Tinamou	Tinamiformes	Tinimidae	Cryturellus	cinnamoneus
Scaled Pigeon	Columbiformes	Columbidae	Columba	speciosa
Scaly-Breasted Hummingbird	Trochiliformes	Trochilidae	phaeochroa	cuvierii
Scarlet Macaw	Psittaciformes	Psittacidae	Ara	macao
Short-Billed Pigeon	Columbiformes	Columbidae	Columba	nigrirostris
Singing Quail	Galliformes	Phasianidae	Dactylorty	thoracicus
Slaty-Breasted Tinamou	Tinamiformes	Tinimidae	Cryturellus	boucardi
Slaty-Tailed Trogon	Trochiliformes	Trogonidae	trogon	Massena
Southern Boat-Billed Heron	Ardeiformes	Ardeidae	Cochlearius	cochlearius
Spotted Wood Quail	Galliformes	Phasianidae	Odontophorus	guttatus
Sumichrast's Blackbird	Passeriformes	Icteridae	Dives	dives
Turkey Vulture	Falconiformes	Cathartidae	Cathartes	aura
Vaux's Swift	Apodiformes	Apodidae	Chaetura	vaui
Vermiculated Screech-Owl	Strigiformes	Strigidae	Otus	guatemalae
Wagler's Oropendola	Passeriformes	Icteridae	Zarhynchus	Wagleri
Western chacalaca	Galliformes	Cracidae	Ortalis	poliocephala
White-Crowned Parrot	Psittaciformes	Psittacidae	Pionus	senilis
White-Fronted Parrot	Psittaciformes	Psittacidae	Amazona	albifrons
White-Necked Jacobin	Trochiliformes	Trochilidae	florisuga	mellivora
White-Throated Robin	Passeriformes	Muscicapidae	Turdus	Albicolis
Yellow-backed oriole	Passeriformes	Icteridae	icterus	chrysater
Yellow-Tailed oriole	Passeriformes	Icteridae	Icterus	mesomelas
Yellow-Winged Tanager	Passeriformes	Thraupidae	Thraupis	Abbas
Yucatan Flycatcher	Passeriformes	Tyrannidae	Myarchus	yucatanensis
Yucatan Poorwill	Caprimulgaformes	Caprimulgadae	Otophanes	yucatanicus

Appendix B. List all 104 birds in each stimuli set

For each bird, we list the approximate number of family members (FAM), ranked mean central tendency (CT) and mean typicality/idealness ratings (TYP) for non-experts (NOV), experts (EXP) and Itza' (ITZ).

	FAM	NOV		EXP		ITZ	
		CT	TYP	CT	TYP	CT	TYP
US bird set							
Acadian Flycatcher	10	13	5	13	1	23	–
Alder Flycatcher	10	18	3	13	20	57	–
American Bittern	10	75	92	42	80	28	–
American Black Duck	26	46	79	8	91	10	–
American Coot	6	74	60	25	75	71	–
American Crow	2	101	5	93	42	98	–
American Kestrel	3	71	53	58	42	57	–
American Woodcock	19	62	66	90	42	66	–
Bald Eagle	11	84	35	76	19	62	–
Barn Owl	6	93	76	79	78	67	–
Barn Swallow	6	30	35	84	25	94	–
Belted Kingfisher	1	27	60	61	100	52	–
Black Tern	10	53	31	86	86	90	–
Black-And-White Warbler	35	6	29	56	1	26	–
Black-Billed Cuckoo	2	16	41	35	68	74	–
Black-Crowned Night-Heron	10	33	63	20	42	15	–
Blue Jay	2	24	7	68	27	75	–
Bobolink	9	23	10	32	59	18	–
Broad-Winged Hawk	11	67	52	48	42	53	–
Brown Creeper	1	45	10	97	75	41	–
Canada Goose	26	61	84	28	68	11	–
Carolina Wren	5	8	14	59	42	37	–
Caspian Tern	10	81	49	85	61	79	–
Cliff Swallow	6	10	18	67	35	63	–
Common Goldeneye	26	44	84	6	42	4	–
Common Grackle	9	63	26	39	1	72	–
Common Loon	1	43	74	12	42	8	–
Common Merganser	26	36	80	3	61	3	–
Common Moorhen	6	77	84	38	104	54	–
Common Nighthawk	2	95	45	103	1	81	–
Common Yellowthroat	35	32	35	56	59	20	–
Cooper's Hawk	11	76	14	51	80	76	–
Double-Crested Cormorant	1	87	98	47	1	78	–
Downy Woodpecker	6	31	18	70	42	87	–
Eastern Bluebird	10	15	1	37	23	50	–
Eastern Kingbird	10	11	41	23	1	49	–
Eastern Phoebe	10	2	18	19	41	43	–
Eastern Screech-Owl	6	95	70	64	74	67	–
Evening Grosbeak	8	28	49	91	42	39	–
Gray Catbird	3	9	18	29	1	51	–
Great Blue Heron	10	66	98	20	38	29	–

(*continued*)

	FAM	NOV		EXP		ITZ	
		CT	TYP	CT	TYP	CT	TYP
Great Egret	10	64	76	26	103	24	–
Great Horned Owl	6	91	87	64	21	67	–
Green or Green-Backed Heron	10	55	87	24	87	35	–
Green-Winged Teal	26	39	66	6	34	9	–
Hairy Woodpecker	6	54	31	70	1	89	–
Herring Gull	10	90	63	74	100	34	–
Hooded Merganser	26	49	104	3	78	12	–
Horned Grebe	3	72	98	17	65	25	–
Horned Lark	1	26	35	50	35	48	–
House Wren	5	5	10	59	1	17	–
Killdeer	4	69	31	69	35	82	–
Least Bittern	10	60	66	43	97	35	–
Little Blue Heron	10	83	103	22	42	22	–
Long-Eared Owl	6	91	80	64	61	67	–
Louisiana Waterthrush	35	20	9	40	1	32	–
Mallard	26	40	89	8	65	7	–
Mourning Dove	2	52	29	77	27	96	–
Northern Bobwhite	5	99	70	98	97	91	–
Northern Flicker	6	58	18	82	65	92	–
Northern Harrier	11	82	48	54	61	73	–
Northern Mockingbird	3	19	31	44	39	46	–
Northern or Baltimore Oriole	2	12	4	31	27	45	–
Northern Rough-Winged Swallow	6	14	35	87	23	86	–
Northern Saw-Whet Owl	6	88	76	62	91	61	–
Orchard Oriole	2	7	2	30	1	56	–
Osprey	11	78	55	78	42	55	–
Pied-Billed Grebe	3	48	101	18	42	5	–
Purple Martin	6	21	18	89	27	80	–
Red-Bellied Woodpecker	6	47	45	70	39	87	–
Red-Breasted Merganser	26	50	92	11	42	14	–
Red-Breasted Nuthatch	2	4	26	95	25	19	–
Red-Headed Woodpecker	6	51	7	70	1	95	–
Red-Shouldered Hawk	11	68	53	48	72	60	–
Red-Tailed Hawk	11	73	57	51	42	59	–
Red-Winged Blackbird	9	29	45	27	1	47	–
Redhead	26	37	92	1	80	1	–
Ring-Billed Gull	10	94	57	74	91	44	–
Ring-Necked Duck	26	38	97	1	97	6	–
Ring-Necked Pheasant	5	103	65	101	80	104	–
Rock Dove	2	85	14	92	87	93	–
Ruby-Throated Hummingbird	1	41	55	102	68	97	–
Ruddy Duck	26	56	92	5	75	13	–
Ruffed Grouse	5	102	80	98	80	103	–
Sandhill Crane	6	86	92	83	1	21	–
Sharp-Shinned Hawk	11	80	18	51	68	42	–
Sharp-Tailed Grouse	5	100	74	98	42	102	–
Short-Eared Owl	6	89	89	62	102	65	–
Sora	6	70	70	88	87	101	–

(*continued*)

	FAM	NOV		EXP		ITZ	
		CT	TYP	CT	TYP	CT	TYP
Spotted Sandpiper	19	97	57	46	72	77	–
Turkey Vulture	1	104	101	94	95	100	–
Upland Sandpiper	19	79	70	45	42	64	–
Veery	10	1	18	33	1	16	–
Virginia Rail	6	65	62	41	96	33	–
Whip-Poor-Will	2	98	49	103	1	99	–
White-Breasted Nuthatch	2	22	26	95	27	38	–
Willow Flycatcher	10	3	10	16	21	27	–
Wilson's Phalarope	19	57	80	55	1	31	–
Wood Duck	26	42	89	10	91	2	–
Wood Thrush	10	25	41	33	27	30	–
Yellow Rail	6	35	41	81	90	84	–
Yellow-Bellied Flycatcher	10	17	14	13	1	40	–
Yellow-Bellied Sapsucker	6	34	35	80	27	85	–
Yellow-Billed Cuckoo	2	59	69	35	80	83	–
Tikal bird set							
Amazon Kingfisher	5	61	95	98	39	68	36
American Sparrow-Hawk/Kestrel	12	28	58	20	21	3	100
Banded Tiger Heron	17	41	95	101	48	97	75
Barn Owl	1	100	62	68	15	26	56
Barn Swallow	14	21	22	35	40	27	75
Barred forest-Falcon	12	45	62	9	1	5	79
Bat Falcon	12	62	27	13	46	6	79
Black & White Wood-Owl	30	96	90	61	85	24	56
Black Hawk-Eagle	40	69	77	14	6	11	79
Black Vulture	5	81	62	48	40	72	69
Black-cheeked Woodpecker	27	33	66	82	83	85	21
Blue Grosbeak	12	57	10	82	30	58	31
Blue Ground-Dove	24	30	52	95	51	37	36
Blue-Crowned Motmot	6	89	95	70	51	79	49
Blue-Crowned Parrot	22	66	36	2	45	38	2
Blue-Winged Teal	40	102	22	104	30	93	33
Brown Jay	24	2	14	22	91	65	90
Brown-Crested Flycatcher	75	19	3	24	30	13	62
Brown-Hooded Parrot	22	77	48	1	80	28	2
Buff-Throated foliage-gleaner	7	10	1	52	19	82	87
Cayenne Kite	40	56	36	23	21	4	100
Cayenne squirrel Cuckoo	12	39	90	69	66	101	68
Chestnut-colored Woodpecker	27	32	84	90	90	41	18
Collared Aracari	3	94	95	58	30	77	26
Collared forest-Falcon	12	72	62	18	66	2	95
Common Black Hawk	40	67	58	7	6	7	79
Common Egret	17	104	41	102	4	87	103
Crested Guan	8	42	102	88	91	69	10
Dusky-Capped Flycatcher	75	25	17	24	85	12	62
Emerald Toucanet	3	94	66	58	51	76	26
Everglade Kite, Snail Kite	40	63	48	32	83	14	79

(*continued*)

	FAM	NOV		EXP		ITZ	
		CT	TYP	CT	TYP	CT	TYP
Ferruginous pygmy Owl	30	38	45	36	85	16	56
Fork-Tailed Emerald	61	59	41	74	4	46	45
Gaumer Swift	10	75	27	67	37	88	96
Giraud's Flycatcher	75	26	17	21	6	19	52
Golden-Fronted Woodpecker	27	11	30	82	91	80	21
Golden-olive Woodpecker	27	9	36	82	30	83	21
Gray potoo	2	5	73	89	80	104	89
Gray's Robin	37	3	21	43	6	45	66
Gray-Breasted Martin	14	8	3	29	37	55	91
Gray-headed Dove	24	50	77	47	62	64	40
Great Black Hawk	40	68	45	7	91	8	79
Great Curassow	8	46	104	91	101	69	11
Great Horned Owl	30	98	66	66	15	32	56
Great kiskadee, Derby Flycatcher	75	34	22	42	101	19	52
Great Tinamou	4	86	77	72	91	43	14
Great-Crested Flycatcher	75	22	14	24	6	17	62
Green Jay	24	15	7	28	85	44	74
Green Parakeet	22	51	17	15	6	36	2
Green-Breasted Mango	61	83	36	74	66	91	50
Groove-Billed ani	12	71	48	65	51	52	32
Hook-Billed Kite	40	53	77	12	74	10	79
Keel-Billed Toucan	3	92	77	60	1	86	24
King Vulture	5	87	102	50	73	78	69
laughing Falcon	12	13	54	11	74	1	100
Lesser Swallow-Tailed Swift	10	75	54	63	27	54	91
Lesser Yellowhead	0	78	90	51	74	75	69
Lineated Woodpecker	27	37	84	86	91	50	18
Little Blue Heron	17	103	77	100	6	99	103
Little Tinamou	4	82	66	93	51	84	28
montezuma Oropendola	33	90	84	37	50	95	51
Mottled Wood-Owl	30	96	90	62	74	23	56
Northern jacana	1	101	88	103	51	102	35
Ocellated turkey	19	91	77	94	91	81	1
Olive-Throated Parrot	22	54	30	15	6	35	2
Ornate Hawk-Eagle	40	43	84	19	48	18	30
Pale-Billed Woodpecker	27	27	95	86	27	50	20
Purple Martin	14	20	10	29	21	57	96
Purple-Crowned fairy	61	85	7	74	6	49	45
Red-Crowned Ant Tanager	30	40	10	39	51	60	67
Red-Crowned Parrot	22	74	41	2	62	34	2
Red-lored Parrot	22	54	36	2	17	30	2
Resplendent Quetzal	9	93	95	80	80	92	29
Ringed Kingfisher	5	47	95	98	25	71	36
Roadside Hawk	40	58	54	9	30	9	79
Rose-Throated Becard	75	4	14	34	74	61	96
Rough-Winged Swallow	14	6	3	29	66	59	96
Ruddy Ground-Dove	24	35	66	53	66	53	36
Ruddy Quail-Dove	24	1	35	40	46	63	43

(*continued*)

	FAM	NOV		EXP		ITZ	
		CT	TYP	CT	TYP	CT	TYP
Rufescent Tinamou	4	52	88	72	74	39	14
Scaled Pigeon	24	16	66	41	51	66	40
Scaly-Breasted Hummingbird	61	64	48	74	25	48	45
Scarlet Macaw	22	60	17	17	62	67	12
Short-Billed Pigeon	24	49	52	54	62	62	40
Singing Quail, long-toed partridge	19	48	73	81	51	40	14
Slaty-Breasted Tinamou	4	65	58	79	61	74	17
Slaty-Tailed Trogon	9	88	22	33	17	98	13
Southern Boat-Billed Heron	17	44	90	97	91	96	75
Spotted Wood Quail	19	29	73	92	91	90	25
Sumichrast's Blackbird	33	17	10	55	21	103	91
Turkey Vulture	5	79	54	48	51	72	69
Vaux's Swift	10	7	3	71	51	56	75
Vermiculated Screech-Owl	30	98	58	64	91	33	56
Wagler's Oropendola	33	83	73	37	66	89	44
Western chacalaca	8	80	41	45	85	94	33
White-Crowned Parrot	22	73	45	2	72	31	2
White-Fronted Parrot	22	70	27	2	44	29	2
White-Necked Jacobin	61	36	66	74	40	46	45
White-Throated Robin	37	12	30	43	30	42	88
Yellow-backed oriole	33	24	30	56	101	21	52
Yellow-Tailed oriole	33	18	30	56	1	21	52
Yellow-Winged Tanager	30	31	7	46	40	25	69
Yucatan Flycatcher	75	22	1	24	19	15	62
Yucatan Poorwill	14	14	22	96	27	100	91

Appendix C. The 16 pairs used as stimuli in the typicality task (half are US birds and half are Tikal birds)

Typicality	Familiality	Pair	US	Tikal
High	Varied	1	Wilson's Phalarope (HI)	Yucatan Flycatcher (HI)
			Common Loon (LO)	Laughing Falcon (LO)
		2	Wilson's Phalarope (HI)	Yucatan Flycatcher (HI)
			N. Mockingbird (LO)	Barn Swallow (LO)
Low	Varied	3	Caspian Tern (HI)	Green-Breasted Mango (HI)
			N. Bobwhite (LO)	N. Jacana (LO)
		4	American Woodcock (HI)	Blue-Winged Teal (HI)
			N. Flicker (LO)	Keel-Billed Toucan (LO)
Varied	High	5	Wood Thrush (HI)	Ornate Hawk-Eagle (HI)
			American Woodcock (LO)	Blue-Winged Teal (LO)
		6	Wood Thrush (HI)	Ornate Hawk-Eagle (HI)
			Caspian Tern (LO)	Green-Breasted Mango (LO)
Varied	Low	7	N. Mockingbird (HI)	Barn Swallow (HI)
			N. Flicker (LO)	Keel-Billed Toucan (LO)
		8	Common Loon (HI)	Laughing Falcon (HI)
			N. Bobwhite (LO)	N. Jacana (LO)

Appendix D. The eight diversity pairs used as stimuli in the diversity task (half are US birds and half are Tikal birds)

Diversity pair	Distance	US	Tikal
1	Close	Short-Eared Owl	Emerald Toucanet
		Barn Owl	Collared Aracari
	Far	Ruffed Grouse	Blue Grosbeak
		N. Flicker	N. Jacana
2	Close	Bald Eagle	Gaumer Swift
		Osprey	Lesser Yellow-Tailed Swift
	Far	Sora	Blue-Crowned Motmot
		American Woodcock	Ocellated Turkey
3	Close	Acadian Flycatcher	Great-Crested Flycatcher
		Yellow-Bellied Flycatcher	Brown-Crested Flycatcher
	Far	Black-Crowned Night Heron	Scaly-Breasted Hummingbird
		Bobolink	Ornate Hawk-Eagle
4	Close	Common Grackle	Slaty-Breasted Tinamou
		American Crow	Little Tinamou
	Far	Killdeer	Resplendent Quetzal
		American Kestrel	Black Vulture

References

Atran, S. (1993). Itza Maya tropical agro-forestry. *Current Anthropology*, *34*, 633–700.

Atran, S. (1995). Classifying nature across cultures. In E. Smith & D. Osherson (Eds.), *Thinking. An invitation to cognitive science* (pp. 131–174), Vol. 3. Cambridge, MA: MIT Press.

Atran, S. (1998). Folk biology and the anthropology of science: cognitive universals and cultural particulars. *Behavioral and Brain Sciences*, *Vol. 21*, 547–611.

Atran, S. (1999). Itzaj Maya folkbiological taxonomy. In D. Medin & S. Atran (Eds.), *Folk biology* (pp. 119–203). Cambridge, MA: MIT Press.

Atran, S. (2001). The trouble with memes: inference versus imitation in cultural creation. *Human Nature*, *12*, 351–381.

Atran, S., Estin, P., Coley, J., & Medin, D. (1997). Generic species and basic levels: essence and appearance in folk biology. *Journal of Ethnobiology*, *17*, 22–45.

Atran, S., & Medin, D. (1997). Knowledge and actions: cultural models of nature and resource management in Mesoamerica. In M. H. Bazerman, D. M. Messick, A. E. Tenbrunsel & K. A. Wade-Benzoni (Eds.), *Environment, ethics, and behavior: the psychology of environmental valuation and degradation* (pp. 171–208). San Francisco, CA: The New Lexington Press.

Atran, S., Medin, D., Ross, E., Lynch, E., Coley, J., Ucan Ek', E., & Vapnarsky, V. (1999). Folkecology and commons management in the Maya Lowlands. *Proceedings of the National Academy of Sciences USA*, *96*, 7598–7603.

Atran, S., Medin, D., Ross, N., Lynch, E., Vapnarsky, V., Ucan Ek', E., Coley, J., Timura, C., & Baran, M. (in press). Folkecology, cultural epidemiology, and the spirit of the commons: a garden experiment in the Maya Lowlands, 1991–2001. *Current Anthropology*, *43*(3).

Atran, S., & Ucan Ek', E. (1999). Classification of useful plants among the Northern Petén Maya. In C. White (Ed.), *Reconstructing ancient Maya diet*. Salt Lake City, UT: The University of Utah Press.

Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure of categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 629–654.

Berlin, B. (1978). Ethnobiological classification. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorization* (pp. 9–26). Hillsdale, NJ: Erlbaum.

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

Berlin, B., Breedlove, D., & Raven, P. (1973). General principles of classification and nomenclature in folk biology. *American Anthropologist*, *74*, 214–242.

Berlin, B., Breedlove, D., & Raven, P. (1974). *Principles of Tzeltal plant classification*. New York: Academic Press.

Boster, J.S. (1986). Exchange of varieties and information between Aguaruna Manioc Cultivators. *American Anthropologist*, *88*, 429–436.

Boster, J. (1987). Agreement between biological classification systems is not dependent on cultural transmission. *American Anthropologist*, *89*, 914–920.

Boster, J. (1988). Natural sources of internal category structure: typicality, familiarity, and similarity of birds. *Memory & Cognition*, *16*, 258–270.

Boster, J. (1991). The information economy model applied to biological similarity judgment. In J. Levine, L. Resnick & S. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 203–235). Washington, DC: American Psychological Association.

Boster, J., Berlin, B., & O'Neill, J. (1986). The correspondence of Jivoroan to scientific ornithology. *American Anthropologist*, *88*, 569–583.

Boster, J., & D'Andrade, R. (1989). Natural and human sources of cross-cultural agreement in ornithological classification. *American Anthropologist*, *91*, 132–142.

Boster, J., & Johnson, J. (1989). Form or function: a comparison of expert and novice judgments of similarity among fish. *American Anthropologist*, *91*, 866–889.

Brown, C. (1984). *Language and living things: uniformities in folk classification and naming*. New Brunswick, NJ: Rutgers University Press.

Bulmer, R. (1974). Folk biology in the New Guinea Highlands. *Social Science Information*, *13*, 9–28.

Cerella, J. (1979). Visual classes and natural categories in the pigeon. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 68–77.

Coley, J. D. (1995). Emerging differentiation of folkbiology and folkpsychology: attributions of biological and psychological properties to living things. *Child Development*, *66*, 1856–1874.

Coley, J., Medin, D., Proffitt, J., Lynch, E., & Atran, S. (1999). Inductive reasoning in folkbiological thought. In D. L. Medin & S. Atran (Eds.), *Folkbiology* (pp. 205–232). Cambridge, MA: MIT Press.

Gigerenzer, G. (1996). The modularity of social intelligence. In A. Whiten & R. Byrne (Eds.), *Machiavellian intelligence* (pp. 264–288), Vol. II. Cambridge: Cambridge University Press.

Hays, T. (1983). Ndumba folk biology and the general principles of ethnobotanical classification and nomenclature. *American Anthropologist*, *85*, 489–507.

Heider, E. R. (1971). "Focal" color areas and the development of color names. *Developmental Psychology*, *4*, 447–455.

Heider, E. R. (1972). Universals in color naming and memory. *Journal of Experimental Psychology*, *93*, 10–20.

Hennig, W. (1966). *Phylogenetic systematics*. Urbana, IL: University of Illinois Press.

Herrnstein, R. (1984). Objects, categories, and discriminative stimuli. In H. Roitblat (Ed.), *Animal cognition* (pp. 233–261). Hillsdale, NJ: Erlbaum.

Hsee, C. K., & Weber, E. U. (1999). Cross-national differences in risk preference and lay predictions. *Journal of Behavioral Decision Making*, *12*, 165–179.

Hubert, L., & Schultz, J. (1976). Quadratic assignment as a general data analysis strategy. *British Journal of Mathematical and Statistical Psychology*, *29*, 190–241.

Hunn, E. (1976). Toward a perceptual model of folk biological classification. *American Ethnologist*, *3*, 508–524.

Hunn, E. (1977). *Tzeltal folk zoology*. New York: Academic Press.

Johnson, K., & Mervis, C. (1997). Effects of varying levels of expertise on the basic level of categorization. *Journal of Experimental Psychology: General*, *126*, 248–277.

Johnson, K., & Mervis, C. (1998). Impact of intuitive theories on feature recruitment throughout the continuum of expertise. *Memory & Cognition*, *26*, 382–401.

Johnson, K., Mervis, C., & Boster, J.S. (1992). Developmental changes in the structure of the mammal domain. *Developmental Psychology*, *28*, 74–83.

Lois, X. (1988). Gender markers as "rigid determiners" of the Itzaj Maya world. *International Journal of American Linguistics*, *64*, 224–282.

López, A., Atran, S., Coley, J., & Medin, D. (1997). The tree of life: universal and cultural features of folkbiological taxonomies and inductions. *Cognitive Psychology*, *32*, 251–295.

López, A., Gelman, S., Gutheil, G., & Smith, E. (1992). The development of category-based induction. *Child Development*, *63*, 1070–1090.

Lorenz, K. (1966). The role of gestalt perception in animal and human behavior. In L. White (Ed.), *Aspects of form* (pp. 157–178). Bloomington, IN: Indiana University Press.

Lynch, E., Coley, J., & Medin, D. (2000). Tall is typical: central tendency, ideal dimensions and graded category structure among tree experts and novices. *Memory & Cognition*, *28*, 41–50.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: implications for cognition, emotion, and motivation. *Psychological Review*, *98*, 224–253.

Mayr, E. (1965). Numerical phenetics and taxonomic theory. *Systematic Zoology*, *14*, 73–97.

Medin, D., Goldstone, R., & Gentner, D. (1993). Respects for similarity. *Psychological Review*, *100*, 254–278.

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., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology*, *32*, 89–115.

Nakao, K., & Romney, A. (1984). A method for testing alternative theories: an example from English kinship. *American Anthropologist*, *86*, 668–673.

Nisbett, R., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: holistic vs. analytic cognition. *Psychological Review*, *108*, 291–310.

Nisbett, R., & Wilson, T. (1977). Telling more than we know: verbal reports on mental processes. *Psychological Review*, *84*, 231–259.

Osherson, D., Smith, E., Wilkie, O., López, A., & Shafir, E. (1990). Category-based induction. *Psychological Review*, *97*, 85–200.

Proffitt, J., Coley, J., & Medin, D. (2000). Expertise and category-based induction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 811–828.

Romney, A. K., Weller, S., & Batchelder, W. (1986). Culture as consensus: a theory of culture and informant accuracy. *American Anthropologist*, *88*, 313–338.

Rosch, E. (1975). Universals and cultural specifics in categorization. In R. Brislin, S. Bochner & W. Lonner (Eds.), *Cross-cultural perspectives on learning* (pp. 177–206). New York: Sage.

Rosch, E. (1977). Human categorization. In N. Warren, *Studies in cross-cultural psychology* (pp. 1–49). London: Academic Press.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: studies in the internal structure of categories. *Cognitive Psychology*, *7*, 573–605.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*, 382–439.

Sloman, S. A. (1993). Feature-based induction. *Cognitive Psychology*, *25*, 231–280.

Smith, E., & Medin, D. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.

Smithe, F. (1966). *The birds of Tikal*. Garden City, NY: The Natural History Press.

Sokal, P., & Sneath, R. (1973). *Numerical taxonomy*. San Francisco, CA: Freeman.

Wolff, P., Medin, D., & Pankratz, C. (1999). Evolution and devolution of folkbiological knowledge. *Cognition*, *73*, 177–204.