A Taxonomic Problem

By J. S. L. Gilmour

AMONG present-day biologists there is widespread interest in, and considerable disagreement on, the proper relationship between taxonomy* and other branches of their science, especially genetics, cytology and ecology. Should the data provided by these branches be incorporated into the existing taxonomic categories, or should new categories be created to meet their needs ?

In discussing these questions, I believe that the purpose and general principles underlying the process of classification have been unduly neglected and that a consideration of these would help materially to clarify the issues involved. A full discussion of these principles may be found in the standard works on logic and scientific method†, and only those points most relevant to the present problem will be dealt with here.

In its simplest terms, classification consists in grouping individual objects (and, ultimately, sense data, expressed as qualities and relations) into classes, so that all the individuals in one class have certain attributes in common. The use of such general names as book, piano, mountain, etc., involves a process of classification no less than the formation of such classes as mammals or flowering plants. Viewed thus, classification is seen to be a necessary stage in the inductive process by which the human mind obtains an ordered knowledge of the universe. Ritchie (op. cit., p. 79) has expressed this concept of classification as follows : ". . . our general knowledge of the external world as expressed by laws of nature is a product of the interaction of two processes, selection from and classification of what is experienced and the discovery of laws relating to the classes".

Any given collection of objects can, of course, be classified in a great number of different ways, depending on the particular attributes chosen as a basis for classification. Thus mankind can be classified on a basis of nationality into Frenchmen, Germans, etc. ; on a basis of occupation into lawyers, doctors, etc. ; on a basis of physical characters into the various so-called 'races', and so on. Further, the choice of particular attributes depends on the *purpose* in view in making the classification. Thus if we wish to investigate the distribution of professions among the inhabitants of different countries, we employ classifications based on the attributes of occupation and nationality, and, from an examination of the resulting classes, we formulate laws relating them, such as, for example, that there is a greater proportion of clergymen in England than in the U.S.S.R.

These considerations apply also to biological classification in general. Fundamentally, the purpose of biological classification is the acquisition of ordered knowledge regarding living things, and logically any grouping of plants and animals should be considered a taxonomic process. Such groups as hydrophytes, succulents, hemicryptophytes, ruderals and alpines are, in this sense, 'taxonomic' groups, no less than the families, genera, species, etc., of "traditional' systematic botany. Why is it, then, that the term taxonomy is usually confined to this latter grouping ? Why are Ranunculaceæ, but not annuals, considered a taxonomic group ? The answer to this question involves a consideration of the difference between a natural and an artificial system of classification.

The term 'natural classification' is a general one, the significance of which is not confined to the grouping of living things. It is usually stated in logic that a system of classification is the more natural the more propositions there are that can be made regarding its constituent classes. For example, a classification of mankind on a basis of nationality is more natural than one based on the initial letter of surnames, because more propositions can be made regarding an Englishman (for example, that he probably speaks English, knows "God Save the King", has a white skin, etc.) than about a man whose name begins with E. Thus a natural classification is one founded on attributes which have a number of other attributes correlated with them, while in an artificial classification such correlation is reduced to a minimum. From the point of view of function, a natural classification can be used for a great variety of purposes, while an artificial one serves only the limited purpose for which it was constructed.

The difference between them can be expressed in another way. The greater number of propositions

that can be made regarding an Englishman depends on the fact that nationality (the basis of the classification) is an attribute connected with a very important factor in human life (namely, the habit of living in territorial groups to which we become emotionally attached), while the initial letter of surnames is not. This point is made clearer if we imagine that a dictator has decided to execute forthwith all those whose names begin with E. This attribute would then become connected with another important factor in human life (namely, the power of a dictator) and immediately the hitherto artificial group of E's would become, at any rate temporarily, a natural one, about which many propositions could be made (for example, that they would soon be dead, were probably frightened, were probably making their wills, etc.). From this rather trivial example, the proposition emerges that a natural classification is one which is based on attributes connected with a factor that has an important influence on the objects being classified : the more important the factor the more natural the classification becomes.

The difference between the two is therefore one of degree only. Indeed, it is impossible to construct a classification that is not to some extent natural. As pointed out by Jevons (*op. cit.*, p. 680), even in so artificial a class as those whose names begin with E, there will probably be an abnormally high proportion of Welshman owing to the presence of a large number of Evanses.

If there is one factor influencing a particular collection of objects which is more powerful than any other, then a classification based on attributes connected with that factor will be more natural than any other. For example, in classifying a number of motor-cars which were manufactured during the last forty years, the most natural classification would be a division based on attributes connected with the period of manufacture, because that is the most important factor determining the attributes of the cars. If, however, there are several factors of nearly equal importance influencing a group of objects, then a number of equally natural classifications is possible. This is the case in the classification of mankind, where, for example, it is difficult to say whether a basis of nationality or annual income gives a more natural classification.

If these conclusions are applied to biological classification as a whole, the reason for the unique position held by traditional taxonomy begins to emerge.

Owing to the method of reproduction and evolution of living things, involving the inheritance of parental characteristics, and to the pre-eminent influence that these factors exert on the attributes of plants and animals, the possibility exists of constructing a series of classifications which are more natural than any others, namely, those based on inherited characters. It is by a constant search for the most natural among these classifications that the present vast structure of traditional taxonomy has been gradually built up. We may say, then, that the unique position held by traditional taxonomy, marking it off from other arrangement of living things, is primarily due to the possibility of constructing a set of peculiarly natural classifications of living things based on inherited attributes.

On this view, a natural classification in biology is a particular example of natural classification in general and not a phenomenon peculiar to living things ; further, a natural classification differs in degree only from other biological classifications. This is in contrast to the concept, widely (though not universally) accepted in post-Darwinian times, of a natural classification as being based on *actual phylogenetic relationship* and thus differing fundamentally from artificial classifications. This view involves the assumption that there exists a criterion of degree of phylogenetic relationship, apart from degree of similarity of attributes, and either that such a criterion should be used in constructing a natural classification, or that degree of similarity of attributes is an invariable indication of such relationship and can therefore be used as a basis for a natural classification.

Now, this criterion of phylogenetic relationship must presumably be the same that leads us to say that two brothers are more nearly related than two cousins—namely, the degree of remoteness of a common ancestor. But in order to judge degree of relationship on this criterion, a knowledge of the genealogy of every *individual* in any group under consideration is essential, and such knowledge, except occasionally in man and in certain domesticated animals and plants, is clearly unobtainable.

Can we, then, fall back on the second alternative and accept degree of similarity of attributes as an invariable indication of degree of relationship ? That such similarity is frequently an indication of relationship is true, but that it is an invariable indication certainly cannot be maintained. For example, two individuals related as cousins may have identical genotypic constitutions, while two sibs may be strikingly different and, further, parallel mutations may occur in two individuals only distantly related phylogenetically but of similar genotypic constitution.

If, then, the criterion of degree of relationship applicable to *individuals* cannot be applied in practice to *groups*, and, further, if similarity of attributes cannot be regarded as a certain indication of such relationship, then it would seem that

the view that a natural classification differs fundamentally from artificial classifications in that it represents phylogenetic relationship cannot be substantiated. A natural classification should rather be regarded, first and foremost, as that arrangement of living things which enables the greatest number of inductive statements to be made regarding its constituent groups, and which is therefore the most generally useful classification for the investigation of living things. Whether or not such a classification does in fact group together individuals who are phylogenetically related is a secondary question which must be answered for any particular case on its merits.

For special purposes, however, additional classifications are essential. As emphasized above, classification is a stage in inductive investigation, and in so far as it is desired to discover laws relating certain attributes of living things, such attributes must be embodied in distinct classifications of greater or less complexity. The need for such additional classifications is obvious when the attributes concerned give an arrangement conspicuously less natural than that of traditional taxonomy, for example, a classification into trees, shrubs, herbs, etc., or on medicinal properties. But when the classification is only slightly less natural than the most natural possible, then its validity becomes obscured and confusion arises between it and traditional taxonomy. Such is the case with many of the attributes that have been investigated during the last thirty or forty years under the headings genetics, cytology, physiology and ecology. For example, a traditional taxonomist may divide a genus into a certain number of species on morphological characters, the result being a good natural grouping. A cytologist may then investigate the same genus and find that, say, sterility barriers in some cases cut right across the taxonomist's groups. If the sterility classification is the less natural of the two, that is, if its groups show a smaller correlation of attributes, it should not, as is sometimes suggested, be dismissed as 'taxonomically useless'; but should be retained as a distinct classification for the purpose of establishing the relationship between sterility and other attributes.

This principle of 'multiple classification' is fundamental to the proper functioning of taxonomy as an instrument for the inductive investigation of living things. How can it be most satisfactorily applied in the present position of biological taxonomy? A few tentative suggestions may be put forward.

The categories and nomenclature of traditional taxonomy should be confined to the most natural classification possible in the existing state of knowledge of any particular group, on whatever attributes it may be based, and such a classification would be the most generally useful for a great variety of purposes, both scientific and non-scientific.

Any exact definitions of the categories within this classification are, in my opinion, an impossibility. The categories, genus, species, etc., are of the same nature as such categories as herd or heap, individual *characters* taking the place of individual animals or stones. A well-known trick in logic known as the 'sorites' illustrates this point. The question is put "Does one stone form a heap?" If the respondent answer "No", it is asked, "Do two stones form a heap?" and so on. Great difficulty is found in deciding when the addition of one stone constitutes a heap. The answer would vary from person to person and also according to the shape and size of the stones. Similarly, in a natural classification, the definition of the categories must vary somewhat according to the mentality of individual biologists and the nature of the material being classified.

If the traditional taxonomy is defined in this way, what should be the aim and methods of the additional classifications necessary for special investigations? Unlike the classification of traditional taxonomy, they should be based on the same attributes throughout, and their categories should have a different terminology from that of traditional taxonomy. A good example of such a system is Danser's classification into commiscuum, comparium and convivium, which is based purely on interfertility criteria[1].

Many other classifications based on data provided by genetics, cytology and ecology have been proposed during this century[2], but these have for the most part either redefined the categories of traditional taxonomy in such a way as to destroy their general usefulness, or have employed mixed attributes as a basis for new categories. It is greatly to be desired that agreement should be reached on the employment of other classifications similar to Danser's (either new ones, or existing ones modified where necessary), based on other attributes. The number and complexity of such classifications are, of course, questions of meeting the ever-changing needs of biological investigation, and no limits either way can be set.

With regard to nomenclatural technique, it would seem possible that, where necessary, the method of employing Latin words with a prefix or suffix to distinguish them from the Latin names of genera, species, etc., which has already been used with success, for example, for the ecological categories of association, etc., might be extended to classifications based on other attributes.

[1] *Genetica*, **11**, 399 (1920).
[2] See Du Rietz, *Svensk. Bot. Tid.*, **24**, 333 (1930).