

PART IV

THE PURPOSES OF BIOLOGICAL CLASSIFICATION

The Purposes of Biological Classification

Norman I. Platnick and Gareth Nelson

The American Museum of Natural History

All biologists use classifications to one degree or another, and those of us who work on classifications use the results of all other biologists to one degree or another, so you might reasonably expect that biologists in general would share some common conception of how classifications should be constructed and how they can be used. Certainly one might expect that all taxonomists, at least, would share such a perspective. But this is not the case; in fact, the theory of taxonomy is at present a very controversial subject, with much of the controversy revolving around the question of how to go about constructing an ideal classification.

Surprisingly, there is relatively little controversy about what properties an ideal classification would have. Admittedly, different terms are used to describe these properties, various workers arguing that classifications should be maximally stable, or maximally useful, or maximally informative, or maximally testable, or maximally refutable, or maximally predictive, and so on. But probably all of these desirable qualities actually refer to a single property of classifications, namely, that what taxonomists do is to hypothesize, generally on the basis of a very small sample of characters, that nature is ordered in a certain specifiable pattern, and that no matter what other characters of a group we might choose to examine in the future, we will find the same pattern again and again, and not a different pattern for each potential source of data.

But what exactly is a taxonomic pattern? We can start with a single species, A. We might observe certain things to be true of it, and we might file this information away, appropriately, in an accumulation of observations about species A. Some of the things, however, would be true of other species besides A. So our observations would be incomplete without a statement specifying that they pertain to A uniquely or, alternatively, to some other species as well. The other species could be few or numerous; let's consider two. Whichever two we choose,

PSA 1978, Volume 2, pp. 117-129

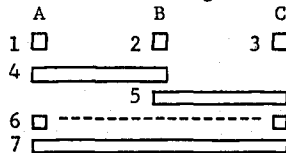
Copyright © 1981 by the Philosophy of Science Association

we might observe some things to be true of A uniquely, some true for B uniquely, and some true for both:



And what do we do with this information? File it away in three accumulations, one for A, one for B, and one for A+B? We could try, but it won't work, for we are actually no better off than before in being able to specify the generality of our observations. Besides, there probably is no accumulation set aside for species A+B.

Consider instead three species. Whatever species we choose, we might observe some things to be true of each, other things to be true of each possible pair, and still other things to be true of all three:



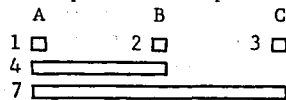
We won't ask what to do with all of these observations, because by now it should be obvious that an impossibly large number of accumulations would be required to handle all the possible information of this kind.

Rather, let's focus on the pattern displayed by the information. The pattern consists of all possible elements. The pattern is random, in the sense that it can be duplicated by tossing three coins in sequence. Every so often they will all come up the same, giving us element 7. If the first coin comes up heads, and the others tails, we would have elements 1 and 5, and so on. For what it's worth, each of the seven elements has an equal probability of coming up, and about eight tosses will produce all seven elements, on the average.

To some persons, this kind of pattern does represent the real world; for them, ultimate reality is chaotic. These persons tend not to become scientists, but if they do, they tend to become frustrated. To other persons, this kind of pattern is unsatisfactory as a representation of the real world; these persons tend to assume that the randomness stems from faulty observations and incorrect hypotheses. The problem is to see order in apparent chaos through critical observation and hypothesis testing. Some of these persons become scientists, and systematists.

One basic postulate of systematics is that what can be learned about any three species exemplifies a pattern of the form: two of the three are more closely related to each other than either is to the third. In the case of species A, B, and C, for example, we might have reason to believe that A and B are more closely related to each other than either is to C. Never mind, for the moment, what exactly is meant by 'related'. We would still have things that are true, each for A, B, and C, and things true for all three (a residue, as it were, of randomness).

But in between we would have simplified the picture:



Interestingly, we would still not have learned that, for example, item 1 really is unique to A, or item 2 to B, or item 3 to C--there still being the possibility that these observations are true for species yet unexamined by us. We would, however, have discovered an element of pattern--item 4--an element of a different kind than the observations that we accumulate. The element of pattern can be restated in a taxonomic, or systematic, form: there is a group, including A and B, but excluding C. If A is a bat, B is a mouse, and C is a robin, the AB group is Mammalia. Now, there is a place--an accumulation, if you like--for information of that kind, and that is systematics in its traditional guise.

So far so good, you might say. But what happened to the other elements, to elements 5 and 6? These elements, to the extent that we have detected them, represent incongruence with pattern 4. If nature is orderly and our hypotheses about characters and their distribution are correct, the incongruence should not exist. How might we decide wherein the problem lies? It is possible that nature, in this case, is not orderly, and that characters distributed like elements 4, 5, and 6 all reflect real (but random) aspects of nature. This position defines a philosophy of classification usually called phenetics, and in its favor it must be admitted that there is no way for us to actually determine whether the disorder that we perceive exists in nature, or only in our own hypotheses. The purpose of science, however, is to discover and explain regularities in nature, and, as Popper ([15], p. 61) has noted, science is predicated not on the metaphysical position that nature is necessarily completely orderly, but on the methodological rule that we are never to abandon the search for order. If we give up the search for regularities, we also give up the game of science. So we may assume that the order really does exist and that we, and not nature, are responsible for the apparent disorder. This assumption defines a philosophy of classification usually called cladistics.

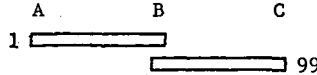
This means that either our original hypotheses about characters distributed like element 4 were wrong, and that our grouping (AB) is therefore wrong, or that our hypotheses about characters distributed like elements 5 and 6 are wrong. In what ways might we have made mistakes about these characters? There seem to be three possibilities.

One possibility is that we might have underestimated the generality of a character, for example, by considering a character to be distributed like element 5 when it is actually distributed like element 7. This could easily happen if species A had that character, but had it only in a modified form, and we failed to recognize the modification as being a form of the same character. This first type of error is usually called plesiomorphy.

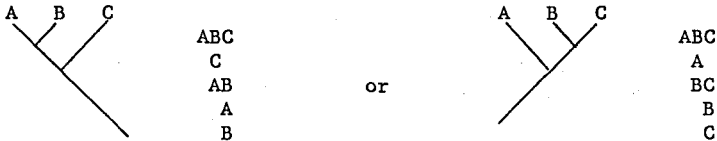
A second possibility is that we might have overestimated the generality of a character, for example, by considering a character to be distributed like element 5 when the attributes of species B and C are actually two different characters distributed like elements 2 and 3. This second type of error is usually called convergence or parallelism.

Finally, there is the possibility that what we have considered a character distributed like element 5 is not a character at all. The most common example of this type of mistake is the use of the absence of some feature as a character. For example, taxonomists used to maintain a group Invertebrata for those animals without backbones, but the absence of vertebrae is not a character (if it were, we'd have to include strawberries in the Invertebrata, since they also lack backbones). The absence of vertebrae in some set of organisms is no more informative than is the absence of Cadillac engines in their stomachs or totem poles on their heads.

The problem can be examined in another way. Suppose that we have three species, and we observe 99 things to be true for species B and C, and only one thing to be true for species A and B:



The information exemplifies two possible patterns, each of which can be represented by a branching diagram, or, alternatively, by a hierarchical classification:



Given this information, is a choice between these two patterns or classifications problematical? This question, believe it or not, causes systematists to part company. One group regards the choice as non-problematical, and opts for the second pattern. The other group regards the choice as being problematical, inasmuch as there are two patterns exemplified by the information. A representative of this second group might state: "These things are all very well so far as they go, but I need to know more about them."

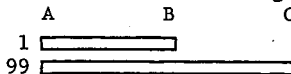
Let's supply this representative with some additional information by mentioning some things true for various vertebrate animals. The lone thing turns out to be lungs, present in both species A and B, absent in C. Of the 99, the list begins as follows:

1. Fins
2. Gills
3. A slimy skin
4. Aquatic habits
5. Lays eggs
6. Cold-blooded
7. No spoken language

and so on. About this time, our representative interrupts: "I think I understand the nature of the problem. But first let me ask you: are you sure that these things which you say are true for B and C are not, in some sense, also true for A?" Well, let's reconsider the list:

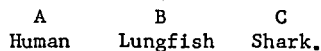
1. Fins? Species A has arms and legs. Could they not be construed as fins of a sort?
2. Gills? Species A, when an embryo, has gill slits and aortic arches. Not gills exactly, but pretty close.
3. A slimy skin? Species A develops internally, in its maternal parent. And while in utero it is a bit slimy.
4. Aquatic habits? Species A in utero is immersed in a miniature sea.
5. Lays eggs? Yes, but internally (and sometimes externally as well, when writing papers).
6. Cold-blooded? Yes, in utero (and all too frequently in later life as well).
7. No spoken language? Yes, when young, but that's the absence of a character anyway.

Suppose, then, that with a little scrutiny, all 99 turn out to be true for all three species. This leaves the following picture:



And what does it add up to? Just a simple illustration that the evidence for classifications is problematical, and deserves to be viewed as such. The observations are interesting in themselves, but what is more impressive is the pattern--the AB component of it--as a general summary of our observations. If we have done our job well, and have discovered the correct pattern, it will be a general summary of whatever other observations we might make, now or in the future. The pattern, therefore, has a truth of its own. Discovering that kind of truth is one task of systematics, and indeed of biology. Why so? Because it is a truth of the real world, or so we may infer. The alternative is to accept the idea that the real world is randomly organized--that the real world contains no truth at all.

By pattern is meant branching diagrams or, alternatively, hierarchical classifications. In this case, our AB observation--lungs--is a defining character of a group, or taxon, of vertebrates--the Teleostomi (the group that includes lungfishes, amphibians, "reptiles," birds, and mammals). Groups or taxa are the kind of items that systematics accumulates, and we found one in our comparison of species A, B, and C. The species? Yes, indeed:



A cladistic classification, then, is one in which the taxa that are recognized are equivalent to the groups united in branching diagrams, or cladograms, by the distributions of unique characters, and the expectations that those characters are just representative of a general pattern. But what exactly is a taxonomic character? The conventional analysis indicates that a character consists of two or more different

attributes (or character states) found in two or more specimens that, despite their differences, can be considered alternate forms of the same thing (the character). For example, two kinds of jaws are found in spiders; some spiders have jaws bearing fangs that move with an up-and-down motion, whereas other spiders have jaws bearing fangs that move with a side-to-side motion. Here, the character is 'orientation of fangs' and there are two character states: vertical and horizontal.

Note that a character is thus a hypothesis, a hypothesis that two attributes which appear different in some way are nonetheless the same (or homologous). As such, a character is not empirically observable; hence any hope to reduce taxonomy to mere empirical observation seems futile. But if alternate character states are in some sense the same, how can they be different? There seem to be only two possibilities: either one state is a modified form of the other, or both are modified forms of a third state. In either case, the 'sameness' that constitutes the character is thus the unmodified state, which all the organisms that show the character share, either in its original or in some modified form.

But suppose that we have recognized a group (spiders) on the basis of two character states believed unique to spiders. In other words, each character has one state found in at least some non-spiders and a second state found only in spiders:

Character X, Y: State 1 (non-spiders)
State 2 (spiders).

Now suppose that we find a new specimen about which we know only that it has one of these unique character states. Can we predict that the new specimen will have the other unique character state as well? No, clearly not; the new specimen might well have a different state of the second character:

Character Y: State 1 (non-spiders)
State 2 (most spiders)
State 3 (some spiders)

if that new character state represents a modified form of state 2, found in all other spiders:

Character Y: State 1 (non-spiders)
State 2 (spiders): Substate a (most spiders)
Substate b (some modified spiders).

In this situation, of course, state 2 is functioning as a character (a unit of 'sameness'), not a character state; just like character Y, it is divided into an original and a modified form. If we agree that for the character (the 'sameness') to exist, character states must be modifications, it is apparent that predictions cannot really function at the level of character states. Because of the ever-present possibility of further modifications, predictions will hold only for characters (i.e., for sets of original plus modified character states).

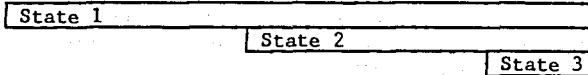
In this sense, then, the concept of a character state is potentially misleading. To view some character X as being composed of three states:

State 1

State 2

State 3

implies that the character states are alternatives, when they are actually additions:



In this case, character X is actually equivalent to state 1 (i.e., it defines a group, all the members of which have state 1, either in its original or some modified form). States 2 and 3 can be regarded as separate characters, for which the same provisions hold. Further, state 1 itself must be a modification of some other character, and itself represents a restricted subset of some other, more general character. Thus, all characters can be seen as modifications (or restrictions) of other, more general characters, and one might envision a great chain of characters (or homologies) stretching from those of complete generality, which are true for all life, on to those true for only a single species.

The implications of this can be seen in a simple example involving tetrapod and non-tetrapod vertebrates. Systematists have long been in agreement that the limbs of tetrapods (amphibians, "reptiles", birds, and mammals) are homologous with the fins of non-tetrapod vertebrates ("fishes"). If we regard fins and limbs as alternative states of a character (paired pectoral and pelvic appendages), we might thereby sort out vertebrates into two groups, Pisces (for those with fins) and Tetrapoda (for those with limbs), and vertebrates were indeed classified in this way for many years. However, one of these groups (Pisces) proved not to be maximally predictive, in that many characters were found that are shared uniquely by tetrapods and some (but not all) "fishes". If, however, the limbs of tetrapods are not only homologous with fins, but are also modifications of fins, the problem disappears. We can recognize that we have two characters; one, fins (or paired appendages) is found in all vertebrates in one form or another; a second, limbs, is found in all tetrapods (including snakes!) in one form or another. And the group Pisces turns out to have been based on the absence of a character: it includes those organisms with fins that also happen to lack modified fins (limbs). But the absence of limbs in sharks, salmon, and lungfishes tells us nothing beyond the fact that they are not tetrapods.

In practice, taxonomists generally have little difficulty corroborating preliminary hypotheses of homology. For example, all arachnologists agree that the two kinds of spider jaws are homologous: they occur in the same position, they have the same mode of development, they're moved by the same muscles, innervated by the same nerves, etc. But then the question arises: are both kinds of jaws independent characters, that is, different modifications of some other kind of jaw (and hence each useful for defining a subgroup of spiders), or is one kind of jaw a modified form of the other? In the latter case, the original (or primitive, or plesiomorphic) form would actually be found in all spiders, in one form or another, and only the modified (or derived, or apomorphic) form would be useful for defining a subgroup within the spiders. Using a character that is plesiomorphic for a

group to try to define a subgroup within the group would be an example of the first type of mistake noted above (underestimating the generality of a character), as well as the third type (if the group is defined by the presence of the plesiomorphic character together with the absence of the apomorphic one).

How might we go about deciding whether one of a pair of homologous characters might be a modified form of the other? Two kinds of tests are available, ontogeny and outgroup comparison. Both kinds of tests depend on the demonstration that one character is more general than the other, that is, that one character is found in more groups than the other [11].

Suppose, for example, that we have two species, one with the character X and another with the homologous character Y. Which character is the more general? As of yet, we can't say. But suppose we study the embryology of these species, and discover that in species 1 character X is found throughout development, whereas in species 2 character X appears early in ontogeny but is subsequently transformed into character Y. We now know that X is the more general character (plesiomorphic), defining a group including both species, and that Y is the less general character (apomorphic), defining a group including only species 2.

Often, however, we don't have ontogenetic evidence of this sort. For example, spider embryologists have found no evidence of ontogenetic transformations between the two kinds of jaws. We might ask, however, what the closest relatives of spiders are, and what kind of jaws they have. There are other characters which indicate that the closest relatives of spiders are a group called amblypygids, found in tropical regions of the world. As it turns out, amblypygids have jaws that work with an up-and-down motion just like one kind of spider jaw. Thus, we have evidence that vertically mobile jaws is the more general, or plesiomorphic, character, and cannot be used to define a subgroup of spiders since it actually defines a group including both spiders and amblypygids. Horizontally mobile jaws, on the other hand, are unique to one group of spiders, and can be used to define that subgroup. This technique of outgroup comparison, unlike the ontogenetic technique, requires a prior hypothesis of what the closest relatives (or outgroup) of the group under study may be (unless, of course, we're willing to survey all other groups of organisms to determine whether they have a homologous character). Obviously, if the hypothesized outgroup turns out not to be the closest relative, we may be lead to an erroneous conclusion. But the outgroup hypothesis can be tested independently, by simply shifting the analysis to a more general level (for example, investigating the relationships of spiders and amblypygids versus all other groups of arachnids, or arthropods).

This view of characters (as hypotheses of homology tested by ontogeny and outgroup comparison) and the significance of their relative generality has implications for views of taxa (species or larger groups). One obvious implication is that the number of taxa which can

be distinguished is a function of the number of characters that have been recognized, since to be diagnosable, a taxon must have a unique set of characters. Another implication is that the tendency (discussed by Hull [6]) to view biological taxa as Wittgensteinian 'cluster concepts' may merely be the result of an inadequate concept of characters. Given a view of the vertebrate character 'limbs' which does not recognize the possibility of character transformation, it might be impossible to treat 'limbs' as a defining character of the Tetrapoda (because snakes might be held to be limbless). Indeed, given such a view, it might be impossible to find any defining characters of the Tetrapoda (and many other groups), and one might conclude that the Tetrapoda is therefore a cluster concept. But taxonomists have not used this limited concept of characters: even when insects were divided into the Apterygota (wingless insects) and Pterygota (winged insects), the fleas (which might superficially appear to be wingless but nonetheless have thoracic features indicative of wings) were always placed in the Pterygota.

At any rate, given that nested sets of characters can be discovered, and that a hierarchical classification can summarize the pattern of nesting, the question arises: what does the classification tell us about the evolution of organisms? And what exactly is the relationship between taxonomy and evolution? The current interest in cladistics dates from the works of Hennig ([4], [5]), particularly as his ideas were applied to biogeography by Brundin [1], although the basic ideas were outlined in detail much earlier by the British ornithologist P. C. Mitchell [10]. Hennig presented his ideas as being the methods of 'phylogenetic systematics', justified them with reference to a particular model of the evolutionary process, and viewed them as techniques of phylogeny reconstruction. The same methods have been treated here not as the methods of phylogenetics *per se*, but as the methods of taxonomy in general, justified not with reference to any particular model of the evolutionary process but only with reference to order in nature. The discrimination of plesiomorphic from apomorphic characters, viewed by Hennig as distinguishing between historically primitive and historically derived characters (i.e., as representing actual evolutionary transformations) appears actually to discriminate only between more general and less general characters. Cladistic methods seem basically to be attempts to find a real order in nature, to discover a real hierarchy [12]. This, of course, has been the goal of taxonomists since long before Darwin, going usually under the name of the 'natural system' (as opposed to artificial systems). As currently understood, artificial systems are keys--useful, to be sure, for the purpose of identifying organisms (i.e., placing them within the taxa recognized by a pre-existing classification); there can be many artificial systems, based on different (even incongruent) characters. Of natural systems, of course, there can be only one, based on nested sets of congruent characters.

So what about evolution, and its role in taxonomy? Presumably, its role is exactly as early evolutionists conceived it, namely, as an explanation for the existence of a natural hierarchic system. Does

that mean that taxonomists must attempt to reconstruct phylogenetic trees (actual evolutionary history) and then somehow chop them up to form classifications? No, not at all; phylogenetic trees stand in the same relation to cladograms as evolution does to the natural system: phylogenetic trees are presumably the cause of cladistic relationships. For any given cladogram, there are numerous possible phylogenetic trees that could have been the cause of the cladogram; for example, a simple three-taxon statement relating species A and B as opposed to C could be the result of at least six different phylogenetic trees [13]. In recent years, it has become apparent that the reconstruction of phylogeny at the level of trees is both beset with numerous theoretical difficulties and basically irrelevant to the process of constructing classifications. Cladograms, not trees, are the necessary and sufficient basis for classifications (for discovering the natural system). Once order has been found in nature, we may assume that the order is the result of evolution, after which natural groups may be viewed as monophyletic, apomorphic characters may be viewed as evolutionary novelties, and degrees of relationship may be viewed as reflecting relative recency of common ancestry. But, as stressed by Hennig, we have no knowledge of evolutionary history that is independent of the existence, and discovery, of a natural hierarchic system. Hence, the investigation of patterns in the distribution of characters (taxonomy) is both independent of, and a necessary prerequisite to, any investigation of evolutionary processes [14].

So much for taxonomy and the biologist. What might it teach the philosopher? While acknowledging that it's probably foolhardy for non-philosophers to try to answer that question, we can't resist making the attempt, for whatever it might be worth. Larry Laudan ([8], pp. 57-61) has argued that the particular views of the philosophy and methodology of science "which a scientist brings to bear in his assessment of theories, have been perhaps the single major source for most of the controversies in the history of science;" further, that a conflict between a scientific theory and a preferred methodology may be resolved by a change in the methodology; and also that one of the effects of the Newtonian revolution on eighteenth century science was to force a change from a predominantly inductivist to a largely hypothetico-deductive view of science. So the question arises: do the controversies within systematics conform to Laudan's view?

The period of the Newtonian revolution was a heady one for taxonomy also. The appearance in 1735 of the first edition of Linnaeus's Systema Naturae, [9], containing an avowedly artificial sexual system of plant classification, sparked a controversy over the differences between, and the relative desirability of, artificial and natural systems; the controversy raged among botanists until at least the publication of A. L. de Jussieu's Genera Plantarum [7] in 1789 [2]. We have suggested here that the major controversies within taxonomy do concern the choice between artificial and natural systems, that the choice involves the acceptance or rejection of incongruence as a feature of the real world, and that one's decision depends not on empirical factors but on one's views of the nature of scientific activity. Artificial systems, with

their unquestioning acceptance of data on characters, even incongruent data, as facts, seem clearly inductivist and hence associated with pre-Newtonian science in Laudan's scheme. The search for a natural system seems to require a hypothetico-deductive, or post-Newtonian, point of view. If so, different methodological norms have indeed been the cause of the major controversies within taxonomy, and Laudan's views seem corroborated.

But there is a problem. The controversy between artificial and natural systems, the controversy between inductive and hypothetico-deductive philosophies, continues today unabated, albeit under the guise of phenetics versus cladistics, and despite claims by early pheneticists for the naturalness of their methods (claims refuted by Farris [3]). So the unanswered question is: why are taxonomists today still fighting an eighteenth century battle? Is this a problem for philosophers of science, or for historians of science? Or perhaps only for sociologists of science? Or is it just a problem for the cladists who still have to do the fighting?

References

- [1] Brundin, L. "Transantarctic Relationships and their Significance, as Evidenced by Chironomid Midges." Kungliga Svenska Vetenskapsakademiens Handlingar 11(1966): 1-472.
- [2] Croizat, L. "History and Nomenclature of the High Units of Classification." Bulletin of the Torrey Botanical Club 72(1945): 52-75.
- [3] Farris, J.S. "On the Phenetic Approach to Vertebrate Classification." In Major Patterns in Vertebrate Evolution. Edited by Max K. Hecht., Peter C. Goody, and Bessie M. Hecht. New York: Plenum Press, 1976. Pages 823-850.
- [4] Hennig, W. Grundzüge einer Theorie der Phylogenetischen Systematik. Berlin: Deutscher Zentralverlag, 1950. (Translated as [5].)
- [5] ----- . Phylogenetic Systematics. (trans.) D.D. Davis and R. Zangerl. Urbana: University of Illinois Press, 1966. (Originally published as [4].)
- [6] Hull, D. "The Principles of Biological Classification: The Use and Abuse of Philosophy." In PSA 1978, Volume 2. Edited by P.D. Asquith and I. Hacking. East Lansing, Michigan: Philosophy of Science Association, 1981. Pages 130-153.
- [7] Jussieu, Antoine Laurent de. Genera plantarum secundum ordines Naturales disposita. Paris: Herissant, 1789.
- [8] Laudan, L. Progress and Its Problems: Towards a Theory of Scientific Growth. Berkeley: University of California Press, 1977.
- [9] Linnaeus, Karl. Systema Naturae. Leyden: J.F. Gronovium, 1735.
- [10] Mitchell, P.C. "On the Intestinal Tract of Birds; with Remarks on the Valuation and Nomenclature of Zoological Characters." Transactions of the Linnean Society of London, Zoology 8(1901): 173-275.
- [11] Nelson, G. "Ontogeny, Phylogeny, Paleontology, and the Biogenetic Law." Systematic Zoology 27(1978): 324-345.
- [12] ----- . "Cladistic Analysis and Synthesis: Principles and Definitions." Systematic Zoology 28(1979): 1-21.
- [13] Platnick, N.I. "Cladograms, Phylogenetic Trees, and Hypothesis Testing." Systematic Zoology 26(1977): 438-442.
- [14] ----- . "Philosophy and the Transformation of Cladistics." Systematic Zoology 28(1979): 537-546.

- [15] Popper, K.R. Logik der Forschung. Wien: Springer, 1934. (As reprinted as The Logic of Scientific Discovery. New York: Harper and Row, 1968.)