



## The Biological Species Concept: A Critical Evaluation

Robert R. Sokal; Theodore J. Crovello

*The American Naturalist*, Vol. 104, No. 936. (Mar. - Apr., 1970), pp. 127-153.

Stable URL:

<http://links.jstor.org/sici?sici=0003-0147%28197003%2F04%29104%3A936%3C127%3ATBSCAC%3E2.0.CO%3B2-%23>

*The American Naturalist* is currently published by The University of Chicago Press.

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/ucpress.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

---

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

THE BIOLOGICAL SPECIES CONCEPT:  
A CRITICAL EVALUATION\*

ROBERT R. SOKAL AND THEODORE J. CROVELLO

Department of Biological Sciences, State University of New York at Stony Brook, and  
Department of Biology, University of Notre Dame

I. INTRODUCTION

A species concept has been a central tenet of biological belief since the early origins of biology as a science. The implications of this term have changed over the years: the fixed, immutable, and sharply distinct entities of the Linnaean period gave way to the more variable and intergrading units of the post-Darwinian era. For many taxonomists before and after Darwin, the species has simply implied the recognition of groups of morphologically similar individuals that differ from other such groups.

Through much of biological history there has been controversy regarding the existence of species in nature. Are species real units in nature? Can the species category be defined objectively? Given an affirmative answer to the above two questions, can real organisms be assigned to one of the nonoverlapping species so delimited? Darwin's work contributed to the recognition of species as real entities. The very title of his book, *On the Origin of Species*, stressed this category. But as Mayr (1959) has pointed out, Darwin himself was so impressed by the variability and intergradation in the material he studied that he considered the term "species" to be arbitrary, not differing in essential features from "variety." Argument regarding these questions has persisted through changing concepts of the biological universe and with increasing insights into the genetic and ecological mechanisms governing the behavior of individuals and populations. The history of these ideas and controversies is reviewed by Mayr (1957), and we shall not enlarge upon it here. Some have considered species as man-made, arbitrary units either because of their philosophical orientation or because of the difficulty of interpreting variable material from widely ranging organisms as consisting of one or more species. These arguments have been countered by evidence of the common-sense recognition of discontinuities in nature even by lay observers (see Mayr [1963, p. 17] for an account of species recognition by New Guinea natives, but see Berlin, Breedlove, and Raven [1966] for a contrary view) and also of species recognition, pre-

\* Contribution no. 3 from the Program in Ecology and Evolution, State University of New York at Stony Brook. This manuscript was submitted to the previous editors and accepted by them prior to the assumption of the coeditorship of this journal by one of the coauthors.

sumably instinctive, by other organisms. Such discontinuities are most easily noted by naturalists who study local faunas and floras, and the species concept derived from such situations has been called the "non-dimensional species concept" by Mayr (1963). But in some taxa, such as in willows, groups generally assigned generic or sectional rank are more easily recognized by local naturalists than are the species.

The apparent necessity to accommodate within one species concept several aspects of organisms led to the development of the so-called *biological species concept* (hereafter abbreviated BSC). These aspects include the variation of characteristics over large geographic areas, changes in these characteristics as populations adapt to environmental challenges or interact with other populations, and the integration of individuals into populations to form gene pools through direct processes, as well as indirectly through their ecological interactions. We shall not trace the development of the concept during the 1930s. Ernst Mayr, recognized as its foremost advocate, has called the BSC a "multidimensional concept" (Mayr 1963) because it deals with populations that are distributed through time and space, interrelated through mutual interbreeding, and distinguished from others by reproductive barriers.

Since its formulation there have been objections to the BSC from a variety of sources and motives. Many taxonomists have ignored it for practical reasons. Some workers (e.g., Blackwelder 1962; Sokal 1962) have charged that the employment of the BSC is misleading in that it imbues species described by conventional morphological criteria with a false aura of evolutionary distinctness and with unwarranted biosystematic implications. In fairness we point out that some supporters of the BSC (e.g., Simpson 1961, p. 149) state clearly the difficulties of correlating phenetic and genetic species criteria even in the same taxonomic group but especially across diverse taxa. Nevertheless, such caveats do not generally affect either taxonomic practice or teaching as it filters down to the level of the introductory courses. These critics also point out that the actual procedures employed even by systematists with a modern outlook are quite different from those implied or required by the BSC. Recent trends toward quantification in the biological sciences and especially emphasis on operationalism in systematic and taxonomic procedures (Ehrlich and Holm 1962; Ehrlich and Raven 1969; Sokal 1964; Sokal and Camin 1965; Sokal and Sneath 1963) have raised fundamental questions about the BSC to discover whether it is operational, useful, and/or heuristic with relation to an understanding of organic evolution.

The general purposes of this paper are: (1) to show, by means of a detailed flow chart, that the BSC is largely a phenetic concept; (2) given the above, to show that the BSC should be at least as arbitrary as phenetic taxonomic procedure; and (3) to explore the value of the BSC to evolution by posing a set of specific questions. Specifically, we shall first review the definition of the BSC and enumerate those of its attributes that require extended discussion and analysis. Next we shall discuss three operations

required for making decisions about actual populations with respect to these attributes of the BSC. Armed with an understanding of these operations, we shall then consider a flow chart of the detailed steps necessary to determine which of a set of organisms under study can be considered to form a biological species.

As a next step we shall note the difficulties of applying the BSC even in the optimal case of complete knowledge regarding the material under study, and examine how problems multiply as knowledge of the organisms diminishes.

Finally, given the difficulties of the BSC as a workable concept for the practicing taxonomist, we shall briefly examine the necessity for such a concept in evolutionary theory, its heuristic value, and the evidence for the existence of biological species in spite of the difficulty of their recognition and definition.

Although our philosophical attitude in systematics is that of empiricism and consequently we are not committed to the existence of biological species, we have approached our task with as open minds as has been possible. We recognize, as must any observer of nature, that there are discontinuities in the spectrum of phenetic variation. The question we have asked ourselves, one which we believe must be asked by every biologist concerned with problems of systematics and of evolution, is whether there is a special class of these discontinuities that delimits units (the biological species) whose definition and description should be attempted because they play an especially significant role in the process of evolution or help in understanding it.

## II. THE BIOLOGICAL SPECIES CONCEPT

The number of species definitions that have been proposed since the advent of the New Systematics and that fall within the general purview of the BSC is very large, but an extended review and discussion of these definitions would serve little useful purpose here. Many are but minor variants of the one to be discussed below, and they share in most ways the problems that we shall encounter with it. We shall employ the classical definition of biological species as restated by Mayr (1963, p. 19) in his definitive treatise. The definition is:

- |                                   |     |
|-----------------------------------|-----|
| Groups of                         | (1) |
| actually                          | (2) |
| or potentially                    | (3) |
| interbreeding                     | (4) |
| populations,                      | (5) |
| which are reproductively isolated | (6) |
| from other such groups.           | (7) |

We have deliberately arranged the definition in the above manner to emphasize those terms or phrases which make separate and important contri-

butions to the overall definition. Let us briefly go through these. We are dealing with *populations* (line 5) whose members *interbreed* (line 4) *actually* (line 2) *or potentially* (line 3). The difficulties of the latter term will be taken up in the next section. There usually is more than one such population (line 1). This group of populations will not exchange genes (line 6) with other interbreeding groups (line 7). This phenomenon is referred to as *reproductive isolation*.

According to Mayr (1963, p. 20) there are three aspects of the BSC: "(1) Species are defined by distinctness rather than by difference." By this he means reproductive gaps rather than phenetic differences (Mayr, personal communication). "(2) Species consist of populations rather than of unconnected individuals; and (3) species are more unequivocally defined by their relation to nonconspecific populations ('isolation') than by the relation of conspecific individuals to each other. The decisive criterion is not the fertility of individuals but the reproductive isolation of populations."

Thus to discover whether a given set of individuals is a biological species in the sense of the above definition we must have information about three essential components of the BSC: (1) that some individuals lack distinctness (*sensu* Mayr) from other individuals and join these in comprising biological populations of interbreeding individuals (this is the meaning by implication of the term "population" in the definition of the BSC); (2) that there is a group of such populations among which interbreeding does, or could, take place (this follows from the "actually or potentially interbreeding" clause of this definition); (3) that this group lacks gene flow with other groups of populations (this covers the "reproductively isolated" portion of the definition). These three aspects of the biological species are worked into the flow chart (fig. 1) presented below.

### III. FUNDAMENTAL OPERATIONS

To ascertain whether a given assemblage of organisms belongs to one or more biological species, three types of operations for grouping organisms and population samples will be found necessary (although only the third is directly implied by the definition given above). The first operation groups organisms by geographic contiguity; the second, by phenetic relationships; and the third, by reproductive relationships. In all these cases there will be some difference in the procedure when the initial grouping is of individuals into subsets (populations), and when these subsets are the basic units being grouped into more inclusive sets (species).

All grouping procedures will of necessity be based on samples of organisms and populations. Only in a minuscule number of instances will we have knowledge of all the individuals about which inferences are being made. This is not necessarily an unsatisfactory state of affairs, but it is important to specify the size of the samples required to estimate parameters of the populations with a desired level of confidence. Also, the use of samples

necessitates that some assumptions be made about the spatiotemporal distribution of individuals and populations.

The grouping operations will frequently refer to the idea of *connectedness*. We shall consider two operational taxonomic units (OTUs; see Sokal and Sneath [1963, p. 121]—individuals or population samples in this context) to be connected if there exists some definable relation between them (geographic contiguity, phenetic similarity, or interfertility, for example). *Minimally* connected sets of such OTUs have at least as many such relations as permit any two OTUs to be connected via any other members of the set. *Fully* connected sets have relations between every pair of members of the set. We use these terms by analogy with their employment in graph theory (Busacker and Saaty 1965).

We shall take up the three types of operations below in the order in which they were introduced.

The first operation groups by *geographic contiguity*. In order to belong to one population, organisms must be within reach of some others, that is, have the possibility of encountering for reproductive purposes other organisms within the same spatiotemporal framework. A first prerequisite for individuals to belong to the same population is that they come from sites which would enable them to be within reach of each other, considering the normal vagility of these organisms or of their propagules. In many cases we can simply assume this when we have samples from one site containing numerous individuals such as are obtained by seining, light traps, or botanical mass collecting. In other cases (especially with large organisms) where single individuals are found at specific sites, we have to be reasonably certain that individuals from separate sites presumed to be within the same local population have intersecting home ranges. In developing a criterion of geographic connectedness among local populations we need to be concerned with the probability of members of one locality visiting members of another one to permit the necessary gene flow required by the model. Again, this will be a function of the distance between localities, the vagility of organisms, and the ecological conditions that obtain between points. Various techniques of locational analysis (see Haggett 1966) can be used for establishing these linkages. We note in passing that the essential information required for this operation is lacking for most taxa. For example, the pollen and seed ranges for most flowering plant taxa are unknown (Harper 1966).

A second operation is the establishment of *phenetic similarity* between individuals within population samples and between such samples from various areas. While the definition of the BSC does not invoke phenetic considerations, it will be shown in the next section that any attempt to apply the definition to an actual sample of organisms will need to resort to phenetics in practice. In the initial stages of a study it may be that sufficient estimates of phenetic similarity can be determined by visual inspection of the specimens. Clearly, when the material is very heterogeneous such an initial sorting of the material into putatively conspecific

assemblages can be profitable. When more refined analysis is indicated, a quantitative phenetic approach is necessary. Here again we need not concern ourselves with the technical details, which are by now well established through the techniques of numerical taxonomy (Sokal and Sneath 1963).

The third operation involves grouping *interbreeding individuals* into population samples and grouping *interbreeding population samples* into larger assemblages. Before discussing this in detail, a semantic digression is necessary. In most relevant texts the term "interbreeding" is not defined precisely or distinguished clearly from intercrossing, interfertility, mating, and similar terms. Recourse to a dictionary is not enlightening. The reader is aware that the very act of mating (i.e., copulation in animals with or without insemination, or pollination in plants, to name only two of the more common mechanisms of sexual reproduction) does not of itself insure the production of viable offspring and especially of fertile offspring. Clearly, the act of mating or the transfer of male gametes toward a female gamete is the single necessary precondition for successful interbreeding, but it does not in itself insure fertile offspring. We shall use the term "interbreeding" to mean crossing between individuals resulting in the production of fertile offspring, but we shall occasionally use the terms "interfertility" or simply "mating" in a similar context.

The only unequivocal, direct basis for forming interbreeding groups is to observe organisms interbreeding in nature. If we wanted to make the definition absurdly rigorous, we would wish to insist that an interbreeding population sample be one where a sufficient number of females from the local population sample is mated with a sufficient number of males in the same sample to insure reproductive connectedness to the required degree. Fertile offspring would have to result from all of these unions. Obviously such observations are unlikely. Even if we were to turn to experiments to answer the question, we could not insist on so complete a test of interfertility, both because the number of experiments would be far too great and because, in most cases such crosses would be impossible, since the biological nature of the organisms precludes more than a single mating (e.g., longevity of mating individuals, incompatibility toward further mates by an already mated female, developmental period of the young, etc.).

Thus, as noted earlier we shall have to resort to samples of field observations or of crossing experiments. The latter raise the often discussed issue of whether laboratory tests of interbreeding should be considered as evidence when contrasted with field observations. Clearly, first consideration must be given to observations of nature as it is. Success in crossing experiments might indicate "potential" interbreeding. In designing crossing experiments as criteria of interfertility, clear instructions must be given on what role these experiments will play and whether the definition to be tested will be satisfied by laboratory crossing experiments or whether field observations are required.

Added to these difficulties is the fact that most of the material systematists deal with is already dead at the time of study and cannot be brought

into the laboratory or experimental garden for crossing purposes. Thus, extensive interbreeding tests are impractical, and one needs to resort to partial or circumstantial evidence on crossing for inference on interfertility. As direct evidence on interbreeding diminishes, the methods become increasingly phenetic. Phenetic information is of value in ascertaining interbreeding relationships only insofar as one may assume that phenetic similarity is directly related to ease of interbreeding. Yet we know that phenetics is an imperfect reflection of interfertility between organisms. In fact, this has been one of the main criticisms of numerical taxonomy by evolutionists.

The above arguments should not be interpreted as insistence on our part for "complete" knowledge of reproductive relationships. Just as one samples in phenetic studies to obtain estimates of phenetic structure of a larger population, so it is entirely justified to test reproductive relationships among only a sample of individuals and make inferences about a larger population. However, both sampling procedures are based on prior phenetic sorting out of specimens and populations. Thus we test reproductive relationships only among organisms likely to be interfertile, and the only way we can recognize these is on a phenetic basis. Therefore, except for the absurdly extreme reproductive test of each organism against every other one—biologically and experimentally infeasible, as well as destructive of the original taxa if it were possible to carry out such a test—reproductive tests based on samples reflect phenetic considerations in choosing the individuals to be tested. Furthermore, we must stress that even if we carried out some crossing experiments we would still need to employ phenetic inference to reason from the results of our limited number of crosses to the larger population sample, to the entire local population living today, and to the entire local population both living and dead.

Depending on the set of reproductive properties chosen by a given scientist, interbreeding will range continuously from complete interbreeding through intermediate stages to total lack of interbreeding. The two properties most often considered are connectedness and success of reproduction. If every individual in a group could interbreed with every other one of the opposite sex, *connectedness* would be complete. But the total number of possible combinations will likely be reduced; that is, some pairs may not be able to interbreed. This could be so for a variety of reasons, directly and indirectly genetic, such as sterility genes, reproductive incompatibilities, behavioral differences, seasonal isolation, etc. We are prepared to accept a sample as connected within itself if each individual is capable of interbreeding with one or more of the opposite sex in such a way that the reproductive relationships would yield a minimally connected graph (Busacker and Saaty 1965) (with  $n + m - 1$  edges, where  $n$  is the number of one sex and  $m$  that of the other), with terminal members being connected to one mate only. Such a minimal interbreeding relationship is unlikely in a large biological sample because it would imply a very complex system of mating types and intersterilities; yet even such a system practiced over



many generations would insure genetic connectedness among its members. A sample whose reproductive relations are less than a minimal connected set should be separated into those subsets which are connected.

But the ability to mate is clearly not enough. Fertile offspring, which have a nonzero probability of survival and of leaving new offspring, must result from such a union. This consideration leads us directly to the second property characterizing interbreeding.

*Success of reproduction* can be expressed as the percentage of fertile offspring resulting from a given mating measured in terms of percentage of eggs hatched, percentage of seed set, litter size, and similar criteria in the  $F_1$  or later generations. The standards set for such criteria and acceptable levels of success will vary with the investigator.

Therefore, members of a local population sample may be considered to interbreed either if they are completely interfertile as defined above or if they are partially interfertile. In the latter case, only samples whose members show at least minimal connectedness and whose average success of reproduction is greater than an arbitrarily established value would qualify.

If organisms are apomicts or obligate selfers, then by their very nature they cannot form biological species (as has indeed been pointed out by proponents of the BSC, e.g., Simpson, [1961, p. 161] or Mayr [1963, p. 27]). If these biological facts are not known to us, they might be suggested by all individuals forming a disjoint set in this step (i.e., no individuals will reproduce with any other individual in the sample). Technically, we should no longer process such samples through the flow chart. However, a useful classification could be arrived at if we ran the individuals of each local sample through the phenetic pathways of the flow chart. We infer this because taxonomists have had no apparent difficulty in describing species by conventional methods in these forms.

Once it has been demonstrated that the individuals *within* each local population sample interbreed, we need only show that there is some gene flow among the samples studied in order to establish interbreeding among them. Once genes from population *A* enter population *B*, (and those from *B* enter *A*), interbreeding among the members of *A* and *B* provides an opportunity for the establishment of the new genes in both populations.

We can conceive of several partially interfertile population samples as a connected set. It would follow that in order to be considered actually interbreeding the several population samples would have to represent at least a minimally connected set of reproductive relationships. Therefore, not every population sample needs to be directly reproductively connected to every other population sample in the study. A *Rassenkreis* is an example of such a situation. These relationships may be somewhat difficult to represent because the paths of connection will have to pass through either the offspring or parents of mates in a zigzag fashion. However, in populations among which there is substantial gene flow, it should be possible to make a chain of connection between any two organisms by going through relatively few ancestral and descendant generations.

The term "potentially interbreeding," which is included in some definitions of the biological species, has never really been defined, let alone defined operationally. It appears to us that the only possible answer one could get to the question of whether two samples are potentially interbreeding is "don't know." At best, one would be reduced to inferences about potential interfertility from phenetic evidence (and we have already seen that this is not too reliable). It is interesting to note that in his latest work, Mayr (1969) has dropped "potentially interbreeding" from his biological species definition.

#### IV. FLOW CHART FOR RECOGNIZING BIOLOGICAL SPECIES

The actual flow chart is shown in figure 1. The various steps in this figure are listed in this section, each followed by an explanatory account of the reasons for the step, the manner in which it could be carried out, and inherent difficulties.

1. *Assemble phenetically similar individuals.*—This preliminary step is important because unless the individuals used for the study are "relatively" similar, it is not reasonable to suppose that they interbreed. Lacking such a procedure, one would be forced to carry out a vast amount of fruitless testing for interfertility. Cottonwoods, aphids, and field mice could all be obtained in samples from the same locality, and, while the subsequent logic of the flow chart should ensure their separation into independent biological species (if we can determine that they are not actually or potentially interbreeding populations), a large amount of unnecessary and most likely impractical work would have to be done to test for fertility between cottonwoods and field mice, for example.

Systematists have appropriately decided not to trouble about this point but to use the relatively great phenetic dissimilarity of such groups of organisms to infer that they would be intersterile if an attempt at artificial crossing were made. Substantial evidence is available, especially from plants, that individuals allocated to different orders, families, or genera are usually intersterile. However, in the vast majority of organisms we may state with certainty that decisions about the presumptive intersterility of two dissimilar individuals or populations are based on phenetics alone. But since phenetic similarity is a continuous variable (as is reproductive interrelationship), it is difficult to designate anything but arbitrary similarity levels above which individuals and populations are potentially interbreeding and hence potentially conspecific and therefore need further testing and below which they are phenetically so different that the likelihood of interbreeding (hence of conspecificity) is small enough to be neglected. In the absurdly extreme instance of cottonwoods and field mice, this phenetic comparison is made instantaneously by the taxonomist without the need for more precise and sophisticated phenetic methods. This step is stressed here mainly to make the logic of the flow chart complete. When assembling similar individuals, dimorphisms and polymorphisms may give rise to

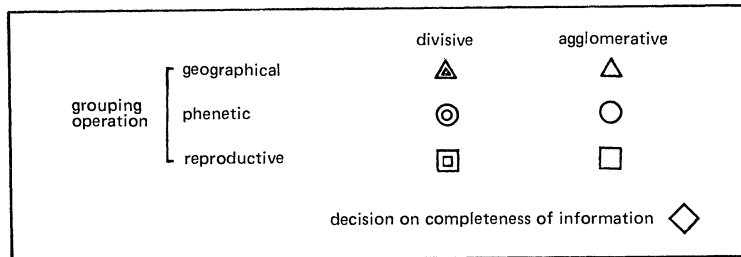
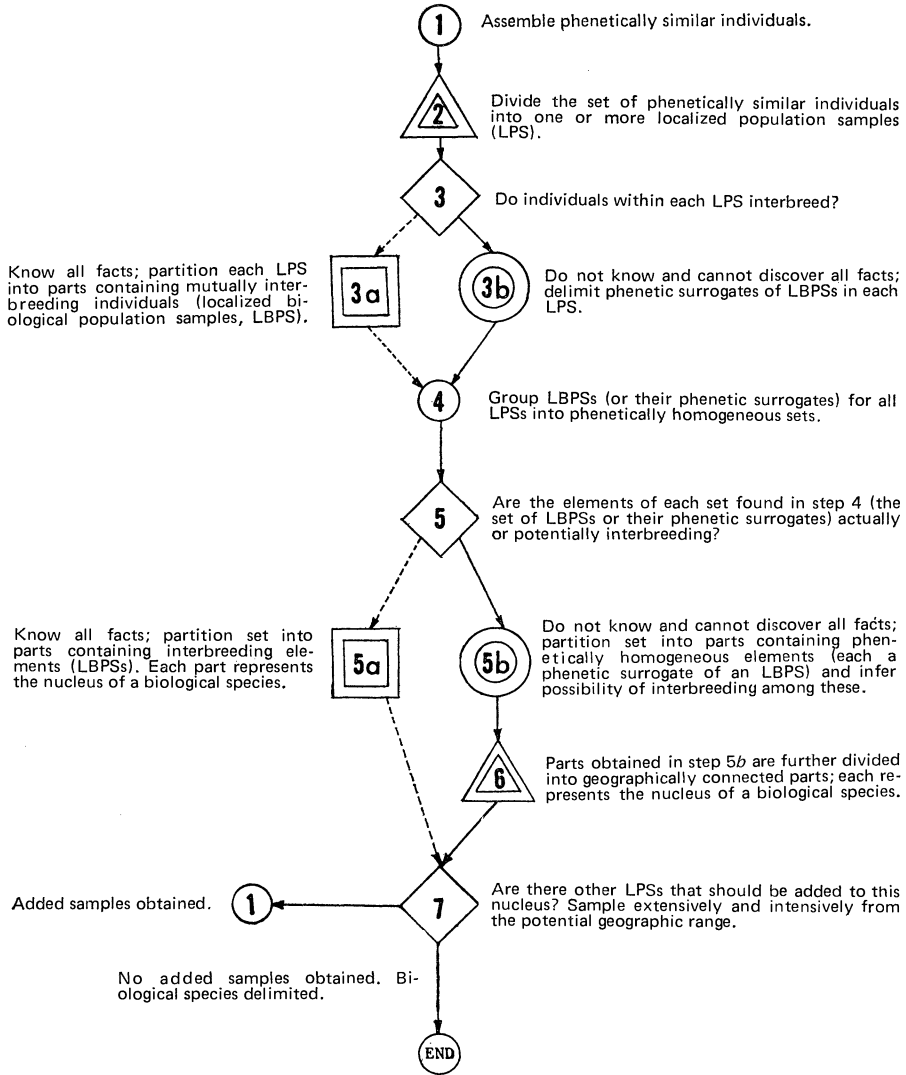


FIG. 1.—Flow chart for determining biological species. For explanation, see text.

practical difficulties, and relational criteria based on knowledge of the biology of the organisms involved may be invoked. Thus, knowing that given caterpillars give rise to given butterflies, we shall associate them, and in cases of marked sexual dimorphism we would wish to associate males and females that appear to form sexual pairs. This can sometimes be done by refined biometric techniques, but where previous knowledge or simple observations suffice these should surely be preferred.

A second point is that step 1 should not be carried out so finely that potential candidates for conspecific status are excluded. Thus the grouping should err by inclusion rather than by exclusion. Otherwise, since the flow chart will not cycle through the original sample again, some of the initial sample of organisms that also belong to the same biological species would be excluded.

2. *Divide the set of phenetically similar individuals into one or more localized population samples.*—The procedure leading to localized population samples is that of grouping by geographical contiguity as defined in Section III. Since we are at the moment concerned with the grouping of individuals to form population samples, we would be unlikely to encounter fixed geographic points from which we can create an interconnected network. Rather, we are likely to obtain a scattering of incidental collection sites from which we must draw inferences about the potential for geographic overlap of the lifetime movement ranges of the individuals concerned.

We shall define a *localized population sample* (LPS) in terms of the natural vagility of the organism. The need for such a definition stems from the biological attributes of populations as integrated gene pools which require that the members of a population be within the geographic range making such integration possible. We use the term “localized,” following the conventions of the statistical geographers (Haggett 1966), rather than the more common “local” population which has certain biological, genetic and ecological connotations that, although hard to define, nevertheless are generally invoked in the minds of systematists. By localized population sample we mean to imply only connection by an external relationship, largely spatial but also temporal and ecological. Unless otherwise qualified, this does not necessarily imply genetic or phenetic similarity among its members.

Gametes or propagules will differ in the distances they travel. A distribution of such distances, if known, could serve as a measure of vagility. The ninety-ninth percentile,  $V$ , gives a near upper limit to the distances travelled. If the largest observed distance between any two members of a cluster being formed is less than  $kV$ , where  $k$  is an arbitrary constant, we may define this cluster as a localized population sample. Problems might arise with uniformly spaced individuals, but such instances invite arbitrary decisions by any procedure. Also, while the samples are likely to be phenetically similar following step 1, we have no assurance that each sample represents one and the same species. Hence the vagilities of the individuals

within each sample are potentially heterogeneous as well. Percentiles other than the ninety-ninth might be employed.

Many times one will not have a distribution of exact locations at which individual specimens have been obtained because the sample will have been collected at one spot or because the collection records for the entire sample refer to one spot or to a broad area. In the former case we are clearly dealing with a sample from a localized population; in the latter we have to make a judicious definition of the area sampled. For instance, if a botanist furnishes only county records and the sample may be from anywhere within a county, the maximum straight-line distance within the county will have to stand for the greatest distance between any two members of the sample.

Of course, in most instances we will not know enough about the biology of the organisms studied to make a useful estimate of  $V$ . We therefore may have to guess at this value by analogy with known similar organisms.

The definition for localized samples given above should perhaps also include other criteria, such as time and ecological factors. The biological species definition as generally stated does not specifically refer to synchronous populations; yet, as has been repeatedly pointed out, the delimitation of species becomes much more complicated if chronistic aspects are also considered. One might very well impose an analogous criterion of chronistic connectedness on the definition and obtain samples localized in both space and time. Restriction to a given general habitat such as crowns of trees or leaf litter could also be imposed to restrict the possibility further, but we do not pursue this subject here.

Each resulting LPS is not necessarily a local population in the conventional biological sense. To be that it would have to be connected not only in the geographical sense but also by interbreeding relationships. The next step in the flow chart will impose this added constraint.

3. *Is each localized population sample, defined by geographic contiguity in step 2 above, also interbreeding within itself; that is, do its individual members interbreed among themselves?*—Localized populations that are not interfertile within themselves cannot make up the elements of a biological species population. In its rigid interpretation, we would have to ascertain whether there is either actual or potential interbreeding within each localized population sample. We have two choices in answering this question: we can either claim to know or hope to find out what the actual interbreeding relations among the organisms are—this leads us to step 3a—or we may decide that the question cannot be answered fully or at all in terms of interbreeding relationships and proceed to make inferences about these from other evidence, usually phenetics (step 3b).

3a) Knowing all the facts about interbreeding interrelationships within each LPS, we may partition it into parts containing mutually interbreeding elements. Each such part represents a *localized biological population sample* (LBPS).

The general criteria for recognizing interbreeding have been given in the previous section and will not be repeated here. The difficulties of testing

even within a limited sample the interbreeding of a sufficient number of members are considerable, and in fact step 3*a* is, for all intents and purposes, impracticable. Even if all necessary crosses were feasible in theory, we have seen that sampling based on phenetics will be required sooner or later for inferences about potential interfertility of some untested members of each LPS and of the larger local population. For this reason a broken arrow leads to and from the grouping operation based strictly on reproductive criteria, 3*a*, to indicate that this is *not* the usual path.

3*b*) When we do not know and cannot discover all the facts regarding interbreeding, we have to delimit at least some phenetically homogeneous subsets in each LPS and infer interbreeding of members of each subset (we may call these subsets phenetic surrogates of an LBPS).

We may assume that markedly dissimilar organisms have already been eliminated in step 1. When eliminating grossly different organisms, one should also take care that polymorphic forms representing sexes, genetic polymorphs, or different ontogenetic or cyclomorphic stages are not excluded. If no obvious differences are present, the establishment of the homogeneity of the individuals within the sample may require sophisticated biometrical analysis. Even then, homogeneity cannot ever be proven. It can only be established that for the set of characters which has been measured the individuals appear to be homogeneous. If a heterogeneity is discovered, as, for example, in the form of a bimodality of a given character or a constriction or discontinuity in character hyperspace, we need to allocate the sampled individuals to the two or more subpopulations thus defined. Intergrades will be troublesome in this context, and final decisions on boundaries of the phenetic groups are bound to be arbitrary.

Another method for grouping the subsamples within the original LPS would be to cluster the organisms by one of the methods of numerical taxonomy. As before, such a procedure is quite arbitrary in terms of the choice of a criterion of homogeneity.

4. *Group the LBPSs (or their phenetic surrogates) for all LPSs into phenetically homogeneous sets.*—This is an agglomerative phenetic grouping procedure and is necessary as a preliminary to the rigorous test by the defined criteria for the BSC. This is so because, following the strict guidelines of the biological species definition of “actually or potentially interbreeding populations,” one would have to test all samples obtained at the various localities for mutual interbreeding. As will be seen later, this is a formidable, if not impossible, task even when the samples are homogeneous within and among LBPSs, so that one may presume that they all belong to the same species. However, at this point in our procedure for determining whether a group of populations constitutes a biological species, we do not as yet know that the separate subsets in the various LBPSs defined in step 3 are similar to such a degree. All we know is that they are homogeneous *within* LBPSs. This does not necessarily mean that they are homogeneous *among* LBPSs.

Markedly different populations will already have been eliminated in step

1 of the flow chart. Thus we would no longer find one LBPS of drosophila and another LBPS of field mice. However, there might well be several species of drosophila, from the same locality, each in a single LBPSs, formed by a partition of one original LPS during step 3. We now must take all LBPSs (subsets from different LPSs) and combine them to form one or more sets whose elements are phenetically closely related LBPSs, regardless of the LBP from which they originated. It should be understood here that LBPSs in this step include not only those samples defined by step 3*a*, when this is possible, but must often include their phenetic surrogates established in step 3*b*.

When the LBPSs comprise two or more phenetically closely related but reproductively isolated groups of samples, this admixture becomes a problem. In most cases, techniques like numerical taxonomy should be able to cluster the populations correctly into those that are phenetically alike and therefore candidates for becoming a biological species, subject to further tests in this flow chart. The criterion of phenetic similarity to be employed is necessarily arbitrary, and for this step to be operational we have to establish phenetic limits. One situation where such an analysis might result in clusters undesirable for the present purpose is with marked geographic variation, possibly related to adaptation to ecological differences. Suppose there were two sibling species distributed over the area. It may well be that samples from reproductively isolated populations showing parallel ecological adaptations may cluster before joining with freely interbreeding samples from ecologically different areas. In such cases, some other form of multivariate analysis that removed the effect of ecological differences from a series of morphological variables would reveal the correct situation.

In summary, in most instances of testing for biological species the preliminary test (step 1) is carried out automatically, often already by the collector who does not bother to pick up animals other than those of the species group he is interested in. Nevertheless, it must be clearly recognized that unless the *phenetic* decisions of steps 1 and 4 are taken, one cannot in practice proceed with the determination of the specific status of these populations.

5. *Are the elements of each set found in step 4, the set of LBPSs or their phenetic surrogates, actually or potentially interbreeding among themselves?*—This question refers to the most important criterion of the BSC. In its rigid interpretation in terms of the definition, we would have to ascertain whether there is either actual or potential interbreeding among individuals of all the population samples obtained for our study. We have two choices in answering this question: we can either claim to know or hope to find out what the actual interbreeding relations among the organisms are—this leads us to step 5*a*— or, we may decide that the question cannot be answered fully or at all in terms of fertility relationships and proceed to make inferences about these from other evidence, usually phenetics (step 5*b*).

5*a*) Knowing all of the facts about interbreeding interrelationships

among elements of this set, we may partition it into parts containing mutually interbreeding elements (LBPSs). Each such part represents the nucleus of a biological species.

As has been pointed out repeatedly by proponents as well as opponents of the biological species definition, it is impracticable to ascertain these facts in most real situations. The difficulties encountered are of many kinds. The only kind of evidence that would unequivocally answer the question posed is direct observations of marked individuals and of their dispersal (or that of their gametes or offspring), plus observations on mating and success of the progeny in the field. Laboratory experiments on interfertility could be carried out but would indicate neither whether such interbreeding would take place in the field nor whether the offspring of such unions would be viable and reproduce under field conditions.

Even if we were to admit the evidence of laboratory tests, or of crossing experiments by botanists in experimental gardens, the number of crosses required would be formidable. With only two reciprocal crosses for any pair of population samples, we would need  $a^2$  tests for  $a$  samples (including controls within samples). Thus, for 10 local populations (a far from adequate number in most modern studies of speciation), 100 crosses would have to be made. Yet, we have no assurance that a single representative of each local sample would suffice to establish the necessary facts. After all, if an incomplete sterility barrier exists between these populations, then certain genotypes representing the population might not be able to cross while others would do so successfully. Doubtless, a more representative subsample of each population sample is needed to arrive at a decision on this matter.

On the other hand, since it was demonstrated—or inferred—in step 3 that the individuals within each sample interbreed, we have already stressed in Section III that we need only show that there is *some* gene flow among the samples being compared in order to establish interbreeding. Again, we need to distinguish between complete interbreeding, which would mean total panmixia or swamping among all population samples (an unlikely occurrence if the samples are reasonably far apart), and partial interbreeding. The latter, again, could depend on *connectedness* between some individuals in different LBPSs, which will govern the amount of gene flow, and *success of reproduction*, which refers to the percentage of fertile offspring from such crosses and the success of these offspring, evaluated by some standard. Arbitrary levels for these parameters must be designated to make the definition operational. We shall not suggest such levels here. In any event, the amount of experimental work and of field observations necessary to obtain answers for step 5a would become staggering and is clearly not practical. Sampling and inferences for the larger population are again phenetically based. For this reason there is once more (as in step 3a) a broken arrow leading to and from this operation.

5b) We do not know and cannot discover reproductive relationships among all of the elements (LBPSs) of this set. We therefore partition it into parts containing phenetically homogeneous elements (phenetic surro-



gates of LBPSs) and infer the possibility of interfertility among these.

This is a phenetic grouping procedure. The type of phenetic connectedness that should reflect whether samples (LBPSs) are actually interbreeding includes a high degree of overall phenetic similarity or the presence of intermediaries (introgression). Both kinds of phenetic evidence are subject to the same arbitrariness associated with the degree to which isolating mechanisms must be present before one can call two samples the same biological species. Here, we have to decide what degree of phenetic similarity must be present before considering two samples members of the same biological species. This will vary, of course, with the particular group under study and most of all with the characters chosen for analysis. As Davis and Heywood (1963), as well as critics of numerical taxonomy (e.g., Stebbins 1963), point out, morphological similarity is not a very accurate reflection of the evolutionary status of biological species. Also, overall similarity may not be the most critical phenetic relationship to be established. Phenetic evidence of introgression may be considered a more important criterion. We shall not discuss the possible procedures in detail here, since our main point is to point out the necessity of inference from phenetic evidence.

Had we been able to follow through on step 5*a* and define parts containing mutually interfertile elements, we could have bypassed step 6 below because we would have met the requirements of the biological species definition. Since we could not rigidly proceed by step 5*a* and had to resort to phenetic evidence in step 5*b*, we should strengthen our inferences by determining the geographical connectedness of these elements as shown in step 6.

6. *The parts of homogeneous sets of phenetic surrogates of LBPSs obtained in step 5b are further divided into one or more parts by geographic contiguity.*—This is done to increase our accuracy in the delimitation of biological species. Criteria of geographic proximity should reflect the likelihood of gene flow occurring between any two populations. Thus localities will be considered connected if some members of one LBPS at one locality have an opportunity to join a similar LBPS at the other locality. Geographical distances in such a model would be modified into ecological distances expressing the probability of propagules from one population entering the other population. We are now in a position to make joint judgments about the biological status of the resulting parts, which are phenetically homogeneous and geographically connected sets, constructed by a technique analogous to that of Gabriel and Sokal (1969) for geographic variation analysis.

It will be obvious that, since the level of phenetic homogeneity designated for assigning LBPSs to the same biological species is arbitrary, as is the accepted degree of geographic connectedness, decisions on membership in a biological species are arbitrary as well. That is, we may occasionally decide to include within the same biological species phenetically homogeneous populations that are not fully geographically connected; and, con-

versely, we may include populations that are phenetically distinct but seem to be fully geographically connected. Since these criteria do not, in any case, meet the formal definition of the BSC, their exact interpretation is not at issue here, unless we wish to infer "potential interbreeding" from them.

Following steps 5*b* and 6 we obtain the intersection of the parts resulting from these procedures. We infer that the elements in such an intersection (LBPSs or their phenetic surrogates) can represent the nucleus of a biological species. If we are prepared to accept the concept of potentially interbreeding populations, then we may simply use phenetic similarity as a criterion and bypass step 6, which implies actual gene flow in the geographic connections defined by its operations. To avoid confusion, this alternative is not shown in figure 1.

We now must ask ourselves whether the delimitation of this particular biological species can be extended to include other local populations. This is done by the final step, which follows.

7. *Are there other LPSs that should be added to the above nucleus?*—This step tests the adequacy of sampling. This question can be answered by further sampling of organisms from newly studied LPSs, starting with step 1 and repeating the entire procedure.

We define two kinds of additional sampling. *Extensive sampling* gathers further samples beyond the spatial limits of previous samples. *Intensive sampling* seeks to sample areas within the spatial limits of previous samples that have not been sampled before. This step will involve phenetic and geographic criteria, since it would be even more impractical to employ fertility criteria here as well. There is little point in going back through the flow chart, since the same information (phenetics and contiguity in distribution) will be used. In this step, as in steps 1, 3*b*, 4, and 5*b*, phenetic considerations will in the end largely delimit the biological species.

#### V. PHENETIC BOTTLENECKS

We now can examine the flow chart as a whole and imagine ourselves running some organisms through it to determine into how many biological species they should be divided. Let us design the optimal case for the systematic study of these organisms by the BSC criteria. Therefore we assume unlimited quantities of live material available from suitably positioned locations throughout the range of the organisms. Since tests of fertility would still require an enormous amount of experimentation, we shall imagine ourselves equipped with an all-knowing computer of unlimited capacity which will provide correct answers for meaningful questions asked of it, obviating experimental tests for interbreeding between pairs of individuals within and between locality samples. To make the situation correspond more closely to the real world about which we wish to make inferences, we shall restrict the computer's performance as follows. It cannot be queried simultaneously about the interbreeding of all individuals

of interest, but it will provide correct replies to sequential questions about relationships between each and every pair of individuals.

Given the above (and assuming that we have agreed on a criterion of interbreeding as discussed in Section III), we should be able to eliminate all steps in the flow chart except those that make critical tests of interbreeding, namely, 3*a* and 5*a*. But we would find that even our phenomenal computer would soon be running overtime providing answers to the millions of questions about interbreeding results of the possible combinations of individuals which we would have to ask. Hence, even in this utopian case we would wish to avail ourselves of steps 1 and 4 for purposes of grouping individuals and populations initially by phenetic likeness so as to cut down on the number of questions about interbreeding that need to be asked. (Thus we shall avoid asking whether an individual cottonwood would cross with an individual aphid.) However, even this timesaving device would not be sufficient. We would still have so many questions to ask about interbreeding, that our patience, if not that of a computer, would soon be exhausted, and we would take certain shortcuts, that is, ask questions about interbreeding of some of the individuals while resorting to phenetic similarities of these with other untested individuals for conclusions about the entire sample. But, having made this concession (i.e., having taken the path of the solid arrows in fig. 1), we are back at steps 3*b* and 5*b*, which we call the *phenetic bottlenecks* because limitations of time will force all studies, even the imaginary optimal study just discussed, into these operations.

Hence, while the definition of the BSC does not involve phenetics, the actual determination of a biological species always will do so, even in the optimal case. As soon as we permit less favorable (and more realistic) conditions to obtain, such as more limited material and no omniscient computer but a hard-working scientist with limited resources and facilities, establishment of biological species from fertility characteristics is entirely quixotic. We are left with what is essentially a phenetic criterion of homogeneous groups that show definite aspects of geographic connectedness and in which we have any evidence at all on interbreeding in only a minuscule proportion of cases.

The above is true for all animal organisms and for most plant organisms, as well. But even in those plant groups where crossing tests (the so-called experimental taxonomy) have been applied, the basic definition of the species is of necessity phenetic because the statements that are made rest on phenetic inferences from the relatively few crosses that have actually been carried out in these groups.

Phylogenetically oriented systematists have pointed out in the past that there are practical difficulties in determining the potentiality of interbreeding in given cases. But, as we have shown here, the concept cannot be used even under optimal circumstances. Simpson (1961, p. 150) has called this a pseudoproblem. He feels that the difficulty of ascertaining whether the definition is met in a given case with a sufficient degree of probability is different from the validity of the concept as such. Yet, as will be discussed

below, there is serious question that the concept is evolutionarily meaningful.

#### VI. DISCUSSION

*The BSC is imprecise in its formulation and inapplicable in practice.*—An obvious conclusion from the above flow chart and analysis is that in practice phenetics plays an essential role at several crucial points in the delimitation of a biological species. This leads to the critical question of the degree to which phenetics reflects interbreeding among individuals and populations. But many examples are known (see Davis and Heywood 1963) where phenetics can only mislead the biosystematist who is seeking the biological species. This ranges from simple polyploidy without phenotypic change and cryptic species, on the one hand, to problems of reactions to the environment, on the other. For example, small flowers are a result of dryness but can also be produced by mutation (Grant 1954). Without subjecting his material to experimental analysis the practicing systematist could not distinguish between these two causes. In other words, the inductive inference that is necessary here is often unwarranted.

Our study of the operations necessary to delimit a biological species revealed considerable arbitrariness in the application of the concept. This is in direct conflict with the claims of nonarbitrariness by proponents of the BSC. We use the terms "arbitrary" and "nonarbitrary" here in the sense of Simpson (1961, p. 115), where "a group is nonarbitrary as to inclusion if all its members are continuous by an appropriate criterion, and nonarbitrary as to exclusion if it is discontinuous from any other group by the same criterion. It is arbitrary as to inclusion if it has internal discontinuities and as to exclusion if it has an external continuity." The degree of sterility required in any given cross and the number of fertile crosses between members of populations, not to mention the necessarily arbitrary decisions proper to the hidden phenetic components of the BSC, make this concept no less arbitrary than a purely phenetic species concept, and perhaps even more so, since phenetics is but one of its components.

Relevant at this point is a contradiction in the use of the BSC regarding hybridization. This is a confusing term because at one extreme some authors call successful crosses between members of two strains a hybrid, while at the other extreme only crosses between members of two species, or between two genera, are hybrids. If a hybrid is produced in nature from two species and there is *any* backcrossing at all, then by a strict application of the BSC the two parents should belong to the same species, even if such hybrids appear in only a small part of the range of the species. But such an application is not usually made, since the investigator has some arbitrary level of frequency of crossing that he will tolerate before assigning the parents to the same species.

One of the prime complaints of the opponents of a phenetic taxonomy has been that it is typological (Inger 1958; Mayr 1965; Simpson 1961).

Whether empirical or statistical typology is an undesirable approach for a classificatory procedure is not at issue here. This question is discussed in some detail by Sokal (1962). In his most recent work on systematics Mayr (1969, p. 67) describes essentialist ideology as synonymous with typology in the following terms: "This philosophy, when applied to the classification of organic diversity, attempts to assign the variability of nature to a fixed number of basic types at various levels. It postulates that all members of a taxon reflect the same essential nature, or in other words that they conform to the same type. . . . The constancy of taxa and the sharpness of the gaps separating them tend to be exaggerated by [the typologist]. The fatal flaw of essentialism is that there is no way of determining what the essential properties of an organism are." However, it should be pointed out that, whether this is desirable or not, the BSC as advanced by its proponents is in itself a typological concept in the above sense. It is typological because it is defined by strict genetic criteria which are rarely tested, and which may not be met by its members (individuals or local populations). We shall examine below the question of whether populations in nature correspond to the biological species type erected by the new systematists. It may well be that the BSC does not reflect a widespread phenomenon in nature but rather represents a theoretical ideal to which existing situations are forced to fit as closely as possible.

It might be claimed that other variants of the biological species definition than the one employed by us could have been shown not to involve unwarranted inferences. However, a careful study of a great variety of such definitions shows this not to be the case. The definition by Emerson (1945)—"evolved (and probably evolving), genetically distinctive, reproductively isolated, natural population"—and that by Grant (1957)—"a community of cross-fertilizing individuals linked together by bonds of mating and isolated reproductively from other species by barriers to mating"—are both prone to the same difficulties. Simpson (1961, p. 153) defined "evolutionary species" as "a linkage (an ancestral-descendant sequence of populations) evolving separately from others and with its own unitary evolutionary role and tendencies." This is so vague as to make any attempt at operational definition foredoomed to failure.

Some plant biosystematists consider the BSC definition we have chosen to be genetic, and not necessarily evolutionary. Some, for example, would maintain that two populations belong to two biological species if they differ in at least one qualitative character and if there exists a certain amount of sterility between them. But this and similar definitions contain the same drawbacks of necessary phenetic inferences and arbitrariness as the concept we have discussed. It still is based in large part on phenetic inferences that may be unwarranted, and it still distorts relationships among populations by lumping them into a smaller number of biological species. The same comments apply to the definition of a biological species as a set of individuals sharing a common gene pool. This last definition may appear to have one advantage over previous ones. It does not demand that local

populations be erected during the process of species delimitation. In terms of our flow chart, steps 2 and 3 would be deleted and subsequent steps reworded. Although this has the "advantage" of reducing the number of necessary steps in the process, this is more than outweighed by the increased amount of inference about gene-pool membership that now must be made from only phenetic evidence, as opposed to inferences made previously from both phenetic and geographic information.

*Some essential questions about the BSC.*—From the above conclusions drawn about the BSC, we see that only in rare instances, such as a species consisting entirely of one small endemic population, is the concept even partly operational in practice. But a nonoperational concept may still be of value. For example, it may be used to generate hypotheses of evolutionary importance. We shall examine several relevant questions for systematists and evolutionists concerned with the BSC. At this time we can do little more than to ask the questions and to suggest possible answers.

1. Is the BSC necessary for practical taxonomy? By practical taxonomy we mean the straightforward description of the patterns of variation in nature for the purpose of ordering knowledge. This is phenetic taxonomy, or perhaps simply taxonomy as Blackwelder (1967) sees it. The BSC is not a necessary part of the theory of practical taxonomy, although the category "species" is. The answer to question one is no.

2. Is the BSC necessary (or useful) for evolutionary taxonomy? This is a more difficult question to answer, since different workers attach different meanings to the term "evolutionary taxonomy." It may mean the relatively less complex task of putting all members believed to be derived from the same ancestral stock into the same taxon, say at the genus, or family, level. Or it may involve detailed (usually phenetically inferred) description of cladistic relationships among taxa at some categorical level. The property of interbreeding may or may not be possessed by all members of the group currently under study. Most evidence for decisions in evolutionary taxonomy (and all evidence above the level of classification where crossing is not possible, e.g., between members of two families) is based not on interbreeding but on phenetics and homologies, whether they are morphological, behavioral, physiological, serological, or DNA homologies. Most work to date, especially on DNA homologies, has involved very dissimilar taxa, such as wheat, corn, pigs, monkeys, and man. Since the biological species does not play an essential role in any of the above work, the answer to question 2 also would appear to be no.

3. Is the BSC valuable as a unique, heuristic concept from which hypotheses valuable for evolutionary theory can be generated at a high rate? It would appear that any evolutionary hypothesis generated in terms of the BSC can also be generated in terms of the less abstract localized population and perhaps generated more easily. Significantly, population genetics, both theoretical and practical, in nature and in the laboratory, concerns itself with the localized population, or a small number of adjacent localized biological populations. There are few if any insights supposedly

obtained from species that cannot be better interpreted at the population level. In fact, some would say that they can be interpreted only at the population level. Nothing is gained by additional abstraction to the species level (except perhaps in efficiency of names), but much is lost, namely, accuracy, for no two localized biological populations are alike. By forcing a large series of them into one biological species we lose the resolution of their differences. The answer to question 3 appears to be no.

4. Is the BSC necessary (or useful) for evolutionary theory? That is, does the general theory of evolution, or any particular evolutionary process, require, or use, the BSC? With respect to the general theory, the answer appears to be no. If we examine the evolutionary situation within some ecosystem, we can generate the same theory based on localized biological populations without grouping sets of interbreeding populations into more abstract biological species. Parenthetically, we may point out that what are probably the most important and progressive books on evolutionary theory that have been published within the last year or so essentially do not refer to the biological species at all. MacArthur and Wilson (1967) in their study of island biogeography, Wallace (1968) in his analysis of evolutionary mechanisms, and Levins (1968) in his theory of evolution in changing environments base their entire discussions on Mendelian populations and hardly mention the BSC. Williams (1966, p. 252) believes that the species is "a key taxonomic and evolutionary concept but [it] has no special significance for the study of adaptation. It is not an adapted unit and there are no mechanisms that function for the survival of the species."

Let us turn to evolution over geological time and consider the birth and death of a presumed biological species. Assume that a certain phenetic form appeared at time  $i$  in the fossil record, subsequently became abundant, and then became extinct at time  $j$ . What does this mean? It means only that certain populations that possessed the given phenotype were able to survive from time  $i$  to time  $j$ . Ignoring polytopic origins, this means that this favorable character combination was transmitted among several localized biological populations. Nowhere does such a process demand that this set of populations be put into one group and that it be called a biological species. This can be done, but it is not essential to evolutionary theory. Of course, it is done for convenience of reference. It orders our knowledge in a certain way, as does grouping organisms into taxonomic species, then into genera, then families, etc. Thus it would seem to us that the biological species is an arbitrary category, which may be useful in given situations but is not a fundamental unit of evolution, except possibly in a case in which there is only one local biological population, and therefore the biological species as a class has only one member.

Furthermore, if we assume a priori that all organisms can be put into some biological species, then we of necessity concentrate on finding such classes. Could it be that the occurrence of well-circumscribed biological species is *not* the rule but the exception, in biology? Although Stebbins (1963) says that 70%–80% of higher plant species conform well to the

BSC, other evolutionists, upon the accumulation of more and more evidence (e.g., Grant 1963, p. 343 ff.) recognize the frequent occurrence of borderline situations.

We do not in any of the above statements imply that reproductive barriers are either nonexistent or unimportant in evolution. Quite clearly they are of fundamental significance. But we do question whether they can be employed to define species and whether emphasis in evolutionary theory should be based on phenomena (including reproductive barriers) pertaining to the species category or to a lower category, the local population.

The answer to question 4 appears to be unclear at best.

*Conclusions.*—If our contention that the BSC is neither operational nor necessary for evolutionary theory is granted, what consequences result for general evolutionary theory? There would be few changes if any in terms of our understanding of speciation mechanisms. For example, the numerous important principles outlined by Mayr (1963) in his treatise on the species would still be relevant even if the term "species" as such were removed and replaced by others referring to phenetically different populations, or reproductively isolated populations, or populations with both properties. The positive aspect of such a procedure would be that evolutionary theory and research would concern themselves more with discovering and describing mechanisms bringing about population changes than with trying to bring organic diversity into an order conforming to an abstract ideal. The emphasis would be on unbiased description of the variety of evolutionary patterns that actually exist among organisms in nature, and of the types of processes bringing about the different varieties of population structure. We believe that in the long run this approach would lead to greater and newer insights into the mechanisms of evolution. Fundamentally this would be so because such an approach would free hypothesis construction in evolution from the language-bound constraint imposed by the species concept. (See Kraus [1968] for a lucid exposition of some of these issues and especially the role of the Whorfian hypothesis.) Even if the Whorfian hypothesis is only partially correct, the very fact that we need no longer put our major emphasis on species definition and description would have a liberating effect on evolutionary thinking. By not tying the variation of individuals and populations to abstract ideals or relating it to a one-dimensional nomenclatural system incapable of handling the higher dimensionality of the variation pattern, we would be led to new ways of looking at nature and evolution.

Having decided that the BSC is neither operational nor heuristic nor of practical value, we conclude that the phenetic species as normally described and whose definition may be improved by numerical taxonomy is the appropriate concept to be associated with the taxonomic category "species," while the local population may be the most useful unit for evolutionary study.

In advocating a phenetic species concept we should stress that, in concert with most numerical taxonomists, we conceive of phenetics in a very



wide sense. All observable properties of organisms and populations are considered in estimating phenetic similarities between pairs of OTUs. These would include not only traditional morphological similarity but also physiological, biochemical, behavioral similarity, DNA homologies (Reich et al. 1966), similarities in amino acid sequences in proteins (Eck and Dayhoff 1966; Fitch and Margoliash 1967), ecological properties (Fujii 1969), and even intercrossability (Morishima 1969). Critics of a phenetic taxonomy have claimed that such a wide definition of phenetics makes the term meaningless, since all possible relationships among organisms are then by definition phenetic. But this is not necessarily so. Similarities over the set of all known properties are surely different from similarities based solely on the ability to produce fertile offspring.

Insistence on a phenetic species concept leads inevitably to a conceptualization of species as dense regions within a hyperdimensional environmental space in the sense of Hutchinson (1957, 1969). Current trends in evolutionary thinking do, in fact, consider this approach to species definition as a more useful and heuristic concept, and, as already mentioned, the existence of apparently "good" asexual species supports this view. However, the establishment of such an environmentally bounded species concept, an idea whose germs can be found in numerous recent papers, is beyond the scope of the present article, which limits itself to pointing out the weaknesses of the generally promulgated BSC.

#### SUMMARY

The term "species" has been a central tenet of biological belief since the early days of biology. But the concepts attached to the term have varied and often were not defined rigorously. The purpose of this paper is to investigate the biological species concept (BSC): to consider its theoretical aspects, how one would actually delimit a biological species in nature, whether such species exist in nature, and whether the concept is of any unique value to the study of evolution.

The classical definition of the BSC is partitioned into its essential components, and some of their aspects and problems are discussed. Three fundamental operations necessary for the delimitation of biological species in nature are described in detail. These are operations based on criteria of: (1) geographic contiguity, (2) phenetic similarity, and (3) interbreeding. Two properties of interbreeding, connectedness and success of reproduction, are defined and discussed.

A flow chart for recognizing biological species is constructed from the definition as given by Mayr. Each step involves one of the three operations mentioned above. Reasons are given for including each step, as well as the inherent difficulties of each. It can be seen that most steps are either largely or entirely phenetic, even in theory. The necessary phenetic steps are termed "phenetic bottlenecks." To test the flow chart, we assume the unrealistic but optimal situation of total knowledge about the interbreeding relations

among sampled organisms. The phenetic bottlenecks remain in this optimal case, and the degree of reliance on phenetic information for the delimitation of biological species increases as we depart from the optimal situation and make it more realistic.

The BSC is found to be arbitrary (*sensu* Simpson) when attempts are made to apply it to actual data in nature, and not only because arbitrary phenetic decisions are a necessary part of the delimitation of biological species in nature.

On asking some essential questions about the value of the BSC to taxonomy and evolution, we find that the BSC is not necessary for practical taxonomy, is neither necessary nor especially useful for evolutionary taxonomy, nor is it a unique or heuristic concept necessary for generating hypotheses in evolutionary theory. Most of the important evolutionary principles commonly associated with the BSC could just as easily be applied to localized biological populations, often resulting in deeper insight into evolution.

Having decided that the BSC is neither operational nor heuristic nor of any practical value, we conclude that the phenetic species as normally described is the desirable species concept to be associated with the taxonomic category "species," and that the localized biological population may be the most useful unit for evolutionary study.

#### ACKNOWLEDGMENTS

We have been fortunate to have benefitted from a critical reading of an earlier draft of this paper by several esteemed colleagues, representing considerable diversity in their attitudes to the "species problem." Paul R. Ehrlich, Richard W. Holm, and John A. Hendrickson, Jr., of Stanford University, James S. Farris of the State University of New York at Stony Brook, David L. Hull of the University of Wisconsin at Milwaukee, and Arnold G. Kluge of the University of Michigan contributed much constructive criticism and helped us remove numerous ambiguities and obscurities. A similar function was performed by many members of the Biosystematics Luncheon Group at the University of Kansas. We are much in the debt of all of these individuals, even in those rare instances where we have chosen not to follow their advice.

Collaboration leading to this paper was made possible by grant no. GB-4927 from the National Science Foundation and by a Research Career Award (no. 5-KO3-GM22021) from the National Institute of General Medical Sciences, both to Robert R. Sokal.

#### LITERATURE CITED

- Berlin, B., D. E. Breedlove, and P. H. Raven. 1966. Folk taxonomies and biological classification. *Science* 154:273-275.
- Blackwelder, R. E. 1962. Animal taxonomy and the new systematics. *Survey Biol. Progress* 4:1-57.

- Blackwelder, R. E. 1967. *Taxonomy*. Wiley, New York. 698 p.
- Busacker, R. G., and T. L. Saaty. 1965. *Finite graphs and networks: an introduction with applications*. McGraw-Hill, New York. 294 p.
- Davis, P. H., and V. H. Heywood. 1963. *Principles of angiosperm taxonomy*. Oliver & Boyd, London. 558 p.
- Eck, R. B., and M. O. Dayhoff. 1966. *Atlas of protein sequence and structure*. Nat. Biomed. Res. Found., Silver Spring, Md. 215 p.
- Ehrlich, P. R., and R. W. Holm. 1962. Patterns and populations. *Science* 137:652-657.
- Ehrlich, P. R., and P. H. Raven. 1969. Differentiation of populations. *Science* 165:1228-1232.
- Emerson, A. E. 1945. Taxonomic categories and population genetics. *Entomol. News* 56: 14-19.
- Fitch, W. M., and E. Margoliash. 1967. Construction of phylogenetic trees. *Science* 155: 279-284.
- Fujii, K. 1969. Numerical taxonomy of ecological characteristics and the niche concept. *Syst. Zool.* 18:151-153.
- Gabriel, K. R., and R. R. Sokal. 1969. A new statistical approach to geographic variation analysis. *Syst. Zool.* 18:259-278.
- Grant, V. E. 1954. Genetic and taxonomic studies in *Gilia*. IV. *Gilia achilleaefolia*. *Aliso* 3:1-18.
- . 1957. The plant species in theory and practice, p. 39-80. *In* E. Mayr [ed.], *The species problem*. Amer. Ass. Advance. Sci. Pub. 50.
- . 1963. *The origin of adaptations*. Columbia Univ. Press, New York. 606 p.
- Haggett, P. 1966. *Locational analysis in human geography*. St. Martin's, New York. 310 p.
- Harper, J. L. 1966. The reproductive biology of the British poppies, p. 26-39. *In* J. G. Hawkes [ed.], *Reproductive biology and taxonomy of vascular plants*. Pergamon, New York.
- Hutchinson, G. E. 1957. Concluding remarks. *Cold Spring Harbor Symp. Quant. Biol.* 22:415-427.
- . 1969. When are species necessary? p. 177-186. *In* R. C. Lewontin [ed.], *Population biology and evolution*. Syracuse Univ. Press, Syracuse, N.Y.
- Inger, R. F. 1958. Comments on the definition of genera. *Evolution* 12:370-384.
- Kraus, R. M. 1968. Language as a symbolic process in communication. *Amer. Sci.* 56: 265-278.
- Levins, R. 1968. *Evolution in changing environments*. Princeton Univ. Press, Princeton, N.J. 120 p.
- MacArthur, R. H., and E. O. Wilson. 1967. *The theory of island biogeography*. Princeton Univ. Press, Princeton, N.J. 203 p.
- Mayr, E. 1957. Species concepts and definitions, p. 1-22. *In* E. Mayr [ed.], *The species problem*. Amer. Ass. Advance. Sci. Publ. 50.
- . 1959. Isolation as an evolutionary factor. *Amer. Phil. Soc., Proc.* 103:221-230.
- . 1963. *Animal species and evolution*. Harvard Univ. Press, Cambridge, Mass. 797 p.
- . 1965. Numerical phenetics and taxonomic theory. *Syst. Zool.* 14:73-97.
- . 1969. *Principles of systematic zoology*. McGraw-Hill, New York. 428 p.
- Morishima, H. 1969. Phenetic similarity and phylogenetic relationships among strains of *Oryza perennis*, estimated by methods of numerical taxonomy. *Evolution* 23: 429-443.
- Reich, P. R., N. L. Somerson, C. J. Hybner, R. M. Chanock, and S. M. Weissman. 1966. Genetic differentiation by nucleic acid homology. I. Relationships among *Mycoplasma* species of man. *J. Bacteriol.* 92:302-310.
- Simpson, G. G. 1961. *Principles of animal taxonomy*. Columbia Univ. Press, New York. 237 p.
- Sokal, R. R. 1962. Typology and empiricism in taxonomy. *J. Theoretical Biol.* 3:230-267.

- . 1964. The future systematics, p. 33–48. *In* C. A. Leone [ed.], *Taxonomic biochemistry and serology*. Ronald, New York.
- Sokal, R. R., and J. H. Camin. 1965. The two taxonomies: areas of agreement and conflict. *Syst. Zool.* 14:176–195.
- Sokal, R. R., and P. H. A. Sneath 1963. *Principles of numerical taxonomy*. Freeman, San Francisco. 359 p.
- Stebbins, G. L. 1963. Perspectives. I. *Amer. Sci.* 51:362–370.
- Wallace, B. 1968. *Topics in population genetics*. Norton, New York. 481 p.
- Williams, G. C. 1966. *Adaptation and natural selection*. Princeton Univ. Press, Princeton, N.J. 307 p.