Popper and Systematics

OLIVIER RIEPEL

Department of Geology, Field Museum, 1400 S. Lake Shore Drive, Chicago, Illinois 60605-2496, USA; E-mail: rieppel@fieldmuseum.org

The philosophy of Karl Popper is frequently appealed to either in support or in defense of theories and methods of systematic biology. Perhaps the best illustration of Popper’s prominence in this field is the collection of papers published in Systematic Biology in June 2001 (de Queiroz and Poe, 2001; Faith and Trueeman, 2001; Farris et al., 2001; Kluge, 2001b).

REFERENCES


In a very general sense, Popper’s philosophy can be seen as a world view, attractive to fellow philosophers, sociologists, historians, politicians, and scientists. Consequently, Popper can be read at different levels. In Objective Knowledge, Popper (1974b:80) appealed to “the critical discussion of competing theories,” which portrays the criterion of testability in its weakest sense (“the criterion of demarcation cannot be an absolutely sharp one” [Popper, 1997:187]); not to assert positive evidence in favor of a cherished hypothesis but to be critical and to seek flaws or mistakes rather than confirmation. Things get more serious when we “look upon the critical attitude as characteristic of the rational attitude” (Popper, 1974b:31), where rationality is the capacity of logical thinking and where Popper recognized “deductive logic as the organon of criticism” (Popper, 1974b:31). For here we cross the bridge from common sense philosophy to a full-blown theory of knowledge. In current discussions of implications of Popper’s philosophy of science for systematics, terms such as explanatory power, degree of corroboration, severity of test, and logical (or absolute) probability have been used in connection with formalisms, which means that today systematics is linked to the full-fledged apparatus of Popper’s hypothetico-deductive (where testable observation statements are deduced from a hypothesis).

De Queiroz and Poe (2001:306) stated that “Popper’s corroboration is based on the general principle of likelihood,” or that likelihood is the foundation of Popper’s concept of corroboration (de Queiroz and Poe, 2001:308), which seems counterintuitive given that likelihood methods have been identified as inductive (Edwards, 1992) or abductive (Sober, 1988). De Queiroz and Poe (2001) were countered by Kluge (2001ib:323), who found likelihood analysis to be verificationist inductive, as opposed to parsimony analysis, which is “practiced in terms of Popperian testability” and hence is falsificationist-deductive (deductive logic does not necessarily exclude a verificationist attitude, as seen in the discussion of the Hempel–Oppenheim model of scientific explanation). Recognizing that there is normally no “demonstrable” Popperian falsification in cladistics, Faith and Trueman (2001:331) decoupled their interpretation of Popperian corroboration from the “popular falsificationist interpretation of Popperian philosophy.” They moved on to recognize supposed corroboration cladistics as merely goodness-of-fit of evidence. Farris et al. (2001) in turn took issue with Faith and Trueman’s “Popper,” debating goodness-of-fit versus data as Popperian evidence.

Here, I do not address issues of methodology, i.e., whether parsimony or likelihood analysis should be preferred, which algorithms should or should not be used in cladistic inference, or whether permutation tail probabilities are meaningful in phylogeny reconstruction. What I am concerned with is whether Popper’s philosophy of science is at all relevant to these methodological discussions. In that context, the following questions must be asked and answered. What is deduction? What is falsification? How do explanatory power, degree of falsifiability, and severity of test relate to degree of corroboration? What is the meaning of Popper’s propensity interpretation of probability? What, if anything, has all this to do with phylogenetic analysis?

A Bit of History

To better understand how Popper’s philosophy can or cannot be brought into a discussion of systematics, it is helpful to review Popper’s place in the history of philosophy of science. Popper was a philosopher, not a scientist. He wrote on the logic of scientific discovery, not on applied physics or on phylogenetic systematics. He wrote about problem solving in science, but he did not himself solve scientific problems (e.g., [Popper 1983:230], wrote of degree of corroboration that it “has little bearing on research practice”). A philosopher, whether approvingly appealed to or criticized, “must be understood immanently, that is, from his own point of view” (Carnap, 1997:a:41). Any characterization of Popper’s philosophy must be a coherent one, or the departure from his original ideas must be made explicit. A good way to achieve a better understanding of Popper’s work is to contrast it with that of his peers, understanding that Popper formulated his ideas in contradistinction to those propagated by the Vienna Circle and in contradiction to his shadow, Ludwig Wittgenstein. It also helps to understand that The Logic of Scientific Discovery (Popper, 1992) was abstracted from a larger work entitled The Two Basic Problems of the Theory of Knowledge (Popper, 1979). These two (interconnected) problems were the problem of induction and the demarcation of science from metaphysics.

Logical positivism is a label that is commonly attached to members of the Vienna Circle and to those of the Unity of Science movement that grew out of it, among which were Carnap and Hempel (Hung, 1992; Carnap, 1997:a). Another prominent logical empiricist, A. J. Ayer, counted Karl Popper in that group as well (Ayer, 1959). The young Popper indeed courted the Vienna Circle but was never admitted (Edmonds and Eidinow, 2001). Although there are profound differences between his philosophy and that of, say, Rudolph Carnap, there also exist profound similarities: “His basic philosophical attitude was quite similar to that of the Circle” (Carnap, 1997a:31). All the authors involved in logical empiricism had one goal in common: the demarcation of science from metaphysics. To achieve this demarcation both camps used the same tactics: they focused on Wittgensteinian truth conditions (truth anchored in logic) more than on the evidence condition (knowledge anchored in observation); science was demarcated from metaphysics on the basis of the syntactic form of statements (the analysis of the grammatical [and logical] structure of sentences) and semantic analysis (the analysis of the meaning of terms) rather than on epistemological grounds (observation and justification; Laudan and Leplin, 2002). Beyond this common ground, Popper’s thinking diverged from the doctrines of the Vienna Circle. Carnap declared theories scientific when they (logically) implied observation statements; for Popper, theories became scientific when they (logically)
entailed the negation of observation statements. Carnap appealed to the principle of verifiability as the criterion by which to establish the meaning of a sentence while declaring metaphysical sentences as meaningless. Popper objected and used his falsifiability criterion to distinguish science from metaphysics. The members of the Vienna Circle dismissed the problem of induction as a pseudoproblem on the grounds that there is no solution for it in the way it is ordinarily conceived (Ayer, 1952). Popper sought a solution to the logical problem of induction instead.

**Syntax, Semantics, and the Problem of Induction**

The opening phrase in Popper’s (1974b:1) book *Objective Knowledge* reads “I think I have solved a major philosophical problem, the problem of induction.” The problem of induction was formulated by David Hume and can be presented as follows. As far as we know, the sun has risen every morning for as long as there have been human beings on planet Earth. Are we justified to conclude from these innumerable observations that the sun will rise again tomorrow? (Whether or not the sun will rise again tomorrow is a contingent fact of nature, not a logical necessity. This can be shown by the fact that the statement “it is not the case that the sun will rise again tomorrow” is not self-contradictory, which it would be if tomorrow’s sunrise were a logical necessity.) Popper’s perception of this problem was that our knowledge of past events, based merely on inductive generalization and on nothing else, constitutes a historical account of those events, and such a historical account has no predictive power that is based in logic (i.e., that is deductive; compare “degree of corroboration as a historical account” below). In other words, the inductive generalization that the sun will rise every day in the future just as it has in the past is, without additional knowledge, a merely historically contingent generalization that expresses no necessity in the natural course of events.

This interpretation changes if Newton’s laws of interplanetary motions are taken into account, for it is the appeal to universal lawfulness that imparts necessity on the natural course of events. The appeal to Newton’s laws establishes what Sober (1988) called a three-place relation: The past experience of the sun rising every morning is explained by Newton’s laws, which in turn allow the prediction that the sun will rise again tomorrow. This prediction is justified to the degree that Newton’s laws are “true,” i.e., they will hold tomorrow and ever after.

Induction has been characterized as a process of reasoning that moves from the particular to the general, whereas deduction is a process of reasoning that moves from the general to the particular. Ever since Peirce (de Waal, 2001), it has been said that abduction is about what could be the case (given the evidence at hand, the butler could be the murderer), induction is about what is the case (emeralds are green, not grue or blue; Goodman, 2001), and deduction is about what must be the case. In other words, induction is (more or less) supported by evidence, but deduction is conclusive because it is based in logic.

Both Carnap and Popper agreed that no finite number of singular observations would suffice to fully verify a generalization. Carnap therefore switched from the concept of verification to that of degree of confirmation (and its complement, degree of disconfirmation; Mayhall, 2002), whereas Popper turned his back on induction and used deduction as the hallmark of scientific reasoning. There is a catch involved with this move, however, because deductive reasoning happens in logical space. Thus, deductive (i.e., logical) relations obtain only between sentences or complexes thereof, not between words and things (nor between observation and justified belief).

A simple form of deductive reasoning is the Aristotelian categorical syllogism: (1) all mammals are vertebrates, (2) all nudibranchs are mammals, therefore (3) all nudibranchs are vertebrates (Weiner, 1999). As long as the premises (1 and 2) are agreed upon, the conclusion (3) must follow necessarily. The deduction is logically valid but factually wrong (unsound). Thus, deduction can deliver truth independent of any epistemic status (independent of observation or justification).

Propositional logic (Weiner, 1999) is based on universal conditional statements such as “(for every A) if A is a dog, then it barks.” If the conditional and the premise that “A is a dog” are true, then it must also be true that “A barks.” This valid deductive argument says nothing about a particular dog, only something about the class of all dogs (which means that barking is a universal, essential, or simply necessary property of the class of all dogs). It is logically not impossible to imagine duckogs that quack (Kripke, 2002). I can say, “(for every A) if A is a duckog, then it quacks,” “A is a duckog,” therefore “A quacks.” Everything concluded so far is based on deduction; no observation is yet involved. But now we can go beyond deduction and appeal to our senses to see whether the conditionals and the premises are true or false (whether there is reason to believe that they are true or false): I have never seen a duckog, but I have seen many dogs, and all those that I have seen barked, never quacked. This A here (to which I point) definitively resembles a dog, and it does not bark right now because it is fast asleep. Although it is highly unlikely that we should encounter a dog in Central Park tomorrow that quacks, it is logically not impossible. However, if you and I can reach an agreement and accept that to the best of our observational powers we have indeed sighted what looked like a dog that quacked, we are logically compelled to accept one of two possible conclusions: either the conditional “(for every A) if A is a dog, then it barks” is not universal, or if “dog” is a name that rigidly refers to As that never quack but only bark, which is what names are supposed to do, then we should not call the animal we sighted a dog; we might call it a duckog.

Cladistic analysis has been labeled deductive (Crother, 2002), but valid deductive inference, as much as a Popperian test, is a complex problem that requires consideration of both the structure of the language used and
the meanings of the terms that are used in that language. Propositions used as a basis for deduction must conform to a certain syntactical structure. The categorical syllogism presented above requires that the subject of the first (major) premise (all mammals [subject] are vertebrates [predicate]) appears as predicate in the second (minor) premise (all nudibranchs [subject] are mammals [predicate]). Similarly, to qualify as a conditional, the statement must be an “if..., then...” statement (or a syntactically equivalent statement such as “...not..., unless...”). Beyond that, semantic analysis of meaning is also required. The meaning of the name “dog” obtains from the attachment of that name to a physical object (the meaning of the proper name Fido obtains from the attachment of that name to my dog to which I point). But how names are attached to things, or more generally how names are used, is governed by rules and conventions. I live in a linguistic community in which the name “dog” is attached only to animals that bark, never to animals that quack. If in that same linguistic community somebody would attach the name dog to an animal that quacks, she would violate the rules and conventions governing the use of the name dog in that community. However, linguistic as well as scientific communities change over time. The name “atom” was attached to the physical world in a different way for Democritus than for Niels Bohr.

Popper believed to have solved the problem of induction by reformulating it. His solution had two parts. First he subdivided the problem into two components, the psychological versus the logical problem of induction. Popper considered the fact that we expect future regularity on the basis of experienced past regularity a psychological component, the roots of which he eventually traced back to evolutionary epistemology (Popper, 1974b). The other component, i.e., the question of whether there could be a rational (i.e., logical) reason to prefer one scientific theory over another, he dubbed the logical problem of induction, and he claimed to have solved that problem by reformulating it in terms of universal propositions. He did this because the syntactical form of universal propositions allows the deduction of the negation of observation statements, i.e., universal theories forbid certain things to happen. The statement “all unicorns have a horn on their forehead” cannot be used as the basis for the deduction “a unicorn is here now, or anywhere,” because the existence of a unicorn in our world of experience can only be established by observation. By contrast, the statement “all ravens are black” can be logically transformed to “there is no one thing that is not black and is a raven,” and that statement in turn can be logically transformed to “no white raven is here now” (Popper, 1989:385). That there should not be a white raven here now is a predictive statement that is obtained on purely logical grounds; no steps have been taken yet to attach that statement to the physical world. But if we accept that we do in fact observe a white raven here and now, then that negation of the negated observation statement contradicts the universal proposition in logical space, and falsification occurs. Logically, this falsification is absolute and conclusive; deduction is about what must be the case. The symmetry that characterizes inductive systems (where a theory is more or less confirmed or disconfirmed) has given way to an asymmetry of falsification based on deductive reasoning (a theory can never be shown to be true but can be shown to be false). Practically, falsification requires that everybody involved accepts that there indeed is a white raven here now (and of course this could be wrong). It must be agreed that this bird is a raven and not some other bird precisely because it is white, and it must also be agreed that this bird is white and not a shade of gray that some may still consider as black. The meanings of “black,” “white,” and “raven” must be fixed, at least to such a degree that a contemporary linguistic and/or scientific community can engage in a meaningful discourse. This, of course, is falsificationism in its strong sense, something cladists have agreed cannot apply to systematics (Farris, 1983), hence the recourse to parsimony analysis. The question to be discussed below is whether parsimony analysis satisfies Popperian falsificationism even in a weak sense.

According to Popper, theories are scientific if they are falsifiable, but falsifiable on the basis of deduction. For this to be possible, scientific theories must have the syntactical form of universal propositions (or must be linked to statements of such a syntactical form). To prevent the immunization of theories, the potential and acceptable falsifying instance must be specified in advance (Popper, 1989:38), such that if it occurs, there is no backdoor open that would allow us to save the theory from falsification by a shift of meaning of the words that report on the observable state of affairs.

Inductive support works symmetrically, confirming or disconfirming theories or hypotheses to a greater or lesser degree. An empirically confirmed hypothesis A disconfirms a rival hypothesis B to the degree to which B is inconsistent with A. So if \( x \) confirms hypothesis A, \( y \) confirms hypothesis B, and \( x \) carries a greater evidentiary weight than \( y \), then A is confirmed and B is symmetrically disconfirmed. In contrast, Popperian falsification works asymmetrically: If it occurs (if it is accepted that it has occurred), it is conclusive. The reason for this asymmetry is that this falsification is based on a contradiction that occurs in logical space. In parsimony analysis, we have alternative hypotheses of relationships, the number of which is determined by the number of terminal taxa. Among those, we choose the one that is supported or confirmed by the largest number of congruent characters. This most-parsimonious hypothesis symmetrically disconfirms alternative hypotheses to the degree that these are inconsistent with the most-parsimonious hypothesis. “Parsimony operates by finding a pattern of relationships that is most consistent with the data” (Kluge and Farris, 1969:7, emphasis added). The criterion of consistency was later replaced by explanatory power, meaning that parsimony maximizes the number of character descriptions that are explained ad hoc as homoplasies (Farris, 1983). This point is perfectly
valid, but it cannot be linked, nor does it need to be linked, to Popperian falsificationism. The meaning of “explanatory power” as used by Farris (1983) is not coextensive with the meaning as Popper used this term. Support, or lack thereof, is also asymmetrical rather than asymmetrical in likelihood analysis.

More recent debate has turned on the issue of whether goodness of fit, the nature of the data, or the improbability of evidence is relevant to Popper’s writings (Faith and Trueman, 2001; Farris et al., 2001). Parsimony is defended in terms of Popper’s explanatory power (minimizing ad hoc auxiliary hypotheses; Farris, 1993) which leads back to Popper’s (1992:145) “principle of parsimony in the use of hypotheses.” Traditionally, parsimony has been defined as “a methodological principle dictating a bias towards simplicity in theory construction” (The Oxford Companion to Philosophy, 1995), which is not the same as Popper’s principle that requests restraint from indulgence in ad hoc auxiliary hypotheses. In the German edition of The Logic of Scientific Discovery, Popper (1976a) used Einfachheit (simplicity) for parsimony in its traditional sense, but Grundsatz des sparsamsten Hypothesengebrauchs for “principle of parsimony in the use of hypotheses.” Popper was dissatisfied with this traditional sense of parsimony (simplicity) and replaced the conventional/methodological rule that the simplest (most parsimonious) hypothesis should be selected among competing ones with his normative rule that the most severely testable hypothesis should be preferred, i.e., the theory with the highest degree of falsifiability. Auxiliary ad hoc hypotheses can serve to immunize a theory from falsification, which is why they must be avoided unless they increase the degree of falsifiability. This argument is tied to Popper’s notion of falsifiability and the requirement not to immunize a theory from falsification; none of this argument applies to parsimony analysis. What parsimony analysis delivers is maximal consistency, which was interpreted as one measure of goodness of fit by Faith and Trueman (2001). In this context, parsimony must be viewed as a model, which can only be justified by reference to the background assumptions that parsimony analysis implies about character evolution (Sober, 1994).

THE HEMPEL–OPPENHEIM MODEL OF SCIENTIFIC EXPLANATION

Both Hempel and Popper championed a model of scientific explanation based on deduction. If cladistics is deductive, it must satisfy the nomological deductive model of scientific explanation originally proposed by Hempel and Oppenheim (as was claimed to be the case by Kluge, 2001a).

According to Popper, a scientific explanation obtains if the state of affairs to be explained is logically entailed by the explanation “the use of a theory for predicting some specific event is just another aspect of its use for explaining such an event” (Popper, 1976b:124). In practice, a theory must be complemented by initial and/or boundary conditions (such as an experimental setup) to yield testable predictions. This kind of scientific explanation has become known as the Hempel–Oppenheim nomological deductive model of scientific explanation (Hempel, 1965; Hempel used this model in a verificationist context, whereas Popper used it in a falsificationist context). This model also addresses the logical relations between sentences (or complexes thereof), and it divides the scientific explanation into two components, the explanans (the sentences that constitute the explanation) and the explanandum (the sentences that describe the state of affairs to be explained). The explanans contains statements of antecedent conditions and general laws. Hempel’s (1965) use of “antecedent conditions” shows that his model of scientific explanation can be rendered in the form of a universal conditional (this is the elliptical form of the N-D model sensu Hempel, 2001). For example, “if this is a magnetized piece of iron (antecedent), then it will attract iron filings (consequent).” The antecedent thus contains the initial or boundary condition, i.e., the magnetized piece of iron, and it makes reference to a universal law, the law of magnetism. Prediction sensu Popper requires deduction, and deduction is a logical operation; so the prediction that iron filings will be attracted by a magnetized piece of iron is possible only because the antecedent logically entails the consequent (by virtue of referring to a universal law). As stated by Hempel (1965:247), “the explanandum must be a logical consequence of the explanans.”

Popper’s falsificationism can now be seen to be linked to the modus tollens form of argument, which is based on the logical entailment of the consequent by the antecedent in a universal conditional statement (this is also the way in which Carnap [1995] rendered universal theoretical laws). If this is a piece of magnetized iron (if P), then it attracts iron filings (then Q); if iron filings are not attracted (if not-Q), then this is not a piece of magnetized iron (then not-P). From this example we obtain “if P, then Q,” “not-Q” “therefore P” (modus ponens), and “if P, then Q,” “not-Q,” “therefore not-P” (modus tollens). This example also shows that the failure to attract iron filings does not quite falsify the law of magnetism in an empirical sense. Maybe the piece of iron we thought was magnetized is in fact not made of iron, maybe the filings that fail to be attracted are not made of iron, or maybe the process of magnetization itself was unsuccessful not because the metal used was not iron but because a faulty procedure was used to magnetize it. These potential problems all relate to the issue of how names or statements can successfully be attached to physical objects or observable states of affairs, and because the attachment of words to things will never be absolutely perfect, there will never be an unequivocal falsification of any universal proposition empirically (i.e., practically). This is not a problem of logic; it is a problem of bridging the (logical; Körner, 1970) gap between words and things. Popper (e.g., 1974c, 1992) always acknowledged that the unequivocal empirical falsification of any theory is impossible and that we live in a layered world of competing theories of varying degrees of universality.
but he also emphasized that his interest was centered on the logic of scientific discovery, not on its empirical fuzziness. However, this is precisely what cladists have to deal with—the fuzziness of the evidence manifest in homoplasy.

The use by Popper of the modus tollens form of argument in his demarcation of science from metaphysics has important consequences for systematics. The modus tollens is a logical argument, and it is decisive because it states necessary conditions: “if $h$, then not-$e$,” “$e$, therefore not-$h$” (this is the form in which Popper used the modus tollens, where $h = \text{a hypothesis, } e = \text{an observation sentence}$). It holds in all possible worlds that obey the same rules of logic. Proponents of likelihood methods must consider that there is no probabilification of the modus tollens (Sober, 1999). Their method is not an appeal to falsification or deductive inference but rather a statement of probabilities for the evidence with and without the hypothesis. Proponents of parsimony analysis must consider that the modus tollens is based on deductive entailment, which is a logical relation, whereas statements that concern historical states of affairs, such as those of systematics and phylogeny, are historically contingent. There is therefore no room for the purely logical entailment of the explanandum by the explanans. There is no antecedent in systematics or phylogeny reconstruction that can logically entail its consequent, i.e., there is no deductive link between a hypothesis of relationships and the character distribution across the terminal taxa it concerns (Sober, 1988). Premises that would allow the prediction (or retrodiction) of how certain species or token fossils are interrelated would require a connection to some sort of generality: “In so far as we are concerned with the historical explanation of typical events they must necessarily be treated as typical, as belonging to kinds or classes of events. For only then is the deductive method of causal explanation applicable” (Popper, 1976b:146; emphasis added). If species or taxa in general are historically unique (i.e., are historical particulars), then deductive reasoning cannot be applied to systematics or phylogeny reconstruction (Sober, 1988). To even consider the applicability of deductive logic to systematics, species would have to be instantiations of some kind or elements of some class, which some have argued might be achieved by considering species as natural kinds.

It is one thing to consider species as natural kinds but another to claim a deductive link between a hypothesis of relationships and character distribution across the species it entails. Kripke (2002) essentially characterized a natural kind as calibrated on standard members that share at least some of a cluster of characteristics: “these properties need not hold a priori of the kind; later empirical investigation may establish that some of the properties did not belong to the original sample, or that they were peculiarities of the original sample, not to be generalized to the kind as a whole” (Kripke, 2002:137).

In biology, this explanation translates into the homeostatic property cluster kinds that are species (Boyd, 1999). The properties that characterize these natural kinds are “homeostatic in that there are mechanisms that cause their systematic coexistence or clustering,” yet “this clustering is the result of incompletely understood mechanisms that govern an embryo’s development” (Wilson, 1999:197f). Wittgensteinian family resemblance (Wilson, 1999:196; see also Glock, 2000:120ff) is just not enough to warrant a deductive link, let alone the incomplete knowledge of initial and boundary conditions of inheritance and ontogenetic development.

Even if descent with modification (Farris et al., 2001; Kluge, 2001a, 2001b) is taken as a universal law, it forbids little more than spontaneous generation. (I believe that spontaneous generation has taken place at least once, such that it is logically not impossible that it could have occurred more than once, hence the talk of life on Mars. “What we call the evolutionary hypothesis…is not a universal law” Popper [1976b:106f] It could be argued that in conjunction with the appropriate initial conditions, it might be possible to explain (by way of prediction or retrodiction) a specific speciation event. However, these initial conditions are so complex that “the properties of new species which arise in the course of evolution” remain “unpredictable” (Popper, in Popper and Eccles, 1977:29).

POPPER AND EVOLUTIONARY THEORY

Popper tied his falsificationism to universal propositions (laws) yet stated that evolutionary theory is no such law. Systematists who appeal to Popper emphasize that Popper acknowledged testability and falsifiability for singular statements as well and therefore claim testability for singular taxonomic statements. Is this claim justified?

Popper frequently referred to evolutionary theory and in that context considered the applicability of his falsificationism. “All ravens are black” is a universal proposition and hence of a syntactical form that characterizes universal laws. By contrast, “all vertebrates have one common pair of ancestors” is not a universal proposition because it does not hold in all possible worlds, even if these obey the same rules of logic. Pegged to time and space in its meaning by virtue of reference to common ancestry, the statement refers “to the vertebrates existing on earth” only, not “to all organisms at any place and time which have that constitution which we consider characteristic of vertebrates” (Popper, 1976b:107; in the latter quote “vertebrates” are presented as a class or a natural kind). Systematists dealing with spatiotemporally restricted entities must be concerned about the testability of singular statements.

Popper was very consistent in his admission that singular statements are in fact testable, but for this to be true, they again need to satisfy certain conditions, i.e., they need to be tokens of the same type. Following C. S. Peirce (de Waal, 2001), we can think of words as signs, and such a sign can denote a quality (red), a specific particular (individual) object (“Fido” is a sign standing for my dog), or a general type (the class [or natural kind] of all dogs). An office furniture
sues is to be found in key to an understanding of Popper’s position on these is-

of the evolution of life on earth” (Popper, 1980:611). The “historical sciences such as paleontology, or the history the accusation of having denied scientific character to a universal law or that it could be tested, but he rejected tested. He also did not state that evolutionary theory is instead said that statements derived from them can be tested. He did not say that the descriptions of unique events can be tested but instead said that statements derived from them can be tested. He also did not state that evolutionary theory is a universal law or that it could be tested, but he rejected the accusation of having denied scientific character to “historical sciences such as paleontology, or the history of the evolution of life on earth” (Popper, 1980:611). The key to an understanding of Popper’s position on these issues is to be found in The Poverty of Historicism (Popper, 1976b:108). Whereas the evolution of life on earth is a unique process, and its description therefore a singular historical statement, this process nevertheless proceeds “in accordance with all kinds of causal laws, for example, the laws of mechanics, of chemistry, of heredity and segregation, of natural selection, etc.” Thus, it is possible to test statements that bear on the process of evolution if these can be linked to any such lawfulness. However, singular taxonomic statements pronouncing phylogenetic relationships do not allow such connections; the known laws of inheritance or natural selection do not allow us to predict or retrodict character distribution across taxa that would falsify hypotheses of relationships. If they did, there would be no need for parsimony or likelihood analysis.

**Degree of Confirmation Versus Degree of Corroboration**

Given that hard falsificationism (falsification in logical space, an asymmetrical all-or-nothing affair) cannot obtain in cladistics (Farris, 1983), cladists have turned to Popper’s notion of degree of corroboration as the method for choosing between competing hypotheses of relationships. To understand Popper’s notion of corroboration, it is necessary to introduce his notion of background knowledge (b). Not everything can be investigated or tested at the same time. Science must partition its problems. Popper referred to the body of scientific knowledge that was not questioned in any particular test situation as background knowledge. Some parts of the background knowledge may be more or less relevant to a specific test situation (b) also contains initial conditions [Popper, 1989:288], and when Popper talked about varying the background knowledge [e.g., Faith and Trueman, 2001:430], he meant the latter, but there is no room in Popper’s philosophy to arbitrarily minimize background knowledge (to maximize the relative empirical content of a hypothesis) as is sometimes done in the cladistic literature (e.g., Kluge, 2001b; see comments of Faith and Trueman, 2001). Not everything can be investigated or tested simultaneously, which is another reason why the ultimate and finite verification or falsification of a natural law, theory, or hypothesis cannot be accomplished in practice. Any testing procedure will make use of background knowledge in some way, and fault might have to be attributed to the background knowledge rather than to the theory under test. For example, I might use a ruler to measure the expansion of a piece of metal as a function of increasing heat. In so doing, I rely on the background assumption that the ruler is inert under varying thermal conditions, but that assumption may be wrong. So I might find that a certain metal does not expand with increasing temperatures, but that result might reflect the fact that the ruler I used expanded to the same degree as the piece of metal I was measuring.

Carnap (1995) invoked the notion of degree of confirmation to characterize the status of a theory or hypothesis. The degree of confirmation increases with the number (and kinds; Nagel, 1971) of its successful instantiations. Carnap thought that over time the degree of confirmation of a theory could increase more and more (Mayhall, 2002), and as the degree of confirmation increased, the degree of disconfirmation would decrease symmetrically. Because inductive inference is tied to a probabilistic approach (Ayer, 1952), an increasing degree of confirmation translates into an increasing probability of a theory. Degree of confirmation therefore makes a
probabilistic prediction as to the future performance of a theory.

In Popper’s system, a theory is said to increase its degree of corroboration as it survives the severest tests. The severity of test will depend on the investigator’s intention (hence the impossibility to fully analyze degree of corroboration in terms of logic; Popper, 1983:236) but also on the testability of a theory that increases with the degree of improbability of a particular test statement. Some theories are more severely testable than others, and those that are most severely testable are the ones that make the most improbable predictions. Testability and with it the corroboration of a theory increase with the improbability of evidence described by statements that can be deduced from the theory. However, the improbability of the evidence that would corroborate or falsify a theory must be established against a standard, which is the background knowledge. For Popper, the class of (negated) observation statements a theory entails constitutes its explanatory power (Laudan and Leplin, 2002; note the difference from Farris’s [1983] use of the term). The empirical content of a theory increases with the increasing number of (negated) observation statements it entails. Falsification of a theory can be seen as a search for counterinstances, i.e., the positive instantiations (i.e., negations) of negated observation statements (Popper, 1983:234, 1989:240). Some of these counterinstances will be embedded in background knowledge, and others will transcend this knowledge. If the empirical content of a theory equals the background knowledge, its relative empirical content is zero. The relative empirical content of a theory increases the more the theory transcends background knowledge. However, the more the theory transcends background knowledge, the more its predictions will be improbable. The more improbable the evidence that is appealed to in the test of a theory, the greater the falsifiability as well as the corroboration of a theory (see Faith and Trueman, 2001) and hence the greater the degree of corroboration if the theory passes the test. Yesterday’s test results are part of today’s background knowledge, so that if we aim for a high degree of falsifiability today, we must test a prediction deduced from the theory that transcends today’s background knowledge and hence yesterday’s test. Thus, we must test a new prediction of a theory, for example by forcing it to make an even more precise prediction than yesterday. In this way we push the theory toward an experimentum crucis (Popper, 1992).

However, a theory that has been variously tested multiple times and that has achieved a high degree of corroboration may become accepted (perhaps because no new ways of further testing it are apparent at that time) and hence relegated to background knowledge. If so, its relative empirical content drops to zero, and its total empirical content is diminished. The degree of falsifiability and the severity of test are then also diminished (in particular because there is no intention anymore to subject the theory to further severe tests). The degree of corroboration drops as a consequence. Hence, what once was consid-
perhaps, that it may be useful as an appraisal of statistical
tests."

**LOGICAL PROBABILITY**

The current debate on Popper and cladistics turns to a
large degree on issues related to the probabilification of
systematics. Whereas systematists mostly think of statisti-
cal probability, Popper distinguished this from logical
probability, a concept that is not easy to understand but is
essential for an understanding of Popper’s philosophy of
science and its bearing on systematics. As noted by Faith
and Trueman (2001:333), “p denotes a logical probability”
in Popper’s formalism for degree of corroboration.

Carnap (1995) claimed that the increasing degree of
confirmation of a theory translates into a greater prob-
ability for it to be correct, but he invoked two differ-
ent kinds of probabilities in this context. The first is
statistical and is expressed in the object language of
science. An observational statement that reports on a state
of affairs is expressed in the object language of science
and in statistical probability. For example, a physicist
may emerge from her lab with a discovery that she an-
nounces in the object language of science: “If p and if
q, then z follows (with a statistical probability of n).’’ So
p and q may specify certain initial or boundary condi-
tions that, under a covering natural law, would result
in z, but that result may obtain with a certain statistical
probability of n only if the covering law is of a prob-
abilistic nature. In other words, if I perform a series of
experiments, I would predict that z should obtain with a
frequency of n (in n percent of the cases), and this predic-
tion should be more closely approximated the longer the
series of experiments is performed. But whether it ob-
tains in this particular experiment I am conducting here
and now is not determined. Statistical probability cannot
relate to a particular event but only to a series of events
(Popper, 1979), which in turn is the reason why Popper
introduced his propensity interpretation of probability
(Popper, 1974a:122).

If the covering law in question is of a probabilistic
nature, the individual instantiation is not determined.
The law still governs the natural course of events, yet
the question remains, to which degree? Statistical prob-
ability is a quantitative relation that governs repeatable
kinds of events (tokens of the same type), which is not
the same as saying that a certain hypothesis (or hypoth-
esized law) is made probable to a certain degree by a
invoked a second kind of probability, the logical prob-
ability, to measure this second type of probability or, in
figurative speech, the truth-value (used here colloqui-
ally, not technically) of a statement rendered in the object
language of science. Borrowing from Wittgenstein, state-
ments rendered in the object language of science mea-
sure the world (Pears, 1988), and for Carnap they do so
in terms of statistical probability. Logical probability is
supposed to measure the truth-value of those object lan-
guage statements of science, but the truth-value of object
language statements cannot be measured or expressed
in the same object language itself. The truth-value of ob-
ject language statements of science must be measured
from the outside, which requires that the measure will
have to be expressed in a metalanguage of science. Figu-
ratively speaking, the metalanguage of science provides
the ruler that is applied to the object language of science
in the attempt to measure its truth-value. If the scientist
pronounces in her object language the statement “if
p and if q, then z follows (with a statistical probability
of n),” the philosopher of science may look at that state-
ment and ask the question, in a metalanguage of science,
“To which degree is it true (are we rationally entitled to be-
lieve: Carnap, 1997b:967, emphasis added) that ‘If p and
if q, then z follows (with a statistical probability of n)?’”
This is figurative speech, however. Technically speak-
ing, logical probability is an analytic statement on the
degree to which stated evidence is logically implied by
believed that it would eventually be possible to express
the logical probability of a theory in numerical terms,
and it is evident that the logical probability of a the-
ory directly corresponds to its degree of confirmation.
The two, indeed, are synonymous (Carnap, 1997a:72).
(Logical probability was also called inductive probabil-
ity [Carnap, 1997a, 1997b; Hempel, 2001] because it was
viewed as the backbone of inductive logic. It is interesting
to compare Carnap’s concept of logical probability to the
prior probability in Bayes’s theorem and to the concept
of entrenchment and the corresponding projectability of
predicates as argued by Goodman [2001].)

Popper did not believe that logical probability (which
he also called absolute, initial, prior, or a priori prob-
ability; Popper, 1983:284) could be expressed numeri-
cally without stepping outside the realm of pure logic
(Popper, 1974b, 1989, 1992; exceptions are the theore-
tical boundary conditions near complete verification or
falsification [Popper, 1989:397]). For him, logical proba-
bility was tied to the explanatory power of a theory in a
deductive framework, hence to its logical range. Popper
used deduction in the establishment of the empirical con-
tent class of a theory, which he considered to be the set
of empirical statements whose negations are (logically)
entailed by that theory (Laudan and Leplin, 2002; see
also Popper, 1983:231). This empirical content of a the-
ory can be pictured as the range or the space of all the
forbidden states of affairs that can be deduced from a
theory (Carnap [1997a:71; see also Glock, 2000:117, 221]
traced the notion of the logical range of propositions to
Wittgenstein’s *Tractatus* [see also Popper, 1992:124]). A
theory that forbids almost nothing (small empirical con-
tent) has a large logical range, a small logical improb-
ability, and hence a large logical probability. The more
“facts” a theory forbids (the greater its empirical content
or explanatory power), the greater its risk of being found
false, hence the smaller its logical range, the larger its
logical improbability, and the smaller its logical probabil-
tity. Corroborability (or falsifiability) is hence the inverse
of logical probability (Popper, 1983:245). Where Carnap
found logical probability to provide support for a de-
gree of credence in the future successful performance
of a hypothesis (Carnap, 1997b), Popper found logical probability to provide a measure for the corroboration (or falsifiability) of a hypothesis. Carnap sought high logical probability for hypotheses, and Popper sought small logical probabilities for hypotheses. Using Sober’s (2001) surprise principle again, we can state that the surprise factor will be greater and hence the degree of corroboration will be higher if a theory with a very small logical (prior; Popper, 1993:244) probability passes a severe test. However, logical probability remains a meta-linguistic expression relative to the object language of science.

Logical probability has long since vanished from the philosophy of science (e.g., Balaschov and Rosenberg, 2002) and for good reasons (e.g., Putnam, 1997; Quine, 2000). Logical probability is built upon the entailment of stated evidence by a stated hypothesis, and if such entailment is not obtained for phylogenetic hypotheses, it certainly does not apply to phylogenetic analysis. The recent systematics literature (e.g., de Queiroz and Poe, 2001; Faith and Trueman, 2001; Farris et al., 2001; Kluge, 2001b) documents extensive use of Popper’s formalisms for degree of corroboration, severity of test, etc. The concept of probability used in these formalisms is that of logical probability, the formalisms therefore are meta-linguistic in nature: \( h \) is a stated hypothesis, \( e \) is stated evidence, \( b \) is the corpus of stated current scientific knowledge that is not questioned when \( h \) is being put to test, and \( p(e, h) \) is the degree by which \( e \) is (logically) entailed by \( h \). In sum, the formalisms of Popper are statements rendered in the language of philosophy, not in the language of science; \( e \) stands to \( h \) not in the relation of an observation to justified belief (hypothesis) but in the relation of logical entailment.

**POPPER’S PROPENSITY INTERPRETATION OF PROBABILITY**

As the debate about the probabilification of systematics in the light of Popper’s philosophy of science continues, it is important to keep in mind that statistical and logical probabilities are only two of several possible interpretations of probability (Nagel, 1971) to which Popper added his own, the propensity interpretation. This interpretation supersedes logical probability (“[propensities] are not merely logical possibilities, but physical possibilities”; Popper, 1988:105; see also Popper, 1983:286) in that it can deal with genuinely indeterminate singular events that seem to evade lawfulness (Popper, 1983:285), such as those addressed by quantum mechanics.

The rise of quantum mechanics dramatically changed the landscape for philosophy of science based on deduction. Deduction is about what must be the case, but given Heisenberg’s uncertainty principle, what must be the case remains indeterminate in its singular (particular, individual) instantiation. Hence, it is Popper’s writing on probability that has attracted the attention of phylogenetic systematists, especially as presented in appendix IX of his *Logic of Scientific Discovery* (Popper, 1992; cf. de Queiroz and Poe, 2001; Kluge, 2001b).

Given Popper’s adherence to deductive logic, such as exemplified in the modus tollens form of argument, he was not at ease with the probabilification of science (Lakatos, 1974). Some support can always be obtained for every theory or hypothesis that is of a probabilistic nature. Statistical theories or hypotheses must hence be made refutable by the conventional determination of criteria of rejection (Lakatos, 1974). These criteria were formulated by Popper (1992) in appendix IX of *Logic of Scientific Discovery*, and they state that the probability of the evidence sought to support a hypothesis must be very small (the evidence must be very improbable; see Faith and Trueman, 2001), but that same evidence must impart a very high probability on the hypothesis under test, and the margin of error (\( \delta \)) must be very small, which means that the hypothesis must make very precise predictions (see Nagel, 1971, for further comments on \( \delta \)). In sum, the evidence must consist of a “statistical report” that asserts “a good fit in a large sample” (Popper, 1992:411).

However, that criterion alone was not enough for Popper. He introduced the propensity interpretation of probability (Popper, 1983, 1989) to be able to deal with singular events (statistical probability can only deal with a series of events) and to remove “certain disturbing elements of an irrational and subjective character” in quantum mechanics (Popper, 1983:351), viz. “the strange role played by ‘the observer’ in some interpretations of quantum mechanics” (Popper, in Popper and Eccles, 1977:26). Propensity means that an object has a disposition to behave in a certain way. Dogs, for example, have a disposition to bark, and radium atoms have a disposition to radioactively decay with an associated half-life of 1,620 years. The example of radioactive decay “is perhaps the strongest argument in favor of what I have called ‘the propensity interpretation of probability in physics’” (Popper, in Popper and Eccles, 1977:26, n. 4). In the example of propositional logic, its conditional can now be interpreted as a disposition statement (Körner, 1959): “(For every A) if A is a dog, then it barks” (then it has a disposition to bark), “A is a dog,” “therefore A barks” (has a disposition to bark). This argument is deductive, specifying a necessary or universal property of dogs, and this property is the disposition to bark. But the deduction does not specify that this particular dog I am pointing at must bark right now. It may not bark because it is fast asleep. The same holds for the radium atom: “(For every A) if A is a radium atom, then it will decay” (it has a disposition to decay with an associated half-life of 1,620 years), “A is a radium atom,” “therefore A will decay,” but maybe not right now. Again, the statement identifies the disposition to radioactively decay as a universal (i.e., necessary) property of the radium atom, but it does not specify which particular radium atom will decay in which particular point in space and at which particular time. These events can only be a matter of (objective) probability. But although the theory of radioactive decay is by its very nature a probabilistic theory, its meaning is not under the propensity interpretation of probability because the disposition to behave in a certain way is a universal property of the atoms that fall...
under that theory. So with his propensity interpretation of probability, Popper was able to maintain universality even for laws that are probabilistic in nature. Popper’s propensity interpretation of probability is thus revealed to ultimately lead back to propositional logic (Popper, 1974c:1132), such that deductive explanations (causal hypotheses) that must be the case assert a propensity of 1 (Popper, 1983:288). As Hempel (2001:212) put it, it is the universal character of the disposition of objects to behave in a certain way “that gives probabilistic laws their predictive and their explanatory force” for objects that fall under these laws. The empirical content of a scientific theory was considered by Popper to correspond to the class of negated observation statements it entails. But this statement asserts nothing else than that “natural laws . . . can never do more than exclude certain possibilities” (Popper, 1976b:139). By the same token, “Propensities may rule out certain possibilities: in this consists their lawlike character” (Popper, in Popper and Eccles, 1977:31, n. 13).

The appeal by systematists to Popper’s writings on probability must take into account the propensity interpretation of probability. Popper (1983:284f) went to considerable length to analyze the meaning of logical probability in a frequentist context, eventually abandoning his early preference for absolute (i.e., logical) probability as he “realized that absolute probability is a special case of relative probability, but not the other way around” (Popper, 1974c:1119). He found his axiomatization of the theory of relative probability stronger and more interesting than the theory of absolute probability, a point that “was decisive for my choice” (Popper, 1974c:1119). From this choice then followed the propensity interpretation as a “measure of possibilities” (Popper, 1983:286). However, that interpretation requires repeatable phenomena, which again are tokens of the same type.

This interpretation of propensity might seem at odds with Popper’s (1974b:276–284) propensity interpretation of evolution, because there would seem to be no room for universality in evolutionary theory. Popper (1976b:108) considered evolution a unique process, but nevertheless one that occurs in accordance with “all kinds of causal laws.” To invoke propensities for the evolutionary process comes at a cost, however, that “the evolution of the executive organs will become . . . goal-directed” (Popper, 1974b:278). What Popper was aiming at is a Darwinian process of variation and natural selection, where behavioral changes that are subject to organismic propensities precede somatic changes and determine the direction of somatic change in a process of directional selection that “simulates” not only Lamarckism, but Bergsonian vitalism also” (Popper, 1974b:284). So the giraffe stretched its neck after all, for without that propensity, “a longer neck would not have been of any survival value” (Popper, 1974b:279). Propensities or dispositions will inevitably be realized given certain circumstances (Glock, 2000:183; e.g., if the burglar startles the sleeping dog, then it will bark), and this inevitability reflects Lamarck’s use of the term besoins (necessities, requirements, needs) in his explanation of adaptive evolution, which Cuvier in a calculated way misrepresented as désires (desires, wishes) (Appel, 1987:169).

POPPER*

Faith and Trueman (2001:331) escaped from large parts of what I have explained because they explicitly decoupled corroboration from “the popular falsificationist interpretation of Popperian philosophy” (rendering it Popper*: Faith, 1999). Thus, the meaning with which Popper used the concept of degree of corroboration cannot be coextensive with the meaning bestowed on that concept by Faith and Trueman (2001; see remarks above on semantic analysis). Recognizing that Popper’s asymmetry of falsification is difficult to realize in empirical space, Faith and Trueman (2001) put emphasis on degree of corroboration instead. To link their argument to Popper’s hypotheticodeductivism, however, they need to show how their notion of corroboration differs from the inductivist notion of confirmation. Popper linked corroboration to testability (Popper, 1983:245) and stated that “scientific tests are always attempted refutations” (Popper, 1983:243). For him, corroboration was thus embedded in a falsificationist context. In a context decoupled from falsificationism, Faith and Trueman (2001) emphasized the “genuineness” of a test that is measured by the improbability of the evidence (i.e., the low p(e, b) value; D. Faith, pers. comm.). Faith and Trueman (2001) capitalized on Popper’s notion that severity of test increases to the degree that evidence e is improbable given background knowledge b in the absence of the hypothesis h (i.e., p(e, b)).

With their argument, Faith and Trueman (2001) recognized the fact that Popperian falsification based on hypotheticodeductivism cannot exist in systematics. Turning the tables, they therefore emphasized the synonymy that Popper allowed for degree of corroboration qua support (Popper, 1983:238ff), and from there follows their emphasis on the necessity for the evidence e to be improbable in the light of background knowledge (b). There is more reason to consider improbable evidence (in light of b) as possibly false than to consider highly probable evidence false, such that the hypothesis h, which possibly accommodates the improbable e, is itself more improbable or has a higher content (as required by Popper, 1983). But once evidence is accepted in support of a given h, they argued, its improbability decreases and with it decreases the improbability and therefore the content of h in support of which e is accepted. This argument is analogous to that of degree of corroboration increasing once a theory is accepted into background knowledge. (The analogy concerns different levels of argumentation. On the one hand there is the relationship of evidence to a hypothesis, and on the other hand there is the relationship of a hypothesis to background knowledge.)

Given that Faith and Trueman (2001) explicitly decoupled their interpretation of Popper’s philosophy from falsificationism, it is interesting to compare their notion of degree of corroboration to Popper’s formalization of degree of confirmation. Popper (1997:222) considered it
his “critical task” to show by way of his falsificationism that “inductive logic [sic!] is impossible.” But Popper commented on his own “positive theory” and provided a definition of degree of confirmation, where $C(x, y)$ is the degree by which $y$ (observation statements) confirms $x$ (theory, hypothesis) (Popper, 1997:221f). Popper’s degree of confirmation is, again, defined in a way that is relativized to the current background knowledge $z$ (as he did for degree of corroboration): $C(x, y; z)$ (degree of confirmation) = $p(y, x · z) - p(y, z)/p(y, x · z) - p(x · y, z) + p(y, z)$. For Popper, $z$ comprises the old evidence, the old and new initial conditions, and accepted theories. Confirmation is bestowed upon a new explanatory hypothesis $x$ by the new observational results $y$, which are not part of the background knowledge $z$ (as is also not the improbable evidence requested by Faith and Trueman, 2001). “That is to say, the total evidence is to be partitioned into $y$ and $z$; and $y$ and $z$ should be so chosen as to give $C(x, y; z)$ the highest value possible for $x$, on the available total evidence” (Popper, 1997:222, n.82; emphasis added). The total evidence is what Carnap (1997b) used to gauge the logical probability of a hypothesis. Popper (1997:222) was satisfied that his definition would meet the condition that the confirmability of a statement—its highest degree of confirmation—equals its content (i.e., the degree of falsifiability). ”Because the degree of falsifiability increases with the degree by which a theory transcends background knowledge, i.e., with the most improbable predictions the theory generates, the degree of confirmability likewise increases with the most improbable of evidence, as was also postulated by Faith and Trueman (2001).

CONCLUSIONS

Many of the issues in systematics that are currently debated with an appeal to Popper are just good common sense. There can hardly be any disagreement with a methodological principle that seeks to minimize ad hoc auxiliary hypotheses to save the phenomena, a principle that is endorsed by every scientist and by Popper (and other philosophers of science). Why disagree with a methodological principle that requires going beyond that sort of supporting evidence that is the cheapest one to be had? However, there is the semantic ambiguity of the notion of a test. Some will consider the testing of a null hypothesis as opposed to alternative hypotheses a legitimate test, others will call the disconfirmation of a theory the result of a test, and all of it is testing, just not in Popper’s sense. Carnap (1997a:59), for example, used testability as an “empiricist criterion of significance” in an explicitly verificationist context: a sentence is confirmable by possible observable events and testable if a method can be specified for producing such events at will. More generally, Popper’s appeal for a critical attitude is commonsensical and endorsed by most if not all scientists. This appeal was Popper’s way of carefully steering his theory of knowledge past the Scylla of dogmatism and the Charybdis of skepticism. But Popper’s theory of knowledge was more. Hume’s predicament was insurmountable: there simply is no way to deduce theory from observation (Quine, 2000). So Popper turned the tables: Start with a theory, and if it is to be scientific it has to allow the deduction of negated observation statements. This is the step in which systematics cannot follow Popper (Sober, 1988). One can use the word “test” in parsimony as well as in likelihood analysis, but it will not correspond to a Popperian test. For Popper, a test is tied to the asymmetry of falsification; by contrast, symmetry has long been noted in whether congruent characters support a certain hypothesis of relationships or falsify alternative potential hypotheses for the same terminal taxa (Bonde, 1974; Gaffney, 1979). There is one easy way to elucidate how far the theory of phylogenetic inference can follow Popper. Phylogenetics must logically demarcate its concepts from those of inductive or deductive inference; it must logically demarcate corroboration from confirmation. However, times move on and meanings change, and it may be absolutely permissible that corroboration in modern systematics no longer needs to be used in the same way Popper used it. But at that point, systematics can no longer be claimed to be Popperian, but rather it is under the influence of the philosophy of its own exponents, as it should be.

ACKNOWLEDGMENTS

The June 2001 issue of Systematic Biology prompted me to read Popper again, an adventure that led to the insights presented in this paper. David Hull unsuccessfully tried to explain some of these issues to me 10 or so years ago. I have also greatly profited from discussions with Dan Faith, Arnold Kluge, Lukas Rieppel, Michael Rieppel, Maureen Kearney, and Peter Wagner. I do not suggest that they agree with my analysis, but often it is through discussions with opponents that one learns most.

REFERENCES


Using a Null Model to Recognize Significant Co-Occurrence Prior to Identifying Candidate Areas of Endemism

AUSTIN R. MAST AND RETO NYFFELER

Institute of Systematic Botany, University of Zürich, Zollikerstrasse 107, 8008 Zürich, Switzerland;
E-mail: amast@access.unizh.ch (A.R.M.)

Recent commentaries on areas of endemism have suggested that these be recognized by the “nonrandom” (e.g., Nelson and Platnick, 1981:56; Morrone, 1994:438) distributional congruence of two or more taxa at some scale of mapping. However, no one has addressed how one might objectively determine the threshold between random and nonrandom co-occurrence for this purpose (Hausdorf, 2002).