



Karl R. Popper, Systematics, and Classification: A Reply to Walter Bock and Other Evolutionary Taxonomists

E. O. Wiley

Systematic Zoology, Vol. 24, No. 2. (Jun., 1975), pp. 233-243.

Stable URL:

<http://links.jstor.org/sici?sici=0039-7989%28197506%2924%3A2%3C233%3AKRPSAC%3E2.0.CO%3B2-U>

Systematic Zoology is currently published by Society of Systematic Biologists.

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/ssbiol.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 and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

- NAPIER, J. R., AND P. H. NAPIER. 1967. A handbook of living primates. Academic Press, New York.
- OXNARD, C. E. 1969. Evolution of the human shoulder: some possible pathways. *Am. J. Phys. Anthropol.* 30:319-332.
- OXNARD, C. E. 1973. Functional inferences from morphometrics: problems posed by uniqueness and diversity among the primates. *Syst. Zool.* 22:409-424.
- ROBINSON, J. T. 1972. Early hominid posture and locomotion. Univ. Chicago Press.
- ROHLF, F. J., AND R. R. SOKAL. 1965. Coefficients of correlation and distance in numerical taxonomy. *Univ. Kansas Sci. Bull.* 45(2):3-27.
- STEYSKAL, G. C. 1968. The size factor in numerical taxonomy. *Syst. Zool.* 17:210.
- WALKER, A. 1973. A reconstruction of the femur of *Australopithecus*. *J. Hum. Evol.* 2:545-555.
- WELLS, L. H. 1973. Forward from Taung. *J. Hum. Evol.* 2:563-566.
- ZUCKERMAN, S. (ed.) 1973. Sir Grafton Elliott Smith and the concepts of human evolution. *Symp. Zool. Soc. Lond.* 33.
- ZUCKERMAN, S., E. H. ASHTON, R. M. FLINN, C. E. OXNARD, AND T. F. SPENCE. 1973. Some locomotor features of the pelvic girdle in primates. *Symp. Zool. Soc. Lond.* 33:71-165.

ROBERT S. CORRUCINI

*Division of Physical Anthropology
National Museum of Natural History
Washington, D.C. 20560*

Karl R. Popper, Systematics, and Classification: A Reply To Walter Bock and Other Evolutionary Taxonomists

Popper (1968a, p. 37) has stated that "*the criterion of the status of a (scientific as opposed to a metaphysical) theory is its falsifiability, or refutability, or its testability.*" This basic criterion of testability is held by the majority of scientific philosophers, Popper differing only in the specific way in which testing is carried out. Bock (1973) has argued that Popper's philosophy should be adopted for biological classification. I will argue that a theory of phylogeny via genealogical descent conforms to the general criterion of testability and the specific methods of testing outlined by Popper (1968a, b) and thus can serve as a basis for classification. I will also argue that the theory and methodology of evolutionary taxonomy as outlined by Mayr (1969, 1974), Bock (1973), and Ashlock (1974) fits neither the basic criterion of testability nor Popper's specific philosophy and thus cannot serve as a basis for deriving classifications which purport to reflect or communicate scientific inferences. I do not claim that a genealogical theory is the only one which conforms to Popper's philosophy, but I maintain that Hennig's (1966) methodology is testable within Popper's philosophy. The basis for this conclusion rests on the relationship between homology and

phylogeny. I will explore this relationship within inductive and deductive systems of hypothesis testing. Certain axioms will be suggested for a phylogenetic system. A restricted definition of homology is defended as best suited for application to estimates of phylogeny. The relationship between homology and phylogeny is then defended as being non-circular. After demonstrating the usefulness of Hennig's (1966) methods within Popper's philosophy, some logical aspects of "classical evolutionary taxonomy" (Bock, 1973) are discussed. Finally, the types of phylogenetic classifications which can exist under Popper's philosophy are discussed.

INDUCTIVE AND DEDUCTIVE HYPOTHESIS TESTING

Induction and deduction play central roles in any science. Induction is the fact or observation gathering process which enables an investigator to pose meaningful hypotheses about the world of nature. Observations comprise our world of experience. Deduction also plays a central role in science for it is through this process that tests are made which corroborate or refute the hypotheses formulated from our world of experience.

Inductive and deductive methods of hypothesis testing are empirical methods of verifying, or corroborating, or refuting hypotheses. Popper (1968b, pp. 27-48) states that inductive methods of hypothesis testing should be rejected because they require adoption of a doctrine of apriorism. In a system of inductive hypothesis testing, attempts to evade apriorism, that is, attempts to evade justification of certain statements as *a priori* true, lead to infinite regression (Popper, 1968b, p. 30). Thus, if we do not regard a universal statement as true based on our world of experience, we would have to justify it by making some inductive inferences. Yet, to make these inductive inferences we must assume some inductive principle of a higher order, and to justify this higher principle we must assume yet another inductive principle of yet an even higher order, and so on to infinity. For example, if we do not take as true a statement that the feet of two species are homologous, we would have to justify the statement by saying that all feet which are attached to the hind leg are homologous. Unwilling to accept this statement as true, we might further qualify our statement by saying that all feet which are attached to the hind leg and are composed of three bones are homologous. To justify this assertion we might add that the muscles of the foot must be arranged in a certain way. Unless we stop and say "I now accept this homology" we will have to go on investigating the structure of the two feet on a finer and finer level, eventually reaching the atomic and subatomic levels, and this eventually leads to the infinite. Appeal to probability (logical, not statistical) does not solve the problem of infinite regression because logical probability statements are themselves based on higher order inductive principles and also lead to infinite regression. Popper (1968a, b) suggests that deductive methods of hypothesis testing do not lead to infinite regression and do not depend on apriorism and thus can provide a suitable separation between science and metaphysics. Such

methods systematically exclude all attempts to *avoid falsification*. The system consists of posing and testing a series of statements or hypotheses. The result of the test can take one of three forms: (1) consistency with the hypothesis, (2) inconsistency, and (3) irrelevancy of the test.

AXIOMS OF PHYLOGENETIC SYSTEMATICS UNDER POPPER'S SYSTEM

Axioms may take two forms, conventions and hypotheses. Conventions are taken as true (not testable) and are to be avoided. Hypotheses which occupy the highest levels of universality within a particular theoretical system are termed the axioms of that system. Hypotheses at a lower level of universality are simply termed hypotheses. It is important to note that Popper conceives of observations as low level hypotheses, not as facts, because he maintains that observations exist only as interpretations of the facts of nature in light of present theories, not as the facts of nature themselves (Popper, 1968b, p. 107).

Three nested axioms occupy the highest levels of universality in the phylogenetic system: (1) evolution occurs; (2) only one phylogeny of all living and extinct organisms exists, and this phylogeny is the result of genealogical descent; (3) characters may be passed from one generation to the next generation, modified or unmodified, through genealogical descent.

HOMOLOGY IN INDUCTIVE AND DEDUCTIVE SYSTEMS OF HYPOTHESIS TESTING

Under the axioms stated above, phylogenetic definitions of homology will suffer from circularity in an inductive system of hypothesis testing unless the homologies are recognized as empirical facts. Because homologies are not empirical facts but hypotheses, the relationship between a homology and a phylogeny would become circular. It is also possible to establish a circular relationship between homology and phylogeny in a deductive system. I maintain that a phylogenetic definition of homology

can be produced in the deductive system which can be applied to phylogenetic hypotheses and classifications in a non-circular way.

There is a problem in selecting a definition of homology to be used within the deductive system because many definitions exist, e.g., phenetic (Sneath and Sokal, 1973), cladistic (Hennig, 1966), evolutionary (Simpson, 1961; Mayr, 1969; Bock, 1969), classical (Owen, 1848), etc. Given the choice among two or more definitions I prefer one which provides the best vehicle or tool for testing the hypothesis that two or more characters are homologous. Popper (1968b, p. 115) has stated:

"a statement x is said to be 'falsifiable to a higher degree' or 'better testable' than a statement y , . . . if and only if the classes of potential falsifiers of x includes the classes of potential falsifiers of y as a *proper subset*."

For example, a phylogenetic definition of homology may be considered more falsifiable than a phenetic definition and therefore preferable if it leads to a hypothesis of homology which includes all the potential falsifiers provided by phenetic comparisons as well as the potential falsifiers provided by phylogeny.

Within this system two (or more) characters are said to be homologous if they are transformation stages of the same original character present in the ancestor of the taxa which display the characters (modified from Hennig, 1966). There are two types of homologous characters: (1) those which are derived from the immediate ancestor of two taxa, and (2) those derived from an ancestor more genealogically distant than the immediate common ancestor. Apomorphic character states are those which are hypothesized to be derived from the immediate ancestral species and absent from earlier common ancestors (two taxa which share them are said to display a synapomorphy). Plesiomorphic character states are those which are hypothesized to be derived from an earlier ancestor and retained in all later ancestors (two taxa

which share a plesiomorphic state are said to display a symplesiomorphy). Finally, those character states that are not derived from the same original state in a common ancestor are said to be non-homologous.

LOGICAL IMPLICATIONS OF A PHYLOGENETIC DEFINITION OF HOMOLOGY IN A METHODOLOGY OF PHYLOGENETIC ANALYSIS

I shall explore the relationship between homology, as defined above, and phylogeny within the methodology of phylogenetic hypothesis testing developed by Hennig (1966).

A hypothesis of homology at a given level of universality contains a *minimum* of two character states, apomorphic and plesiomorphic. Whether a character state is plesiomorphic or apomorphic is relative depending on the level of universality (Fig. 1) of the phylogenetic hypothesis to which it belongs as a *proper subset*. Homologies can be tested only at the level of universality at which they are hypothesized to exist as synapomorphies because the best test of homology is common ancestry. Supposed homologies, therefore, necessarily take on the characteristics of axiomatic conventions at the level of universality at which they exist as symplesiomorphies. Hence, the only valid test of homology under this system is to hypothesize that the supposed homology is a synapomorphy.

Homologies themselves exist at different levels of universality depending on the number of potential falsifiers they contain. If two hypotheses of homology have classes of potential falsifiers which do not intersect they may be thought of as independent of each other. The character states of a transformation series have classes of potential falsifiers which do intersect and comprise no more than a single potential falsifier of a phylogenetic hypothesis. Different transformation series have non-intersecting classes of potential falsifiers. Although they may not be compared as to their suitability as potential falsifiers of a phylogenetic hypothesis, they can be compared as to their

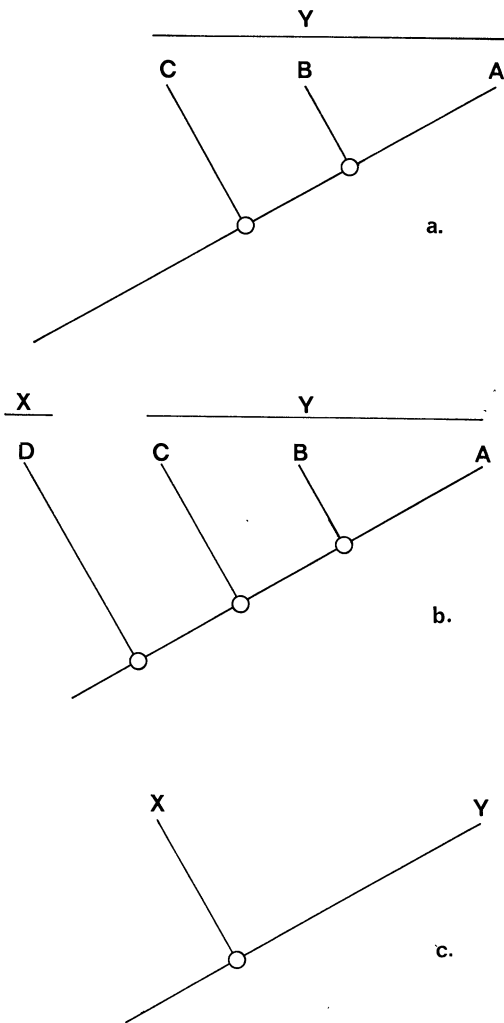


FIG. 1.—Relative levels of universality. Phylogeny “b” exists at a higher level of universality than phylogeny “a” because it has the potential falsifiers of “a” plus an extra and independent set of falsifiers associated with taxon D. Phylogeny “c” exists at the same level of universality as phylogeny “b.”

relative strength for corroborating or rejecting that hypothesis. An hypothesis of synapomorphy which has a greater number of potential falsifiers may be said to exist at a higher level of universality than another hypothesis of synapomorphy which has fewer potential falsifiers and the former

would have more weight than the latter *when they conflict*.

Under the theoretical system developed by Popper the mechanism by which an original hypothesis is formulated is not important. But in the practical world of systematics most hypotheses are generated from our world of experience, inductively. An investigator usually avoids trivial hypotheses by experience and a clear formulation of the objectives he has in mind. This may be thought of as the “pre-hypothesis” stage of systematics and is not scientific, strictly speaking, under Popper’s philosophy.

Once a hypothesis of homology is formulated from the world of experience it is tested in two phases: by its own set of potential falsifiers and by a set of potential falsifiers of the phylogenetic hypothesis to which it belongs as a proper subset (i.e., it is tested by other hypotheses of synapomorphy through the testing of the phylogenetic hypotheses which they corroborate). Both phases of testing must be done under the rules of parsimony, not because nature is parsimonious, but because only parsimonious hypotheses can be defended by the investigator without resorting to authoritarianism or apriorism.

In the first phase of testing (= attempting to falsify), any potential falsifiers thought to form proper subsets of the hypothesis of homology may be used (without reference to a phylogeny of which it may be a proper subset). Most of the potential falsifiers are morphological similarities and dissimilarities between the characters compared. This does not imply acceptance of a phenetic definition of homology such as that applied by the numerical school of taxonomy. Nor does it imply that only morphological attributes may be used. Rather, any observable differences and similarities between characters, including phenetic criteria, may be used to test an initial proposition that the characters compared are worthy of consideration as possible homologies and thus worthy of consideration as

possible falsifiers of phylogenetic hypotheses. Thus, this phase may be thought of as a tentative initial test which may be useful in sorting out those hypotheses which are not worthy of consideration.

Various criteria which may be applied during this phase of testing have been discussed by Simpson (1961), Hennig (1966), Mayr (1969), Sneath and Sokal (1973) and others. Any applicable criterion that was not used to formulate the hypothesis in the first place may be used during this phase in an attempt to falsify the hypothesis. This avoids circularity. For example, if similarity of topographic position is used to hypothesize homology, corroboration must be achieved by testing the hypothesis with some other criterion. Popper's philosophy does not provide for absolute falsification or absolute corroboration. Every species, and indeed every individual of a species, differs in some morphological respect from every other species. Any investigator can demonstrate differences and similarities between structures. The question of "how different" non-homologous characters are, and "how similar" homologous characters are becomes the *opinion* of the investigator. Hypotheses of homology based on overall similarity can always be rejected because no two structures are exactly similar. But, that hypothesis of homology based on morphological comparisons which has been rejected the least number of times relative to other possible hypotheses is to be preferred over these other hypotheses. Another way of looking at this process may be summarized thus: a hypothesis is proposed that two characters are homologous, and that hypothesis of homology inherently carries certain predictions. For example, it might predict that the structure will continue to be similar at finer and finer levels of morphological comparisons, or perhaps two rather dissimilar structures can be traced back to the same embryological structure. The greater the number of these predictions fulfilled by the hypothesis, the stronger will be the hypothesis of homology.

I think it is important to precision of methodology that some form of testing be done at this lower level of the problem rather than simply applying characters that look similar to a phylogenetic hypothesis. A good heuristic rule is that any chance to test should be taken. I do not suggest that morphological testing proves homology. Instead, testing may eliminate certain ill-founded ideas of homology at an early stage in the investigation, thus strengthening phylogenetic statements with which other homologies are associated. Finally, establishment of testing on this level may prove useful in further testing of incongruent synapomorphies as discussed later.

The highest level of corroboration during this phase of testing is failure to refute the characters as synapomorphies, that is, failure to observe any differences between the states of the characters *as formulated*. Another level might be reached in which two of three character states are not exactly similar as stated, but they are more similar to each other than either is to the third. Thus, we might like to investigate the possibility that these two characters are part of a transformation series (with one state plesiomorphous and the other apomorphous, although no initial judgment need be made as to which of the states is apomorphous). Lack of general and specific similarity of the characters of the organisms might be reason to reject the hypothesis of homology. If the hypothesis survives this round of testing it may be thought of as an unrejected morphological hypothesis of homology.

Several examples might be discussed. I might, for example, formulate a character state "spines present in the dorsal fin" and find that two of three taxa I am considering have spines while the other does not. It makes no difference that one species has three spines while the other has twelve because the character state specifies only the presence of spines and not their number. I would, based on this observation, be unable to reject a hypothesis of synapo-

morphy and might like to corroborate this hypothesis by checking the fine structure of the spines and their ontogeny. If the hypothesis is corroborated, then the alternate state, without spines on the dorsal fin, would assume the alternate plesiomorphous state in the transition series. This is because it is assumed (as a convention) that the dorsal fin itself is homologous in all three species at a level of universality that the investigator does not wish to consider. In comparing the foot of *Australopithecus* and *Homo* I might reject the structure of the foot as being synapomorphic because it is not exactly similar in the two genera. But, I might also observe that while there is not exact correspondence, these two structures are much more similar to each other than either is to the foot of *Pan*. Thus, we might suspect that they are parts of a transformation series. Failure to find synapomorphy in this example might be a result of our inability to formulate the correct hypothesis at the correct level of universality, but it is valuable nevertheless in pointing out characters and structures which may be potential synapomorphies or autapomorphies. As a last example, we might reject a hypothesis of homology between the anal fin modifications of osmerid and poeciliid fishes because of a lack of similarity in anything except their being anal fins (that is, although both are modified anal fins, they are modified in different ways).

TESTING PHYLOGENETIC HYPOTHESES AND THEIR PROPER SUBSETS

From an unrejected hypothesis of synapomorphy the investigator can proceed in two ways. (1) If there is no previous hypothesis of relationship the investigator may generate a hypothesis via the inductive process, and the synapomorphy logically becomes a proper subset of the phylogenetic hypothesis. *This does not provide a test of either the homology or the phylogeny.* It simply provides the investigator with a hypothesis of phylogenetic relationships

with which he can proceed to the second step. (2) If a hypothesis of phylogenetic relationship already exists, then the hypothesis of synapomorphy can be applied to this phylogeny as a test of that phylogenetic hypothesis and its proper subsets (that is, the synapomorphies which corroborate the phylogeny). It is a valid test only if its potential falsifiers do not intersect the potential falsifiers already present in the phylogenetic statement, i.e., it must be independent. Two outcomes are possible: (a) it fails to refute the hypothesis of phylogeny as stated and thus becomes another proper subset of the hypothesis; or (b) it refutes the hypothesis and its proper subsets.

If (a) is the result, the hypothesis has "proved its mettle" (Popper, 1968b, pp. 265-268), or it has been corroborated. The greater number of hypotheses of synapomorphy that are applied to a hypothesis of phylogeny without falsifying it, the more strongly corroborated is the phylogenetic hypothesis. The best corroborated hypothesis is preferred over its alternates.

If (b) is the result, the phylogenetic hypothesis has been falsified unless, of course, the test can be shown to be invalid. If the "incongruent synapomorphy" is shown to be either (1) a symplesiomorphy or (2) a non-homology, then it can be rejected as a valid test of the phylogenetic hypothesis, i.e., it would not be a valid refutation of the phylogenetic hypothesis.

In Figure 2 a hypothetical example of rejection of a character state as a synapomorphy is shown. In Figure 2a and 2b, phylogenies "a" and "b" were generated by induction from hypotheses that characters "1" and "2" are synapomorphies of their respective phylogenies. Both phylogenies have therefore been rejected. If the axiom of only one true phylogeny in nature is to be upheld then at least one of these phylogenies and its associated synapomorphy must be false. It is possible, of course that both are false. In attempting to refute both hypotheses (Fig. 2c) we find three

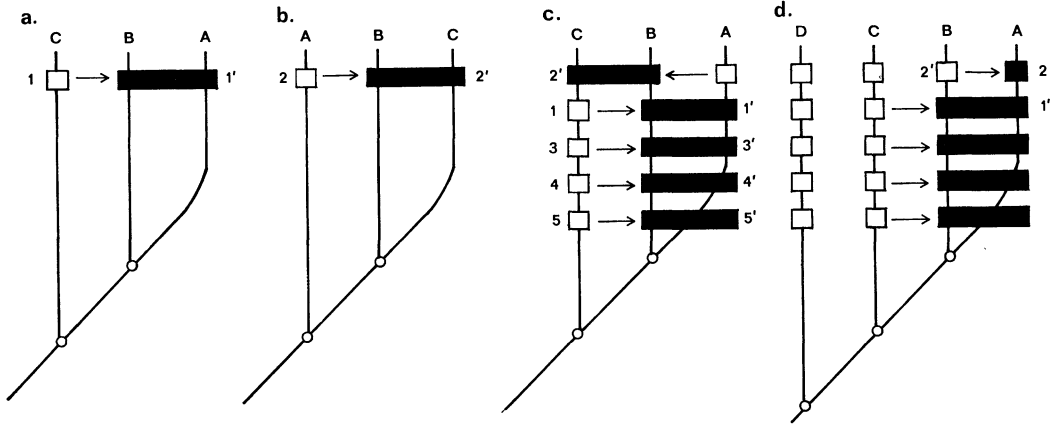


FIG. 2.—Rejection of an “incongruent synapomorphy” using the holomorphological method of Hennig (1966). Hypothesized synapomorphies and autapomorphies are shaded black while the alternate symplesiomorphies are open squares. See text for explanation.

additional proper subsets for phylogeny “a” but no additional subsets for phylogeny “b.” We may begin to suspect that synapomorphy “2” is not a valid synapomorphy and therefore not a valid test with which we can reject phylogeny “a.” Casting about in our world of experience we come up with another taxon (D, Fig. 2d) which has character state “2” but none of the other synapomorphous character states shared by taxa C and B. By adding taxon D to our analysis we have *raised the level of universality* of the phylogenetic hypothesis and we have demonstrated that it is more parsimonious to consider “2” as a symplesiomorphy at our original level of universality than to consider it a synapomorphy. It is interesting to note that the alternate character state “2” now becomes an autapomorphy of taxon A (or a synapomorphy if, from our world of experience, we know that taxon A is composed of 2 or more species) and that “2” becomes a *possible* synapomorphy at a still higher level of universality.

We might also reject “2” as a synapomorphy without raising the level of universality of the phylogenetic hypothesis by raising the level of universality of the morphological testing. For example, further ontogenetic investigation may show that

the structures expressed as “2” in the adult develop from different germ layers in the two taxa and thus cannot be considered homologous (although the germ layers might be).

Hennig (1966) presents several criteria which help the investigator to decide which of the alternate character states in a transformation series is plesiomorphous and which is apomorphous. It is my contention that application of such criteria will automatically raise the level of universality of the phylogenetic hypothesis. Thus, “2” might be determined to be the plesiomorphous expression because it appears earlier in development than its alternate “2’.” Application of this criterion automatically raises the phylogenetic level of universality by calling on our experiences with animals and/or plants outside the group of immediate interest. Calling on such experiences is perfectly valid as long as these experiences can be tested. It is important that such statements be made clearly so that authority is called upon only as a reservoir of specialized knowledge and not as apriorism.

BOCK, HOMOLGY, AND CLASSIFICATION

Bock (1973, pp. 386–389) discusses various aspects of homology, namely, the

criteria by which homologies are corroborated or falsified, and the relationship between homology and classification.

Bock (1973, p. 387) states: "similarity between features is the only criterion by which homologues can be recognized." This concept, says Bock, leads to difficulties in distinguishing similar but non-homologous characters from similar and homologous characters: "Unfortunately, the methods by which homologues are recognized and distinguished from nonhomologous features have low resolving powers." Now, much depends on what Bock means by the word "recognize." If recognize means to be aware of the similarity between two characters and thus to be aware that a potential homology exists, then I would agree that similarity is the sole recognizing criterion of homology. In other words, how can one be aware of a potential homology without perceiving similarities? But, Bock gives no means of testing our perceptions. On the other hand, if Bock uses recognize as a synonym for corroborate, then I must disagree. In this case, Bock would be saying that similarity is the only criterion for corroborating or falsifying hypotheses of homology. Similarities are important, but similarities are neither the only, nor the most severe test for homology. Hypotheses of homology may be tested with other hypotheses of homology. The process is reflected in building and testing hypotheses of phylogeny and their associated synapomorphies. Thus, homologies are *potential falsifying hypotheses* (Popper, 1968b, p. 87) of other independent homologies and the phylogenies with which these homologies are associated. Similarity is important in providing corroboration for the hypothesis of homology, thus qualifying the hypothesis as a falsifying hypothesis. Such a process is not circular, but is the process of reciprocal illumination (see Hull, 1967, for the difference). It provides a deductive test for a phylogenetic hypothesis of homology.

I have tried to show that my concept of the relationship between homology and

phylogeny/classification is non-circular. In contrast, I find Bock's (1973, p. 389) concept of the relationship between homology and classification distinctly circular: "Classifications are deduced on the basis of previously established homologies, . . . , which then serve as falsification tests for the classification." If one is to follow Popper, then one cannot say that evidence used to formulate a hypothesis can also be used to corroborate the hypothesis. That is circular inductive testing. Bock approaches a deductive test when he states (1973, p. 389): "more severe falsification tests may be undertaken by careful study of certain features." But, Bock has missed an important point: *no test existed in the first place.*

EVOLUTIONARY AND CLADISTIC CLASSIFICATIONS

It now appears, based on the recent papers of Mayr (1974) and Ashlock (1974), that evolutionary taxonomists accept Hennig's (1966) methodology as best suited for testing cladistic hypotheses. They contend, however, that evolutionary classifications convey more information about the classified organisms than do strictly cladistic classifications. On the surface, we might conclude that since evolutionary classifications supposedly contain more information, they would be preferred over cladistic classifications. This might especially be concluded under Popper's philosophy because the evolutionary classification would be the "bolder" hypothesis. But, as Popper has stated (1968b, p. 267): ". . . it is not so much the number of corroborating instances which determines the degree of corroboration as the severity of the various tests to which the hypothesis in question can be, and has been, subjected. But the severity of the tests, in turn, depends on the degree of testability, and thus on the simplicity of the hypothesis." I hasten to add that being simple does not imply being simplistic. The real question is: are evolutionary classifications testable?

Mayr (1969, 1974), Bock (1973), and Ashlock (1974) have stated that some form of genetic similarity, or inferred genetic similarity, should be incorporated into the phylogeny and thus into the classification of the taxa studied. The incorporation of these data is the fundamental difference between evolutionary and cladistic classifications and thus is central to the discussion. Problems associated with the application of these data to the modification of cladograms must lead to problems in testing the resultant classifications. One major problem is the lack of a precise and specified methodology in applying a measure of genetic similarity to the actual modification of a cladogram (Rosen, 1974). Those methods actually used seem very vague, even "arty" (Hull, 1970). Another major problem is the loss of information when phenetic and cladistic systems are combined (Hennig, 1966; Griffiths, 1972; Sneath and Sokal, 1973; Cracraft, 1974). Both Bock (1973) and Mayr (1974) concede this when they talk of maximizing the two "semi-independent" variables of phenetic (or genetic) and cladistic information. Can one maximize two semi-independent variables without losing some information from both? A third difficulty lies in the assumption that phenetic divergence may be equated with some determinable amount of genetic divergence. That this is not necessarily the case has been pointed out by Crowson (1970), and Rosen (1974).

Ashlock (1974) has outlined a general methodology for producing evolutionary taxonomic classifications. While he avoids specific methods of modifying cladograms, he does suggest that evolutionary phylogenies and classifications are produced in a two step process. First, a cladistic analysis is performed. This is followed by an anagenetic analysis. The result supposedly gives a measure of relative evolutionary change. Ashlock (1974, p. 96) states: "as the cladogram was established on the basis of these relatively few characters (i.e., synapomorphies or autapomorphies), anagenetic anal-

ysis would attempt to place as many other characters on the dendrogram as possible." He makes it clear that these are plesiomorphous or ambiguous character states left over from the cladistic analysis. The problem is that such character states are *untestable hypotheses*. They can only be made testable by raising the level of universality of the phylogeny to the level at which each exists as a synapomorphy or autapomorphy. Thus, Ashlock suggests that including untestable hypotheses of homology into a testable hypothesis of cladistic relationships results in a better classification. Can such a procedure result in a better hypothesis of classification? Is a hypothesis of classification derived from such a methodology testable? No, and I submit that such an analysis would obscure the only testable elements included in it, the cladistic relationships.

Would we not be better off if we raised the level of universality of the problem? And if we raised the level of universality to the ultimate degree, by inclusion of every known taxon, then where would these "anagenetic" character states exist? They would not exist at all because every character state of every character would already be incorporated into the analysis at the level at which it exists as an apomorphous character state, or, it would be so ambiguous as to defy analysis. Why should we incorporate, or try to incorporate, such anagenetic elements into our analyses? Cannot the description of taxa include all of the character states, apomorphous and plesiomorphous, that the investigator feels are important to the full understanding of the taxa he studies? Such a procedure would save the testability of the hypothesis of classification and permit the investigator to list those character states he or she feels are biologically meaningful.

Besides the difficulties outlined above, evolutionary taxonomy has an additional characteristic which seems to nullify Bock's contention that the theory of evolutionary taxonomy is consistent with "Popper's basic

ideas" (Bock, 1973, p. 382). The nullification of Bock's contention stems from the lack of the additional potential falsifier which is necessary to test evolutionary classifications as they are put forth today.

A classification purporting to show more than cladistic relationships would be preferred under Popper's philosophy if that classification had all the potential falsifiers of genealogical descent plus at least one additional falsifier independent of genealogical descent. If there is no independent falsifier then the falsifiability of a statement of classification, $f(cl)$, would equal the falsifiability of a statement of phylogeny, $f(phy)$. Or, $f(cl) = f(phy)$. A logical implication of Bock's line of thought is that $f(cl) > f(phy)$ and that overall genetic similarity provides the additional falsifier. But, is this an independent falsifier? Bock (1973) and Mayr (1974) have termed genetic similarity and phylogeny semi-independent variables. Is genetic similarity even a semi-independent variable? I suggest that genetic similarity is either the product of descent from a common ancestor or that it results from convergence at the genome level. That is, it is either a proper subset of genealogical descent or has nothing to do with phylogenetic relationships at all (except to produce error). So, genetic similarity is like any other kind of similarity, it is either apomorphic, plesiomorphic, or nonhomologous. Thus, it cannot provide the additional falsifier. I conclude that $f(cl) = f(phy)$. A logical consequence of this conclusion is that classifications and phylogenies must mirror each other and that the method of falsifying a classification is to refute the phylogeny with which it is associated.

The difficulties embodied in that part of evolutionary taxonomy which differs from phylogenetic systematics lead to a general lack of testability of evolutionary classifications. Claims that evolutionary taxonomy can exist within Popper's philosophy are, in my opinion, invalid. Until such time as evolutionary taxonomists demon-

strate the testability of their classification, we may take the advice of Karl Popper when he states (1968b, p. 277): "those theories which are at too high a level of universality, as it were (that is, too far removed from the level reached by the testable science of the day) give rise, perhaps, to a 'metaphysical system.' In this case, even if from this system statements should be deducible . . . which belong to the prevailing scientific systems, there will be no *new* testable statements among them, which means that no crucial experiment can be designed to test the system in question."

SUMMARY WITH A RECOMMENDATION

(1) Although the relationship between phylogenetic homology and phylogeny may be circular within an inductive hypothesis testing philosophy of science, it is not circular within the deductive hypothesis testing philosophy advocated by Popper (1968a, b).

(2) The terms apomorphic and plesiomorphic (and their derivatives) convey precise concepts which are logical derivations of a phylogenetic definition of homology. As such, they should be substituted for the word homology in systematic studies. This distinguishes the concept of homology used in the study from all other concepts of homology. It also makes the conditional phrase of Bock (1969, 1973) unnecessary.

(3) Only synapomorphies can be used to test hypotheses of phylogeny, and a synapomorphy which corroborates a phylogeny becomes a proper subset of that phylogeny.

(4) Tests of phylogenetic hypotheses are valid if the potential classes of falsifiers of the synapomorphy used to test the phylogeny do not intersect the potential classes of falsifiers which are already proper subsets of the phylogenetic hypothesis.

(5) Production of a phylogenetic hypothesis via induction does not constitute a valid test of the phylogeny or a test of the synapomorphy used to generate the phylogeny.

(6) Hypotheses of synapomorphy which refute a phylogeny also refute all hypotheses of synapomorphy which form proper subsets of the rejected phylogeny. Such a test is valid unless the supposed synapomorphy is demonstrated to be either a plesiomorphy or a nonhomology at the original level of universality of the phylogenetic hypothesis. This demonstration can only be accomplished by raising the level of universality of the problem.

(7) The phylogenetic hypothesis which has been rejected the least number of times is preferred over its alternates.

(8) The classical evolutionary classification system advocated by Bock (1973), Mayr (1974), and Ashlock (1974) is invalid under Popper's philosophy and will remain invalid because concepts such as genetic similarity, phenetic similarity, adaptive breakthrough, and evolutionary divergence are not independent of genealogical descent and cladistic relationships. Thus, they cannot independently alter a classification based on genealogical relationships. Attempts to alter these relationships will lead to apriorism because of the vagueness of the methodology employed. Thus, this system is rejected in favor of a system wherein classification mirrors phylogeny.

ACKNOWLEDGMENTS

I accept full responsibility for the conclusions presented but I do not claim originality for the ideas. This paper has grown from dialogues with my colleagues at the American Museum of Natural History and their views have significantly influenced my own. Drs. Donn E. Rosen and Gareth Nelson were particularly helpful in critically evaluating my ideas. Both they and Michael K. Oliver read and improved earlier versions of

the manuscript. Financial support from the American Museum of Natural History and City University of New York is gratefully appreciated.

REFERENCES

- ASHLOCK, P. D. 1974. The uses of caldistics. *Ann. Rev. Ecol. Syst.* 5:81-99.
- BOCK, W. J. 1969. The concept of homology. *Ann. N.Y. Acad. Sci.* 167:71-73.
- BOCK, W. J. 1973. Philosophical foundations of classical evolutionary classification. *Syst. Zool.* 22:375-392.
- CRACRAFT, J. 1974. Phylogenetic models and classification. *Syst. Zool.* 23:71-90.
- CROWSON, R. A. 1970. Classification and biology. Heinmann Educational Books, Ltd., London.
- HENNIG, W. 1966. Phylogenetic systematics. Univ. Illinois Press, Urbana.
- HULL, D. L. 1967. Certainty and circularity in evolutionary taxonomy. *Evolution* 21:174-189.
- HULL, D. L. 1970. Contemporary systematic philosophies. *Ann. Rev. Ecol. Syst.* 1:19-54.
- MAYR, E. 1969. Principles of systematic zoology. McGraw-Hill, New York.
- MAYR, E. 1974. Cladistic analysis or cladistic classification? *Z. Zool. Syst. Evolut.-forsch.* 12: 94-128.
- OWEN, R. 1848. Report on the archetype and homologies of the vertebrate skeleton. Rep. 16th Meeting British Assoc. Adv. Sci.:169-340.
- POPPER, K. R. 1968a. Conjectures and refutations: the growth of scientific knowledge. Harper Torchbooks, New York.
- POPPER, K. R. 1968b. The logic of scientific discovery. Harper Torchbooks, New York.
- ROSEN, D. E. 1974. Cladism or gradism?: a reply to Ernst Mayr. *Syst. Zool.* 23:446-451.
- SIMPSON, G. G. 1961. Principles of animal taxonomy. Columbia Univ. Press, New York.
- SNEATH, P. H. A., AND R. R. SOKAL. 1973. Principles of numerical taxonomy. W. H. Freeman and Co., San Francisco.

E. O. WILEY

*Department of Ichthyology
The American Museum of Natural History
New York, New York 10024*