The Origins of Taxonomy

A review of its historical development shows why taxonomy is unable to do what we expect of it.

Peter H. Raven, Brent Berlin, Dennis E. Breedlove

By current taxonomic standards, there are probably about 10 million species of organisms in the world, of which we have in the past 218 years described, at some level, 10 to 15 percent. For more than 99 percent of the described species, we know nothing more than a few morphological facts and one to several localities where they occur. The human population of the world, currently 3.7 billion, is growing at a rate which, if maintained, would lead to a doubling of the present size in 35 years. Yet it is far from certain that the world can support even present population levels indefinitely. Pollution on a world scale is increasing so rapidly that organisms are undoubtedly already becoming extinct at a high rate. For example, tens or even hundreds of thousands of kinds of new synthetic molecules are being dumped into the sea continuously; in almost all instances, their effects are unknown.

In view of the available taxonomic manpower and the enormous rate of extinction that will characterize the next century, it is doubtful that even 5 percent more of the world's organisms can be added to our inventory before the remaining 80 percent becomes extinct. Those familiar with tropical environments, such as Paul Richards (*1*), predict that there will be *no* undisturbed tropical rain forest anywhere in the world by the end of the present century. The tropical rain forest is the richest area in the world in species of plants and animals, and by far the most poorly known. The vast majority of the estimated million species of mites, of the quarter million to million species of ichneumonid wasps, and the unknown numbers of flies will never be seen by a taxonomist, to say nothing of even more obscure groups such as the Rickettsiae, which are of very high economic importance.

In this article, we shall attempt to show why our present taxonomic system is, in the face of the job for which it has responsibility, inadequate. Being basically a Renaissance codification of folk taxonomic principles made on the implicit assumption that the number of organisms to be dealt with would perhaps be 25,000 to 50,000, it is incapable of doing what we expect of it. Before our inventory of world organisms can be conducted in a way that is truly meaningful and productive for scientific advance, we need to find new standards for recording information about organisms in a readily retrievable form.

Development of Taxonomy

Man is by nature a classifying animal. His continued existence depends on his ability to recognize similarities and differences between objects and events in his physical universe and to make known these similarities and differences linguistically. Indeed, the very development of the human mind seems to have been closely related to the perception of discontinuities in nature. In view of this, the study of folk taxonomic systems, which has received a great deal of interest in recent years, has a high significance in interpreting the logical processes going on in our own minds, as well as in understanding the application and utility of the taxonomic systems themselves.

We have (*2*) reviewed evidence suggesting that there are a number of general characteristics common to the folk taxonomic systems that have been studied in peoples as diverse as the Tzeltal-speaking Mayans of southern Mexico, the Hanunóo of the Philippines, the Cantonese-speaking boat people of Hong Kong, the Guaraní of Argentina, the Navajo, and many others. The principles common to all these folk taxonomic systems, and to others that have been studied in sufficient depth, are as follows.

1) In all languages, recognition is given to naturally occurring groupings of organisms. These groupings appear to be treated as psychologically discontinuous units in nature and are easily recognizable. They will be referred to here as *taxa*.

2) These taxa are further grouped into a small number of classes known as taxonomic ethnobiological categories. These categories, definable in terms of linguistic and taxonomic criteria, seem to number five: *unique beginner, life form, generic, specific, varietal.*

3) The five taxonomic ethnobiological categories are arranged hierarchically, and taxa assigned to each rank are mutually exclusive.

4) The taxon found as a member of the category *unique beginner* is often not labeled linguistically by a single expression; that is, the most inclusive taxon, for example, *plant, animal,* is rarely named.

5) Taxa that are members of the category *life form* are invariably few in number, ranging from five to ten, and these include a majority of all named taxa of lesser rank.

6) In most folk taxonomies, taxa that are members of the category *generic* (*3*) are more numerous than life form taxa, but are nonetheless finite in number, usually about 500. Some particularly aberrant generic taxa—for example, cacti, pineapple, cassowary, pangolin, platypus—or those that are of great economic importance and interest may be unaffiliated; that is, they are not included in one of the life form taxa.

7) Specific and varietal taxa are, in general, less numerous than generics. Characteristically, they exist in sets of few members within a single generic. Sets of more than two members tend to refer to organisms of major cultural importance, and sets of 20 or more members inevitably do. Specific and varietal taxa can be recognized linguistically in that they are commonly labeled in a binomial or trinomial format that includes the name of the generic or specific to which they belong.

8) *Intermediate* taxa are those that are immediately included in one of the

Dr. Raven is director of the Missouri Botanical Garden and professor of biology at Washington University, St. Louis. Dr. Berlin is associate professor of anthropology, at the University of California, Berkeley. Dr. Breedlove is assistant curator of botany at the California Academy of Sciences, San Francisco.

major life form taxa and that immediately include taxa of generic rank. They are invariably rare in natural folk taxonomies, and, when evidence has been presented that unambiguously demonstrates their existence (4), they are not linguistically labeled. We have referred to them as *covert categories*.

Equivalence of Folk and Early Published Taxonomies

Early biologists, such as Theophrastus, and later the herbalists merely wrote down folk taxonomic systems sharing completely the characteristics just enumerated. Both systems included relatively few specifics and varietals in relation to the number of generics. As early writers had no ready means of duplicating and distributing their works, they still confined themselves to a limited number of generics in each domain. Generic plant names in languages that have been fairly well studied range from 250 to 800 forms, regardless of the richness of the environment in which the people live. Animal names fall in the same range for generics (5). In all cases, including early published taxonomic works, the size of the basic set of generics appears to be controlled by the number of categories into which the known organisms can be divided in a culturally significant fashion. We have shown earlier that organisms of high cultural significance are apt to be subdivided into more categories, in terms of Western biological nomenclature, than are those of lower cultural significance (6). In an operational system, especially one that is strictly verbal, the number of names cannot be multiplied beyond meaningful limits; many speakers of the language must be familiar with each name that is passed down from generation to generation as part of that language. As there are more and more names, the names become less and less useful.

The lack of names for intermediate categories in folk and early written taxonomies is probably related to these same influences on language. For a given name to be assigned to a category including two or more generics, this intermediate category would have to be one that was often a subject of discussion and a point of reference. This is, however, unlikely because, by the very nature of a folk taxonomy, all of the culturally significant properties of the organisms included in the system are known to the user. Given the contexts in which organisms are discussed and classified, there is generally nothing to be added to this knowledge by grouping the organisms into more inclusive categories. Indeed, if there is a finite number of generic and other names for organisms that can be transmitted verbally in a language, intermediate categories would have to be of very great interest indeed to receive explicit labels.

Linnaean Taxonomy and Its Predecessors

With the invention of movable type in the mid-15th century, men began to aspire to more and more comprehensive listings of the kinds of organisms. It began to seem worth while to attempt to describe and name the many kinds of plants and animals being found in this age of exploration, although the biological dividends were at first modest. As travel to distant lands became easier and easier, biologists began to realize that the plants and animals of different regions often differed greatly. Thus, a herbal written by Dioscorides in Greece was not apt to provide an adequate or useful description of the plants of central Germany some 15 centuries later. The emphasis was still on generic names, and, by the close of the 17th century, Tournefort, in the face of rapidly accumulating numbers of species, had stabilized the concept of the genus as central for biological taxonomy. He defined 698 genera (7), a number still consistent with the number of generics recognized in folk taxonomic systems (5). Tournefort reviewed the genera that had been proposed and tried to put them in a logical system (8).

The situation was much the same at the time of Linnaeus; thus Cain (9) has pointed out that this 18th-century taxonomist founded his binomial system on the conception of a limited number of memorizable genera. Linnaeus felt that the genera must be separate, distinct units with distinct names, and natural so that all would agree on their limits. But Linnaeus had to deal with more genera than Tournefort. By 1737 he included 935 in his *Genera Plantarum* (10), and in all of the editions of the work and its supplements he presented diagnoses for 1336. The age of exploration had resulted in the discovery of an uncomfortably large number of genera, as defined by the logical principles of Tournefort. Nevertheless, Linnaeus recognized almost no named taxa above the generic level, even though he perceived and discussed many such groupings. He presumably saw no need to give them names, since he still considered his genera to be limited in number and memorizable.

One interesting similarity between the taxonomy of Linnaeus and the folk taxonomic systems on which it was based has been pointed out by Walters (11). In folk taxonomies, organisms of high cultural significance are apt to be subdivided into more generics, in terms of standard biological nomenclature, than are those of lower cultural significance (6). In just the same way, Linnaeus recognized many genera, each with few species, in plant families such as Brassicaceae and Apiaceae—plants of high cultural utility that already had Greek or Latin generics associated with them. In contrast, he named few genera, each with many species, in families such as Cyperaceae that were of low cultural utility and that had not been treated extensively by his predecessors. These historical trends are still reflected in the level at which genera are recognized in the respective families at present.

In broad outlines, then, the system of Linnaeus was a codification of the folk taxonomy of a particular area of Europe, as has been stressed by Walters (12). It differed in principle from the unwritten folk taxonomy of the Tzeltal speakers of Chiapas, Mexico, only in the somewhat larger number of generics that was included and in the vastly greater number of specifics into which they were divided. Both of these trends were facilitated by the invention of movable type some 300 years earlier and made possible psychologically by the expectation that the works would be widely distributed and useful.

But Linnaeus had to concern himself with relatively few kinds of plants and animals. Had he even suspected the existence of the tens of thousands of genera to be discovered in the tropics and other poorly known portions of the world in the 19th and 20th centuries, he might not even have attempted the synthesis of plant and animal classification that he achieved so well. In an effort to solve this difficulty, various post-Linnaean authors built up the curiously deep and cumbersome taxonomic hierarchy that is characteristic of modern classification schemes. They added and named categories such as family, order, and phylum, and hoped that these at least, being few in number, would be more comprehensible than genera.

Some of the better marked and larger families of flowering plants had been recognized as distinct units for many centuries, but the first synoptical treatment of the families of plants was presented by de Jussieu in 1789 (*13*). De Jussieu recognized exactly 100 families of plants, a number that could scarcely have resulted from anything but a conscious effort to produce a strictly limited number of memorizable units (*11*). He was trying to solve the dilemma to which the burgeoning numbers of genera diagnosed by Linnaeus, who had died in 1778, had given rise. Many of his families were divided subsequently as more came to be known of the plants, and in 1920 Gundersen (*14*) was able to write casually, "The number of families of vascular plants is generally considered as about 300." A few years later, Hutchinson (*15*) pointed out that the delimitation of families is "very much a matter of taste and personal idiosyncrasy" and to advocate the recognition of about 400 for the flowering plants. How curiously these numbers resemble the numbers of generics recognized in a very diverse sample of folk taxonomies (*7*)!

Problems for Modern Taxonomy

A folk taxonomic system is designed, not for information retrieval, but for communicating about organisms with those who already understand the nature of the organisms being discussed. These organisms and their culturally significant features are part of the active ethnobiological knowledge of most adult speakers of the language. It would therefore be meaningless to ask a Tzeltal speaker what properties the taxa indicated by his set of generic names have in common. Only when the classificatory system is extended to hundreds of thousands of poorly known organisms do we begin to ask for a "definition" of genera and species. Confronted with this difficulty, the human mind is all too ready to accept spurious generalities such as the "biological species concept" or the earlier assumption, current around the year 1900, that whole groups of related plants were merely environmental modifications of one another. It would be comforting to find that evolution had produced a series of identical units about which we know a great deal. We know in fact that it has not, but are still loathe to give up our generalities, which we hope will help us understand the 10^7 kinds of organisms estimated to occur in the world.

In other words, the taxonomic system we use appears to communicate a great deal about the organism being discussed, whereas in fact it communicates only a little. Since, in the vast majority of instances, only the describer has seen the named organism, no one with whom he is communicating shares his understanding of it. The basis of communication present in folk taxonomic systems dealing with a limited number of organisms is lacking, although we still believe implicitly in its existence. Our system of names appears to achieve a reality which it does not in fact possess.

If we are truly interested in recording information about organisms in a retrievable form, then we need a system in which the points of reference are stable and the observations are recorded in a standard format. When a species is transferred from one genus to another, its name changes with serious loss to the information retrieval capabilities of the system as a whole. Yet such shifts in generic position are one of a few mechanisms we have available to register our changing appraisal of the relationships of a particular species. In fact, the current taxonomic system is hopelessly inadequate as an information retrieval device, and it must be supplanted with one allowing the characteristics of organisms to be handled and retrieved in a much more efficient manner.

The Future

We have as yet named only about 15 percent of the world's organisms and have no real chance of adding many to the total before the rest become extinct. Despite this, we cling to the naive view of Renaissance man and assume that the extension of those folk principles deeply rooted in our collective psychology is the only appropriate way of dealing with this diversity. We implicitly assume that we know as much about a mite from the Amazon Basin as we do about the mallard duck, and reflect this assumption in our presumably scientific but actually folk system of naming the two kinds of organisms. We continue to conceal or lose most of our systematic data when we make taxonomic decisions about organisms, since we tend to focus on the process of classification rather than on the information being gathered.

This general line of reasoning leads to important considerations for the planning of future research. For example, it is often argued that, if we know about the systematics of a particular group of organisms in detail, we will be better able to utilize them in biological control programs and the like. This assumption has almost never been realized in fact. What we have achieved in biological control, pathology, and allied fields has been almost entirely the result of ad hoc studies of the problems when they become of interest and direct application of the results of these studies.

Taxonomic work has helped us only to a limited extent in understanding the functioning of ecosystems, a problem that is of crucial importance for human survival. Knowledge gained about a particular kind of organism outside of the ecosystem being studied will often be totally inaccurate for the area we are trying to understand and, thus, misleading, even if it is available. At best, equivalencies in the naming of organisms from different ecosystems may provide a rough idea of what to expect, and no more. Clearly, if we are to advance more rapidly in this vital field, new taxonomic methodology seems to be needed.

Completing a world survey of all organisms is patently impossible; moreover, it has not been demonstrated to be of more than a very limited practical, and of virtually no theoretical, importance. It might be possible and desirable to complete the world survey of flowering plants, butterflies, or fleas; but it is clearly out of the question for many other groups and should not be taken as a priori justification for the funding of any proposed scientific work. When surveys of groups, particularly tropical groups, are undertaken, the appropriate depth of the survey proposed should always be taken into account. Perhaps blocking out families or genera would be adequate for many groups; there may be no need whatever to describe and give names to the species. The "stockpiling" of specimens, especially from poorly known areas of the world, continues to appear worth while, since, although we cannot guess to what uses these specimens might be put in 200 years, we can be certain that most of the species involved will be extinct.

High priority should also be given (i) to taxonomic work that utilizes "unusual" characters or a broad spectrum of characters; (ii) to the accumulation of information about organisms which does not seem to have direct taxonomic applicability; and (iii) to the search for original ways of looking at the structure of nature, including new methods of presenting "taxonomic" information.

One of the most significant trends in modern systematics has been the development of electronic data processing equipment. When information about organisms is entered into a data bank, it can be summarized in any way that might be desired, including the construction of taxonomic systems. The information that goes into the construction of these systems need not be lost, however, once the systems have been constructed; it is still available for retrieval or recombination in other ways. We can continue to have direct access to the actual information that is available about a particular kind of organism and need not generalize so crudely from the fact that it happens to have been considered a species by one or more taxonomists. Since the discontinuities in the biological universe were the great proving ground for the human mind, which seems to have developed largely as a device for dealing with them effectively, we cannot predict what we may learn when we bring these discontinuities into sharper focus with the aid of modern technology.

Summary

There are approximately 10 million kinds of organisms in the world, of which we have described some 15 percent. The rapid growth of the human population will cause most of the remainder to disappear from the earth before they are seen by a taxonomist. These facts suggest a more rigorous application of priorities in systematic biology as well as a careful review of the principles upon which our taxonomic system is based.

Folk taxonomies all over the world are shallow hierarchically and comprise a strictly limited number of generic taxa ranging from about 250 to 800 forms applied to plants and a similar number applied to animals. These numbers are consistent, regardless of the richness of the environment in which the particular people live. Very few specific and varietal taxa are recognized in folk taxonomic systems. Until the invention of movable type in the mid-15th century, written taxonomies were simply records of the folk taxonomies of particular regions. Subsequently, with the possibility for the wide distribution of books, it began to seem worth while to attempt to describe and name all species of plants and animals in the world. By the year 1700, 698 genera of plants were recognized; and by the year 1778, some 1350 genera, including tens of thousands of species. In 1789 de Jussieu interpolated the family as a higher level taxonomic category in an attempt to reduce the number of important units in the system to a memorable number. The family is still the focal point in systems of angiosperm classification at present, several hundred families being recognized.

Problems with the taxonomic system stem largely from the fact that it is not designed as an information retrieval device. In folk taxonomies, names are given to organisms and these are used to communicate about the organisms with others who already know the culturally significant properties of the organisms being discussed. In dealing with the vast numbers of organisms that exist, we tend to overemphasize the process of classification and the decisions it involves at the expense of the information about the organisms that we are supposedly accumulating. Frequent changes in names exacerbate the difficulties of the system and render it still less useful for information retrieval.

With modern electronic data processing equipment, it has become possible to record information about organisms, to retain this information in a data bank, and to utilize it for various purposes, including the construction of various taxonomic systems. The invention of high-speed electronic data processing equipment is seen as analogous to but more important than the invention of movable type in the history of systematic biology. By using such equipment to its full potentialities, we should be able to achieve a qualitative improvement in our perception of the living world.

References and Notes

1. P. W. Richards, *The Tropical Rain Forest* (Cambridge Univ. Press, 1952).
2. B. Berlin, D. E. Breedlove, P. H. Raven, *Amer. Anthropol.*, in press.
3. H. H. Bartlett [*Bull. Torrey Bot. Club* **67**, 349 (1940)] seems to have been the first to point out that the concept of the genus had priority to that of the species in folk taxonomy, a position amply supported by subsequent studies.
4. B. Berlin, D. E. Breedlove, P. H. Raven, *Amer. Anthropol.* **70**, 290 (1968).
5. Generic plant names for languages that have been fairly well studied range from 250 to 800 forms; for example, Tzeltal 475, Hanunóo 822, Cantonese 267, Navajo 306, Seminole 250, Hopi 350. Animal names fall in the same range for generic classes; for example, Tzeltal 311, Hanunóo 461, Tzotzil 233, Navajo, insects only, 105, Maya 281 (see *2*).
6. B. Berlin, D. E. Breedlove, P. H. Raven, *Science* **154**, 273 (1966).
7. J. P. de Tournefort, *Institutiones Rei Herbariae* (Paris, 1700).
8. W. T. Stearn, Introduction to facsimile reprint of C. Linnaeus, *Genera Plantarum* (1754). *Historiae Naturalis Classica* **3**, v (1960).
9. A. J. Cain, *Proc. Linnean Soc. London* **169**, 144 (1958).
10. C. Linnaeus, *Genera Plantarum* (Uppsala, ed. 1, 1737).
11. S. M. Walters, *New Phytol.* **60**, 74 (1961).
12. ———, *Preslia* **34**, 207 (1962).
13. A. L. de Jussieu [*Genera Plantarum* (Paris, 1789)] recognized exactly 100 families in the first synoptical use of this category.
14. A. Gundersen, *New Phytol.* **19**, 264 (1920).
15. J. Hutchinson, *The Families of Flowering Plants* (Macmillan, New York, 1926), vol. 2.
16. We are grateful to the National Science Foundation for support through grants GS 383, 1183, 2280, and GB 7949X. We thank Richard S. Cowan, Richard W. Holm, Roy L. Taylor, and S. M. Walters for their comments on this article.