

Rainbows in retrospect: L.A.S. Johnson's contributions to taxonomic philosophy

David L. Hull

Hull, David L. (Department of Philosophy, Northwestern University, Evanston, IL 60208) 1996. *Rainbows in Retrospect: L.A.S. Johnson's Contributions to Taxonomic Philosophy* *Telopea* 6(4): 527–539. Thirty years ago, L.A.S. Johnson wrote an extensive evaluation of the principles and techniques of numerical (or phenetic) taxonomy. Ten years ago he returned to these topics and this time included the principles of cladistic taxonomy as well. In this paper I re-examine Johnson's criticisms of phenetic and cladistic taxonomy to see how well they stand up to the test of time. On the main they stand up very well indeed.

Introduction

In 1968 in a Presidential Address to the Linnean Society of New South Wales, Johnson published a blockbuster of a paper criticizing numerical taxonomy. At the time I wrote at the top of my reprint of this paper 'The most stimulating paper that I've ever read.' Two years later this paper was reprinted in *Systematic Zoology* with an Addendum. Thereafter, Johnson remained all but silent on taxonomic methodology and philosophy until 1987 when he participated in a symposium at the 14th International Botanical Congress in Berlin (Johnson 1989). In his 'Rainbow's End,' Johnson presented a highly sophisticated view of systematics, not to return to these topics again for almost two decades. Why the long hiatus? Johnson (1989: 95) himself explains:

Discussions of methodology in science tend to be much more prolix and less profitable than those of what we might loosely call fact, or of theory as it relates to the linking of facts and extracting generalizations about their relations to each other, whether in a causal framework or simply a correlative one. Far too much time, I believe, has already been spent on the methodology of taxonomy and indeed of phylogeny.

Although my own contributions to systematics have been entirely within the context of taxonomic philosophy and methodology, I am forced to agree with Johnson about the relative value of methodological discussions to matters of fact and theory (Hull 1979: 419). As in the case of political systems, many scientific systems may look good in principle, but when it comes to applications, they are disasters. Putting a particular method to work in order to see what the results are is the ultimate test of any methodology, and Johnson was in a position to do just that.

Roughly thirty years have passed since Johnson and I started writing on systematics. Throughout this period we have found ourselves in basic agreement. We both agreed with the numerical taxonomists that systematics practice should be made as explicit, objective, quantitative, and repeatable as possible but were skeptical of what we took to be the overly empirical philosophy underlying numerical taxonomy. Johnson also presented specific criticisms of some of the mathematical techniques suggested by the numerical taxonomists. Finally, when cladistics arose, we had pretty much the same reaction to it. It seemed too 'rigid.' We also were equally put off by the later development of 'pattern' or 'transformed' cladistics, once again because we had our doubts about the extremely empirical philosophy that seemed to underlie it.

From the 20:20 perspective of hindsight, how do our views look today? On which issues did we turn out to be right? On which issues do we now seem to have missed the mark? More importantly, what effects did these two periods of heightened activity in taxonomic philosophy and methodology have on systematics? In the following pages I deal with Johnson's early evaluations of numerical taxonomy as well as his later criticisms of cladistics.

What's in a name?

Robert Sokal, Peter Sneath, Paul Ehrlich and others decided in the late 1950s that systematics, as it had been practiced for generations, was too subjective, impressionistic and downright messy. They decided that taxonomic method had to be made more explicit, quantitative, objective and repeatable. In short, they wanted to eliminate what Simpson (1961) termed the 'art' in systematics. The founders of this school of systematics preferred to call themselves 'numerical taxonomists,' to emphasize the increased role that they saw for computers and various mathematical techniques in systematics, but they also formulated a general theory about proper scientific method, the sort of methodology that they thought necessary if systematics was to become quantitative. In particular they repeatedly warned about allowing a priori speculation to enter into the early stages of classification, and among the most dangerous sorts of a priori speculations were those that concerned the evolutionary process or possible phylogenetic relationships. Systematists should limit themselves to observable features of organisms, at least in the early stages of classification. As a result of their emphasis on phenotypic characters and their anti-theoretic stance, their opponents termed them 'pheneticists' or more accurately 'numerical pheneticists.'

As my use of quotation marks in the preceding discussion indicates, I have some doubt as to exactly what the school of systematics initiated by Sokal and Sneath should be termed — numerical taxonomy, phenetic taxonomy, or numerical phenetics. In the past decade or so, several schools of thought have arisen in the humanities in which concern with language swamps any interest in the non-linguistic world. For these folks, it seems that the term used to refer to the AIDS virus is vastly more important than the development of a vaccine or treatment for this terrible disease. These 'postmodernists' seem to think that we can chat our way to solutions to the world's problems. Once we rework language to eliminate sexist, racist, homophobic, etc. connotations and implications, all will be well. Socially sensitive and politically correct language will rule the day.

Numerical taxonomy and phenetics

As can easily be inferred from my characterization of these various schools of thought in the humanities, I am not especially taken with them. Even so, need I convince systematists that names do make a difference? For example, Sneath (1995: 281) sees an 'ambiguity in the meaning of the term "numerical taxonomist" between the original broad sense of those who use any quantitative computer methods and the narrow sense of a discernable group of systematists who practice numerical phenetics.' If numerical taxonomy is defined in terms of the use of quantitative techniques, then it has been extremely successful, because the taxonomic literature is now filled with algorithms and computer programs. However, even though Camin and Sokal (1965) published one of the earliest papers on using computers to infer phylogeny, numerical cladistics is not seen as a branch of numerical taxonomy. Instead, fair or not, it is viewed by most systematists as a branch of cladistics. Numerical taxonomists would

like the huge literature on numerical cladistics to count as part of numerical taxonomy, but thus far they have not been very successful in this regard.

Although many systematists objected to the 'numerical' aspect of numerical taxonomy just because it was numerical, Johnson did not. He raised objections to certain of these techniques but did not reject the basic goal of making systematics as quantitative as possible. He merely concluded that taxonomic evaluations could not be made entirely quantitative. In his own taxonomic work, Johnson used the comparative method but not in a formal, mathematically rigorous way. As he remarked in 1972, he himself had not used taximetric methods because the kind and quantity of data needed for such studies were not available. However, he hoped that 'within the next decade it will indeed be feasible to carry out taximetric analysis both by variable-strategy phenetic techniques and by the use of phyletic (cladistic and perhaps patristic) models' (Johnson 1972: 12).

Because Johnson clearly understood the quantitative techniques being devised, such numerical taxonomists as Sokal and Sneath took his criticisms seriously. Johnson was not a mathematical Luddite. In fact, Sokal and Sneath would have very much liked to persuade Johnson to join with them in their efforts to improve taxonomic principles. They failed. Johnson joined no 'school' of taxonomy but tended to his Eucalypts.

Certainly one of the goals of numerical taxonomy was to make systematics as quantitative as possible, but this group of systematists also set out a basic philosophy that sounded quite hostile to what they termed 'a priori speculation.' As Sokal and Sneath (1963: 55) put this position:

A basic (and very controversial) attitude of the proponents of numerical taxonomy is the strict separation of phylogenetic speculation from taxonomic procedure. Taxonomic relationships between taxa are to be evaluated purely on the basis of the resemblances existing *now* in the material at hand. The relationships are thus *static* ... or *phenetic*, as we now prefer to call them.

At the time, Sokal and Sneath were interpreted as opposing a priori speculation entering into the classificatory process, at least in the early stages of classification. The unequal weighting of characters because of their presumed phylogenetic importance was their usual target (Sokal & Sneath 1963: 16, 34, 50), but they also questioned 'speculation on phyletic relationships based on neontological evidence ... since there is no way of being certain which embryonic features do and which do not reflect that actual phylogeny' (Sokal & Sneath 1963: 24).

Later Sneath and Sokal (1973: 6, 23) repeated their objections to a priori weighting but added that 'phenetic similarity can be based on equally or unequally weighted characters as long as the operation for obtaining the similarity has been defined explicitly by the investigator' (Sneath & Sokal 1973: 29). They did not object to weighting just so long as it was done in an explicit, testable, quantitative way. In a recent retrospective evaluation of numerical taxonomy, Sneath (1995: 285) finds it 'perverse to to imply [that] phenetics is theory free or that phylogeny requires no models of evolution (points that have regrettably been misunderstood by philosophers of science)' — including this one.

I must plead guilty. I really did think that pheneticists objected to letting theoretical speculations about such things as the connection between ontogeny and phylogeny enter into classification, especially in the early steps. At least, I shared this misinterpretation with Johnson. When Johnson (1989: 96) returned to taxonomic philosophy, he remarked that the 'pure phenetic approach to taxonomy is now moribund, or indeed to many of us, quite dead.' If Sneath is right, phenetic taxonomy was never born in the first place. No one ever held any of the beliefs usually attributed

to the pheneticists. All that pheneticists ever intended was that systematists should pay attention to data and not assume that they know which taxa are phylogenetically related to which prior to running the necessary studies. In no way can systematics be theory free even in the earliest stages. On this interpretation, phenetics is anything but controversial.

However, several participants at a conference I attended on the species concept in Cardiff, Wales in the spring of 1995 seemed to be as hostile as ever to theories entering into the early stages of classification — theories of any kind. In Great Britain, at least, phenetics in the sense of theory-free classifications still has some enthusiastic supporters. If these systematists represent a larger group, phenetics is far from dead. I also feel compelled to note that, after a few initial polite responses to the papers being presented, the proceedings became more lively, reminiscent of the good old days when the New York code of conduct prevailed.

Cladistics and cladistics

To complicate matters further, one branch of cladists has been interpreted as sharing the anti-theoretic stance frequently attributed to the pheneticists, a group that is commonly termed 'pattern cladists.' Johnson (1989: 95) dismisses transformed cladistics as 'meaningless.' Once again, terminology is controversial. Initially cladists did not want to be termed 'cladists' (Nelson 1971), and later 'pattern cladists' objected to being singled out to be grouped together under a new name. However, several of Hennig's most productive and original descendants reject his method of reciprocal illumination and argue, as did ideal morphologists and pheneticists before them, that systematists must begin with observations and nothing but observations. Any intrusion of theoretical speculation, especially speculation about phylogeny or the evolutionary process, in the early stages of classification is hopelessly 'unscientific.' The abstract of Platnick's (1979: 537) 'Philosophy and the Transformation of Cladistics' serves as a good early summary of this view:

Although Hennig presented cladistic methods by referring to a model of the evolutionary process, neither the value nor the success of the methods is limited by the value or success of that evolutionary model. Dichotomous cladograms can be preferred simply on the basis of their maximal information content, without reference to speciation mechanisms. Because only the interrelationships of diagnosable taxa (those with unique sets of apomorphic characters) can be investigated, questions about whether speciation can occur without branching, or whether species become extinct at branching points, are irrelevant to cladistic practice. The distinction between plesiomorphic and apomorphic character states depends not on the reconstruction of actual evolutionary history, but on the discrimination of more general from less general characters; groups based on plesiomorphy are defined by the absence of characters and are therefore artificial. Hence, cladistic methods are not the methods of phylogenetics *per se*, but the methods of natural classification in general; phylogenetic conclusions are an extrapolation from hypotheses about natural order.

As Platnick sees it, cladistics does not consist in the methods necessary to produce phylogenetic classifications of the sort that Hennig had in mind but are the methods of natural classification in general. However, as wrong as Hennig was on so many points, Platnick (1979: 538) insists that cladistics has not been transformed because Hennig's 'methods for analyzing data and constructing classifications from them, remain essentially unchanged,' and Hennig's methods are the essence of his

phylogenetics. Without realizing it, Hennig had set out the general principles of natural classification. Cladistic methods discover timeless, static patterns, and patterns are necessarily prior to the processes that produce them. 'Pattern analysis is, in its own right, both primary and independent of theories of process, and is a necessary prerequisite to any analysis of process' (Nelson and Platnick 1981: 35).

Although Hennig did not distinguish clearly and consistently between cladograms and trees, Nelson and Platnick (1981: 141–42) do. Cladograms have no necessary connection to evolution. Cladograms represent patterns within patterns, and this concept of patterns within patterns is an 'empirical generalization largely independent of evolutionary theory, but, of course, compatible with, and interpretable with reference to, evolutionary theory. The concept rests on the same empirical basis as all other taxonomic systems (the observed similarities and differences of organisms).'

Nelson and Platnick (1981: 142) go on to note that two of the basic elements of cladistic analysis (relationship and monophyly) are 'definable only with reference to the branching diagram, and carry no necessary evolutionary connotation.' However, the concept of patterns within patterns is 'not wholly independent of evolutionary theory' because its third element (synapomorphy) is commonly interpreted only in connection with evolution. But, if 'synapomorphy' is defined 'purely as an element of pattern — a unit of resolution, so to speak,' then cladistics becomes the 'general theory of taxonomy of whatever sort' (see also Nelson & Platnick 1981: 165). Parallel observations hold for the term 'homology' as well (Nelson & Platnick 1981: 159).

Thus, 'cladistics' in the broad sense concerns the general methods of discovering patterns within patterns, while 'cladistics' in the narrow sense is limited to the study of those patterns that arise through the evolutionary process. With respect to cladistics in the broad sense, Nelson and Platnick (1981: 324) note that some persons 'may think it strange to use words beginning with "clado-" in a sense divorced from evolution and phylogeny' but go on to argue that such usage is justified given the etymology of the term.

Reference to 'cladistics' in a broad and narrow sense may sound reminiscent of Sneath's (1995: 281) distinction between 'numerical taxonomist' in the broad and narrow sense, as it should. 'Numerical taxonomy' refers both to the use of quantitative techniques in systematics and to numerical phenetics. 'Cladistics' refers both to the general principles of the recognition of patterns within patterns and more narrowly to the applications of these principles to phylogeny.

As in the case of pheneticists, cladists reject the view that they ever held that classification can be or should be theory free or theory neutral. From the start, Nelson and Platnick (1981: 301) acknowledged that a character is a 'theory that two attributes which appear different in some way are nonetheless the same (homologous). As such, a character is not empirically observable, and the hope of pheneticists to reduce taxonomy to mere empirical observations seems futile' (see also Platnick 1979: 542, 1985: 88).

By now, it should be clear that the notion of theory-free classification is far from clear. Both the pheneticists and the pattern cladists claim never to have held the anti-theoretic views attributed to them. On this score, Johnson and I are both equally guilty because both of us thought that pheneticists and pattern cladists had something against letting 'theories' enter into the classificatory process, especially in the early stages. As it turns out, the issue is merely which theories are to enter into the classificatory process right from the start. As Platnick (1985: 88) sees it, 'phenetics is no more theory-free than is cladistics — it's just based on a different theory' (for more recent views on pattern cladism, see Grande & Rieppel 1994).

In 'Rainbow's End' Johnson paid equal attention to both the numerical and the phenetic aspects of 'numerical phenetics.' In this paper I discuss only his philosophical objections. This decision on my part tends to work against the numerical taxonomists, making them look much less successful than they actually were, but I am a philosopher, not a mathematician. Any evaluations I might make of Johnson's mathematical criticisms of numerical taxonomy would be derivative at best. In his later comments, Johnson (1989) criticized both the particular computer programs devised by cladists and the philosophical stance known as pattern cladism. Once again, I am forced to limit myself to his philosophical criticisms. Although I went to IBM school in 1955, my computer skills are decidedly below those of the average teenager today. To complicate matters further, no one seems to have ever held the philosophical views that Johnson and I once criticized. Were we as confused as later authors claim?

Overall similarity and general purpose classifications

The main message of Johnson's presidential address was that there is no one correct classification of plants and animals — no crock of gold at the end of the rainbow — and the search for such a classification is futile and misguided. One of the chief errors made by early pheneticists was believing that something properly termed a 'general purpose' classification reflecting 'overall similarity' was possible. Johnson (1968: 11) traces this belief to two philosophical views — operationism and British empiricism:

The background to this way of thought is the 'operational' approach of logical positivism, a more far-reaching anti-metaphysical philosophy than empiricism but, like empiricism, of obvious appeal to the scientific mind.

According to operationism all scientific terms, even the most theoretical, are to be defined totally and exclusively by the operations used in their application; e.g., the general concept of length in physics is to be defined in terms of meter sticks, light triangulation, etc. Empiricists want to ground all knowledge in observations. Not only must we *begin* all scientific investigations with observations and nothing but observations, but also all scientific knowledge must be *justified* in terms of observations and nothing but observations. The appeal of both of these tenets to scientists is clear. As Johnson sees it, scientists really do need explicitly stated ways of applying their concepts, and observations do play a necessary and important role in science. The issue is the 'nothing-but' interpretation of these tenets.

Johnson zeroes in on the correlative notions of overall similarity and general-purpose classification as one of the weakest parts of phenetic taxonomy. He argues at great length and with considerable skill that both notions are 'metaphysical' in the sense that neither can be operationally defined in even the weakest sense. Nothing out there in nature answers to the name 'overall similarity.' Any set of objects can be described in indefinitely many ways. Although we can limit ourselves to certain attributes and regularities if we so choose, such choices are inherently arbitrary. In short, *'there is still no parametric value of similarity'* (Johnson 1968: 22; see also 1989: 96). As a result, the notion of a general purpose classification is a metaphysical delusion (Ehrlich & Ehrlich 1967, Ghiselin 1969).

For a group of scientists who were so concerned to be as hard-nosed and observation-based as possible, the existence of such a clearly metaphysical notion as overall similarity at the heart of phenetic taxonomy is disconcerting. Pheneticists expressed extreme doubts about our ability to reconstruct phylogeny. For most groups we have little if any fossil evidence, including the Proteaceae as Johnson and Briggs (1963: 22–26, 1975: 94) readily admit. Even in those cases in which we have a reasonably rich fossil record, no unique ordering into phylogenetic trees is possible. Numerous alternative

trees are always possible, and to make matters worse, given any one tree, several alternative ways can be found to classify the groups that make it up.

These worries about phylogeny are genuine, but certainly no more genuine than worries about discerning overall similarity (or affinity) and reflecting it in a general-purpose classification. If anything, reconstructing phylogeny is *more* operational than determining overall similarity. As Johnson (1968: 24) put this point:

We may not *know* the details of phylogeny but (unless we reject biological evolution) we must accept that they *exist uniquely* in space and time, and therefore form a concrete basis for concepts of phylogenetic relationship, however defined. In contrast, the notion of 'affinity' is subject to unlimited variation and any claim for a firm basis for it must be metaphysical.

To be sure, reconstructing phylogenies is extremely difficult, but at least the goal is clear. In the case of overall similarity, we do not even know what it is we are attempting to approximate, even if it does flit like a ghost before the mind's eye.

Looking back over thirty years at Johnson's critique of overall similarity and general-purpose classifications, I find his objections to be as well taken now as they were then. Although some numerical taxonomists came to see problems with these notions quite early, Sokal (1985: 5, 16; 1986: 423) and Sneath (1995: 284) continue to view overall similarity and general-purpose classifications as part of the fundamentals of numerical taxonomy, albeit in a somewhat modified form. As Sneath (1995: 284) observes:

The power of overall similarity measures to construct taxonomic groups, to determine evolutionary relationships, and for identification has been amply borne out, *even if somewhat different forms of similarity may be needed for different purposes* (emphasis added; see also Sneath & Sokal 1973: 28, 107, 109).

From the preceding quotation, one might conclude that Sneath has abandoned the notion of a single measure of overall similarity for numerous special-purpose measures of similarity, a position in keeping with an abandonment of the notion of a general-purpose classification for numerous special-purpose classifications (Ehrlich & Ehrlich 1967).

Vicious circles and reciprocal illumination

As Johnson (1968: 11) points out, any scientific inquiry must either begin from unsubstantiated first principles or be led into an infinite regress. The whole point of axiomatization is to formulate first principles by which all the rest of the system can be derived, and these first principles receive no substantiation within this system. If they are to be justified at all, these justifications must come from the outside. But this insoluble problem arises in the context of the rational reconstruction of science. Scientific *practice* is quite a different matter. Johnson agrees with Popper and Hennig that there is no one place where all scientists can and must begin. Scientists begin wherever they happen to find themselves and then proceed in a sort of feedback process in which improvements in one area lead to improvements in another area, and another area, and so on until this process reflects back on the original contribution, leading to an improvement in it.

For example, systematists working on a particular group may have extensive data on Recent species but little fossil evidence. Some of these systematists may decide to delve into the fossil record to see if they can find anything to supplement their knowledge of present-day species. Others might try out different computer programs to see what happens. Several different programs may focus on the same problem; e.g.,

the patterns produced by one suite of characters are consistently different from the patterns discernible using all the other characters. Hence, they might look much harder at the discordant characters to see if they might have been individuated incorrectly. Embryological investigations might show that these suspect characters have been misidentified. Probably the most difficult aspect of cladistic analysis is to distinguish primitive from derived characters and to organize them into transformation series.

The trouble with the views of scientific method set out by Popper (conjectures and refutations) and Hennig (reciprocal illumination) is that they portray scientific method as being messy. There is no one preferred way to begin scientific investigations. You start out wherever you happen to be and try one damned thing after another. If one line of research doesn't work out, try some other avenue. It may turn out to be a dead end too, or you might succeed in organizing a large area of science into a single system. In short, there is no such thing as *the* scientific method. As Johnson (1972: 12) characterized his own method:

Since we accept, as indeed most pure pheneticists do, that the characters of organisms which are important to them, and to us, are determined largely by their evolutionary history, we inevitably become involved in partial circularity of argument if we base our classifications themselves to some degree on phylogenetic considerations and interpretations. The building of such partially phylogenetic classifications involves some positive feedback from conclusions to argument, and hence incurs stern disapproval from those who seem to think that scientific investigation and interpretation should depend on simple elementary logic.

Numerical taxonomists were not the first to see that messy lines of inference run the danger of becoming circular. However, the only alternative is to have a Cartesian 'straight line' methodology in which first one does A, totally and perfectly, and then one does B, just as totally and perfectly, and so on. No errors ever get introduced into science. Hence, none ever have to be eliminated. Although Johnson admits that sometimes systematists have lapsed into arguing in circles, he insists that traditional phylogenetic methods themselves are not inherently circular.

Back in 1968 I agreed with Johnson, and I still do. But as I mentioned previously, systematists such as the pheneticists and pattern cladists, who seemed to have been arguing for straight-line scientific methods that exclude all theoretical speculations in the early stages of classification, now claim that they have been seriously misinterpreted. Although Sokal and Sneath (1963: 22) rejected Hennig's method of reciprocal illumination as one more example of the 'much-condemned vertical construction of hypothesis upon hypothesis,' later they set out a method that looks very much like reciprocal illumination:

It should be clear that generalizations about the taxa cannot be made before one has recognized the taxa; that taxa cannot be recognized before the resemblances between organisms are known; and that these resemblances cannot be estimated before organisms and their characters have been examined (Sneath & Sokal 1973: 5).

Thus, systematists must start with examining organisms and their characters, then proceed to estimating resemblances between organisms, then to the recognition of taxa, and finally to generalizations about taxa. This procedure looks very much like a straight line methodology, but Sneath and Sokal (1973: 5) go on to acknowledge that 'some of these steps may be in effect combined in certain computational methods, or the whole procedure may be repeated a second time for some special reason.' However, whenever such repetition occurs, the 'order of the steps within the procedure cannot be changed without destroying the rationale of the classificatory process' (Sneath & Sokal 1973: 5).

I am afraid that I am at a loss to figure out what Sneath and Sokal intend. If a systematist has made generalizations about taxa at the end of his first run through of the steps that Sneath and Sokal list, is he allowed to use these generalizations when he returns to step one? If so, then this method looks like the much condemned vertical construction of hypothesis upon hypothesis. If not, then all the systematist is doing is junking the results of his first run through and starting all over again from scratch.

Pattern cladists took as part of their inspiration the writings of Karl Popper. (I suspect that *Systematic Zoology* is the only journal that includes reviews of all of Popper's major works.) One of Popper's fundamental positions is that scientists need not and cannot begin their investigations solely on the basis of theory-free observation statements. Even the most observational of terms include all sorts of theoretical assumptions built into them.

For example, Johnson (1976: 160) lists the sorts of morphological features that he used to discriminate species in the genus *Eucalyptus*. They include:

... more-or-less gross characters of the bark, leaf-shape, opposite and decussate versus 'alternate' (actually pseudoalternate and still essentially decussate) phyllotaxy, shapes and sizes of peduncles and of flower parts (especially the operculum and the anthers) and details of size and shape of the fruits.

Anyone familiar with the history of botany is aware that what counts as bark, leaves, peduncles, opercula, anthers and fruits, not to mention various sorts of phyllotaxy, are not matters of simple observation. No ordinary person could simply look at a tomato and see that it is a fruit. Very complicated arguments and theoretically committed distinctions go into defining the notion of what is or is not a 'fruit.'

Defenders of theory-free classification respond at this point that the characters listed by practicing systematists may well be highly sophisticated, theory-impregnated constructions but that these concepts can be analyzed into absolute simples which are, in the relevant sense, theory free. Ultimately the entire observational basis of science can be replaced by statements such as 'Red spot now.' Of course, terms like 'red' and 'now' depend on physics, but for biological classification such dependence is unproblematic. As seductive as this position may seem, it has proven to be a total failure. In *Der logische Aufbau der Welt*, Rudolf Carnap (1928) proposed to reconstruct the entire world of our experience in terms of absolute observational simples. However, he was unable to produce a single theory-free observation statement that satisfied even himself.

Pattern cladists now claim that they never thought that theory-free observation terms exist, numerous misleading statements notwithstanding. Quite obviously, the view that pheneticists and pattern cladists have been trying to enunciate is extremely subtle, possibly too subtle to be expressed in any natural language. Or just possibly, they have changed their minds on this topic. Because I was taught in my early days in biology that one of the strengths of scientific investigation is that it can force you to change your mind, I continue to be puzzled by how resistant scientists (not to mention philosophers and theologians) are to admitting that an earlier view that they held may well be mistaken, especially if it is fundamental to their entire world view. 'Oops' is not a popular word in scientific publications.

Sokal and Sneath are willing to re-evaluate and even abandon some of their early views (e.g., matches asymptote and non-specificity), but overall-similarity and general purpose classifications are quite another matter. Either they must remain untouched or else be transformed surreptitiously. Cladists are willing to abandon several of Hennig's principles but his basic methodology of three-taxa statements remains inviolate. The 'theories' that underlie cladistics are not theories about what is or is not a *fruit*, but what is or is not a *character*.

Operationism revisited

Johnson agrees that scientific concepts must be made as operational as possible. Operationism as a philosophy of meaning can be shown to be impoverished (Hull 1968). It cannot begin to do what it is supposed to do, even in those areas of science where it arose — relativity theory and behavioral psychology. If operationism won't work in any other area of science, what makes biologists think that it can work in biology? However, a more limited sense of operationism is absolutely central to science. In looking back on my own work, I am most uneasy about my blanket rejection of operationism. Sure, as a general philosophy of meaning, it won't do, but did I really think that the pheneticists were attempting to produce a general theory of meaning? If philosophers really want to contribute to our understanding of science, we might use our time more wisely by studying how scientists actually do operationalize their concepts rather than working out the logical consequences of operationism as a philosophy of meaning. The trouble is that very little can be said about this process that is in the least bit *general*, and philosophers are concerned to uncover *generalizations* about science.

Johnson is in a much stronger position than philosophers in this respect. He rejected operationism as a philosophical program but could exemplify in his own work the ways in which scientific terms can be operationalized. For example, Johnson (1972: 18) coined the term 'umbellaster' to refer to the basic unit of inflorescence in Eucalypts. He then went on to explain how to recognize such units, in part through a diagram. Does it follow that all attempts to operationalize scientific concepts must include diagrams? Obviously not. The problem with saying anything general about how scientists operationalize their concepts is that this process is highly particularized. You can list example after example, and that is about it. Perhaps the methods that scientists use to operationalize their concepts exhibit some interesting regularities. So far I have not been able to discern any.

Did the pheneticists really think from the start that all systems of classification, including their own, depend on various theories, including theories of homology, theories of information-rich groupings, theories of general causes, and theories of history (Sneath 1995: 285)? I think not. At the very least, their repeated protests against 'theories' and 'speculation' served as a rallying cry for the pheneticists. 'Your work rests on the shifting sands of theories and speculations, while our work rests solidly on observable fact.' If the pheneticists had from the start emphasized that their views differed from those of evolutionary systematists only in degree, not in kind, they are not likely to have caught the imagination of young systematists. Cladists carry on in this time-honored tradition.

Conclusion

As sceptical as Johnson is about the long-term usefulness of all the methodological and philosophical discussions that have been carried on primarily in the pages of *Systematic Biology* (formerly *Systematic Zoology*), I would like to say a few words in their defense. For one thing, they have shown that some ideas that initially looked very appealing just don't pan out. For example, the notion that out there in nature there is something termed 'overall similarity' has repeatedly seemed quite plausible to a lot of people, including scientists. If the numerical taxonomists have done anything, they have shown that this belief is illusory. Several of us have argued against this position, but the actual failure of so many bright, hardworking systematists to deliver on the promise of a general-purpose classification based on overall similarity is even more convincing. If after forty years, they have not been able to do it, then probably it cannot be done.

Generations of systematists have also thought that a fairly direct and unproblematic relation exists between phylogenetic trees and hierarchical classifications. The problem was only in reconstructing the tree. Given the tree, dividing it up into indented taxa is just busywork. I think all of the back and forth over cladism has finally disabused systematists of this appealing, though mistaken view. Couldn't this literature have been made more efficient and succinct? To be sure, early confusions of trees with cladograms might well have been avoided, but science in general is not an extremely efficient process. All the false starts and dead ends are intrinsic to science. We all tend to think that at least *we* know what *we* mean, but time and again we don't really understand what we said until we discover what others thought we said. Science does not consist in the definitive pronouncements of infallible authorities but numerous conversations among extremely fallible human beings.

Time and again systematists have found the notion of a theory-free classification appealing. The claim that systematists can start only where they happen to be sounds too haphazard and perilous. There *must* be some one place where all systematists *must* begin. What better place to begin than observations? Observations as the one and only starting place for systematics looks attractive because observations are closest to the world that we are trying to classify, but as close as observation statements are to the real world, they can be mistaken. Mistaken observations can introduce error into science just as readily as mistaken theories. Although the preceding conclusion does not come through loud and clear in the taxonomic literature, it does seem to be one of the messages to be learned from the disputes that have taken place in systematics over the past thirty years.

In retrospect, on what issues were Johnson and I right — or at least on the winning side? Both Johnson and I objected to the notions of overall similarity and general-purpose classifications based on overall similarity. Today numerical taxonomists no longer advocate such a view, although they play down the importance of any changes they may have made on this score. Both Johnson and I interpreted the numerical taxonomists as advocating a phenetic philosophy which precluded 'theories' entering in the classificatory process, especially in the early stages. We both argued against such a notion. In addition, we both recognized a distinct group of cladists termed pattern cladists who we thought held views similar to those of the pheneticists with respect to the role of theories in classification.

In the mean time, both numerical taxonomists and cladists have distinguished between several different sorts of 'theories.' Some of these theories can enter into classification from the start (e.g., the theory of the character); others not (assumed phylogenetic relationships). Who was right on this score? The issues are so tangled that I cannot say. Perhaps the apparent misinterpretations of phenetics and pattern cladists helped to point out places where greater clarity was needed. Perhaps such criticisms introduced unnecessary confusions. Who knows?

With respect to operationism I think that I was not sufficiently sensitive to the issues involved. As a philosophical thesis, I think that operationism won't do, but as a call to connect our theoretical claims with some sort of observational base, I think that it is more than appropriate; it is necessary. Perhaps the meaning of theoretical terms cannot be totally cashed out in terms of observational consequences, but the more the better.

Johnson and I had similar reactions to cladism in its early days. It looked too 'rigid.' We both preferred more 'reasonable' classifications that result from a mixture of various relations. As time has marched on, we are increasingly in the minority. Both numerical taxonomists and cladists urged that systematists pick *one* relation to express in their classifications and apply it *consistently*. Better one relation expressed clearly and unambiguously than several relations expressed only impressionistically.

The chief difference between the numerical taxonomists and cladists on this score is that the numerical taxonomists wanted to represent overall similarity in their classifications while the cladists wanted to represent cladistic relations. As it turns out, overall similarity disappeared under analysis, while the notion of cladistic relations became increasingly clear.

Johnson (1989: 103) distinguishes between two scientific attitudes:

Some scientists are analysts, strongly influenced by recent philosophies of science and concerned to demonstrate their purity of method, however inadequate the method may be in its coverage of the phenomena of nature. Others are synthesists, less concerned with rigour or the appearance of it, but certainly not less concerned with truth. The latter are interested in forming a picture of what really happens, or happened, in the light of all reasonably reliable evidence that they can bring to bear.

Both sorts of scientists are needed if science is to progress. Johnson (1968: 41) clearly sees himself as of the second sort.

Acknowledgement

Thanks are owed to an anonymous referee who read and commented on this paper.

References

- Camin, J.H. & Sokal, R.R. (1968) A method for deducing branching sequences in phylogeny. *Evolution* 19: 311–326.
- Carnap, R. (1928) *Der logische Aufbau der Welt* (Meiner: Leipzig).
- Ehrlich, P.E. & Ehrlich, A.H. (1967) The phenetic relationships of the butterflies. *Systematic Zoology* 16: 301–317.
- Ghiselin, M.T. (1969) The principles and concepts of Systematic Biology. Pp. 45–55, in C.G.S. Sibley (ed.), *Systematic Biology* (National Academy of Sciences: Washington).
- Grande, L. & Rieppel, O. (1994) Summary and comments on systematic patterns and evolutionary process. Pp. 227–255, in L. Grande and O. Rieppel (eds.), *Interpreting the Hierarchy of Nature: From Systematic Patterns to Evolutionary Process Theories* (Academic Press: New York).
- Hull, D.L. (1968) The operational imperative: sense and nonsense in operationism. *Systematic Zoology* 17: 438–457.
- Hull, D.L. (1979) The limits of cladism. *Systematic Zoology* 38: 416–440.
- Johnson, L.A.S. (1968) Rainbow's End: the quest for an optimal classification. *Proceedings of the Linnean Society of New South Wales* 93: 8–45.
- Johnson, L.A.S. (1970) Rainbow's End: the quest for an optimal classification. *Systematic Zoology* 19: 203–239.
- Johnson, L.A.S. (1972) Evolution and classification in *Eucalyptus*. *Proceedings of the Linnean Society of New South Wales* 97: 11–29.
- Johnson, L.A.S. (1976) Problems of species and genera in *Eucalyptus* (Myrtaceae). *Plant Syst. Evol.* 125: 155–167.
- Johnson, L.A.S. (1989) Models and reality: doctrine and practicality in classification. *Plant Syst. Evol.* 168: 95–108.
- Johnson, L.A.S. & Briggs, B.G. (1963) Evolution in the Proteaceae. *Australian Journal of Botany* 11: 21–61.
- Johnson, L.A.S. & Briggs, B.G. (1975) On the Proteaceae — the evolution and classification of a southern family. *Botanical Journal of the Linnean Society*: 83–182.
- Nelson, G. (1971) 'Cladism' as a philosophy of classification. *Systematic Zoology* 20: 373–376.
- Nelson, G. & Platnick, N. (1981) *Systematics and Biogeography: Cladistics and Vicariance* (Columbia University Press: New York).

- Platnick, N. (1979) Philosophy and the transformation of cladistics. *Systematic Zoology* 28: 537–546.
- Platnick, N. (1982) Defining characters and evolutionary groups. *Systematic Zoology* 31: 282–284.
- Platnick, N. (1985) Philosophy and the transformation of cladistics revisited. *Cladistics* 1: 87–94.
- Simpson, G.G. (1961) *Principles of Animal Taxonomy* (Columbia University Press: New York).
- Sneath, P.H.A. (1995) Thirty years of numerical taxonomy. *Systematic Biology* 44: 281–298.
- Sneath, P.H.A. & Sokal, R.R. (1973) *Numerical Taxonomy* (W.H. Freeman: San Francisco).
- Sokal, R.R. (1985) The principles of numerical taxonomy: twenty-five years later. Pp. 1–20 in M. Goodfellow, D. Jones & F.G. Priest (eds.), *Computer-Assisted Bacterial Systematics* (Academic Press: London).
- Sokal, R.R. (1986) Phenetic taxonomy: theory and methods. *Annual Review of Ecology and Systematics* 17: 423–442.
- Sokal, R.R. & Sneath, P.H.A. (1963) *Principles of Numerical Taxonomy* (W. H. Freeman: San Francisco).

Manuscript received 23 August 1995
Manuscript accepted 14 December 1995