A Measure of the Degree of Correspondence of Folk to Scientific Biological Classification
Author(s): Eugene Hunn
Source: *American Ethnologist*, Vol. 2, No. 2 (May, 1975), pp. 309-327
Published by: Wiley on behalf of the American Anthropological Association
Stable URL: http://www.jstor.org/stable/643820
Accessed: 21-03-2015 18:14 UTC

a measure of the degree of correspondence of folk to scientific biological classification[1]

EUGENE HUNN—*University of Washington*

Folk science investigates the nature of human knowledge and the relationship between man's knowledge of his environment and his adaptation to that environment. A fundamental issue remains controversial: Does folk science involve mental processes comparable to those of modern science and, if so, to what degree? Until recently anthropologists have stressed the dissimilarities. During the early decades of this century it was generally believed that "primitive mentality" was based largely on principles contrary to those of science. More recently, cultural relativists have asserted that each culture was free to compartmentalize its experience of the world as it wished, independent of objective constraints. In Sapir's words, "the worlds in which different societies live are distinct worlds, not merely the same world with different labels attached" (1921:209). Recent work with color terminology (Berlin and Kay 1969) and in folk biology, however, sharply qualify this position. The impressive similarities between the scientific classification of organisms and folk bio-taxonomic systems revealed by this latter research suggest that folk science is appropriately named. To clarify this issue it is important to describe more precisely the degree to which folk and scientific descriptions of the natural environment stem from universal cognitive processes, on the one hand, and from unique cultural and environmental conditions, on the other.

I propose, here, techniques for evaluating the degree to which any folk bio-taxonomic classification agrees with the scientific classification as applied to the same set of organisms. The two types of bio-systematics do not, of course, produce isomorphic, i.e., identical, classifications nor are they totally independent of one another. I will argue, rather, that folk bio-taxonomies tend to be isomorphic to a subsystem of the scientific system. That is, all the categories and their consequent relations of a folk taxonomy will, with high probability, have counterparts in the scientific system. On the other hand, many of the scientific taxa and relations will have no counterpart in a given folk system. To the extent that this is true, the scientific system of biological classification may serve as an etic grid rather than simply as a convenient language for glossing exotic lexical items. The broad comparisons the measure proposed here is intended to facilitate may

The paper begins by discussing the utility of new techniques for measuring the degree to which various folk systems of biological classification approximate the scientific zoological or botanical classification. One such technique is described and applied to the Tzeltal classification of mammals. Statistics provided by the proposed measure provide a test of the hypothesis that there is a psychologically fundamental and logically general set of folk biological taxa. The measure is formally defined.

demonstrate that all human beings do live in "the same world" of living things and that all men meet that world with a consistent logic. I will suggest a formula for measuring the degree to which this criterion of limited isomorphy is approximated for any given folk biological domain. To exemplify this technique I will compare the Tenejapa Tzeltal taxon *čanbalam*$_2$ 'mammal' point for point with the scientific class Mammalia (the evidentiary detail supporting the data presented here may be found in Hunn [1973:276-319]).

Tzeltal mammal classification

The class Mammalia is represented in the Mexican state of Chiapas, of which Tenejapa is one small part, by 174 species (according to Hall and Kelson 1959). Of these, sixty-seven are bats; many other species are unfamiliar to the people of Tenejapa due to geographical and habitat restrictions. A reasonable estimate of the number of species of mammals which the Tenejapa Tzeltal have had some opportunity to classify is between sixty-five and eighty-five. The Tzeltal term *čanbalam* is nearly identical in extension to the scientific term Mammalia in one sense, but like the folk English term "animal" it can also refer to any macroscopic, actively motile living thing. Thus I will distinguish between the broader reference, *čanbalam*$_1$ 'animal,' and the more restricted usage with which we are concerned here, *čanbalam*$_2$. *čanbalam*$_2$ includes all known mammals except human beings, bats, and the armadillo. It includes no other organisms (Hunn 1973:161-164, 278-279). Seventy-five consistently named and/or well defined sub-categories are recognized (see Appendix A). Fifty-five (72 percent) of these taxa correspond exactly to some taxon of the scientific system. However, let us examine a few examples of taxa which fail to correspond in this way.

The Tzeltal deny that human beings (*kirsiáno* > Spanish, *cristiano* 'Christian') are *čanbalam* in either sense of the term. However, they are quite cognizant of an affinity between men and monkeys and attribute to both a common origin (Hunn 1973:281). There is no named taxon inclusive of men and monkeys; this would deny man his unique status. Man alone possesses a 'soul' (*č'ulel*) and the capacity for speech. Bats (*so¢'*) are 'animals' but not 'mammals.' Though they fly like birds, only one of more than twenty informants queried considered bats to be a kind of *mut* 'bird.'[2] Though anomalous, bats are thought to be closely related to shrews (*yaʔalbe*, which is a kind of *č'o* 'small rodent,' which in turn is included in *čanbalam*$_2$). A story detailing the transformation of shrews into bats is widely repeated in Tenejapa (Hunn 1973:288-289). The armadillo (*mayil tiʔbal*) is also anomalous. Though it goes "on all fours" like 'mammals' (as do lizards and frogs, etc.) and has a nose and tail very similar to those of the Virginia opossum (*ʔuč*), it lacks hair, having a "ridged" back like a turtle (*šʔahk'*). It should be noted that no extraordinary ritual significance is attributed to these "anomalous" animals (cf. Douglas 1966). It seems that the Tzeltal are not greatly concerned with the ambiguity they find at this level of their bio-taxonomy.

In the Tzeltal system *č'o* 'small rodent' includes shrews (*yaʔalbe*) in addition to representatives of several rodent families. Shrews belong to the order Insectivora, not the Rodentia, while several rodents such as the Mexican porcupine (*č'iš ʔuhčum*), the paca (*halaw*), squirrels (*čuč*), and the pocket gopher (*ba*) are excluded from the category *č'o*. There is no neat correspondence between this Tzeltal taxon and any scientific taxon. However, the Tzeltal recognize *yaʔalbe* as the most deviant sub-category of *č'o* and are aware of the fact that it is the only kind of *č'o* which is carnivorous.

Three Tzeltal names (*yašal č'o* 'gray mouse,' *sakʔeal č'o* 'white-mouth mouse,' and *k'alel č'o* 'bright mouse') are used in reference to several species of mice of the genus

Peromyscus. These terms are not synonymous; the three categories so named are distinguished by differences in the color of the mice's fur. This attribute happens to correlate poorly with the species distinctions recognized by zoologists. None of these named categories corresponds with a scientific taxon. Yet the Tzeltal certainly recognize the genus *Peromyscus*, treated here as a "covert generic"[3] taxon, for these three terms are never applied to mice of any other genus (see Table 1).

Table 1. Naming patterns of *Peromyscus*.*

	yašal č'o	*sak ʔeal č'o*	*k'alel č'o*	Total	Other Names
P. boylii	3	8	0	11	0
P. mexicanus	20	19	5	44	1
P. zarhynchus	7	6	7	20	2
Peromyscus sp.	1	1	2	4	2
Total *Peromyscus*	31	34	14	79	5
Other genera labeled *č'o*	0	0	0	0	38

*This table is adapted from Hunn (1973:293). Identifications total 84 while *Peromyscus* specimens total 78. A few specimens were not identified, but some were identified by more than one informant.

The two species of peccaries (native animals related to the pig) are known as *wamal čitam* 'bush pig.' Two kinds of *wamal čitam* are widely recognized. It is clear from informants' descriptions that *niwak wamal čitam* 'large bush pig' refers to the adults of either species while *bahk'al wamal čitam* 'group-of-four-hundred bush pig' refers to the young of either species. Yet informants deny that one is a growth stage of the other. This "confusion" may be explained by reference to the fact that peccaries are no longer found near Tenejapa and by the fact that adults and young of both species often forage in separate groups (Alvarez del Toro 1952).

These examples indicate that the differences between the Tzeltal classification of mammals and that of the scientist are readily comprehensible, i.e., we have no indication that radically different principles underly the two systems. Furthermore, the similarities are striking.

defining an adequate measure of taxonomic correspondence

To describe precisely how "radically different" folk and scientific classifications of organisms are, or how "striking" the "similarities" between them, requires what I propose to call a Coefficient of Dissimilarity. The specification of this coefficient involves a choice among measures. The choice is not entirely arbitrary. First, we want a measure which promises to be of general applicability to folk bio-taxonomic systems. Second, the measure must be explicit; any two individuals analyzing the same set of data should arrive at the same result. Finally, the measure should reflect as closely as possible our intuitive notions about degrees of similarity and difference.

Berlin, Breedlove, and Raven (1966) calculated the percent of folk "specific" taxa, i.e., terminal taxa, in one-to-one correspondence with scientific species (using a sample of 200 Tzeltal botanical terms). They reported a low figure of 38 percent (reduced to 14 percent for native species). On the basis of this less than compelling figure, they concluded that there was no necessary correspondence between folk and scientific systems of botanical classification. They argued that this lack of correspondence is due to the "special purpose" nature of folk systems as opposed to the "general purpose" nature of the scientific system.

Bulmer (1970:1072-1073) took issue with this early conclusion of Berlin et al., but he provided no alternative to their method of comparison. Bulmer estimated that for the Karam of New Guinea, "only about 60% of the terminal taxa applied to vertebrates, let alone invertebrates, appear to correspond well with zoological species" (1970:1075). The percentages of correspondence in both instances did not give strong support for Bulmer's assertion that "there is a *conceptual* correspondence between the great majority of terminal taxa applied by Karam and the species recognized by zoologists" (1970:1076). In short, this method of statistical comparison does not closely reflect "intuitive notions about degrees of similarity and difference."

Recently Berlin (1973:267-269; Berlin, Breedlove, and Raven 1974:120-122) has applied a modified version of the index cited above. The new measure avoids the most serious limitation of the earlier version by selecting "folk generic taxa" (defined in Berlin, Breedlove, and Raven [1973:216-219]) rather than terminal folk taxa as the units of comparison. Berlin then calculated the percent of folk generic taxa which correspond in a one-to-one fashion with scientific species. Additional tabulations distinguished over-differentiated taxa—folk generic taxa which subdivide a scientific species—and under-differentiated taxa of two types, those which include more than one species of a single scientific genus (type 1) and those which include species from more than one scientific genus (type 2). The substitution of generic for terminal folk taxa in the calculation raised the percentage of "basic" folk taxa in one-to-one correspondence to scientific species from 38 percent (for a random sample) to 61 percent (for the total sample [Berlin 1973:269]). The revised measure thus better reflects the intuition that folk classification does not differ radically from the scientific, as in Berlin's current view: "There is at present a growing body of evidence that suggests that the fundamental taxa recognized in folk systematics correspond fairly closely with scientifically known species" (1973:267).

However, I believe Berlin's revised measure is not yet adequate in two respects. First, restricting the comparison to folk generics and scientific species ignores a significant proportion of the data: most notably, (1) those folk generics which correspond perfectly with scientific taxa above the species level, as in the case of folk English "bat," and (2) those "folk specifics" which correspond in a one-to-one fashion with scientific species, such as folk English "grizzly bear." Berlin justifies this limitation as follows:

> One of the difficulties in any comparison concerns the units of analysis to be considered. In the case of Western systematics, the selection of the basic unit is straight forward—it *must* be the species. In folk systematics, it now appears useful to focus on the folk genus as the primary unit. The folk genus, it will be recalled, is the smallest linguistically recognized class of organisms that is formed, as the folk zoologist Bulmer has succinctly stated, " . . . by multiple distinctions of appearance, habitat, and behavior." These two units, then, the scientific species and the folk genus, will be those selected as the basic taxa to be examined in *any* comparison of the folk and scientific systems of classification (1973:267-268; italics mine).

It is not the case that the scientific species *must* be selected. Scientific species are certainly the basic units of the scientific classification, due to their unique logical status *vis-à-vis* other taxa. However, this unique status derives directly from evolutionary theory, *viz.*, species are genetically isolated populations (Mayr 1970:12-13). The theory of folk systematics is not predicated on such considerations. Thus with respect to the correspondence of folk to scientific taxa, the scientific species is not necessarily the basic unit. Rather, folk taxa are directly comparable to the "arbitrary" non-species taxa in terms of their logical basis (cf. Hunn 1973:98-100). I will propose a measure which makes no restriction on the scientific category or rank of comparable taxa.

Second, the reliable application of Berlin's revised measure depends on the "proper" application of the distinction between folk generic taxa and other folk taxa. I have argued elsewhere that this distinction is not without ambiguity (Hunn 1973:88-92). Thus different investigators might arrive at different figures for the same set of data (cf. the difficulty of applying the familiar anthropological post-marital residence typology [Goodenough 1956]). In addition, Berlin is not quite correct in equating Bulmer's "folk species," to which Bulmer is referring in the quotation from Berlin cited above, with his own folk genera. Berlin's folk genus is recognized primarily by reference to the type of lexeme which labels the taxon (Berlin, Breedlove, and Raven 1973:218). Bulmer's folk species is defined with reference to the multiplicity of observable differences which characterize it (1970:1072). These two categories of taxa broadly overlap but are not, in general, the same (cf. Hunn 1973:106-114). Thus it should not be assumed that one-to-one correspondence to scientific taxa is to be expected only of Berlin's folk generics. In short, Berlin's revised measure can be refined.

The numerical taxonomic literature might be expected to provide a more refined approach to the measurement problem posed here. Numerical taxonomy attempts to provide an explicit methodology for bio-systematics. Rather sophisticated "measures of congruence" have been developed to evaluate alternative taxonomic approaches to the classification of a given set of OTUs (Operational Taxonomic Units). Unfortunately these measures are not appropriate for the special requirements of folk taxonomic data. Farris (1971) reviews many of these measures. All indices rely on direct or indirect comparison of "character sets." However, as argued elsewhere (Hunn 1973:8-9), folk biological taxa, in general, are not defined by reference to verbalizable feature contrasts. These measures are not "of general applicability to folk bio-taxonomic systems."

Folk taxa are defined *vis-à-vis* scientific taxa by comparing the folk and scientific names for each organism of a representative sample of familiar organisms. Small samples are quite adequate to establish the simple correspondence of a folk to a scientific taxon (cf. Hunn 1973:36 ff). Thus the proportion of all folk taxa under consideration which correspond to scientific taxa (a folk taxon which includes, as inferred from sample data, all and only those locally occurring organisms included in a single scientific taxon is said to correspond to that scientific taxon) is a measure which can be readily applied to folk biological data. This measure avoids a significant limitation of the Berlin-Bulmer statistics, i.e., the restriction of the comparison to scientific species, on the one hand, and to folk terminal or generic taxa on the other. There is no *a priori* rationale for judging the folk English taxon "raccoon" (= the species *Procyon lotor*) as corresponding more closely to the scientific system than "bat" (= the order Chiroptera). As will be seen, these restrictions systematically lower the proportion of correspondence.

Though the proportion of all folk taxa which correspond to scientific taxa of whatever categorical rank meets the requirements of an adequate measure, it is insensitive to degrees of correspondence. For example, it would be useful to distinguish a folk taxon which included some but not all species of a single scientific genus from a folk taxon which grouped assorted species from a variety of higher level scientific taxa and to distinguish both from a "random" collection of organisms. (Berlin's distinction between two types of underdifferentiated taxa is a first step in this direction, but it is not logically exhaustive.) Those folk taxa which fail to correspond to a scientific taxon may vary a great deal in the degree to which they approximate scientific groupings. To measure the degree of correspondence more accurately we need to define a weighted measure. What follows is a description of a means to that end.

a weighted measure of taxonomic dissimilarity

First we need to delimit the folk taxonomy to which the measure will apply. This is done by choosing a unique beginner. For present purposes the unique beginner may be any taxon of the folk biological system. The folk taxonomy thus defined consists of the unique beginner and all taxa which are included in that unique beginner.

It should be noted that a taxonomy implies several facts. First, taxa are *sets of things*, in this case, organisms. Second, subordinate taxa are *subsets* of superordinate taxa. Thus horses cannot be both mammals and birds. Furthermore, if a Mallard is a kind of duck and ducks are kinds of birds, then a Mallard must be a kind of bird. The structures underlying both folk and scientific taxonomies may be represented as tree-like graphs (i.e., dendrograms, as in Figure 1). These graphs consist of *vertices* which represent taxa. The vertices are connected by *directed edges* which, proceeding downwards, indicate relationships of strict and immediate set inclusion. The unique beginner has no incoming edges. Every other vertex has exactly one incoming edge. A vertex may have no outgoing edges, in which case it is a *terminal taxon*, but if it has any, then it must have at least two (Kay 1971:882; Theorem 10).

As noted above, the measure proposed here involves comparing a folk taxonomy with a part of the scientific taxonomy. It is necessary to pare the scientific taxonomy down to size. This is done by a series of *reduction rules*. First, I exclude all scientific taxa which have no representatives in the local flora or fauna. Such taxa are of various degrees of inclusiveness, from species to entire phyla. It is not particularly surprising that the Tzeltal

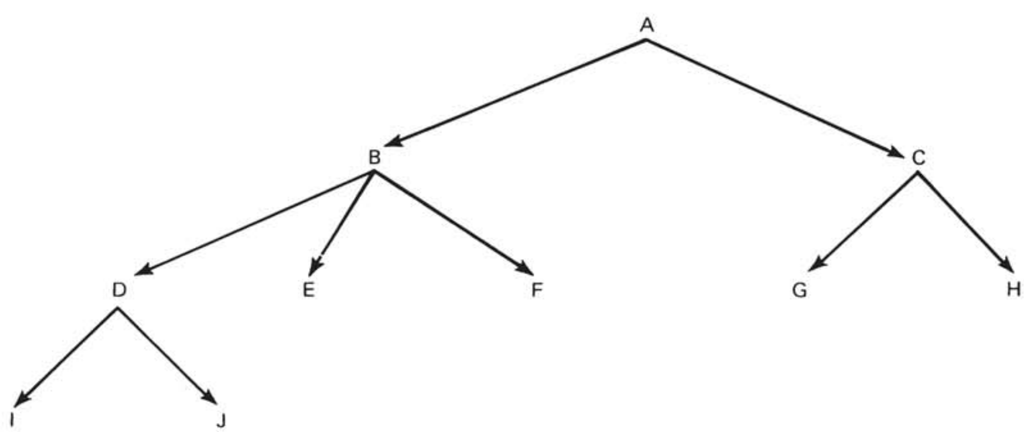

Figure 1. A taxonomic structure represented as a tree-like graph.*

*A graphic representation of a taxonomic structure involves vertices (taxa) indicated by capital letters and directed edges (relations of strict and immediate inclusion) indicated by arrows. Note that there is a unique path from the unique beginner, A, to all other taxa. Note also that every taxon which is further subdivided is subdivided into at least two subtaxa. Taxa which are not further subdivided are terminal (i.e., E, F, G, H, I, J).

do not deal with *Felis leo*, the African lion, or the Coelenterata, a phylum composed almost entirely of marine organisms such as the jellyfish. The exclusion of a taxon from the local scientific taxonomy implies the exclusion of all its subordinate taxa as well.

The second reduction removes from consideration all scientific taxa that contain no organisms contained in the unique beginner of the folk system under consideration. The unique beginner in this example is *čanbalam*$_2$. All organisms included in the taxon so labeled are also included in the scientific class Mammalia. Thus all scientific taxa not included in the class Mammalia are excluded. By this reduction we need no longer worry about birds, fish, or worms. At this stage we have what may be called a *truncated* and *local* scientific taxonomy, local by virtue of the first reduction and truncated by virtue of the second.

A third reduction collapses all taxa which are exactly equivalent in membership to a superordinate taxon. This follows from the specification that a minimum of two descending edges is required of all non-terminal vertices in the graphic model of a taxonomic structure. Take, for example, that section of the scientific taxonomy which includes the two local opossum genera. The first two reduction rules leave us with the following sequence of taxa: class Mammalia, infraclass Metatheria, order Marsupialia, family Didelphidae, then the two genera. The third reduction rule deletes the order and the family. The family Didelphidae is the only family of marsupials found locally and is thus equivalent in extension to the order. Likewise, the order Marsupialia is the only order of the infraclass Metatheria with local representatives. The infraclass is retained because it contrasts with the infraclass Eutheria, which includes the remaining mammals. The two local opossum genera, *Didelphis* and *Philander*, are also retained because neither is equivalent to the family which includes them both.

The final reduction eliminates all scientific taxa below the level necessary to specify the content of the terminal folk taxa. For example, the Tzeltal regularly recognize only a single kind of *yaʔalbe* 'shrew.' It is thus a terminal taxon. The order Insectivora is represented in southern Mexico by one family (the shrews, Soricidae), of which two genera containing several species are found locally. All these local species are classified indiscriminately as *yaʔalbe*. Since all and only the local representatives of the insectivores are so classified, I delete all the scientific taxa subordinate to Insectivora. On the other hand, the several local species of the genus *Felis* (cats) are distinguished in Tzeltal. Thus no further deletion of scientific taxa is possible. What is now left of the scientific taxonomy may be described as *terminated, collapsed*, truncated, and local. It is the *reduced* scientific taxonomy.

We may now determine the degree to which each folk taxon fails to correspond to the scientific classification. For each folk taxon trace downward in the reduced scientific taxonomy, following every branch which includes some organism also included in the folk taxon. For example, "tree" in folk English includes a portion of two major botanical taxa, the cone-bearing plants and the flowering plants. It is thus necessary to follow both cone-bearing and flowering plant branches. We need not consider such taxa as ferns and mosses, however, because they lead to nothing which is called a tree.

Proceed downward until a taxon is reached which is a subset of the folk taxon. If a single scientific taxon is found which includes everything included in the folk taxon and nothing which is included in any other folk taxon, the corresponding scientific and folk taxa are equivalent. All such folk taxa merit a zero degree of dissimilarity. An example from Tzeltal is *čuč* 'squirrel' which both includes and is included in the family Sciuridae of the reduced scientific taxonomy.

If several scientific taxa are found which jointly partition the folk taxon, retrace

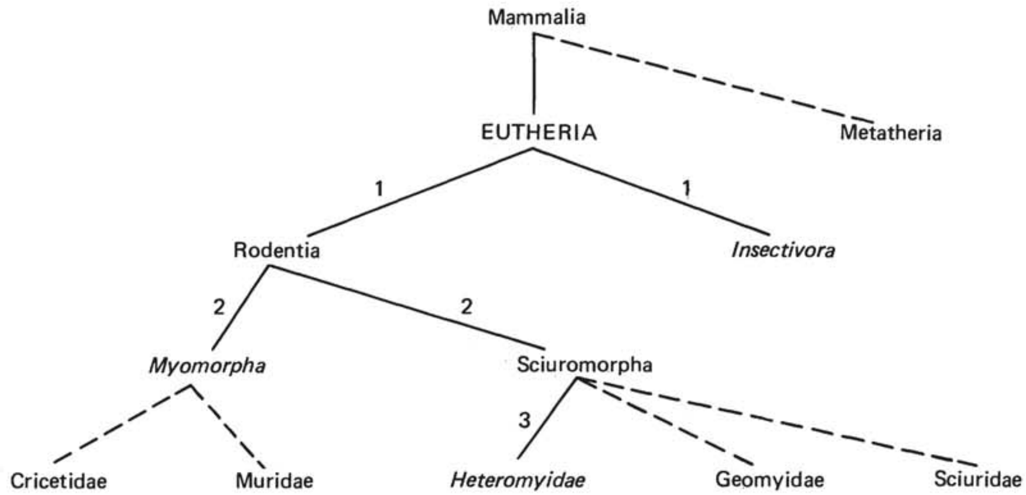

Figure 2. Measuring the degree of the folk taxon *č'o*.*

*The italicized scientific names indicate the maximal subsets of *č'o*. The lowest common denominator of *č'o* is written with capital letters. Solid lines indicate edges traced in measuring this taxon. Dotted lines are not traced. The numerals beside the solid lines count the number of steps between the lowest common denominator and each maximal subset. The largest such number is the degree of dissimilarity of the taxon, in the case of *č'o*, three.

upward in the scientific taxonomic tree from each to the first taxon which includes all such scientific taxa. This taxon may be referred to as the *lowest common denominator* of the folk taxon. If we count the intervening edges of the longest path retraced, we have an index of the degree to which the folk taxon is out of correspondence with the scientific system.

The Tzeltal category *č'o* 'small rodent' provides an example (see Figure 2). First we trace downward from Mammalia. We pass the infraclass Eutheria, then continue to both Insectivora and Rodentia, since members of both orders are included in the taxon. On the rodent branch we must continue past the suborder Sciuromorpha to the family Heteromyidae (the pocket mice), since other families of that suborder are not included in *č'o*. We need not continue past the second rodent suborder, the Myomorpha, since all local representatives of the Myomorpha are included in *č'o*. We end up with three scientific taxa which are subsets of *č'o*. Therefore we must retrace our steps until we find the lowest common denominator, which in this case is the infraclass Eutheria. A degree of three is assigned to *č'o* because the longest retrace path involves three steps (i.e., Heteromyidae–Sciuromorpha–Rodentia–Eutheria).

A third possibility must be considered: all or part of the folk taxon may be equivalent to but a portion of a *terminal* scientific taxon, i.e., it may be impossible to find a set of scientific taxa which partition the folk taxon. The Tzeltal taxa *yašal č'o*, *sakʔeal č'o*, and *k'alel č'o* illustrate this contingency. The lowest common denominator in each case is the genus *Peromyscus* represented locally by at least four species. Yet no species is entirely included within any one of these folk taxa (refer back to Table 1). We are left without scientifically recognized subsets as starting points for counting the length of the retrace paths. This dilemma is resolved by declaring that the "distance" between the subset of the species which is included in the folk taxon and the terminal scientific taxon is one. Thus these three folk taxa are all of degree two, the distance between the common genus and a proper subset of each species. In the same fashion we may count as degree one all folk taxa which are equivalent to a proper subset of a single terminal scientific taxon, e.g., *kašlan ¢'i?* 'Ladino dog,' which is equivalent to a portion of the species *Canis familiaris.*

The unweighted Coefficient of Dissimilarity, *D*, of the entire folk taxonomy defined by the unique beginner *čanbalam*$_2$ is then calculated as the proportion of taxa of degree one or more to the total number of folk taxa in the taxonomy.

$$D = \frac{\#T^1 + \#T^2 + \ldots + \#T^n}{\#T}.$$

#indicates the cardinality, or number of members of, the set T^k of all folk taxa of degree k, and T is the set of all folk taxa of whatever degree.

The weighted Coefficient of Dissimilarity, D', is derived from the unweighted measure by multiplying the number of taxa of each degree by the value of the degree itself. Thus

$$D' = \frac{1(\#T^1) + 2(\#T^2) + \ldots + n(\#T^n)}{\#T} \quad \text{or} \quad D' = \frac{\sum_{i=0}^{n} i(\#T^i)}{\#T}.$$

This weighting provides recognition for the intuition that *kašlan ¢'i?* 'Ladino dog' (degree one) departs less from the scientific system than does *č'o* 'small rodent' (degree three). D for the domain defined by *čanbalam*$_2$ is .28, D' is .39, since there are fourteen taxa of degree one, five of degree two, two of degree three, of a total of seventy-six folk taxa (see Table 2).

Table 2. Calculation of D and D' for *čanbalam*$_2$.

$\#(T^0) = 55$		$0\ \#(T^0) = 0$	
$\#(T^1) = 14$		$1\ \#(T^1) = 14$	
$\#(T^2) = 5$	$D = \frac{14 + 5 + 2}{76}$	$2\ \#(T^2) = 10$	$D' = \frac{14 + 10 + 6}{76}$
$\#(T^3) = 2$	$= 0.28$	$3\ \#(T^3) = 6$	$= 0.39$
$\#(T) = 76$			

The lowest possible value of the unweighted coefficient, D, is zero, indicating a perfect correspondence between the folk taxonomy and the sub-system of the scientific taxonomy generated by the reduction rules. The highest possible value is one, indicating that no folk taxon corresponds exactly to any scientific taxon. The weighted measure, D′, also has a minimal value of zero, again indicating a perfect fit. However a very high value of D′ might be in the neighborhood of ten.[4] Such a value would be expected if organisms had been randomly assigned to taxa. The weighted measure with its wider range of values provides a more sensitive index of the degree of fit. However, the upper bound of D′ is a function of the depth of the reduced scientific taxonomy which varies both within and between domains. In order to gain the advantage of both a weighted and a "normalized" measure (i.e., one which varies between 0 and 1) it is necessary to recompute the degree of dissimilarity for each taxon such that it is expressed as a proportion of the maximum possible value. This maximal value is equal to the length (i.e., number of edges) of the longest path in the reduced scientific taxonomy from the unique beginner, through the folk taxon in question, to the lowest level of the scientific taxonomy (see page 325). Normalized D′ (i.e., D″) then is equal to the sum of the degrees of dissimilarity calculated in this fashion for each folk taxon, divided by the number of folk taxa. For *čanbalam*$_2$

$$D'' = 5.19/76 = .07$$

or 7 percent of the maximum possible dissimilarity. The values obtained in this example clearly indicate that the Tzeltal have paid close attention to the structure of "reality." The ultimate significance of these numbers depends, of course, on subsequent comparisons with coefficients calculated according to this technique for other domains and for comparable domains from other cultures.

evaluation of the measure

As noted above, a measure should meet at least three criteria. How well does the present weighted measure qualify according to each criterion? First, a measure should be of wide utility. The reduction rules by which a specific subsystem of the scientific taxonomy is selected for comparison with each folk system assures that the variations in local floras and faunas are taken into account. These same rules also minimize the effect of variations among scientific classifications due to competing schools of thought among systematists or to differential degrees of detail applied to a classification. I have performed a brief test of this factor by recomputing D′ using only the required taxonomic categories, i.e., class, order, family, genus, and species, ignoring the several intermediate levels which may be poorly defined. The value of D′ changes only .02, reduced from .39 to .37. The effect of utilizing the more generalized scientific taxonomy as opposed to the finely detailed version is a slight loss of sensitivity of the measure. Thus this technique should be useful even in the absence of a fully developed scientific classification. Furthermore, the nature of the technique permits its application to biological domains of any size.

The second criterion demands an explicit procedure to insure replicability. The procedures outlined can be precisely defined in set theoretical terms, as is done in Appendix B. Thus once the folk system has been adequately described and a standard version of the scientific taxonomy agreed upon, the results are predetermined.

Finally, the measure should accurately reflect our intuitions concerning what it purports to measure. I argued above that the statistic first employed by Berlin, Breedlove, and Raven (1966) and Bulmer (1970) was deficient in this respect, systematically underestimating the degree to which folk and scientific bio-taxonomic systems correspond. The summary data in Table 3 prove this point for the present example. Subgeneric taxa correspond rather poorly with scientific taxa in this example, but such taxa are disproportionately represented among the terminal taxa of a folk system, the basis for the first measure. It is also clear from Table 3 that Berlin's revised measure (1973), which is based on folk generic taxa, avoids this earlier source of bias. Thus it is not surprising that shifting from the terminal-species index to the generic-species index raises the percentage of correspondence from 38 to 61 for the Tzeltal plant domain. However, avoiding selective attention to folk or scientific taxa of any particular rank, the Coefficient of Dissimilarity indicates a much higher figure than either of Berlin's calculations, i.e., 93 percent. Does the plant taxonomy correspond less well to the

Table 3. Distribution of dissimilarity by type of folk taxon.

	Degree of Dissimilarity				
Type of Taxon	0	1	2	3	Totals
Generic*	42	2	0	0	44
Subgeneric	13	12	5	0	30
Supergeneric	0	0	0	2	2
Totals	55	14	5	2	76

*Berlin, Breedlove, and Raven (1973) define an hierarchical series of folk bio-taxonomic categories as follows: varietal < specific < generic < intermediate < major life form < unique beginner. They describe procedures for deciding to which category a taxon belongs. Varietal and specific taxa are grouped here as subgeneric taxa. The remaining non-generic taxa are supergeneric.

scientific system than the domain treated here, or is the difference due to the contrast in measures employed? If the latter, which is the more meaningful measure? These questions require a closer examination of the data. The obvious first step is to apply Berlin's revised measure to the Tzeltal 'mammal' data (see Table 4). We may conclude that the Tzeltal 'mammal' taxa correspond more closely than the plant taxa to the scientific sorting (75 percent as opposed to 61 percent). The remaining difference (75 percent as opposed to 93 percent) is due to differences in the measures. Which, then, is the more meaningful figure? The answer depends on precisely what each purports to measure. I claim that the Coefficient of Dissimilarity is a "measure of the degree of correspondence of folk to scientific biological classification." Berlin's generic-species index is designed to demonstrate "that the fundamental taxa recognized in folk systematics correspond fairly closely with scientifically known species" (1973:267). Our goals are different, thus our measures

Table 4. Tzeltal 'mammal' taxa classified according to Berlin's revised index (1973).

	Folk Generics	Folk Specifics
One-to-one correspondence	33 (75%)	11 (37%)
Overdifferentiated	0 (0%)	13 (43%)
Underdifferentiated	11 (25%)	1 (3%)
Type 1	5 (11%)	1 (3%)
Type 2	6 (14%)	0 (0%)
Not classifiable	0 (0%)	5 (17%)*
Totals	44 (100%)	30 (100%)

*These five specific taxa cannot be classified because they both overdifferentiate (by splitting a species) and underdifferentiate (by combining portions of several species). These cases are discussed above (pp. 310-311). Berlin does not consider this contingency.

are not the same. However, I contend that the measure proposed here retains an advantage over Berlin's technique even with respect to his own stated goals.

The present measure, applied to the data at hand, supports a key hypothesis suggested by both Bulmer and Berlin, namely that there is a subset of folk taxa which is psychologically fundamental in folk systems of biological classification. Berlin identifies this as the set of "generic" taxa (Berlin, Breedlove, and Raven 1973:216). Bulmer distinguishes a "hard core of lower order groupings in any [folk biological] taxonomy" which are "multi-purpose, multi-dimensional" and logically natural groupings from taxa "at the peripheries" which may be special purpose categories defined by a strictly limited set of attributes (1970:1087). If we examine the 'mammal' domain, we are struck by the fact that the major portion of the dissimilarity is attributable to either the most general categories or to the most specific (see Table 3). The two taxa with the highest degree of dissimilarity (3) are those which Berlin would call supergeneric, i.e., *čanbalam*$_2$ and *č'o*. Three of the five taxa with the next highest degree (2) are exceptional in being defined by a single attribute, fur color, and are thus what Bulmer and I call varietals. Seventeen of the nineteen taxa of degree one or two have names of binomial form, according to Berlin a characteristic of specific (subgeneric) as opposed to generic taxa. Thus there is a subset of the taxa of this domain characterized by the near-perfect correspondence of its members to scientific taxa, and this subset corresponds rather closely to the generic taxa of Berlin's system and to Bulmer's "hard core" taxa. However, the tabulation indicates that while most generic taxa are fundamental in this sense (forty-two of forty-four), not all fundamental taxa are generic (forty-two of the fifty-five taxa of zero degree of dissimilarity are generic, thirteen are specific taxa). A significant proportion of specific taxa (43 percent) might best be included in the "hard core" defined by Bulmer. Berlin's measure *assumes* that generic taxa are the "fundamental taxa." The present measure *demonstrates* that generic taxa exhibit the near-perfect correspondence with scientific taxa which Berlin attributes to the fundamental set of taxa but indicates, further, that this property is not exclusive to the generic set.

A closer look at the data suggests an additional caution with respect to Berlin's technique. He selects the scientific species as his basis for comparison. The scientific species within a limited geographic range is the finest taxonomic distinction which can be made. Berlin's technique thus implies that folk systematists will tend to make distinctions comparable to the finest distinctions of the scientist. This ignores a significant consideration: of the folk taxa in one-to-one correspondence, how many correspond to a scientific species which has no congeners in the local area? Correspondence to such "isolated" species provides no justification for equating this distinction with the finest distinctions of the scientist. Twenty-five of the thirty-three (75 percent) Tzeltal generic 'mammal' taxa which Berlin would describe as in one-to-one correspondence with scientific species in fact correspond to such "isolated" species, species with no close relatives known to the Tzeltal. Only eight are actually distinguished from congeners at the species level. A more meaningful test of Berlin's contention is provided by those Tzeltal generics which apply in situations where species of a single genus might be distinguished. The Tzeltal taxa correspond to single species in only eight of seventeen (47 percent) of such test cases (see Table 5: H, I, J). It seems that Berlin's technique does not measure precisely what it purports to measure. Furthermore, the above analysis indicates that because "isolated" species are more likely to correspond to folk generic taxa than non-isolated species, Berlin's index is susceptible to a gratuitous source of variation, *viz.*, the frequency of "isolated" species in the scientific co-domain. The Coefficient of Dissimilarity will not be affected in this fashion.

Table 5. Correspondence of Tzeltal generic taxa to scientific taxa of various types.

A.	To no single scientific taxon	2	(5%)
B.	To a scientific taxon of supraspecific rank represented locally by more than one species	9	(20%)
C.	To an "isolated" scientific species	25	(57%)
D.	To a single species which is not "isolated" (i.e., more than one species of the scientific genus is known to the Tzeltal)	8	(18%)
	Total	44	(100%)
E.	One-to-one correspondence (C + D)	33	(75%)
F.	Underdifferentiated (A + B)	11	(25%)
G.	Degree of dissimilarity = 0 (B + C + D)	42	(95%)
H.	Situations of polytypy (B + D)	17	
I.	One-to-one correspondence in polytypic situations [D/(B + D)]	0.47 of H	
J.	Underdifferentiation in polytypic situations [B/(B + D)]	0.53 of H	

One final qualification is essential. The measure proposed here claims to measure the similarity of two independent classification schemes. This claim is limited by the extensional logic employed to define taxa. Taxa are defined simply as sets of organisms (cf. Kay 1971). Obviously, the mental correlate of a taxon is not a set of organisms but must be a *rule*, since human beings have no difficulty recognizing new instances of a taxon. I argue elsewhere (Hunn 1973:8-9; 84-85; Hunn n.d.) that this simplification is required by the fact that most folk taxa are semantic primitives. The proposed measure does not directly compare the rules of categorization employed by folk and scientist to classify organisms. It simply compares the sortings which result. Yet it is highly unlikely that incomparable rules of categorization will produce very similar sortings repeatedly. Thus the measure proposed, if applied to an adequate comparative sample, can tell us whether or not folk and scientist employ comparable cognitive strategies in classifying living things. In addition, the method suggested here will enable students of folk biology to formulate and test specific hypotheses concerning the relationship between the categorizing activity of the human mind and a variety of conditioning factors, be they linguistic, cultural, or ecological.

appendix a: the Tzeltal taxa *re* mammals

Following is a list of all Tzeltal taxa which classify mammals. All taxa included are consistently recognized. Each is referred to by its primary name. Alternative names and less consistently recognized taxa are treated in Hunn (1973). Tzeltal taxa follow the linear order of the scientific taxa to which they correspond, except that all subordinate taxa follow their superordinate taxon. Indentation of folk names indicates the folk taxonomic category of each: the unique beginner (a "major life form" following Berlin) is at the left margin, *č'o* (an "intermediate" taxon) is indented one space, all "generic" taxa are indented two spaces, subgeneric taxa are indented four spaces. Portions of some Tzeltal names are included in parentheses. These portions are optional. The three covert taxa are bracketed.[5] The glosses cited are English names for approximately equivalent scientific taxa. In the event that no such gloss is available, a literal translation of the Tzeltal name is cited as indicated by an asterisk. The taxonomic category of the lowest common denominators and maximal subsets is indicated as follows: genus names are italicized and capitalized, species names are italicized but not capitalized, superfamilies end in *-oidea*, families in *-idae*, subfamilies in *-inae*, all others are labeled. The degree of dissimilarity is as described in the text.

Tzeltal Name	Gloss	Lowest Common Denominator /Maximal Subsets	Degree /D"
čanbalam₂	'mammal'	Mammalia (class) Metatheria (infraclass) Eutherian orders Pilosa (infraorder) Ceboidea	3/0.38
ʔuč	'Virginia opossum'	*Didelphis*	0/0
huyum ʔuč	'fat opossum'*	*Didelphis* subset of *D. virginiana*	1/0.33
ba¢'il ʔuč	'true opossum'*	*Didelphis* subset of *D. virginiana*	1/0.33
ʔuč č'o	'four-eyed opossum'	*Philander*	0/0
ba¢'	'howler monkey'	*Alouatta*	0/0
maš	'spider monkey'	*Ateles*	0/0
¢'uh¢'uneh čab	'anteater'	Pilosa (infraorder)	0/0
t'ul	'rabbit'	Lagomorpha (order)	0/0
kašlan t'ul	'European rabbit'	*Oryctolagus*	0/0
ba¢'il t'ul	'true rabbit'*	*Sylvilagus*	0/0
čuč	'squirrel'	Sciuridae	0/0
yašal čuč[5]	'gray squirrel'*	*Sciurus* *S. aureogaster,* *S. griseoflavus*	1/0.17
¢'eh (čuč)	'Deppe's squirrel'	*S. deppei*	0/0
pehpen čuč	'flying squirrel'	*Glaucomys*	0/0
ba	'pocket gopher'	Geomyidae	0/0
č'uypat ba	'white-banded gopher'*	Geomyidae subset of *Heterogeomys hispidus*	1/0.20
ba¢'il ba	'true gopher'*	Geomyidae subset of *H. hispidus*	1/0.20
č'o	'small rodent'	Eutheria (infraclass) Insectivora (order) Myomorpha (suborder) Heteromyidae	3/0.43
yaʔalbe (č'o)	'shrew'	Insectivora (order)	0/0
k̇iwoč č'o	'pocket mouse'	Heteromyidae	0/0
sabin č'o	'rice rat'	*Oryzomys*	0/0

Tzeltal Name	Gloss	Lowest Common Denominator /Maximal Subsets	Degree /D"
moin teʔ č'o	'vesper rat'	*Nyctomys*	0/0
sin (č'o)	'harvest mouse'	*Reithrodontomys*	0/0
[*Peromyscus*]		*Peromyscus*	0/0
yašal č'o	'gray mouse'*	*Peromyscus* subsets of *Peromyscus* spp.	2/0.29
sakʔeal č'o	'white-mouth mouse'*	*Peromyscus* subsets of *Peromyscus* spp.	2/0.29
k'alel č'o	'bright mouse'*	*Peromyscus* subsets of *Peromyscus* spp.	2/0.29
čitam č'o	'cotton rat'	*Sigmodon*	0/0
hseʔ teʔ (č'o)	'tree-cutter rat'*	Cricetidae *Neotoma, Tylomys*	1/0.17
karánsa (č'o)	'black rat'	*Rattus*	0/0
šlumil č'o	'house mouse'	*Mus*	0/0
čiš ʔuhčum	'Mexican porcupine'	Erethizontoidea	0/0
hulaw	'paca'	Cavioidea	0/0
¢'iʔ	'dog'	*Canis familiaris*	0/0
polisía ¢'iʔ	'German shepherd'	*C. familiaris* breed	0/0
kašlan ¢'iʔ	'Ladino dog'*	*C. familiaris* subset of *C. familiaris*	1/0.17
ba¢'il ¢'iʔ	'true dog'*	*C. familiaris* subset of *C. familiaris*	1/0.17
ʔok'il	'coyote'	*Canis latrans*	0/0
waš	'gray fox'	*Urocyon*	0/0
meʔel	'raccoon'	*Procyon*	0/0
kohtom	'coatimundi'	*Nasua*	0/0
ʔuyoh	'kinkajou'	*Potos*	0/0
sabin	'long-tailed weasel'	*Mustela*	0/0
sakhol	'tayra'	*Eira*	0/0
pay[5]	'skunk'	Mephitinae	0/0
č'in pay	'spotted skunk'	*Spilogale*	0/0
ʔihk'al pay	'hooded skunk'	*Mephitis*	0/0
lempat pay	'hog-nosed skunk'	*Conepatus*	0/0
haʔal ¢'iʔ	'river otter'	*Lutra*	0/0
balam	'jaguar'	*Felis onca*	0/0
¢ahal čoh	'ocelot'	*F. pardalis*	0/0
¢is balam	'margay'	*F. wiedii*	0/0
ʔik' sab	'jaguarundi'	*F. yagouaroundi*	0/0
šawin	'house cat'	*F. cattus*	0/0
¢emen	'tapir'	Ceratomorpha (suborder)	0/0
búro	'donkey'	*Equus assinus*	0/0
kawáyu	'horse/mule'	*Equus* *E. caballus* *E. caballus* X *E. assinus*	1/0.17
[mule][5]		*E. caballus* X *E. assinus*	0/0

Tzeltal Name	Gloss	Lowest Common Denominator /Maximal Subsets	Degree /D″
[horse] [5]		*E. caballus*	0/0
čitam	'pig'	Suidae	0/0
merikáno čitam	'American pig'*	Suidae subset of *Sus scrofa*	1/0.20
škoen čitam	'low-slung pig'*	Suidae subset of *S. scrofa*	1/0.20
baȼ'il čitam	'true pig'*	Suidae subset of *S. scrofa*	1/0.20
wamal čitam	'peccary'	Tayassuidae	0/0
niwak wamal čitam	'large peccary'*	Tayassuidae subsets of *Tayassu* spp.	2/0.33
bahk'al wamal čitam	'numerous peccary'*	Tayassuidae subsets of *Tayassu* spp.	2/0.33
teˀtikil čih	'deer	Cervoidea	0/0
yašal (teˀtikil) čih	'white-tailed deer'	*Odocoileus*	0/0
ȼahal (teˀtikil) čih	'brocket'	*Mazama*	0/0
(tunim) čih	'sheep'	*Ovis*	0/0
tenȼun	'goat'	*Capra*	0/0
wakaš	'cattle'	*Bos*	0/0
merikáno wakaš	'zebu'	*Bos taurus* breed	0/0
búro wakaš	'donkey cattle'*	*Bos* subset of *B. taurus*	1/0.17
baȼ'il wakaš	'true cattle'*	*Bos* subset of *B. taurus*	1/0.17

Taxa excluded from *canbalam*$_2$ (not included in the computation):

soȼ'	'bat'	= Chiroptera (order
kirsiáno	'human being'	= *Homo sapiens*
mayil tiˀbal	'armadillo'	= *Dasypus novemcinctus*

appendix b: formal treatment

The starting point for this description is Kay's axiomatic treatment of taxonomic structures (1971:881-885). Kay's definitions of Taxonomic Structure (Def. 1), the Unique Beginner (Axiom 2), a Terminal taxon (Def. 11), and taxonomic Level (Def. 13) are assumed here and will not be explicated. The symbol $\subset$ is used below to indicate the relation, "is a proper subset of"; $\subseteq$ is used to indicate the more general case, "is a subset of."

(1) Definition: *O* is the set of all organisms. The membership is not restricted by time or place.

(2) Definition: *S* is a set of scientific taxa such that each member taxon is a set of organisms, i.e., the scientific taxon $s_i \subset O$, $s_i \neq \emptyset$. The binary relational structure $\langle S, \subset \rangle$ is an open Linnean structure (cf. Gregg 1967; Kay 1971:885-887). For convenience, S will be referred to as the Scientific Bio-taxonomy or Scientific Taxonomy.[6]

(3) Remark: It is useful to distinguish the ideal scientific taxonomy, i.e., a *closed* Linnaean structure such that the terminal taxa of S constitute a partition of O, from any existing scientific taxonomy, which classifies only *known* organisms. This contrast between ideal and real creates the "problem of monotypy" (Kay 1971:886).

(4) Axiom: There is a set O′, O′ $\subset$ O, of organisms known to science. The unique beginner of the scientific taxonomy s_I = O′.

(5) Remark: There is no single scientific taxonomy. For the purposes of the measures defined here, any scientific taxonomic opinion which classifies a relevant subset of O′ may be defined as the scientific taxonomy in question.

(6) Definition: F is a set of folk taxa such that each member taxon is a set of organisms, i.e., the folk taxon $f_i \subset O$, $f_i \neq \emptyset$. The binary relational structure $\langle F, \subset \rangle$ is a taxonomic structure. F will be referred to as a Folk Bio-taxonomy or Folk Taxonomy.

(7) Remark: The unique beginner of a folk taxonomy, f_I, is as defined by Kay. However, the unique beginner may be defined with reference to any subsystem of the entire folk taxonomy as long as that subsystem is a taxonomic structure and given that $f_I \subset O'$.[7]

(8) Definition: There is a set O_L, $O_L \subset O'$, of all organisms known to the informant/s from whom the folk taxonomy was elicted (i.e., the local set of organisms), and $f_I \subseteq O_L$.

(9) Definition: The Local scientific taxonomy $Lo(S)$[8] $= \{s_i \ni s_i \in S, (s_i \cap O_L) \neq \emptyset\}$.

(10) Definition: The Truncated scientific taxonomy $Tr(s) = \{s_i \ni s_i \in S, (s_i \cap f_I) \neq \emptyset\}$.

(11) Definition: The Collapsed scientific taxonomy $Co(S) = \{s_i \ni s_i \in S$, such that there is no $s_j \in S, s_i \subset s_j$, and $(s_i \cap O_L) = (s_j \cap O_L)\}$.

(12) Definition: The Terminated scientific taxonomy $Te(S) = \{s_i \ni s_i \in S$, such that there is no $s_j \in S, s_i \subset s_j$, and $s_j \subseteq f_i$, where f_i is a terminal folk taxon$\}$.

(13) Definition: the subsystem of the scientific taxonomy used to calculate the Coefficient of Dissimilarity of a folk taxonomy (i.e., the reduced scientific taxonomy) is $S' = \{s_i' = (s_i \cap O_L), s_i \in Lo \cdot Tr \cdot Co \cdot Te(S)\}$.

(14) Theorem: The binary relational structure $\langle S', \subset \rangle$ is a taxonomic structure. Proof: S' is finite since S is finite and $\#S' < \#S$. $S' \neq \emptyset$ since $(s_I \cap O_L) \in S'$. This follows from the fact that $(s_I \cap O_L) \neq \emptyset$, $(s_I \cap f_I) \neq \emptyset$, and there is no s_j, $s_I \subset s_j$. $\emptyset \notin S'$ by definition 10. It remains to demonstrate that Axioms 2 and 3 of Kay (1971) hold for S'. Since $(s_I \cap O_L) \in S'$, Axiom 2 holds. Axiom 3 holds, since by Definitions 11 and 12, $c(s_i \cap O_L)$, $s_i \in Lo \cdot Tr \cdot Co \cdot Te(S)$, is either null or a partition of $s_i \cap O_L$.

(15) Definition: The Lowest Common Denominator of a folk taxon $Lcd(f_i) = lcd_i \ni lcd_i \in S'$, $f_i \subseteq lcd_i$, and there is no $s_i' \in S'$ such that $f_j \subseteq s_j' \subset lcd_i$.

(16) Theorem: There is a unique $lcd(f_i)$, $f_i \in F$. Proof: first we show that $Lcd(f_i) \neq \emptyset$. Since $f_i \subseteq f_I$, and $f_I \subseteq O_L \subset O' = s_I$, $f_i \subset s_I$. Furthermore, $s_i' \subseteq s_I$, $s_i' \in S'$. Thus $Lcd(f_i)$ is either some s_i, $s_i \subset s_I$, or $Lcd(f_i) = s_I$. Next we show that $Lcd(f_i)$ is unique. Assume the contrary: $Lcd(f_i) = \{lcd_i, lcd_j \ni lcd_i, lcd_j \in S'\}$. Then either $lcd_i \subset lcd_j$, $lcd_j \subset lcd_i$, or they are disjoint (by Kay's Theorem 16). If $lcd_i \subset lcd_j$ or $lcd_j \subset lcd_i$, the assumption contradicts Definition 15. If they are disjoint, either $f_i \not\subset lcd_i$ or $f_i \not\subset lcd_j$ (by the properties of the relation of set inclusion).

(17) Definition: The Maximal Subsets of a folk taxon $Ms(f_i) \supseteq \{ms_i \ni ms_i \in S', ms_i \subseteq f_i$, and there is no $s_j' \in S'$ such that $ms_i \subset s_j' \subseteq f_i\}$.

(18) Remark: The level of a taxon $L(t_i)$ is as defined by Kay. For our purposes it is only necessary to note that $L(s_i')$ is calculated with reference to S'.

(19) Definition: If there is a set/s of organisms $q_i \ni q_i \subseteq f_i$, $q_i \subset s_j'$, $s_j' \not\subseteq f_i$, where s_j' is a terminal taxon in S', then $q_i \in Ms(f_i)$ and $L(q_i) = L(s_j') + 1$.

(20) Theorem: The member sets of $Ms(f_i)$ constitute a partition of f_i. Proof: the members of $Ms(f_i)$ are disjoint by Definitions 11, 17, 19. It remains to demonstrate that if $o_i \in f_i$, $o_i \in O$, either $o_i \in$ some ms_i or $o_i \in$ some q_i. If $o_i \in f_i$, then $o_i \in lcd_i$. Since $o_i \in lcd_i$, then $o_i \in$ some $s_j', s_j' \subseteq lcd_i$, and s_j' is a terminal taxon in S'. Either $s_j' \subseteq$ some ms_i, in which case $o_i \in$ some ms_i, or $s_j' \not\subseteq ms_i$, in which case $o_i \in$ some q_i (by Defintion 19).

(21) Definition: The Degree of Dissimilarity of a folk taxon $D(f_i) = \max [L(ms_i, q_i)] - L(lcd_i)$, $ms_i, q_i \in Ms(f_i)$.

(22) Definition: The degree of dissimilarity of a folk taxonomy (weighted) $D'(F) = \sum_i D(f_i) / \#F$, $f_i \in F$.

(23) Definition: The normalized degree of dissimilarity of a folk taxon $D''(f_i) = D(f_i) / \max [L(s_i', q_i)]$, $s_i', q_i \subset lcd_i$.

(24) Definition: The normalized degree of dissimilarity of a folk taxonomy (weighted) $D''(F) = \sum_i D''(f_i) / \#F$, $f_i \in F$.[9]

notes

[1]This is a revised version of a paper presented at the Kroeber Anthropological Society meetings, June 1972. I am indebted to John Atkins, Brent Berlin, William Geoghegan, and Paul Kay for their assistance, to the National Institute of Mental Health and the Latin American Studies Center, Berkeley, California, for financial assistance, and especially to my Tenejapan assistants, *hȼiak mentes ȼaʔpat* and *hšoš kusman ȼ'uhkin*. The data cited were collected between April and December 1971 in the conservative Tzeltal-speaking *municipio* of Tenejapa, located approximately 15 km NE of San Cristóbal de las Casas in the central highlands of Chiapas, México. The Tzeltal orthography follows Berlin (1962).

[2]The query was as follows: *mut bal te soȼ'e* 'Is the bat a bird?' A variety of animal names may be substituted for the first and last terms of the Tzeltal question. A question frame useful for eliciting information about defining attributes of Tzeltal taxa is, *bi yuʔun la sbilin mut* 'Why is it named bird?' Other terms may be substituted for *mut* 'bird.' A typical response to the question just cited is, "They are called birds because they have feathers, two wings, two legs, and they fly on their two wings."

[3]Berlin, Breedlove, and Raven (1968) describe techniques for discovering covert or unnamed taxa. However, Berlin (1972) argues that such taxa are typically found only on the level of "intermediate" taxa or the "unique beginner." "Generic" taxa are defined primarily by reference to the form of the names applied to them (Berlin 1972:52-58). Thus a "covert generic" is something of an anomaly. However, the three subordinate taxa, *yašal č'o, sakʔeal č'o*, and *k'alel č'o*, appear to be defined by reference to a single attribute. They do not, therefore, exhibit the multidimensional character which may be the psychological hallmark of generic taxa. The superordinate taxon, *č'o*, includes several obviously generic taxa and thus cannot be considered the generic.

[4]The maximal value of D' is determined by the number of levels which may occur in a particular reduced scientific taxonomy. Though upwards of twenty levels may be named in the total zoological taxonomy, it is unlikely that more than ten would remain after the reduction rules are applied. Such a maximal value might obtain if one were to define a taxon as the set of all organisms encountered on a given day.

[5]According to a recent taxonomic revision (James Patton, personal communication), a single species, *Sciurus aureogaster*, includes all organisms referred to as *yašal čuč*. This folk taxon is assigned a degree of 1 here, because Hall and Kelson (1959) recognize two species referrable to that folk taxon. *polisía ȼ'iʔ* and *merikáno wakaš* are assigned a degree of zero since they obviously refer to well defined breeds. It is likely that further research might justify assigning degree zero to several other such taxa. Though breeds find no place in official zoological classifications, they are logically and phenotypically, if not evolutionarily, comparable to scientific taxa. *pay* 'skunk' is assigned a zero degree despite the fact that Hall and Kelson (1959) do not cite the subfamily Mephitinae. This taxon is widely recognized. Horses and mules are treated as covert specifics. Neither taxon is named, but the two sexes of each are named, and the Tzeltal are well aware that the mule, male or female, is a hybrid offspring of horse and donkey.

[6]In Kay's formal treatment, a taxonomy is defined as "a ternary relational structure" which involves a taxonomic structure, a set of names, and a mapping of those names to the set of taxa of the taxonomic structure. Thus I use the term "taxonomy" loosely here to refer either to a taxonomic structure or to the set of taxa which is one element of a taxonomic structure (Kay 1971:884, Def. 31).

[7]I find it necessary to assume that a folk biological taxonomy, or some definable sub-system of such a taxonomy, does not include non-organisms, e.g., spirit beings. Though I find no inevitable exceptions to this rule in the Tzeltal data, it is possible that exceptions may occur. The measure suggested here is well defined only with reference to a common set of organisms classified by both folk and scientist. The asymmetry of the measure takes into account the fact that there are many organisms which the scientist classifies but the folk do not. Those organisms do not exist for a given folk. This measure, however, presents no solution to the possibility that there may be "organisms" which exist for a given folk but not for the scientist.

[8]Lo(S), Tr(S), Co(S), and Te(S) define sets. The first element of each symbol, e.g., *Lo*, defines an *operator* which, in the case of Lo(S), maps from the set S to the set Lo(S) or to the null set. The reduced scientific taxonomy is defined as the product of these four operators (Definition 12). It might as well be defined as the set intersection of the four sets generated respectively by the four operators.

[9]I would like to thank Paul Kay for suggesting the idea of a normalized weighted measure.

references cited

Alvarez del Toro, Miguel
1952 Los Animales Silvestres de Chiapas. Tuxtla Gutiérrez: Instituto de Ciencias y Artes de Chiapas.

Berlin, Brent
1962 Esbozo de la Fonología del Tzeltal de Tenejapa, Chiapas. Estudios de Cultural Maya 2:17-36.
1972 Speculations on the Growth of Ethnobotanical Nomenclature. Language in Society 1:51-86.
1973 Folk Systematics in Relation to Biological Classification and Nomenclature. Annual Review of Ecology and Systematics 4:259-271.
Berlin, Brent, Dennis E. Breedlove, and Peter H. Raven
1966 Folk Taxonomies and Biological Classification. Science 154:273-275.
1968 Covert Categories and Folk Taxonomies. American Anthropologist 70:290-299.
1973 General Principles of Classification and Nomenclature in Folk Biology. American Anthropologist 75:214-242.
1974 Principles of Tzeltal Plant Classification: An Introduction to the Botanical Ethnography of a Mayan-Speaking People of Highland Chiapas. New York: Seminar Press.
Berlin, Brent, and Paul Kay
1969 Basic Color Terms: Their Universality and Evolution. Berkeley: University of California Press.
Bulmer, Ralph N. H.
1970 Which Came First, The Chicken or the Egghead? *In* Echanges et Communications, Mélanges offerts à Claude Lévi-Strauss à l'occasion de son 60ème anniversaire. Jean Pouillon and Pierre Maranda, Eds. The Hague: Mouton. pp. 1069-1091.
Douglas, Mary
1966 Purity and Danger. London: Routledge and Kegan Paul.
Farris, James S.
1971 The Hypothesis of Nonspecificity and Taxonomic Congruence. Annual Review of Ecology and Systematics 2:277-302.
Goodenough, Ward H.
1956 Residence Rules. Southwestern Journal of Anthropology 12:22-37.
Gregg, John R.
1967 Finite Linnaean Structures. Bulletin of Mathematical Biophysics 29:191-206.
Hall, E. Raymond, and Keith R. Kelson
1959 The Mammals of North America. Volumes 1 and 2. New York: The Ronald Press.
Hunn, Eugene S.
1973 Tzeltal Folk Zoology: The Classification of Discontinuities in Nature. Ph.D. dissertation. University of California, Berkeley.
n.d. Cognitive Processes in Folk Biology: The Identification of Gulls. Unpublished manuscript. Language-Behavior Research Laboratory, University of California, Berkeley.
Kay, Paul
1971 Taxonomy and Semantic Contrast. Language 47:866-887.
Mayr, Ernst
1970 Populations, Species, and Evolution. Cambridge, MA: Belknap Press. *An abridgement of* Animal Species and Evolution, 1963.
Sapir, Edward
1921 Language: An Introduction to the Study of Speech. New York: Harcourt, Brace.

Date of Submission: December 14, 1973
Date of Acceptance: March 24, 1974